Designing Mobile Interactions: The continual convergence of form and context

Volume 1

Kjeldskov, Jesper

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

Volume I

Designing mobile interactions

the continual convergence of form and context

Jesper Kjeldskov

Designing mobile interactions

the continual convergence of form and context

Jesper Kjeldskov

Volume I

Designing mobile interactions - the continual convergence of form and context, Volume I

Printed by UniPrint, Aalborg University

Denne afhandling er af Akademisk Råd ved Det Teknisk-Naturvidenskabelige Fakultet ved Aalborg Universitet antaget til forsvar for doktorgraden i naturvidenskab. Forsvaret finder sted på Aalborg Universitet i Auditorium 0.1.95, Selma Lagerlöfs Vej 300, den 29. maj 2013 kl 13.00.

Aalborg den 31 januar, 2013

Eskild Holm Nielsen
Dekan

This thesis has been accepted by the Faculty of Engineering and Science at Aalborg University for public defence in fulfillment of the requirements for the doctoral degree in science The defence will take place at Aalborg University in Auditorium 0.1.95, Selma Lagerlöfs Vej 300, on 29 May 2013 at 1.00 pm.

Aalborg, 31 January, 2013

Eskild Holm Nielsen
Dean

Table of contents

Preface

At the end of 2010 more smartphones were, for the first time, being sold worldwide than personal computers, hailing the coming of the ”post-PC” era. This enormous uptake of mobile computers has had a huge impact on the way we perceive and use these technologies in our work and private spheres. Interactive mobile systems and devices have become functional design objects that we care deeply about the look, feel and experience of, and that we orchestrate in concert with a plethora of other computing technologies in our everyday lives. If such systems and devices are to be successful, they need to be designed to fit into the greater whole, or digital ecosystem, of other devices, systems and services that are part of the contextual richness of the world that surrounds us. This is, in my opinion, not achieved well through traditional methods of user-centered design and usability engineering. Instead, it calls for designerly approaches to interaction design that help us create new desired practice, design for wholes rather than just focussing on the parts, and deal with the often ill-defined and changing goals emerging from the process.

As a particularly important source of inspiration for developing such approaches to interaction design, I have always been fascinated by thoughts and practices in the discipline of architecture, especially the practice of embracing a contextual view and approach. The relationship between interaction design and architecture has been addressed previously by others, but in my opinion, there is still much to be learned from architectural design about how to think about and do interaction design. Looking at the design of mobile interactions as a continual convergence of form and context is an attempt to provide this designerly approach to thinking and doing interaction design, inspired by thoughts and practices in architecture. Through this approach, I seek to explore the view that most activities are unbounded and situated in dynamic contexts, and that the relationship between context and form is therefore a continually changing one requiring that design is inherently cyclic, able to deal with emergent and changing goals, and about construction of context as well as form.

My interest in mobile computing and telecommunication began twenty years ago, around 1992, when I first had the occasional opportunity to borrow a laptop computer and a portable car-phone while on the road with a band. Although they were bulky, heavy and expensive to use, I got the immediate feeling that these kinds of technologies were going to play a major role in our everyday lives in the future. I often spent a lot of time away from home: staying with family, at the library, commuting on public transport; or on the road with bands: waiting at music venues and cafés, backstage, wherever we were lodging that day or weekend. In those days I was not really interested in computers – simply because they did not fit into the context of my everyday life. With the availability of laptops and mobile phones that all changed. Suddenly I found my mobile and nomadic life empowered by technologies that in earlier forms had been of little value to me away from my desk at home. But what really had an impact on my thoughts about the potentials of these new technologies at the time was not just the many ideas about what I could possibly do with them, it was more the realisation that, in essence, my mobile and nomadic life was not unique or unusual. In fact, it was very similar to the way most people spent a lot of their life: away, on the move, in-between, temporarily somewhere. This made me realise that mobile computing and telecommunication would probably be embraced by billions of people, and that this would change the world forever. Still, I underestimated the actual impact.

Over the two decades since, I have watched the evolution and enormous uptake of mobile systems and devices with great interest, enthusiasm and astonishment. I have, like so many others, embraced the potentials of mobile computing and the beginnings of the post PC era, and it has indeed changed my life. From my first mobile phone, the Nokia 2110 in 1995, which could actually fit in a pocket, and that I managed to forward emails to using SMS before texting became a common means of communication. To the Nokia 3110 in 1997, which had a graphical user interface that was actually a pleasure to use. To my first iPod in 2002 that allowed me to free myself from my CD collection and take my favourite music with me around the world. To my first camera phone in 2002, which had terrible picture quality but still captured a lot of moments that I keep in my photo collection today. To my first Apple iPhone in 2008, which completely changed my use of the Internet and became my primary device for managing email overnight. To my first iPad in 2010 that I bought with no particular purpose in mind, but which quickly replaced my use of a laptop at home.

My research interest in mobile interaction design began around 1999/2000 when I first started working with application design for the PalmPilot and with the potentials of Internet access on mobile devices through WAP. At the time, it quickly became clear to me that although mobile technology was maturing at a very fast pace, there were still major obstacles ahead of us in terms of interaction design that had to be addressed before non-specialists would be able to make use of it. Looking for inspiration, I decided to attend the Mobile HCI 2001 workshop in Lille, France. Back then this had not yet been established as a self-contained conference but ran as a small 1-day event as part of the IHM-HCI conference. The workshop was a great motivator for my research. I found every paper presented interesting and inspiring, and I really enjoyed the drive and enthusiasm of this new and vibrant research community. Returning from the conference I was hooked on mobile interaction design and began writing on a submission for the next conference straight away. The resulting paper, “Just-in-place information” became my first publication on mobile interaction design, and the notion of indexical interaction design addressed in it set the foundation for my future collaboration with Steve Howard and Frank Vetere’s group at The University of Melbourne. It also later led to an early career research grant in 2004, and went on to be published in a leading HCI journal.

The Mobile HCI workshop in 2001 also led to an interest in mobile interaction design evaluation, which I began investigating together with Mikael B. Skov and Jan Stage. Together with Jan I also investigated system analysis techniques for mobile software development.

Later in 2001, on Mikael’s suggestion, I visited the Interaction Design Group at The University of Melbourne in Australia for the first time and gave a talk about my ideas for mobile interaction design research. This visit led to a long-lasting (continuing) research collaboration where I was fortunate enough to spend four semesters in Australia over the years that followed, working on projects such as TramMate, Mediating Intimacy, Customers of the Future, eSpective, and Indexing to Situated Interactions, which are all featured in this thesis. It was also during this time that I met Connor Graham, and we compiled our Mobile HCI research methods survey article for the Mobile HCI 2003 conference. It was also while in Melbourne that I met my wife, the lovely Jeni Paay.

Over the years, I have worked particularly closely with Steve, Frank, Mikael, and Jeni. Together with Steve and Frank I investigated the use of mobile technologies for mediating close personal relationships. Steve also spent 6 months in Aalborg in 2005, during which my thinking about interaction research significantly matured, and the groundwork for much of the research reported in this thesis was laid (over several bottles of red wine). Mikael and I went from working with evaluation techniques to focussing more on interaction design and studying the broader use of mobile and pervasive technologies in various domestic settings (over several bottles of beer). Together with Jeni, I developed the concept of indexical interaction design further, and began exploring the use of sketching and other general design techniques in mobile interaction design, drawing on her background in architecture, (over several bottles of champagne). These collaborations have all been essential for the research results presented in this thesis.

Around 2006, Steve and I began discussing the need for mobile interaction design research to cast a wider perspective on the orchestration of multiple devices, rather than just looking at the design and use of individual artefacts in isolation. This interest in a broader perspective led me to accept an offer in 2007 to lead a research group in Sydney, Australia, where we would be able to work in this area. Here I was extremely fortunate to meet and work closely with Kenton O'Hara, who joined the team a few months after me. Working with Kenton led to renewed inspiration and a new level of depth and detail in my work and thinking, which greatly influenced the way I subsequently developed and wrote this thesis, after returning to Aalborg.

The thesis presents the core of my research contributions in the area of mobile interaction design. It consists of a position summary about the design of mobile interactions as a continual convergence of form and context. This is followed by the research methods survey from Mobile HCI 2003 mentioned earlier, and twenty-one other research articles divided into five parts. The first part is about studying and analysing aspects of context relevant for mobile interaction design. The second part is about the process of designing and building interactive mobile systems that are grounded in their context. The third part is about techniques for studying the user experience of mobile interaction design in context. Part four, slightly different in nature from the previous three parts, describes five prototype systems and is about the use of context in the implementation of concrete interactive systems. Finally, part five is about describing and understanding the relationships between interactive mobile systems, users, and their context.

As described, the research presented in the thesis is the result of several collaborations with colleagues and students between 2001 and 2010. Most of the research was developed and carried out at the Department of Computer Science at Aalborg University and the Department of Information Systems at The University of Melbourne. I wish to thank these institutions for their support and, in particular, thank my closest colleagues and co-authors over the last decade: Jeni Paay, Mikael B. Skov, Steve Howard, Jan Stage, Frank Vetere, Connor Graham, Kenton O'Hara, and Sonja Pedell. My collaborations with Mikael B. Skov, Steve Howard, and Jeni Paay have been particularly inspiring and productive, and have greatly informed the thinking leading to this thesis. I also thank, in particular, Lars Mathiassen and Steve Howard for detailed feedback and valuable discussions on early drafts. I also want to thank Erik Frøkjær for ongoing fruitful discussions of my work, and Ellen Christiansen for introducing me to the works of Christopher Alexander. I also want to thank the Masters students in Human-Computer Interaction that have co-written papers presented in the thesis: Rahuvaran Pathmanathan, Claus M. Christensen, Klaus Kjeldsen, Niels Husted, Jacob Nørskov, and Kenneth Pedersen.

The position summary in Chapter 1 was developed and written between early 2010 and 2012. Being back in Aalborg gave me the opportunity to revisit my own research and think about it in a broader perspective while working in a positive, fun and supporting environment. For this I thank my colleagues at the Department of Computer Science, especially the members of the Information Systems group/Centre for Socio-Interactive Design: Ivan Aaen, Peter Axel Nielsen, Jeni Paay, John S. Persson, Jeremy Rose, Mikael B. Skov and Jan Stage, as well as the group's Ph.D. students during this period of time Anders Bruun, Lise T. Heeager, Karsten Jahn, Kenneth Nielsen, Dimitrios Raptis and Henrik Sørensen.

Finally, I thank my wife Jeni for putting up with me, and for giving me the space for thinking and writing needed to produce this thesis.

March 2012 *Jesper Kjeldskov*
Department of Computer Science
Aalborg University
Denmark

Position summary

Chapter 1. Designing mobile interactions

Table of contents

Chapter 1

Designing mobile interactions
- the continual convergence of form and context

Jesper Kjeldskov

Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world

– Albert Einstein

Abstract. This thesis presents a contextual approach to designing contemporary interactive mobile computer systems as integral parts of ubiquitous computing environments. Interactive mobile systems, services and devices have become functional design objects that we care deeply about. Although their look, feel and features impact our everyday lives as we orchestrate them in concert with a plethora of other computing technologies, these artefacts are not well understood or created through traditional methods of user-centred design and usability engineering. Contrary to more traditional IT artefacts, they constitute holistic user experiences of value and pleasure that require careful attention to the variety, complexity and dynamics of their usage. Hence, the design of mobile interactions proposed in this thesis transcends existing approaches by using the ensemble of form and context as its central unit of analysis. As such, it promotes a designerly way of achieving convergence between form and context through a contextually grounded, wholeness sensitive, and continually unfolding process of design.

1. INTRODUCTION

"Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world" (Einstein 1931). In computing and interaction design today, imagination is every bit as important to advance our knowledge and practices as it was to science in the 1930's. Without imagination and creativity we are not able to move beyond how we think and do today, towards the thinking and doing of tomorrow. This is the timeless way of designing, and it is my starting point for looking at the design of mobile interactions. How can we imagine thinking and acting differently in order to enable ourselves to make future generations of interactive computer systems and devices fundamentally better than the ones we have now? Our current landscape of interactive technologies has itself grown out of paradigmatic shifts in the way we thought about computer systems and did systems development in the past. These shifts brought computing into areas like the workplace, the home office, and into the private sphere, and it made computing about things like work support, collaboration, communication, media consumption, and social networking. But where do we go from here? How can we, once again, reach beyond our presently established ways of thinking and doing, and actively advance the design of interactive computer systems of tomorrow? In this thesis I am going to address this question by revisiting my research contributions within the area of interaction design for mobile computer systems.

One of the things that makes mobile computing an interesting topic of research and design is that the area is strongly driven by innovation, characterized by rapidly evolving use, and has enormous market potential and growth. New technologies are constantly being developed, new use domains are constantly being explored, and successful new ideas and applications reach millions of users. In fact at the end of 2010 more smartphones were, for the first time, being sold worldwide than personal computers, with more than 100 million units shipped in the last three months of that year alone. Reflecting this dynamic and rapidly evolving nature of the area, the industrial lead position has been passed on several times within only a decade, from Palm to Nokia to Apple, and possibly soon to Google, and is most likely to be passed on again in the future. This obviously motivates researchers and designers to keep innovating and developing new technology and applications. A primary driver of mobile technology development has been the enormous uptake of interactive systems and devices for work as well as for leisure. Mobile phones have long been something almost everyone owns at least one of and uses extensively for personal purposes and not just for work. With Internet and multimedia-enabled phones such as the Apple iPhone and phones running Google's Android operating system, smart phones have now firmly reached this mass market too and are no longer something exclusively for a small elite of business professionals. The uptake of mobile technology in our work and private spheres has had a huge impact on the way we perceive and use these technologies. They are no longer just computers on batteries. They have become functional design objects, which we care deeply about the look, feel and experience of, and that we juggle in multitude in our everyday lives. Hence, work in the area of mobile computing has rapidly evolved from being strongly an *engineering* profession to being, at least, equally strongly a *design* profession where the contextual user experience of interactive mobile systems and devices, and the digital ecosystems they are forming, is of utmost importance. This presents the field of mobile interaction design with new challenges forcing us to seek beyond user- or technology-centred approaches.

Being a design profession is different from being an engineering profession. It involves careful engineering, but where engineering is about implementing a solid solution to a problem, design is about also understanding and defining that problem in the first place, and exploring a wide range of solutions before choosing which one to implement. This requires different methods and techniques than the ones taught at engineering schools. It requires techniques that sit at the intersection between technology and art – that support the process of exploring a problem and a design space by generating, communicating and reflecting on ideas, and facilitate choosing between multiple paths of possible solutions. Traditional usability engineering and user-centred design approaches do not facilitate such opportunity seeking ideation and elaboration, but are better at supporting decision making for further reduction and specification of a particular solution.

Mobile computing is a relatively new field of research with little more than three decades of history. During its lifetime it has expanded from being primarily technical to now also being about usability, usefulness and user experiences. This has lead to the birth of the vibrant area of *mobile interaction design*[1] at the intersections between, among others, mobile computing, social sciences, human-computer interaction, industrial design, and user experience design. However, the field of mobile interaction design is still young and immature. Growing out of the "Mobile HCI" community of the early 2000s, it has survived infancy and become an acknowledged part of the established research area of computing with a notable presence in mainstream HCI literature and with its own conferences and journals. But it still doesn't have a strong and unified identity. There is no well-defined methodological and theoretical base for the design of mobile interactions, or even a catalogue of best practices, and there are no well-defined goals or benchmarks for good mobile interaction design research. This is not to say that there is not a lot of good mobile interaction design and research taking place. There is indeed. It is, however, rather fragmented, and rather than an organised community it can, like the field of interaction design generally, better be characterised as "being composed of a number of roving tribes who occasionally encounter one another, warily engage, and, finding the engagements stimulating, remain open to other encounters" (Erickson 2006 p. 301). The advantage of this might be a high level of autonomy, but the disadvantage is less than optimal collective accumulation of knowledge and impeding our ability to leap forward in a pace beyond small incremental steps of each individual piece of research.

This thesis is an attempt to respond to these new challenges by suggesting a holistic approach that rethinks and ties together central activities of interaction design into an ongoing process evolving around the central concept of converging form and context. By convergence I simply mean the combination of two or more things that in concert make up something new that is bigger than the sum of the contributing parts. In my understanding of form and context I subscribe to the fundamental view promoted

1) In line with Preece et al. (2002) I use the term mobile interaction design rather than mobile human-computer interaction because I believe it reflects the emerging research community better. Interaction design includes human-computer interaction, as well as other disciplines, but emphasises an overall goal that is more constructive in nature and that acknowledges mobile computing as a design profession.

by Christopher Alexander in his 1964 "Notes on the Synthesis of Form". According to Alexander, design is "an effort to achieve fitness between two entities: the form in question and its context" (Alexander 1964 p. 15). "Form" is the response to a situation, or problem, whereas "context" defines or frames this situation or problem. Hence, in this use of the term, form does not just mean physical shape, but unites *shape, look, function* and *content*. Using this conceptual optic the design of mobile interactions is about considering the ensemble of particular forms (i.e. interactive mobile systems) in relation to their context (i.e. users, technology, settings, activities etc.).

Figure 1. Form and context in the design of mobile interactions

Following Alexander's line of thought, when we deal with the process of design "the real object of discussion is not the form alone, but the *ensemble* comprising the form and its context" (Alexander 1964 p. 16, italics added). Secondly, but often overlooked, the quality as a whole of a designed form-context ensemble can be influenced by changing the form, but also by changes in the *context*. In direct continuation of this, the views promoted in this thesis build on the belief that the design of mobile interactions should embrace the potentials of designing for wholes, rather than individual parts, and that the notion of form-context ensembles provides a suitable higher-level unit of analysis for such transcension of focus and scope beyond user- or technology-centeredness.

The work presented in this thesis aspires to contribute to the accumulation of a theoretical and methodological body of knowledge about mobile interaction design. It is grounded in a series of specific studies involving interaction design for concrete mobile systems and devices, and uses this foundation to 1) shape a holistic perspective on the process of mobile interaction design in which the main activities of studying, analysing, designing and building interactive mobile systems evolves around the central unity of form-context convergence, and 2) provide insight about each of these individual activities of mobile interaction design, and their contribution to the unfolding whole.

This first chapter presents the background and my thinking about the design of mobile interactions and summarizes the work included in the following 22 chapters. As a starting point, in section 2, I trace the history of mobile computing through seven distinct phases, and introduce the discipline of interaction design. This is followed by a description of the most notable existing design approaches within this field. In section 3, I describe the starting point for my own research on the design of mobile interactions in the early 2000s, and outline the five research topics that I have worked with. I then describe two emerging challenges for the design of mobile interactions: 1) transcending beyond the dichotomy of people- or technology- oriented research and design, and 2) widening the scope beyond the individual mobile device and an individual user's interaction with it. In response to these, in section 4, I discuss the need for doing, and thinking about, interaction design in *a designerly way* rather than in a traditional scientific way. I do this by drawing on thoughts in the literature on design research largely from outside the field of interaction design, such as Schön's notions of problem-setting and reflection-in-action (1983), Pepper's notion of contextualism (1942), Pierce's notion of abductive reasoning (1931-58), and Alexander's notion of unfolding wholeness (2002-05), and showing how these thoughts can help enrich the way we think about mobile interaction design. This is followed by a critical discussion of the established user-centred design model, leading to a series of proposed changes. In section 5, I then present and discuss, as an alternative, the design of mobile interactions as a matter of continual convergence of form and context. In section 6, I present an overview of my individual research contributions, and relate these to the central activities, transitions and outcomes of the design of mobile interactions as a continual convergence of form and context. Finally, section 7 summarizes the main conclusions of my work.

2. BACKGROUND

To ground my work on the design of mobile interactions I will first briefly trace the history of mobile computing. The purpose of this is to map out the origins of this field of research and design, show how it is continually evolving, and illustrate the influence of careful and innovative mobile interaction design at different points in time. This is followed by an introduction to the discipline of interaction design and its established design approaches.

2.1. Mobile Computing

Mobile computing is a significant contributor to the pervasiveness of computing resources in modern western civilisation. In concert with the proliferation of stationary and embedded computer technology throughout society, mobile devices such as cell phones and other handheld or wearable computing technologies have created a state of ubiquitous and pervasive computing where we are surrounded by more computational devices than people (Weiser 1991). Enabling us to orchestrate these devices to fit and serve our personal and working lives is a huge challenge for technology developers, and "as a consequence of pervasive computing, *interaction design* is poised to become one of the main liberal arts of the twenty-first century" (McCullough 2004, italics added).

The field of mobile computing has its origin in a fortunate alignment of interests by technologists and consumers. Since the dawn of the computing age there has always been technological aspirations to make computing hardware smaller, and ever since computers became widely accessible there has been a huge interest from consumers in being able to bring them with you (Atkinson 2005). As a result, the history of mobile computing is paved with countless commercially available devices. Most of them had short lifespan and minimal impact, but others significantly pushed the boundaries of engineering and interaction design. It is these devices, and their importance, that I wish to emphasize here.

The history of mobile computing can be divided into a number of eras, or waves, each characterized by a particular technological focus, interaction design trends, and by leading to fundamental changes in the design and use of mobile devices. In my view, the history of mobile computing has, so far, entailed seven particularly important waves. Although not strictly sequential, they provide a good overview of the legacy on which current mobile computing research and design is built.

1. Portability
2. Miniaturization
3. Connectivity
4. Convergence
5. Divergence
6. Apps
7. Digital ecosystems

The era of focus on *Portability* was about reducing the size of hardware to enable the creation of computers that could be physically moved around relatively easily.

Miniaturization was about creating new and significantly smaller mobile form factors that allowed the use of personal mobile devices while on the move. *Connectivity* was about developing devices and applications that allowed users to be online and communicate via wireless data networks while on the move. *Convergence* was about integrating emerging types of digital mobile devices, such as PDAs, mobile phones, music players, cameras, games, etc., into hybrid devices. *Divergence* took an opposite approach to interaction design by promoting information appliances with specialised functionality rather than generalized ones. The latest wave of *apps* is about developing matter and substance for use and consumption on mobile devices, and making access to this fun or functional interactive application content easy and enjoyable. Finally, the emerging wave of *digital ecosystems* is about the larger wholes of pervasive and interrelated technologies that interactive mobile systems are increasingly becoming a part of.

Portability

The first mobile computers, the precursors to present time's laptops, were developed in the late 1970s and early 1980s inspired by the portability of Alan Kay's Dynabook concept from 1968 (Kay 1972). The Dynabook concept was originally thought of as a machine for children, but observant entrepreneurs, such as the founder of GRiD Systems, John Ellenby, quickly realised that the starting point for something that innovative would have to be "the customer with the most money and the most demanding need" (Moggridge 2007).

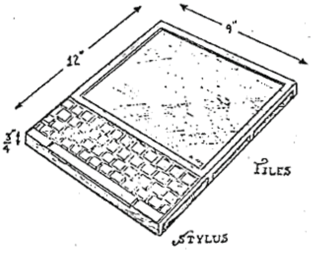


Figure 2. Alan Kay's Dynabook: "a personal computer for children of all ages" (Kay 1972)

The first laptop computer was the GRiD Compass 1101 designed by Bill Moggridge as early as 1981 in response to the design brief of fitting within half the space of a briefcase (Moggridge 2007, Atkinson 2005). The Compass had a 16MHz Intel 8086 processor, 256K DRAM, a 6-inch 320x240 pixel flat screen display, 340kb bubble memory, a 1200 bit/s modem, weighed 5 kg and ran its own graphical operating system called GRiD OS. It was primarily sold to the U.S. government and was, amongst others, used by NASA on Space Shuttle missions during the early 1980s, and in combat. The GRiD Compass featured a stunning forty-three innovative features in its utility patent, including the flat display and hinged screen. The first portable computer to reach real commercial success, however, was the suitcase-style Compaq Portable from 1982, which as the first official IBM clone could run MS-DOS and standard PC programs. In 1988 Grid Systems also developed the first tablet computer, the GRiDpad, initiated and led by Jeff Hawkins who later designed the first PalmPilot and founded Palm Computing.


GRiD Compass 1101 (1981)


Compaq Portable 1 (1982)


GRiDpad 1910 (1989)

Figure 3. Mobile computers in the 1980-90s

In terms of design longevity and impact, Bill Moggridge's work on the first laptop computer and Jeff Hawkins' work on the GRiDpad illustrates the value of careful and well-considered interaction design in mobile computing. The GRiD Compass was superior in terms of its design and performance for a decade. It defined the folding design still used in today's laptops 30 years later, and its basic form factor was not surpassed until the Apple PowerBook 100 introduced the, now standard, clamp-shell design and integrated pointing device in 1991. The basic design of the GRiDpad paved the way for tablet computers and handheld devices such as the Apple Newton, the PalmPilot, and even the iPad over 20 years later.

Miniaturization

By the early 1990s the size of computer hardware had reached a point that allowed radically new and smaller form factors of mobile computers to evolve and emerge on the market. These predominantly handheld devices were labelled palmtop computers, digital organizers, or "Personal Digital Assistants" (PDAs). PDAs differed from laptop PCs by being truly mobile and something that the users could operate while actually moving around physically. They were not thought of as alternatives to desktop or laptop computers but rather as small and lightweight supplemental devices for busy businessmen spending some of their time away from their PC. The first PDA was the Apple Newton from 1992. In 1997 the first PalmPilot was introduced, and in 2000 Compaq released the iPAQ Pocket PC. Whereas the focus of laptop computing was predominantly on portability and mobile access to documents and applications available on desktop computers, palmtop computing introduced an additional focus on applications and interaction styles designed specifically for mobile devices and mobile users.


Apple Newton (1992)


PalmPilot (1997)


Psion 5 (1997)


Compaq iPAQ (2000)

Figure 4. Mobile computers in the 1990-00s

The PDA generation of mobile devices represented a number of distinct interaction design choices and form factors. Most notably they introduced the combination of a relatively small touch-sensitive screen and a separate pen (or stylus) for user interaction. Using the stylus the user could interact with content directly on the screen and enter text via an on-screen keyboard or through handwriting recognition software. Other interaction design innovations included function buttons for accessing pre-defined applications and functions, navigation keys for operating menus, and the "one-click" dock for synchronizing with a stationary computer and for charging. While the Psion series 3 to 5 replicated a "laptop in miniature" design, the Newton, PalmPilot and iPAQ all represented a fundamentally new mobile computing form factor where the majority of the device's surface was used for its display. In terms of interaction design process, the PalmPilot in particular was a product of careful and detailed rethinking of the emerging class of handheld computers; what they should look and feel like, what functions they should perform, and how they should perform them. As an example, the creator of the PalmPilot, Jeff Hawkins, later explained how he carried blocks of wood with him in different sizes and shapes until he had reached the perfect physical form for the device (Bergman and Haitani 2000).

With the emergence of PDAs came also new categories of applications developed specifically for mobile devices and users. The devices each had their own operating systems, optimized for their particular screen sizes and input capabilities, and a suite of standard applications for calendars, contacts, note taking, and email. Adding to this, a wide range of 3rd party applications soon became available for purchase, or, as something new, downloadable via the Internet. By the late 1990s application development specifically for mobile devices was an acknowledged research area and profession, and in 1998 the first international workshop on Human-Computer Interaction with Mobile Devices (Mobile HCI'98) was held in Glasgow specifically addressing the emerging challenge of interaction design and user experiences for mobile devices, systems and services.

Connectivity

The third wave of mobile computing had its origins in wireless telecommunication. As early as 1973 a Motorola team led by Martin Cooper developed and patented a handheld mobile phone concept that led to the first commercial mobile phone small enough to be carried, the DynaTAC 8000X, in 1983.

Figure 5. The first handheld cell phone: Motorola DynaTAC 8000X (1983)

In the 1980s and early 1990s mobile phones were not really considered to be computers. However, with the introduction of the digital GSM mobile phone system in 1991, which also included the Short Message Service (SMS) communication component, the complexity and functionality of handsets began evolving rapidly. So did the uptake of mobile phone technology by the broad population worldwide. This meant that mobile phone developers were suddenly faced with a huge challenge of interaction design not only for making phone calls, but also for handling contacts, calendars, text-based messages, and browsing the Internet. In the late 1990s interaction design for mobile phones was unarguably dominated by the work at Nokia, which led to a series of groundbreaking handsets. The challenges of the time were to design for tiny low-resolution displays and for input capabilities limited to a 12-key numeric keypad along side a small number of function and navigation keys. One of the first mobile phones explicitly resulting from a careful process of interaction design at Nokia in the 1990s was the Nokia 3110. It introduced a simple graphical menu system and the "Navi-key" concept for simplifying user interaction – an interaction design that reached the hands of more than 300 million users through subsequent Nokia handsets (Lindholm 2003). In 1999 the basic interaction design of the Nokia 3110 was extended with T9 predictive text for SMS messaging, games, customisable ring tones, and changeable covers for the extremely successful Nokia 3210.

In the late 1990s, the enormous, and completely unexpected, uptake of the Short Message Service (SMS) inspired attempts to bring the Internet to mobile handsets too. This led to the development of the Wireless Application Protocol (WAP) allowing simplified websites to be viewed on small displays and paving the way for Internet access on mobile devices. The first mobile phone to feature a WAP browser was the Nokia 7110. In response to the need for scrolling through long WAP pages it also featured the first "Navi-roller" thumb wheel. As an interesting example of interaction design, the 7110 also featured a *spring-loaded* cover concealing the keypad, which was inspired by the film The Matrix where the main character uses an earlier Nokia phone modified by the film's production crew to have this functionality. "Life imitating art" (Wilde 1889) you could say. WAP, however, never lived up to its expectations due to slow data transfer and poor usability (Ramsay and Nielsen 2000, Nielsen 2000, Kjeldskov et al. 2002) and was soon superseded by access to the real web on mobile devices. Nevertheless, mobile phone design in the 1990's had a fundamental and lasting impact on the future of mobile computing to come.

Nokia 3110 (1995) Nokia 3210 (1999) Nokia 7110 (1999)

Figure 6. Three mobile interaction design milestones: Navi-key, T9, and WAP

Convergence
One of the most interesting eras of mobile computing began when different types of specialized mobile devices began converging into new types of *hybrid* devices with fundamentally different form factors and interaction designs. The first phase of this was the emergence of “smart phones”, which combined the functionality of a PDA with that of a mobile phone. The development of smart phones involved exploration of a wide range of form factors and interaction designs and led to a series of innovative solutions. Many of these involved designs where the physical shape of the device could be changed depending on what the user wanted to use it for. Other designs, like the Blackberry, introduced a “wide-body mobile phone” form factor with a PDA size display and a miniature QWERTY keyboard in place of the traditional 12-key numeric keypad. The first smart phone that as well as making phone calls could also be used for calendars, addresses, notes, e-mail, fax and games was the IBM Simon from 1992. It had no physical buttons but only a touch screen, which could be operated with a finger or a stylus.

IBM Simon (1992) Nokia 9000 (1996) Ericsson R380 (2000) Blackberry 5810 (2002)

Figure 7. Smartphones exploring different physical form factors and interaction styles

The second phase of convergence combined mobile phones with various rich media capabilities, such as digital cameras, music players, video recording and playback, and television and radio reception. Whereas smart phones were attractive for business professional’s work activities and productivity, multimedia phones were attractive for everyday people’s leisure, fun and socialising.

The most notable example of convergence for leisure was the invention of the camera phone. The first mobile phone to feature a digital camera was the Sharp J-SH04 from 2001. It was only available in Japan through the i-mode mobile Internet service, but the rest of the world soon followed. Two years later more camera phones were sold than digital cameras, and in 2006 half the world’s mobile phones had a built-in

Sharp J-SH04 (2001) Nokia N-Gage (2003) Nokia N90 (2005) Sony Ericsson W600 (2005)

Figure 8. Converged mobile devices: camera-phones, game-phone and walkman-phone

camera – making Nokia the biggest brand of digital cameras and forcing prominent brands such as Minolta and Konica out of the camera business. By 2009 there were more than 1.9 billion camera phones in existence, and mobile phone photography had already had a huge social impact through new ways of capturing and sharing photographs over the Internet (cf. Kindberg et al. 2005, Gye 2007). Whereas early camera phones were clearly phones with cameras, novel interaction design led to several converged devices truly blurring the boundaries between the two (Murphy et al. 2005/chapter 23). As an example, it can be hard to tell if the Nokia N90 is a phone or a camcorder. Another converged functionality to become widely available on mobile phones was the ability to listen to digital music. Most notably Sony re-launched its successful "Walkman" brand of the 1980s in the shape of the converged Sony Ericsson W600 in 2005. With the W44 multimedia phone from 2006, they even went a step further and extended video and music playback with the ability to watch and listen to digital TV and radio. Convergence also led to the creation of hybrid game-phones like the Nokia N-Gage with form factors resembling handheld game consoles.

The fundamental driver behind the trend of convergence is that mobile user experience is proportionally related to the functional scope of interactive mobile devices and systems: "more means more" (Murphy et al. 2005/chapter 23). As a consequence, convergence has often been criticised for generating weak general solutions with usability comparable to the Swiss army knife: clumsy technology with a wide range of functions, none of which are ideal in isolation (see e.g. Norman 1998, Bergman 2000, Buxton 2001). However, in my view the real strength of convergence should not be sought in the simple availability of several functions implemented in the same device. Rather it should be found in the potential creation of something new and hybrid that facilitates use that wasn't possible before, like for example taking pictures and sharing them immediately with your friends, browsing the Internet on your phone, or purchasing music directly on your iPod.

Divergence

Contrasting the convergence approach, the trend of *divergence* suggested a single function/many devices or "information appliance" approach where each device is "designed to perform a specific activity, such as music, photography, or writing" (Bergman 2000). The driving force behind this line of thought is that having a wide range of good specialized tools is better than a general one that does not perform any task particularly well. Specialized tools facilitate optimisation of functionality over time and refinement of well-known paradigms of use. The fundamental view promoted by the trend of divergence is that mobile user experience is *inversely* proportionate to the functional scope of interactive mobile devices and systems: "less is more" (Murphy et al. 2005/chapter 23).

The 2000s saw the emergence a wide range of diverged mobile devices dedicated to do one specific task really well, particularly mobile music players, video players and games. Of course functionally dedicated mobile devices were not a new phenomenon as, for example, early mobile devices such as pocket calculators, cell phones, GPS receivers, digital cameras, and PDAs could unarguably be classified as information appliances too. But what was interesting about the trend of divergence in the early 2000s was that it

Apple iPod (2001) Archos Gmini (2004) Sony PSP (2004) iPod Nano (2010)

Figure 9. Specialised mobile media and gaming devices

was a deliberate interaction design choice and not a technological necessity. Probably the most legendary example of an information appliance was the Apple iPod from 2001. Although not the first mobile digital music player, its interaction design, including the integration with iTunes and later the iTunes Music Store, fundamentally changed global music consumption and purchasing behaviour. Although most mobile phones on the market in the mid 2000s were able to play MP3 files, people still preferred to carry an additional device, the iPod, for playing their music. It provided a better user experience for that particular task, and the device itself had become a popular fashion item. In late 2011 the total number of iPods sold had exceeded 300 million units. Other diverged mobile devices included video players like the Archos Gmini from 2004, the Sony PSP game and video console, and later versions of the iPod extended with video playback capability and multi-touch interaction design.

The interaction design challenge of a diverged mobile device is considerably different from that of a converged one because its functional scope is much narrower. However, as diverged devices are by definition typically used in concert with a plethora of other interactive devices and systems unknown to the designer, there is a huge interaction design challenge in supporting good and flexible integration and "convergence-in-use" (Murphy et al. 2005/chapter 23).

Apps

In June 2007 Apple launched the iPhone. Like many of its contemporaries this was a converged mobile device functioning as a camera phone, portable media player, and Internet client with e-mail, web browsing, and high-speed wireless network connectivity. However, rather than being just another incremental step in the evolution of converged mobile devices, the iPhone represented a significant rethinking of the design of mobile interactions and a series of notable interaction design choices. It featured a large high-resolution capacitive multi-touch display with simple gesture capabilities such as swiping and pinching, and departed completely from the predominant use of physical keys and a stylus for text entry and interaction. Instead of navigating large and deep hierarchies of menus, the user experience was much more fluid and aesthetic, and the phone was both extremely easy and pleasurable to use. The iPhone also featured a number of embedded context sensors, which changed the orientation mode of the display depending on how it was held, as proposed in a UIST conference paper by Hinckley et al. (2000), and changed the mode of the phone application when held close to the face during a call. The later inclusion of GPS and a digital compass extended this "context-awareness" capability to also enable location-based services.

On the software side, the iPhone's web browser actually made it possible to access web content on a mobile device with a positive user experience, and many soon described handling email on the iPhone as favourable compared to its desktop counterparts. Dedicated applications provided direct access to watching video content from YouTube and purchasing music from the iTunes Store. In concert, this meant that people actually started using their mobile device as a *preferred* gateway to the Internet, rather than as a last resort. Consequently iPhone OS dominated the total amount of mobile web traffic worldwide by mid 2009 (Admob 2009). In addition to this, data and media content could be integrated seamlessly with the user's other devices and computers at home or at work through cloud computing services such as MobileMe in a seamless way never seen before in mobile interaction design, illustrating initial steps towards the creation of *digital ecosystems* of mobile and stationary computer systems connected through the Internet.

The iPhone completely redefined mobile computing and set new standards for mobile interaction design and user experiences that other companies, such as Google and HTC, still struggled to match up to 4 years later with the Android mobile operating system and associated online application store. In many ways, the iPhone was the device that mobile interaction design researchers had envisioned for a decade, and its enormous uptake worldwide, with over 310 million iOS enabled devices sold by the end of 2011, confirms that we were indeed right in our speculations about what people would want to do with mobiles – if only we could provide them with a good enough interaction design and user experience. The biggest impact of the iPhone, however, was not only in the interaction design of the device itself and in the high quality of its native applications. As it turned out, it was in the creation of an interaction design that provided users with easy access to a vast and unprecedented amount of *applications* for their mobile device.

In 2008 Apple launched the online "App Store" which provided a mechanism by which iPhone users could easily download, and pay for, third-party application content directly from their mobile device. These Apps span a range of functionalities, including social networking, productivity tools, personal utilities, games, navigation, and advertising for movies and TV shows. For creating this application content, an iPhone OS software development kit (SDK) was released for free along with a business model where Apple handles payments and distribution while leaving App creators with 70% of the profit. By 2012 more than 25 billion Apps had been downloaded from a selection of more than 500.000, making this hugely profitable for both Apple and for the individual third-party creators of particularly popular Apps, which in return has motivated the creation of even more application content. As an indication of the incredible size of this business, third party mobile software developers generated a total income of $2 billion by selling their products through the Apple App store in less than three years. Contrary to developing mobile applications in Java 2 Platform, Micro Edition (J2ME) or Qualcomm's Binary Runtime Environment for Wireless (BREW), developing in iPhone SDK involves no need for customizing applications for a vast range of different handsets, which means that more time can be spent on application design. Also, in sharp contrast to the generally horrific mobile phone user interfaces for installing especially J2ME software, the iPhone provides not only a supply chain and billing model out-of-the-box but also an application shopping user experience that is *positive in itself*. Hence, prior to the iPhone, downloading and installing software onto a mobile phone or PDA was something only technology-savvy people would do. Today this is common practice for millions of users of all ages and computing experience.

As an interesting effect of the iPhone-approach to mobile interaction design, improving the hardware specification of devices was suddenly surpassed in importance in favour of improving the software that is available for them. This is evidenced in the pace and scope of software developments and updates compared to equivalent hardware ones. This is an important shift within the design of mobile interactions, and indicates that a level of stability has been reached in terms of physical form factors and basic input and output capabilities, in favour of a focus on *applications and content*.

Figure 10. The Apple iPhone and iPad (2007 and 2010)

Apple's success with the iPhone lead to a third endeavour within mobile computing, the iPad, which was released in April 2010. Initial media reaction was mixed, but commercial uptake was unprecedented, and the iPad was sold in over 2 million units in its first two months, reaching 15 million units sold by the end of the year. While Microsoft's explicit interaction design approach for PDAs and tablets had long been to replicate their desktop OS (Zuberec 2000), Apple took the opposite approach with the iPad tablet and based it on iPhone OS rather than MacOSX. This was a surprising move for many, admittedly including myself, but had the effect of reinterpreting, and subsequently redefining, the so-far troubled category of "tablet computers" into a new category of *mobile devices,* that are not just laptops without keyboards. Although criticised for being a closed system, the strength of the iPad lay in the user experience created through its meticulous interaction design, which invited the already growing community of iPhone interaction designers and application developers to explore the tablet form factor. Until then, nobody had cared to create web or native application content for tablets (Chen 2010), but with the iPad, tablets suddenly became one of the most interesting and promising mobile platforms on Earth. Consequently, after only one year of existence, there were more than 65.000 Apps available for the iPad.

Digital ecosystems

As we move into the second decade of the new millennium the challenges of mobile computing and interaction design continue to evolve. The technical capabilities of our mobile devices have improved significantly to the point where factors such as screen real estate, input capabilities, processing power, network speed and battery lifetime are much less of an issue than only half a decade ago. At the same time, we have also become rather skilled at designing for relatively small screens and for the different input capabilities of mobile devices. So much that millions of ordinary people are actually able to download and use the applications being developed, and are even willing to pay for some of them. To a large extent, therefore, we have now successfully solved the majority of problems facing mobile interaction researchers and designers in the past. However, as the history of all areas of computing have shown us, it is highly unlikely that we have reached an end point. As in the past, the technology and interaction design we are witnessing today is just the starting point for the continuing evolution of the technology and interaction design of tomorrow. But what are then the challenges and opportunities for the design of mobile interactions to come? What will the next wave of mobile computing be about?

Fuelled by the enormous interest and uptake of "post-PC devices" like smart phones and tablets by the general population, it is not unreasonable to speculate that a major platform shift away from desktop computing is imminent. Mobile devices are becoming more and more important and widespread. They will soon be the dominating point of access to the Internet, and in combination with the growth of cloud computing they will soon dominate peoples' use of computational power. Importantly, what we are witnessing here is not just the development of even smarter smart phones with improved abilities to imitate desktop PCs in miniature. It is a radical evolution of a major computing platform for new applications allowing us to do things that couldn't be done before. This may well be a genuine paradigm shift for mobile computing and mobile interaction design.

In my belief, the next wave of mobile computing and interaction design is going to be about the creation of *digital ecosystems* (Miller et al. 2010) in which mobile computing plays a central role in concert with other ubiquitous computing resources. This challenges us to move beyond considering interactive mobile devices, systems and services as entities that can meaningfully be designed and studied in isolation from the larger use context or artefact ecologies (Jung et al. 2008, Bødker and Klokmose 2011) that they are a part of. Yes, mobile computers, in various forms, play hugely important roles in most peoples' everyday lives, but they are not the only technologies and artefacts we make use of at home or at work, or in the space between. Most people use multiple mobile devices for different purposes, but they also use a multitude of stationary or embedded computer systems, at work, at home, in our cars, or in the city around us. In concert this makes up a rich digital ecosystem of interactive devices, systems and services often referred to as ubiquitous or pervasive computing, in which mobile computing is a central, but not the only, component. The challenge of designing mobile interactions in such ubiquitous and pervasive information societies is to facilitate the creation of interactive devices, systems and services that fit well into this ecosystem of other devices, systems and services, as well as into the rich new use practices, for work and leisure, created by these technologies and their users. Like any other type of ecosystem, understanding, creating and maintaining digital ecosystems requires a holistic perspective on the totality and ecology of the system at play, and not just detailed views on each of its individual components. As argued earlier, it is my position that this cannot be achieved well through a technology- or a user-centred approach, but requires a change of the unit of analysis towards one that continually includes both these viewpoints.

The digital ecology wave of mobile computing will build on the achievements of previous eras within hardware miniaturization, connectivity, new form factors, input devices, interaction styles, applications, convergence, divergence and content, but it will broaden the scope to include the wider context of use and an explicit sensitivity for the contextual factors that influence the user experience. It is going to be about creating interactive devices, systems, and services that respond to the broad and diverse aspects of human life, and that not only provide utility and are easy to use, but also provide pleasure and value, and fit naturally into peoples' complex and dynamic lives of constantly changing settings and situations. My position here is that this will best be achieved through careful attention to the details, richness and dynamics of *form-context ensembles* during all phases of mobile interaction design and system development: from initial domain studies and analysis through creative design, implementation of actual interactive systems, evaluation studies and analysis of their outcomes.

2.2. Interaction design

I will now turn my attention towards the notion of interaction design. The term interaction design was coined by Bill Moggridge and Bill Verplank in the late 1980s, and is about "designing interactive products to support the way people communicate and interact in their everyday and working lives" (Sharp et al. 2007 p. 8), or more broadly about "the design of everything that is both digital and interactive" (Moggridge 2007 p. 660) with particular attention to its subjective and qualitative aspects. In other words, it is about creating life and work enhancing user experiences through the design, development, construction and implementation of interactive products, devices, systems and services.

Today, interactive products are typically computer-based. This means that interaction design is relevant within all disciplines, fields and approaches that concern themselves with research and design of computer-based systems for people. Hence, alongside design practices such as graphic and industrial design, academic disciplines such as psychology and sociology, and multi/interdisciplinary fields such as human-computer interaction and information systems, interaction design also involves the technical academic disciplines of computer science and engineering. However, interaction design differs from each of these practices, disciplines and fields by having a different, overall, focus and purpose. It is concerned with the *totality* of the user experience of interactive products and with all of the factors that may contribute to their successful creation. When we design computer-based interactive systems, we are not just designing how it appears but also how it behaves. We are designing how people and technology interact (Moggridge 2007). As described by Winograd (1997), doing interaction design can in many ways be compared to doing architecture. The architect is concerned with people and their interactions within the building being created. For example, does the space fit the lives or work styles of the family or business that is going to inhabit it? Does the flow within and between rooms work well? Are functionally related spaces in close proximity? And so on. Supporting the work of the architect, engineers are concerned with the structural soundness and construction methods of the building, and knowledge from other disciplines, such as human factors and social sciences may also influence the architect's ability to create functional and liveable spaces. Just like a good architect understands these other relevant disciplines, so does a good interaction designer. However, just like there is a difference between designing and building a house there is also a difference between designing an interactive product and engineering its software (Sharp et al. 2007 p. 9).

This analogy also illustrates how interaction design differs from the fields of human-computer interaction and human factors. Whereas HCI and HF have traditionally focussed on the relatively narrow study of human-machine interaction and the major factors surrounding this in order to optimize user interfaces and overall system performance, interaction design is "about shaping our everyday life through digital artefacts – for work, for play, and for entertainment" (Smith 2007 p. xi). In this way, interaction design "has cast its net much wider, being concerned with the theory, research, and practice of designing user experiences for all manner of technologies, systems and products" (Sharp et al. 2007 p. 10). The goal, and the challenge, of interaction design is to make powerful computing technology fit into the peoples' work and private lives rather than forcing peoples' lives to fit technology. In order to achieve this goal, solutions produced by interaction designers must be appropriate to their context (Smith 2007).

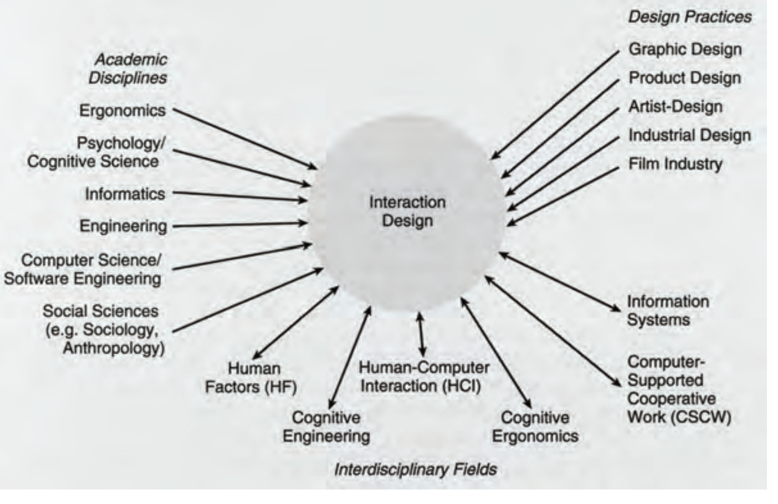


Figure 11. Academic disciplines, design practices and research fields concerned with Interaction Design (Sharp et al. 2007 p. 10).

There is a difference between interaction design *practice* and interaction design *research*. Interaction design practice is about creating concrete interactive systems and solutions to a particular design problem by applying the knowledge and approaches of the discipline. Interaction design research is about creating and improving this knowledge and these approaches, which is often done through studies of interaction design practice. Hence, it can be argued that interaction design is merely a practice and that it is based on the cumulative sum of research in all of its contributing disciplines and research fields, such as human-computer interaction, industrial design, psychology, sociology, computer science, etc. as depicted in figure 11 (Sharp et al. 2007 p. 10). The risk of such a view, however, is that it presents interaction design as a research-less discipline and distributes interaction design research on a large number of individual disciplines relevant for, but external to, interaction design itself. Thinking about disciplinary borders, which I will return to later, this view essentially keeps interaction design within the confinements of multidisciplinarity without enabling the potentials of cross-disciplinary collaboration in the creation of new types of knowledge and methodological approaches. The opposite view, which I share and promote in this thesis, is that interaction design is, or rather *should be,* a discipline not only with its own practice, but also with its own research agenda and challenges, and that interaction design practice and research are both inherently interdisciplinary.

There is also a difference between being an interaction designer and being an interaction design researcher. Not all interaction designers are skilled researchers, and not all interaction design researchers work as interaction designers. However, given the fundamentally creative and design-oriented nature of the discipline of interaction design, it is my fundamental and strong belief that being a good interaction design researcher requires skills and talents as an interaction designer too.

Mobile interaction design

Mobile interaction design is an area of interaction design that is concerned specifically with the creation of user experiences with interactive products, devices, systems and services that are not stationary but that people can take with them. It is enabled by advances in mobile computing – as described earlier – that have allowed designers and system developers to conceive interactive products that are small enough to be carried with us, held in our hands, or even worn, while also providing computational power and network capabilities sufficient enough for enabling useful and attractive interactive systems and services. This includes handheld and wearable devices, PDAs, mobile phones, smart phones, portable digital media players, handheld games, etc. as well as the software applications and services that run on these devices or can be accessed from them. However, mobile interaction design is not only facilitated and driven by advances in computer science and engineering. It is also increasingly advanced by our ability to develop new use practices for mobile computing and to include and appropriate available and emerging mobile computer and network technologies into new and innovative interactive products and solutions. Hence, we have long gone beyond the "anytime anywhere" mobile computing hype of the late 1990s and grown much more sensible aspirations to develop "mobiles that work at the right time, and that know their place – that fit in" (Jones and Marsden 2006).

The challenges of mobile interaction design have changed and evolved over time as new technologies were developed and new use practices emerged. Early mobile interaction design dealt with the physical design of portable computers. This evolved into a focus on input devices and interaction styles suitable for handheld operation and mobile use. For mobile phones the interaction design challenge has primarily been a matter of reducing physical size while optimizing the use of limited display real estate and the standard 12-key numeric keypad for more and more possible applications. With the emergence of functionally hybrid and more complex devices, the interaction design challenge became about developing new forms and shapes of devices as well as developing new types of applications available on them, without making the devices (even) harder to use. For the growing range of functionally dedicated mobile devices like digital cameras and media players the interaction design challenge became about facilitating peoples' "orchestration" of all these devices, and their content, in increasingly complex ecosystems of interactive computer systems and digital data.

Today, the challenge of designing mobile interactions is very much about the development of *software applications*. The physical device form factor appears to have stabilized, for some time at least, on the basic size, shape and interaction capability introduced by the Apple iPhone in 2007, which has remained unchanged for more than five years and been replicated by all major handset producers. This has shifted focus towards downloadable and purchasable third party application content available for these devices, in the form of relatively small "Apps" with highly specialised functionality, designed not only by large software corporations but also by small companies and even individuals, including students. By early 2011 more than 500.000 third party applications were available from the Apple App Store and more than 450.000 were available from Google's Android Market. In less than four years more than 25 billion Apps were downloaded for the iPhone and iPod Touch, with another 10 billion downloaded for Android. These are staggering numbers. However, although a lot of interesting and innovative new mobile applications are appearing in Google's and Apple's online stores every day, and application developers and interaction designers world wide are pushing the boundaries of what mobile computer devices are being used for, the state of current mobile application design can be compared to the state of the web in the mid 1990s[2]. There is a lot of excitement and interest, the development tools are easily accessible, and there is a huge audience of potential users. Exceeding the potentials of the web in the mid 1990s there are even well established digital supply chains and mechanisms for micro-payments. But as with the web 15 years ago, I dont think we have yet seen or understood the significance and scope of the impact that third party application design for mobile devices will have on all aspects of our lives, for work as well as for leisure.

Research impact on practice

Much of the future impact of mobile computing envisioned above will be driven by skilful and creative design of mobile interactions conceived by entrepreneurial developers and designers who understand how to create useful and enjoyable utility and user experience that fits the users needs, desires and contexts of use. Unfortunately, however, the current research-based literature on mobile interaction design does not provide as much foundation as we probably could for these developers and designers to base their innovations and interaction design on, nor much methodological guidance on how

2) In 2006, Jones and Marsden (2006) described the state of mobile application design as being comparable to the state of the web in the early 1990s. With the release of the Apple iPhone and App Store, mobile application design took a huge leap forward, comparable in significance and scope to the effects of the first Netscape Navigator web browser in 1994.

to approach the process. Whereas there are a lot of research-based books about user interface and interaction design for desktop applications and web sites there is not yet a lot of equivalent literature available about mobile interaction design. Although mobile computing has a history of approximately three decades, and interaction design has played an important role throughout about 2/3 of this history, only one good general textbook, by Matt Jones and Gary Marsden (2006), has been published on the topic to date. And although this book is indeed a good starting point for addressing the particular challenges of *mobile* interaction design, it still doesn't have the completeness and depth of equivalent human-computer interaction and interaction design primers such as Laurel (1990), Shneiderman (1998), Preece et al. (1994), Winograd (1996), Raskin (2000), Dix et al. (2004) Benyon et al. (2005) Lauesen (2005), Bagnara and Smith (2006), Preece et al. (2002), and Rogers et al. (2011). This is potentially an opportunity missed for large-scale real-world impact on mobile interaction design *practice* in respect to the massive amount of good interaction design research that has been done within the field over the last decade and a half. While it might indicate that the area of mobile interaction design still hasn't stabilized enough for general guidelines, principles, methods and techniques to evolve, it also demonstrates an opportunity, and a need, to push forward on developing such foundational work further.

Several of the textbooks that *do* exist on aspects of interaction design for mobile devices, systems and services, such as Helal et al. (1998), Weiss (2002) Ballard (2007), Fling (2009) and Frederick and Lal (2010), essentially target application development for particular and very specific classes of devices and software platforms, and address ephemeral technical limitations such as particular operating systems, low screen resolution, reduced processing power, limited memory and poor bandwidth. While unarguably useful when designing for these exact platforms, the weakness of such types of works is that they are almost *too* practical. They are highly vulnerable to technological advances and therefore quickly rendered irrelevant as new devices and platforms emerge. As a consequence they usually end up as short-lived and overly specific user interface guidelines tied to a specific point in time, and not as generally applicable and timeless principles for interaction design. Distilling the essence of these works – the higher-level challenges and solutions that apply beyond specific devices and platforms – would be useful for moving the field of mobile interaction design forward. But such work has not yet been done systematically and in depth.

As a step in the right direction though, a different class of textbooks on mobile interaction design is the collection of case study-like accounts for successful and influential design solutions, such as Eric Bergman's "Information Appliances and Beyond" (2000), Lindholm et al.'s "Mobile Usability: how Nokia changed the face of the mobile phone" (2003), parts of Bill Moggridge's "Designing Interactions" (2007), and Bondo et al's "iPhone User Interface Design Projects" (2009). These writings aim to capture universally important lessons learned from the experience of actual mobile interaction designers. They provide interaction design as well as methodological insight about influential solutions and how they came about. The potential weakness of *these* works, however, is that they easily end up being anecdotal and difficult to transfer into present time's design challenges. To support such transfer and transcendence of knowledge, we must provide not only the case study accounts, but also analysis across these case studies that elevates our learning from the concrete and specific level to the abstract and general. This accumulation of an abstract and general body of knowledge is probably better suited for a design researcher than for a design practitioner, and it is what the work presented in this thesis aspires to contribute to.

Multi- and interdisciplinarity

Interaction design, whether it is mobile or not mobile, is a field of research that involves several disciplines and works across disciplinary boundaries (Sharp et al. 2007 p10). It is widely accepted that research across disciplines is difficult, and consequently, in practice, a lot of such research "actually works at the level of being *multidisciplinary* (or pluridisciplinary): where a group of researchers from different disciplines cooperate by working together on the same problem towards a common goal, but continue to do so using theories, tools, and methods from their own discipline and occasionally using the output from each other's work" (Rogers et al. 2005). Today, mobile interaction design is an example of such a *multidisciplinary* area. This means that it involves a mixture of disciplines, such as computer science, engineering, human factors, psychology, sociology, design, etc. and that these disciplines each contribute to a composite body of knowledge about the design of mobile interactions. Being multidisciplinary means that the challenges of mobile interaction design are approached from different perspectives and with different competences, and that research is therefore diversified and broad. This ensures outcomes that span widely, from new technological endeavours to exploration of new use domains. However, the problem of being multidisciplinary is, by definition, that each discipline retains its own identity, methodologies, assumptions and aims, and that these are not changed or influenced by the other disciplines within the multidisciplinary relationship. Although multidisciplinary research involves several disciplines "each discipline makes as separate contribution" (Moore and Lottridge 2010). This non-integrative mixture of disciplines basically means that there is a shared interest in the topic of mobile computing from within multiple disciplines, but that each of these disciplines treats the common topic of interest in their own way and with their own focus, as defined and guided by their individual school of thought.

The idea of mobile interaction design has appeared in different disciplines at around the same time, but cooperation between these multiple disciplines is largely "mutual and cumulative but not interactive" (Augsburg 2005: 56). This underlying separation between disciplines can be seen in the general observation that although the field is concerned with the same overall topic of *interaction* between people and technology, most specific research is in fact focussed primarily on one or the other. It is either technology- or user-centred, and as pointed out by Rasmussen (2007) when making such clear-cut distinction between technology- or user-centred approaches, a valuable dialectic between the two tends to disappear at the cost of possible synthesis of the two opposing interests and forces at a higher level (Nonaka and Toyama 2002, Dahlbom and Mathiassen 1993). As we observed in an empirical study comparing these two disciplinarily different approaches to the development of two similar interactive mobile systems (Kjeldskov and Howard 2004/chapter 7, Jones and Marsden 2006 pp. 88-89) applying either view in isolation has a notable negative impact on the quality and completeness of the produced outcomes. Producing well functioning and usable mobile interaction design would be supported better by explicitly combining and integrating technology- and user-centred approaches.

In contrast to multidisciplinarity, *interdisciplinarity* crosses the traditional boundaries between disciplines or schools of thought as new challenges, needs and professions emerge, and blends the involved disciplines including their identities, methodologies, assumptions and aims. It *connects* and *integrates* several academic schools of thought or professions in the pursuit of a common task, along with their specific standpoints, and is not just different disciplines pasted together but rather an "integration and synthesis of ideas and methods" (Moore and Lottridge 2010). On the basis of this, interdisciplinary research derives "novel concepts, methods and theoretical frameworks through the melding of concepts, methods and theoretical frameworks coming from different disciplines" (Rogers et al. 2005). Interdisciplinary research areas often emerge from mutual beliefs that traditional disciplines are insufficient for addressing an important topic on their own, or in a simple non-integrative mixture with each other, due to the topic's multi-faceted, or even transdisciplinary, nature where a unity of knowledge is needed across disciplines, or even beyond them. Where multidisciplinarity approaches a problem space using a *coordinated* effort from *distinct* methodological foundations, interdisciplinarity approaches a problem space using an *integrated* effort from *combined* methodological foundations (Blevis and Stolterman 2009). Hence, the main difference between multi- and interdisciplinarity lies in the way research is conducted, and as a consequence of this also in the types of outcomes produced.

The position that I put forward here is that in order to better inform the creation of interactive mobile devices, systems and services in the future, mobile interaction design needs to evolve from being a multidisciplinary research field to becoming an interdisciplinary one. This position echoes the message from Steve Jobs since early 2010 that the key to Apple's position in the "post-PC era" of interactive mobile systems and devices such as the iPhone and iPad is credited to the explicit belief at Apple that technology alone is not enough, but that "it's technology married with liberal arts, married with the humanities, that yields us the result that makes our hearts sing" (Jobs 2011). According to Jobs, this explicit interdisciplinarity sits "at the intersection between technology and liberal arts", where technical sciences facilitate the creation of "extremely advanced products from a technology point of view" (Jobs 2010) and contemporary liberal art disciplines, such as literature, philosophy, history, science, and design facilitate making them "intuitive, easy to use, fun to use, so that they really fit the users - the users don't have to come to them, they come to the user". This interdisciplinarity is what drives Apple's ability to develop new mobile interaction design that repeatedly pushes the boundary of what is technically possible and, at the same time, is almost immediately embraced globally and soon taken for granted by millions of people.

As pointed out by Rogers et al. (2005) it is not problematic to use the terms multi- and inter-disciplinarity interchangeably if simply referring to collaboration between people from different disciplines working on a common problem. However, depending on the underlying rationale for the collaborative activity, the two can have rather different meanings. Put simply, "bringing together a group of experts from different disciplines or professions to contribute to a single project, which would not be able to be accomplished by any one profession alone" is different from when "a group of researchers from distinct disciplines try to generate novel concepts and integrate different levels of explanation" (Rogers et al. 2005). Here, the former denotes multidisciplinarity where each researcher contributes to the project with unique expertise, whereas the latter denotes interdisciplinarity where new research questions are dealt with. Achieving interdisciplinarity can be very difficult. Whereas multidisciplinarity can be done through coordination of research efforts, there are many more obstacles to the "cross-fertilization" of ideas required in interdisciplinarity, including incommensurability of concepts, dissimilar units of analysis, variation in world view, etc. (Rogers et al. 2005). This raises the fundamental question of when interdisciplinarity is needed and, and how it then can be achieved in practice.

Basically, interdisciplinarity is desirable when reaching a point where the constraints of ones own discipline prevents any further significant progress, and researchers are forced to work in the outer periphery of their field and, in doing so, are having to forge new ones (The Royal Society 1996). According to Rogers et al. (2005) there are two types of impetus leading to such circumstances motivating input from several disciplines: cases where "an existing problem has simply become too large for a single discipline to cope with", and cases where "something external to the disciplines has forced itself on their attention". Good examples of these two types of cases are the attempt to develop a comprehensive cognitive science program, which held interdisciplinarity as an ideal, and the evolution of the fields of HCI and CSCW, where applied interdisciplinarity was motivated by technological advances within computing. However, for both two cases the project of forging interdisciplinarity faced significant challenges, and it can be questioned how successful the endeavors were. For cognitive science, a major limitation was that the key issues of the "interdiscipline" could still meaningfully, and sometimes advantageously, be studied within a single existing discipline (cf. Norman 1990). For HCI and CSCW a major limitation has been to break away from the multidisciplinary mindset of simply dividing up the joint challenge into coordinated applied disciplinarity, and to really tie together and make mappings across concepts from different disciplines towards the development of unified theory (Bannon 1992). Hence according to Rogers et al. (2005) "the jury is still out as to whether either HCI or CSCW have in fact been able to achieve any significant level of interdisciplinarity".

Modifying the unit of analysis

Mobile interaction design shares properties with both cases above. As with HCI and CSCW applied interdisciplinarity is motivated by technological advances, and as with cognitive science the ideal of interdisciplinarity is motivated by the fact that the full scope of the field is impossible to grasp from one perspective alone. But how can we then avoid repeating the modest success of interdisciplinarity within these related fields? According to Rogers et al. (2005) a possible key to achieving interdisciplinarity lies in explicitly transcending beyond current disciplinary dogma and units of analysis. "If true interdisciplinarity is ever to take off, then what is needed is a paradigm shift whereby a whole set of new issues and research questions are framed that force new ways of conceptualizing and working" (Rogers et al. 2005). This facilitates what they call “reconceptualizing the domain of interest through using a modified unit of analysis” whereby the scope can be broadened while still allowing the use of existing concepts and theory. As an example, the distributed cognition approach (Hutchins 1995) extended established cognitive science’s focus on properties and processes inside a single person’s mind to a system of cognitive systems involving several actors and their environment studied through “cognitive ethnography”. Although such broadening of scope and change in level of abstraction is difficult and precarious, the benefit is that it has the potential to reveal phenomena that go across, and cannot be reduced to, existing units of analysis.

In relation to Bannon’s (1992) concern that the goal of developing interdisciplinary unified theory within HCI and CSCW is fundamentally flawed due to the inherent incommensurability of theory, concepts, traditions, perspectives etc., the significance of the more modest approach of changing the unit of analysis is that theoretical developments from such endeavours can advance our understanding of the field of interest through extending, adapting and integrating *existing concepts and theories.* This can be seen as advantageous over creating completely new concepts and theories, which entail a risk of unintentionally disregarding legacy of previous accomplishments and achievements within the individual contributing disciplines, and potentially separating from valuable epistemological and methodological inheritance.

Classifying research as either multi- or interdisciplinary can sometimes be overly simplistic, as in reality there is a continuum between the two, “from multidisciplinary work with sharp boundaries between the disciplines at one end to the holistic approach of interdisciplinarity at the other” (The Royal Society 1996). This also means that going straight to complete holistic interdisciplinarity is usually not possible but requires evolution through several stages of involved disciplines becoming increasingly integrated as the shared problem is explored, developed and defined. Redefining the unit of analysis can be seen as a way of stimulating such stepwise evolution towards interdisciplinarity by taking its offset in the knowledge and methodologies of existing disciplines, and aiming at developing better understanding of broader phenomena that are, in essence, new to all of those disciplines, and therefore also suitably peripheral to established belief systems.

In order to support the research field of mobile interaction design evolving from being a multidisciplinary area of research to becoming an interdisciplinary one, a change in the basic unit of analysis towards one with a wider scope that transcends the individual contributing disciplines’ focus on either users or technologies may be called for. Similarly to Hutchins’ (1995) broadening of scope towards “cognition in the wild”, a good candidate for such modified unit of analysis, I argue, is the broader and more holistic phenomena of *form-context convergence*. Supporting this position, I am going to briefly turn my attention towards the role of context in the design of mobile interactions.

The role of context

Since the early days of mobile computing and mobile human-computer interaction, the use contexts of interactive mobile systems and devices have often been highlighted as being particularly important for system developers to “be aware of” and “take into account” when designing and building interactive mobile systems, and when evaluating and studying their use (cf. Johnson 1998, Rodden et al. 1998, Brown et al. 2000). Mobile use contexts have been described as being particularly challenging compared to, for example, the use contexts of traditional stationary office systems due to their highly dynamic, complex, and indeed mobile, nature. It has also often been suggested that when using an interactive mobile computer system, other activities in the surrounding context are often more important than the actual interaction with and use of the system itself – walking down the street, socialising in a bar or café, or attending to a patient in a hospital.

There are many different definitions of context, and the debate on what constitutes context for mobile computing, and what role it plays, is ongoing. Early works within mobile computing referred to context as primarily the location of people and objects (Schilit and Theimer 1994). In more recent works, context has been extended to include a broader collection of factors such as physical and social aspects of an environment (McCullough 2004, Dourish 2004, Bradley and Dunlop 2002, Agre 2001, Dey 2001, Abowd and Mynatt 2000, Schmidt et al. 1999a, Crabtree and Rhodes 1998). Dey (2001) characterizes context as “any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and the application themselves.” Although this definition is quite complete, it is not very specific about what type of information could in fact be used to characterize such a situation. In contrast to this, Schmidt et al. (1999a) present a model of context with two distinct categories: human factors and physical environment. Human factors consist of the three categories: information about the user (profile, emotional state, etc), the user’s social environment (presence of other people, group dynamics, etc), and the user’s tasks (current activity, goals etc.). Physical environment consist of the three categories: location (absolute and relative position, etc.), infrastructure (computational resources, etc.), and physical conditions (noise, light, etc.). This model provides a good catalogue of specific contextual factors to complement broader definitions like the one by Dey (2001). Other works are not as comprehensive in their coverage of different contextual factors but go into detail about one or a few. In the works of Agre (2001) and McCullough (2004) particular importance is given to physical context consisting of architectural structures and elements of the built environment, for example, landmarks and pathways. In the works of Dourish (2001b, 2004) particular importance is given to social context including interaction with, and the behaviour of, people in an environment. Dourish (2004) also states that context cannot be defined as a stable description of a setting, but instead arises from, and is sustained by, the activities of people. Hence it is continually being renegotiated and redefined in the course of action. These works provide us with additional contextual factors of particular relevance to mobile computing in context, and with the knowledge that what defines context is in itself contextually dependent.

The purpose here is not to challenge these existing definitions of context proposed in the literature. Instead, I subscribe to the definition by Dey (2001) and to the fact that several dimensions of context exist, and that the relevance of each of these for a particular interactive system or use situation is itself dependent on the specific situation.

What *is* important here is the role that "context" can potentially play as a suitable central and mediating concept, or boundary object (Star and Griesemer 1989), in a holistic and interdisciplinary approach to designing mobile interactions. The context of mobile computing is something that several individual disciplines within mobile interaction design are concerned with, and that has influenced the shaping of methodology, technology and theory within and across the field's internal disciplinary boundaries. These different disciplines have each approached the challenge of contexts differently, and have yielded different types of responses.

In domain studies of mobile computing, where context plays an obvious central role as essentially the phenomenon under scrutiny, the challenge has been partly to understand theoretically what use contexts are and how they can be described, and partly to study empirically what characterizes specific use contexts of interest, and how the phenomenon of context can be studied and analysed in ways that generate such understanding. This has led to a body of theoretical and socio-technical research building largely on methods and theories from sociology, anthropology, and phenomenology (e.g. Luff and Heath 1998, Dourish 2001b, Dourish 2004, Dey 2001, Ling 2001, Perry et al. 2001, Fortunati 2001, Green et al. 2001, Agre 2001, McCullough 2004, Chalmers 2004, Aoki et al. 2009, Kostakos et al. 2009), as well as my own work in chapters 3-6 (Paay and Kjeldskov 2005, Paay and Kjeldskov 2008a, Kjeldskov et al. 2004b, Kjeldskov and Stage 2006).

In systems development and design for mobile computing the challenge of context has primarily been about creating an appropriate fit between systems and context and how this can be supported structurally through new, or modified, systems development and design methods. While relatively very little has been published on this topic, there is an emerging body of methodological research building largely on methods and theories from information systems, software engineering and human-computer interaction (e.g. Sharples et al. 2002, Mikkonen et al. 2002, Hosbond 2005, Paay 2008, de Sá and Carrico 2009, Paay et al. 2009a), as well as my own work in chapters 7-10 (Kjeldskov and Howard 2004, Paay et al. 2009b, Vetere et al. 2005, Kjeldskov and Stage 2012).

In usability evaluation for mobile computing the challenge of context has primarily been to understand its role in relation to the scope, richness and validity of empirical findings and how usability tests can be carried out in contextually realistic settings through use of new or modified methods and techniques. This has led to a growing body of empirical research building largely on methods and theories from usability engineering. These include, for example (Brewster 2002, Betiol et al. 2005, Hagen et al. 2005, Kaikkonen et al. 2005, Nielsen et al. 2006, Rogers et al. 2007, Reichl et al. 2007, Oulasvirta 2009, Oulasvirta and Nyyssonen 2009, de Sá and Carrico 2011), as well as my own contributions in chapters 11-14 (Kjeldskov and Stage 2004, Kjeldskov et al. 2005, Kjeldskov and Skov 2007a, Høegh et al. 2008) and also Kjeldskov et al. (2004a)

In implementation of mobile computing the challenge of context has largely been about capturing, formalizing, and modelling this attribute in computational data models, how to make sense from such models, and how to use them in the construction of *context-aware* systems that are responsive to their surroundings. This has led to an extensive body of technical research building largely on methods and theories from computer science (e.g. Schilit and Theimer 1994, Crabtree and Rhodes 1998, Schmidt et al. 1999a, 1999b, Cheverst et al. 2000, Dix et al. 2000, Chen and Kotz 2000, Hinckley and Horvitz 2001, Dey 2001, Jameson 2001, Jones et al. 2004, Edwards 2005, Hinckley et al. 2005), as well as my own work in chapters 15-19 (Kjeldskov and Skov 2007b, Kjeldskov and Paay 2005, Kjeldskov and Paay 2006, Kjeldskov et al. 2010, Skov et al. 2012, Kjeldskov et al. 2012a).

In user experience research for mobile computing the challenge of context has been to understand what impact rich and dynamic user contexts have on peoples' experience of using technology, and to describe how this user experience can be improved. This has led to a body of theoretical, conceptual, and design-oriented research building on methods and theories from a wide range of disciplines from sociology and psychology to cognitive science, computer science, human-computer interaction and computer-supported cooperative work. These include, for example (Abowd and Mynatt 2000, Cheverst 2001, Palen et al. 2000, Weilenmann 2001, Bradley and Dunlop 2002, Brown and Randell 2004, Bardram 2009, Little and Briggs 2009, Benford et al. 2009, Karapanos et al. 2009, Lindley et al. 2009, Rowland et al. 2009), as well as my own work in chapters 20-23 (Paay and Kjeldskov 2008b, Kjeldskov and Paay 2010, O'Hara et al. 2011, Murphy et al. 2005).

This is not to say that context is a new phenomena appearing on the research agenda with the emergence of *mobile* computing. Context has indeed been an important concept within human-computer interaction and interaction design since the *second wave* or *paradigm* of HCI (Bødker 2006, Harrison et al. 2007). The first wave of HCI was a mixture of engineering and human factors focussing on optimizing human-machine fit. The second wave was largely based on cognitive science focussing on the simultaneous processing of information in machines and in the human mind, but this also involved a strong focus on the use of interactive computing systems in the context of the workplace. However, as pointed out by Bødker (2006) while there was lot of discussion about the intricate concept of context in second wave HCI, this research achieved little in terms of defining and operationalising it in a way of any real significant value to HCI and interaction design. In the *third* wave, focus has broadened further towards a post PC ubiquitous and pervasive information society where computer technology has spread "from the workplace to our homes and everyday lives and culture" (Bødker 2006). This means that context is now an *elemental* concept that we not only need to *define* well, but also need to *understand* better in terms of its complexity, significance, and influence on peoples' experience of technology in use, in order to inform technology design.

Mobile interaction design is positioned within the second and third waves of HCI. It grew out of the second wave, but the tremendous uptake of mobile computing by the general population subsequently played a contributing factor in the creation, force and velocity of the third wave by enabling some of the completely new potentials and patterns of computing technology use that we are witnessing globally today.

2.3. Design approaches

As I have allured to a number of times in the previous sections, mobile interaction design is, broadly speaking, characterized by two different approaches to design: a user- and a technology-centred one. This duality reflects the field's strong roots in the discipline of human-computer interaction on one side, and computing on the other, and an associated difference in primary interest amongst researchers and designers in *people* or in *systems*.

User-centred design

User-centred design (UCD) is a design philosophy and overall methodological framework for conceiving interactive systems, which can be traced back to Henry Dreyfuss' 1955 book *Designing for People* (Dreyfuss 1955). It is about designing interactive computer systems from the user's point of view emphasising people rather than technology (Norman and Draper 1986), and it does this by discovering unmet user needs and responding to these through design. Traditionally, user-centred design and research follows an iterative cycle that consists of the four central stages of studying, designing, building and evaluating interactive systems (cf. Preece et al. 2002, Sharp et al. 2007, Harper et al. 2008). The strengths of this framework is that it is simple, captures some of the essential components of an interaction design process, and provides an overall structure of how to organize them in relation to each other. Historically, another strength of the model has been that it promotes an iterative "prototyping" approach to systems development rather than a linear "waterfall" approach.

The user-centred design philosophy has several branches of more specific methods and associated philosophical views on people and design, such as Prticipatory Design (PD), Usability Engineering (UE) and Contextual Design (CD). Participatory Design (see, for example, Ehn & Kyng 1987, Bødker 1996, Kensing & Blomberg 1998) has its roots in Scandinavia where it was developed as "Cooperative Design" as a part of several research projects in systems development in collaboration with trade unions dating back to the 1970s and 1980s (see, for example, Bødker et al. 1987, Ehn & Kyng 1991). Participatory Design puts particular emphasis on the active involvement of users, and other stakeholders, in the design process in order to ensure the usefulness of the produced outcomes. This is typically done through an action research process where researchers/designers and users cooperate closely and iteratively, and both gain from the relationship. Participatory Design has a clear underlying political dimension of emancipation and democratization, and can as such be described as one of the more radical human-oriented approaches within user-centred design. Some of the key principles of Participatory Design are that users of technology should be respected as experts in their own domain, and recognized as prime sources of innovation. The design task is approached broadly with focus on people, practices, and technology embedded in specific organizational contexts. As a consequence of this, Participatory Design is heavily field based, with researchers and designers spending significant amounts of time "in the wild" rather than in the lab.

While Participatory Design takes user involvement very seriously and elevates user needs, satisfaction and human well being far beyond anything else, a key critique of this form of user-centred design is that it leads merely to incremental design improvements grounded in current practice rather than to design leaps that can facilitate the formation of new practices. Consequently, and somewhat ironically, from a design perspective Participatory Design could therefore be criticised for being fairly conservative rather than progressive. The essential reason for this is that users are not necessarily good designers. They are experts in their own domains, and possess important knowledge for informing design, but they usually know very little about technology and design, and are not trained in projecting into the future and envisioning new technologies and designs. This, it can be argued, is better done within the expertise of trained designers or engineers. As described by Jakob Nielsen, “users are not designers, so it is not reasonable to expect them to come up with design ideas from scratch. However, they are very good at reacting to concrete designs they do not like or that will not work in practice” (Nielsen 1993 p. 88-89).

This perspective on the role of users in user-centred design is fundamental to the Usability Engineering approach (see, for example, Nielsen 1993, Preece et al. 1993, Spool et al. 1999, Rosson & Caroll 2001). Usability Engineering emerged as a distinct area of research and practice in the mid- to late 1980s. As indicted by its name, it originates from the more engineering-oriented parts of human-computer interaction with many usability engineering researchers and practitioners having a background in either computer science or a sub-area of psychology such as perception, cognition or human factors. Usability engineering emphasises the making of user interfaces with high usability or user friendliness as defined by, for example, the ISO 9241 standard about the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in a specified context of use. This is done by iterating through phases of user requirements gathering, prototype development and usability evaluation against a set of quality metrics. Evaluation is done either by usability experts through, for example, heuristic inspection (Nielsen & Molich 1990) or by testing systems with users (Rubin 1994). Although field based usability testing has been given some attention, usability engineering is heavily laboratory based, with researchers and practitioners spending most of their time in controlled environments rather than in the wild – even when testing mobile computing devices and applications (see for, example, Weiss 2002).

Although usability engineering implies user involvement, a key criticism of this approach is that it reduces the roles of these users to being informants for requirements and test subjects for designs. Consequently they are typically far removed from the actual design and development process, and only brought in very early in the process, before any design has been done, and/or very late when many design decisions have already been made. Another critique of usability engineering in relation to interaction design is that it focuses too exclusively on one particular quality of a design – its usability – and on assessing and making recommendations about where to improve this, rather than on the broader user experience of the design, in all of its contextual complexity, and facilitating the creation of life- and work-enhancing interaction design matching this context.

In attempt an to overcome some of these limitations, Contextual Design seeks to broaden the perspective of the usability engineering process to explicitly include the context, and to inform design more strongly rather than merely assessing and, sometimes, providing recommendations for redesign. Contextual Design is a specific method within the user-centred design philosophy developed by Beyer and Holtzblatt (1998) that accommodates field studies in system development by combining ethnographic methods for data gathering relevant to a particular product, rationalizing workflows, and designing human-computer interfaces. Although parts of the method can fruitfully be applied to non-work related domains, the contextual design method is in essence about supporting work activities through structured and contextually grounded user-centred design. To put it bluntly, it is not a method of interaction design but a good method for doing user-centred usability engineering. Contextual Design provides a detailed account for what exactly to do, when and how, in a user-centred design process. It does, however, not provide an overall framework for integrating the different disciplines of interaction design. Although it has a strong component of fieldwork and contextual analysis of work and organizations, it is weaker in terms of supporting design and evaluation. Both of these are approached as activities of usability engineering. Contextual Design also doesn't integrate well the technical aspects of interaction design in working with prototypes and exploring actual technology. On the positive side, one of the absolute strengths of Contextual Design is that it is very solution oriented, which makes it attractive as a methodological foundation in industry and research organizations doing more short term and solution-oriented ICT work.

Technology-driven design

Contrasting the user-centred approach, another overall philosophy is to let *technology* drive the process of design. Technology-driven design is about letting the possibilities of new technologies, or existing ones in new combinations, inform what we design and build. Put on the edge, technology-driven designs are not solutions to problems expressed by users but instead solutions to technical problems looking for possible use. They are propositions envisioned by designers or researchers based on their knowledge of what *can* be done, and not in response to what people or users have asked for. In his provocative and highly debated March 2010 column of the ACM Interactions magazine "Technology First, Needs Last", Donald Norman argues that technology-driven design is invariably responsible for all conceptual breakthroughs in modern history. Grand inventions, such as flush toilets, plumbing, electric lighting, cars, airplanes, and ICT, Norman argues, happen because they have been made possible technologically, and not because people have asked for them or necessarily needed them beforehand. Technologists create new designs *because they can.* Products, applications, use and especially *needs,* evolve slowly afterwards (Norman 2010).

Technology-driven design is inherently a trial-and-error based approach in which new ideas are developed, tested, and then explored further or scrapped. Methodologically this makes it a risky and time consuming approach relying on the team's capabilities of reasoning through intuition, experience, deduction and induction with no guarantee of a useful outcome or solution being produced (Kjeldskov and Graham 2003/chapter 2, Danis and Karat 1995). On the other hand, however, this fundamental reliance on the competence of the researchers and designers strongly facilitates the utilization of the precious source of creativity and innovation that these exact people represent. They are typically experts in their area of technology and therefore able to look beyond the technological horizon that "users" without this insight normally have.

Although not a formally established paradigm (like user-centred design) technology-driven design has always been implicitly present as a fundamental approach within the area of computing. According to van den Ende and Dolfsma (2004) neither the development of digital computers after 1960 nor the convergence of computing and communication technologies after 1990 can be explained by rising demand or newly discovered user needs. Instead it was the development of new technological knowledge that was the enabler of new types of computers such as the PC and new communication applications such as the Internet. Had this knowledge been available earlier, those technologies would most likely have been developed and adopted then. Hence, just like any other area of technology, computing has partly emerged and evolved through a process of technology-driven design based on an interest to explore what would be technically possible more than what would be useful to do. This is simply an established way of doing research within an emerging area of technology (Kjeldskov 2003). Before we can begin contemplating broader phenomena like use, usability, and user experiences of new technology, this technology simply needs to be available in some concrete and functioning form. Technology-driven design provides this, and should therefore not be ignored as a driver, or enabler, of interaction design. On the contrary one could argue that it is in fact a necessary precondition for user-centred design.

Faced with progress where processing power has increased exponentially for half a century, and networking capability, display technology and input devices also follow a significantly fast paced evolution, Jones and Marsden (2006) argue that it is very hard *not* to let technology drive the development of interactive mobile systems and devices. They argue that while user-centred design approaches stay clear of thinking about technology until the user needs have been identified and described – in order not to be distracted by detailed design consideration – when designing mobile interactions there appear to be advantages in allowing the potentials of technology to play a more central role in the process. This view is shared by Rogers et al. (2002) who argue that looking at technologies themselves can be a valuable source of inspiration for design, and by Danis and Karat (1995) who argue that technology-driven design facilitates a dual goal of advancing technology *as well as* creating new usage benefits, while user-centred design primarily focus only on the latter. In the latest issue of ACM Interactions, the importance of technology-driven research is also addressed by Kritina Höök, arguing that it is time for the CHI community to reconnect with the potentials of emerging technologies and, as in the past, "shape interactions based on a deep, well-cultivated understanding of technological capacities" (Höök 2012).

At the intersection between users and technology

In light of the discussion above, if we choose to see interaction design as a matter of innovating and creating new ways of using technology, one can argue that a central shortcoming of the UCD approach is its strong focus on users. Undoubtedly, hearing such a statement from an HCI research academic is bound to cause some raised eyebrows, but bear with me here. Over the last couple of years it has become more and more evident that some companies, most notably Apple, are able to develop, produce and sell millions of novel interactive computer products with groundbreaking interaction design, high utility and usability, and premium user experiences *without* following a user-centred design approach as prescribed by our primary textbooks. How is this possible? *What is it that those companies do that is missing from the established literature on the topic?* It appears to me that the problem originates in UCD being so good at grounding the interaction design task in *current practice* that it impedes our ability to break free from this and imagine *future practices* that are fundamentally different. User-centred design focuses on discovering unmet user needs, but in successfully doing so these needs are reinforced rather than questioned and challenged.

As a consequence, UCD has a track record of resulting in small incremental improvements rather than in fundamental breakthroughs or radical new inventions (Verganti 2010). As discussed by Norman (2010) such improvements are often highly valuable, and in fact where the most frequent gains come from, *but* this type of value adding is very different from the "success by innovation" that prompted the questions about Apple's success. Contrasting the incremental improvements facilitated by for example UCD, Norman continues to argue that conceptual breakthroughs are driven exclusively by new inventions in technology in response to which user needs arise much later. His examples are many, including the automobile, the airplane, radio, TV, the computer, and, of course, mobile phones. Although I agree with Donald Norman about the shortcomings of UCD raised in his column, I do not agree with the conclusion that we should "leave it to the technologists" to "get the grand ideas running". In fact, Normans own argument of new technology initially failing until someone other than the inventor comes along and envisions new usage, indicates to me that there is something else at play here. Instead, I believe that what we are looking for lies in the middle ground between the two. Between fundamental technology invention and incremental user-centred development, or "at the intersection between technology and liberal arts" (Jobs 2010) as often illustrated by Steve Jobs in the later years (figure 12). What we are looking for is exactly what Norman (2010) does not extract from his own examples: the radical innovation created by *merging new technological possibilities with visions of future practice* – without asking the users what they want, but by making radical proposals about it.

This is precisely what companies like Apple are good at doing. In Apple's innovation process "insights do not move from users to Apple but the other way around. More than Apple listening to us, it's us who listen to Apple" (Verganti 2010). However, these radical proposals are not created by chance or from intuition of a visionary guru but from very precise processes and capabilities (Verganti 2009). They require a solid understanding of both technology and of users – perhaps even understanding users more than they understand themselves – and for the latter some of the techniques used within UCD, such as ethnographic studies, are still valuable. Not "to discover hidden, unmet needs" (Norman 2010) but to deeply understand the context that we are designing for.

Figure 12. Steve Jobs on Apple's success at designing interactive products: "We've always tried to be at the intersection of technology and liberal arts, to be able to get the best of both, to make extremely advanced products from a technology point of view, but also have them be intuitive, easy to use, fun to use, so that they really fit the users – the users don't have to come to them, they come to the user" (Jobs 2010).

3. OPPORTUNITIES FOR MOBILE INTERACTION DESIGN RESEARCH

I am now going to change focus and look at the opportunities for mobile interaction design in the early 2000s that motivated my research in this area.

My own research in interaction design for mobile computers began in 2001. At that time, Mobile HCI and interaction design research was still very much in its infancy as an academic research area. Widely commercially successful devices had only been around for about a decade, and leading conferences had only a few years of history behind them. As a consequence only a small body of knowledge existed about this emerging research topic in terms of methodology, interaction design, and real world use, and no coherent sets of methods and techniques for mobile interaction design had yet been established. Driven by the saturation and increasing technological maturity of mobile devices throughout society, there was, however, a huge interest in the new interaction design possibilities of this fast expanding area of computing. This inspired me to carry out a comprehensive literature survey of mobile interaction design research. The purpose was to provide a snapshot of state-of-the-art and current practices, and through this identify shortcomings and opportunities for future research directions. The literature survey was presented at Mobile HCI in Udine in 2003, the second time this conference ran as a full and independent international event (Kjeldskov and Graham 2003/chapter 2), and contributed strongly to bringing the issue of research methodology on the agenda within mobile interaction design. We reviewed 102 articles published between 2000-02 in the most central outlets of its time and classified them in terms of their research method and purpose, as described in table 1 (inspired by Wynekoop and Conger 1990). This provided a picture of how mobile interaction design research was being done, and for what intent, and brought to attention a number of trends characterizing the field, and a number of hidden assumptions influencing its focus and approach. The distribution of articles on research methods and purpose is illustrated in figure 13 clearly showing two peaks, a number of notable but less frequent groupings, and a large number of gaps.

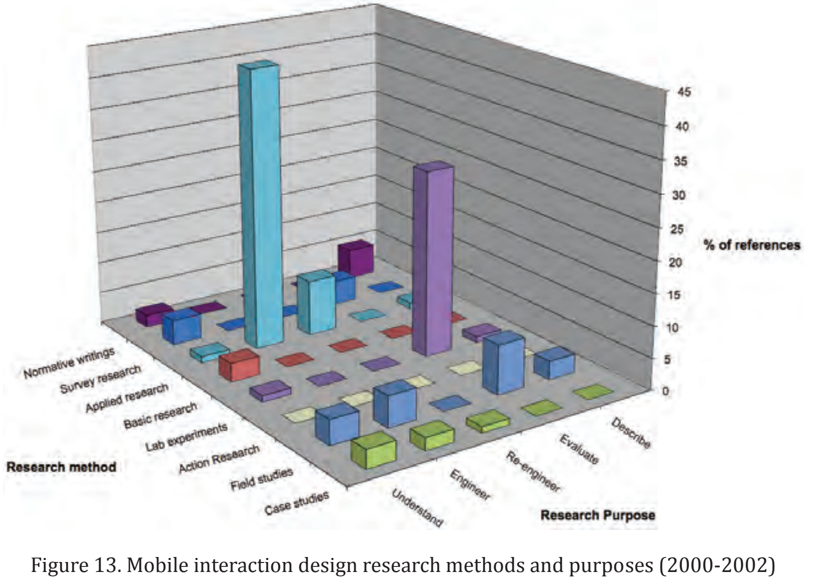


Figure 13. Mobile interaction design research methods and purposes (2000-2002)

Table 1. Overview of research methods and purposes

Method	**Case studies**	Intensive empirical investigations of contemporary phenomena within small size entities such as groups, organizations, individuals, systems or tools in real-life context with the researcher distinct from the phenomena being studied
	Field studies	Characterized by taking place in "the real world" covering a range of qualitative and quantitative approaches from ethnographic studies of phenomena in their social and cultural context to field experiments where independent variables are manipulated
	Action research	A method through which researchers not only add to the body of scientific knowledge but also apply that knowledge to the object of interest through intervention and participation in the activity being studied
	Lab experiments	Characterized by taking place in a controlled environment created for the purpose of research or in dedicated laboratories allowing a detailed focus on specific phenomena of interest with a large degree of experimental control
	Survey research	Informs research gathers large amounts of data through various techniques such as questionnaires and interviews from a known sample of selected respondents assumed to be independent of their environment
	Applied research	Builds on trial and error on the basis of reasoning through intuition, experience, deduction and induction. Typically the desired outcome of an applied research process is known while means of accomplishing it is not. This makes applied research very goal oriented.
	Basic research	Characterized by trial and error based development of new theories and the study of well-known problems to which neither solutions nor methods are known, relying on the competences of the researcher
	Normative writings	Cover the body of "non-research" writings about phenomena of interests such as concept development writings organizing ideas for stimulating future research, presentation of truth describing ideas that seem intuitively correct, and descriptions of applications.
Purpose	**Understanding**	The purpose of research focusing on finding the meaning of studied phenomena through, for example, frameworks or theories developed from collected data.
	Engineering	The purpose of research focused towards developing new systems or parts of systems, for example an interaction technique for a mobile device, or a mobile application or device.
	Re-engineering	The purpose of research focusing on improving existing systems by redeveloping them such as, for example, adapting a web browser to a small display.
	Evaluating	The purpose of research assessing or validating products, theories or methods, for example, the usability or user experience of a specific application, or a theory of interaction.
	Describing	The purpose of research focusing on defining desirable properties of products, for example, an interactive mobile guide system, or mobile interaction design method.

The survey of mobile interaction design research in the early 2000s to a large extend served as a guiding light for my own research in this area in the decade that followed. Not as a collection of ideals or exemplars for what to do and how to do it, but rather as a provocative prompt for exploring their opposites, inspired by Philip Agre's Critical Technical Practice (Agre 1997). But before I outline the specific research topics that my colleagues and I explored in response to the survey, I will briefly describe and discuss some of the trends and assumptions it revealed.

3.1. Trends and assumptions

The literature survey revealed a strong bias towards applied research for engineering and laboratory experiments for evaluation, as shown in figure 13. Put simply, mobile interaction design research in the early 2000s was dominated by building new systems in a trial-and-error manner, and evaluating them in laboratory settings – if evaluating them at all. There was very little going on in terms of trying to understand the phenomenon of mobility itself in relation to interaction design and technology use, and to use such insight when designing and building actual interactive systems. Nor was much attention given to the role of context in relation to understanding, building or evaluating interactive mobile systems (Kjeldskov and Graham 2003/chapter 2). In essence this echoed a fundamental segregation between use- and technology-centeredness depending on whether the involved researchers were primarily interested in *people* or *systems,* as discused earlier. On a more general level, it became apparent that methodology seemingly played a very small role. The approaches taken often remained unexplained, their suitability unchallenged, and their limitations and alternatives not discussed.

Based on our more detailed analysis of what types of research and purposes were missing, or largely underrepresented, we cautioned that the bias towards trial-and-error building of interactive systems, evaluations only in the lab, and the lack of research for understanding design and use in real world contexts, would limit the quality and scope of the body of knowledge about mobile human-computer interaction being accumulated, and thereby inhibit the advancement and impact of the research field in the future. In particular we found from our analysis that three underlying, and unfortunate, assumptions appeared to be characterizing mobile interaction design research at the time:

1. We already know what to build
2. Context is not important
3. Methodology matters very little

In the following I will briefly take a critical look at these three assumptions and their potential implication for research.

We already know what to build

The prevalent approach of applied research for engineering indicated an assumption at the time that we already knew what systems to build and what problems to solve, such as limited screen real estate, limited means for interaction, and limited network bandwidth. We just didn't know yet exactly how to build these systems and how to solve those problems – but the solutions existed out there and were just waiting to be uncovered. Only very little research addressed the more fundamental questions of what is useful and what is perceived problematic from a user-perspective, and evaluations focused on technical functionality rather than context-centred and user-centred issues. Given the young age of the research field we argued that this assumption could hardly be true and that, on the contrary, young emerging research fields such as this particularly require research addressing such fundamental issues. Continuing to do research on the basis of the assumption that we already know the problem would, in it self, make it very difficult to set this assumption aside and identify the more fundamental challenges at hand.

Context is not important

The limited focus on real-world studies indicated an assumption that context was not really important for what we build, and that interactive mobile computer systems are by definition suitable solutions. Building and evaluating interactive systems on the basis of applied research and laboratory experiments also results in very concrete conclusions about very specific solutions. These conclusions can be difficult to generalize and therefore it can be difficult to elevate our learning from the systems we develop, and study in use, to an abstract level where knowledge can be transferred to other design cases, technologies, domains, users, purposes, etc. This limits the research field's ability to move forward at a pace beyond incremental steps from one specific design to the next. Hence, in our opinion, the assumption, that building and evaluating systems by trial and error is better than grounding engineering, evaluation and theory in user-based studies, seriously weakened the mobile HCI and interaction design research field at the time.

Methodology matters very little
The observation that only very few studies were based on an explicitly described and considered methodological foundation indicated an assumption that methodology mattered very little in mobile HCI and interaction design research. We presented this supposition as a particularly problematic one because it is well-known that the choice of method influences the results subsequently produced (Myers 1997). Problem solving by applied research is, for example, often viewed as a rather poor method because it demands huge efforts by researchers and often "translates into poor performance because they require search of a large space of possibilities" (Wynekoop and Conger 1990). In relation to mobile interaction design research, a lack of fundamental critical reflection on methodology, and a sensitivity to its importance, would impede our ability to identify the limitations of our work and inhibit the breakthroughs in design and use discovered through deliberately looking at and doing things different than before.

Opportunities
Part of the reason for the bias towards applied research for engineering combined with laboratory experiments for evaluation is that this is a natural place to start when exploring a new field of emerging technology. Before we can study and understand phenomena like use contexts, usability, and user experiences of new technology, we need this technology to be available to us in some concrete and functioning form. However, if a field of *emerging* technology is to evolve into a field of *applied* technology, it is important not to get stuck in research methodologies where solutions are created and put to use by trial-and-error rather than grounded in real world context. Another part of the reason, I believe, is that rather than mobile technologies not being ready for studies in natural settings, the body of mobile HCI researchers at the time were not really ready for natural settings research. In the early 2000s only very few studies had been published that used natural setting research methods within mobile HCI and interaction design. Consequently only very few examples existed for others to be inspired by and follow, and the whole debate about doing research slightly differently had not even started. Adding to this, it was still very unclear exactly how to make use of methods like field studies, case studies, and action research in mobile HCI and interaction design, and what value they might bring to a specific project. Hence, in the early 2000s even the *multi*-disciplinarity of mobile HCI and interaction design was not yet strong. Like other areas of emerging computing technologies it was very much a technical research area dominated by electrical engineering and computer science. In terms of methodology this meant that methods and techniques from, for example, social science, the humanities, and the arts naturally did not yet have a strong presence in the minds and traditions of the dominant

mobile HCI researchers at the time, and prominent researchers from those adjacent fields were not yet working with mobile HCI and interaction design themselves.

Apart from describing and discussing current practice in mobile interaction design research, the survey also outlined a variety of opportunities for doing future research methodologically differently. In particular, the noticeable lack of *field studies* presented an obvious opportunity to use this method for exploring rich real-world technology use cases, contexts and user needs to gain deeper understanding of these. In particular, we suggested to learn more from disciplines that have struggled with the study of similar "slippery" phenomena using techniques such as ethnography. We also proposed that engineering and re-engineering should be informed more by field studies of context and users. Finally, we promoted the opportunity for systematic investigation of field studies for evaluation, as mobility and context can be difficult to emulate in a laboratory.

The lack of *survey* and *case study* research also presented notable opportunities. In the field of information systems, for instance, these approaches are used widely to collect large amounts of data from, for example, a large segment of actual end-users of an interactive system, enabling much greater power of generalization. More case studies within mobile HCI and interaction design could increase learning from existing interactive systems within real-world contexts. This would also enable close scrutiny of specific phenomena in specific settings, which would enrich the collective knowledge in the discipline and enable key issues to be described and understood in depth. The very limited amount of *basic research* indicated an opportunity for the development of theoretical frameworks, and application of existing ones from other disciplines, for describing and understanding mobile interaction design and use. Finally, we argued that the complete absence of *action research* pointed to both the lack of a well-established body of theoretical research within the discipline, and to an unwillingness to roll out mobile systems uncertain to succeed and taking a long time to implement and evaluate. This was perhaps not surprising, given the cost of such technology and the associated implementation overhead at the time. Nonetheless, such studies of practice and intervention was, again, an opportunity to develop new kinds of knowledge in the discipline.

Only a few years after the literature survey, mobile interaction design research had already started to change. The methodological opportunities proposed were indeed taken up by a lot of our colleagues, and today, a decade later, the research field has matured considerably and is making use of a much wider palette of research methods in interesting combinations, and for a much wider range of purposes. This trend was confirmed by a follow-up survey reviewing all research articles concerning the design of mobile interactions published in top outlets in 2009 (Kjeldskov and Paay 2012). From this survey it is apparent that the research field of mobile interaction design has grown substantially in the last decade and is now a substantial part of mainstream HCI and interaction design research. Out of the 246 full and short papers in the Proceedings of the annual ACM Conference on Human Factors in Computing Systems (CHI) that year, almost a fifth concerned human-computer interaction with mobile systems or devices. It is also clear that there has been an increase in the level of empirical research, and that a more diversified set of methodologies for this has evolved. For example, the use of *field studies* have notably changed and diversified into at least three noteworthy sub categories of *field ethnographies, field experiments,* and *field surveys*.

At the same time, however, there is also clear evidence that the underlying segregation into camps primarily interested in *people* or in *systems* persists. The first aims primarily at understanding mobile user experiences theoretically and conceptually, and the second aims primarily at building new mobile systems and evaluating them in use. This segregation of course stems from the multi-disciplinarity of the research field, but maintaining such a divide sadly sustains the unfortunate implicit assumption that in mobile interaction design research people and technology can, or perhaps even *should*, be studied separately. In turn, such an assumption can be partially responsible for researchers in the people- and technology- oriented camps continuing to investigate the same types of questions and problems as before, rather than defining and exploring new ones in closer collaboration. Within both approaches *users* play an important role, but in the first they are the *objects* of the research, while in the second they are research *subjects*. According to Rasmussen (2007) such clear-cut distinction tends to cause the potentially fruitful dialectics between the two approaches to disappear. If one of the two approaches is considered 100% good and the other 100% bad, from either side of the divide, then one is destined to subsume the other. Dialectic thinking, on the other hand, encourages us to develop a synthesis at a higher stage of the opposing interests, as also discussed by Dahlbom and Mathiassen (1993). This is not simply a matter of finding a balance between the two but about transcending beyond opposing views and shaping a new unity at another level (Nonaka and Toyama 2002). Hence, in order to continue informing the creation of better interactive mobile devices and systems, a closer integration of these two approaches, the user- and the technology-centred one, is still needed.

3.2. Impact on my research

The literature survey described above had a strong impact on my own research in the decade that followed. In particular my colleagues and I attempted to transcend the "people or technology divide" by scoping our work on the design of mobile interactions to be contextual in nature, and aiming to understand the *interplay* between people, technology and context. We did this by working with five particular research topics:

1. Studying and analysing use contexts
2. Using context to inform interaction design
3. Developing new methods for evaluation in context
4. Exploring context-awareness
5. Understanding user experiences in context

These topics are reflected in the five corresponding parts of this thesis.

Studying and analysing use contexts
The first topic addressed the importance of understanding context in mobile interaction design research. Although the literature survey had indicated an assumption that context mattered very little in the field of mobile interaction design, several key researchers had argued for the opposite and stressed the importance of understanding this complex phenomena properly (see, for example, Luff and Heath 1998, Brown et al. 2000, Agre 2001, Dourish 2001b, Greenberg 2001, Perry et al. 2002, Chalmers 2004, Tamminen et al. 2004, Blom et al. 2005). My colleagues and I shared this view, and believed that one of the keys to successful design of mobile interactions lay in understanding, profoundly, what that context is for a particular design task, what dimensions of context are important, what role they play, how they interact and influence each other, etc.? This required empirical and analytical research, prompting us to explore how different aspects of context for mobile interaction design could be studied empirically, analysed conceptually, and presented in a form that could eventually be used to inform design. Our studies of use contexts took the shape of field studies of people, environments, and work activities for the purpose of informing prototype design.

Chapters 3-6 in Part I contain four selected publications from this research.

Using context to inform interaction design
The second topic addressed a methodological gap in terms of how to inform and support the process of designing mobile interactions in a way that was both structured yet supportive of creativity and innovation in design. The ambition of this topic was to move beyond design and development by trial-and-error, and instead introduce some level of methodological guidance for our exploration of new mobile technologies as sought by, for example Krogstie et al. (2004). The literature survey had revealed that mobile interaction design research was very much about building new systems "because it was possible", and although we agreed that a lot of this work was interesting and could possibly be applied to some kind of use, we also believed that this could be done in a much more informed way. However, generally very little had, and still has, been published on this topic. Our view on the matter was that in order to make use of our understanding of context in the exploration of a design space we need to develop ways of transferring empirical and theoretical understanding into the interaction design activity, and develop ways of grounding design ideation and exploration in such input. This required research of creativity that looks in detail at actual processes of interaction design where different methods and techniques have been applied for this purpose. Our studies of such design processes took place as case studies of interaction design projects with the purpose of describing methods, approaches and techniques that worked particularly well – on their own, or in concert.

Chapters 7-10 in Part II contain four selected publications from this research.

Developing new methods for evaluation in context
The third topic responded directly to our suggestion to explore new ways of evaluating mobile interaction design in context as an alternative to the prevailing use of context-deprived laboratory experiments. Already at the first workshop on human-computer interaction for mobile devices in 1998 researchers and practitioners had been

encouraged to investigate further into the criteria, methods, and data collection techniques for usability evaluations of interactive mobile systems (Johnson 1998). By 2002 several studies had confirmed this need (Graham and Carter 1999, Pascoe et al. 2000, Brewster 2002, Rantanen et al. 2002) but methodologically it was still an under-investigated topic. Our position was that in order to realistically assess the quality of mobile interaction designs we need to evaluate them under conditions that are appropriately representative of the future use context. This involves a challenge of trade-offs between factors such as ecological validity, experimental control and data quality, and required empirical research of the applicability and systematic comparisons of the relative strengths and weaknesses of new and existing methods and techniques. Our own work on the topic aimed at systematically investigating the pros and cons of different approaches to empirical evaluation of interactive mobile systems, adapting techniques from ethnography on one side and traditional usability testing on the other, and to learn when, where and how to evaluate interactive mobile systems. The research took its offset in the hypotheses that field studies were a valuable approach because of their ecological validity, but that laboratory experiments could also be improved by simulating aspects of context. On the basis of this, we wanted to explore and refine techniques for evaluating in the field, and to develop and study techniques for recreating contextual factors in the lab. This was done through a combination of laboratory experiments, field studies, and quasi experiments across these settings.

Chapters 11-14 in Part III contain four selected publications from this research.

Exploring context-awareness

The fourth topic embraced the idea of making mobile systems *context-aware* by exploring dynamic use context as means of input in interaction design (Dey 2000). Context-awareness for mobile devices was a relatively new idea at the time (introduced by Schilit and Theimer in 1994) and although there had been quite a bit of work done on the technical side of sensing context, very little had been done on the use of such data in interaction design. Studies like Hinckley et al.'s (2001) use of embedded tilt sensors for automatically changing between portrait and landscape mode, and Cheverst et al.'s (2000) use of location in mobile tourist guides, which are now common features in interactive mobile systems and devices, had only just been published in the research literature, and little knowledge still existed about their actual use. Our position was that in order to inform the creation of contextually informed mobile interaction design solutions, we needed to experiment with functional systems exploring the implementation of design ideas and emerging technologies in practice. This required technical research working with prototypes as vehicles for exploring feasibility, usability, utility, and user experience of various interaction mechanisms, information representations and system architectures.

Chapters 15-19 in Part IV contain five selected publications from this research.

Understanding user experiences in context

The fifth topic grew partly out of the other four and partly out of the lack of basic research for understanding phenomena within mobile interaction design conceptually and theoretically. From the other four research topics we learned a lot about context, how to design contextually grounded interactive systems, how to make interactive systems adapt to context, and how to evaluate them. However, the insight gained from our studies of technology in use often remained rather specific to the designs that we studied. Looking at related research this seemed to be a general pattern. Only a limited body of knowledge existed that explained, theoretically, the user experience of mobile interaction designs in context. This required theoretical research that developed conceptual understanding of the relationship and interplay between users, technology and their context, and motivated, amongst others, the Indexical Interaction Design project from 2004-06, the eSpective project from 2006-08, and my participation in the Blended Interaction Spaces project from 2008-10.

Chapters 20-23 in Part V contain four selected publications from this research.

3.3. Emerging challenges

Having explored the five topics above, what then are the challenges emerging today? As I described in the introduction, it is my opinion that the enormous uptake of mobile devices, and the role that they have come to play in our lives, means that mobile computing has evolved from being strongly an engineering profession to being, at least, equally strongly a design profession. Hence, interaction design is today of greater importance for the continuing development of mobile computing than ever before, and there is a need to ensure that our overall approaches to the way that we think about and design mobile interactions are up for the challenges that lay ahead. In order to do this I argue that we may benefit from exploring, more profoundly, *designerly ways* of thinking and doing, and from widening our scope, more significantly, to look at the contextual user experience of interactive mobile systems and devices and the digital ecosystems they are forming.

In my view, the first emerging challenge for designing mobile interactions is to *transcend beyond the dichotomy of people- or technology- oriented research and design.* Continuing such divide we are at risk of missing the holistic nature of the mobile interaction design challenges currently at hand, as mobile technologies have matured considerably and now pervade almost every aspect of our lives. What is instead needed is to "reconceptualize the domain of interest through using a modified unit of analysis" (Rogers et al. 2005) – creating a shift in focus where new and shared problems are framed in a way that force new ways of thinking and operating, while still allowing the use of existing concepts and theory. The second emerging challenge is to *widen the scope beyond the individual mobile device and an individual user's interaction with it*. This initial perspective has been researched in depth for over a decade, in both artificial and natural settings, and is now understood quite well. What is *not* understood very well is mobile interaction design within the even larger context, or wholeness, of everyday life and the use of other technologies, that interactive mobile systems have become a part of.

Sections 4 and 5 in this chapter describe my response to these questions.

As I will demonstrate in sections 4 and 5, transcending beyond people and technology, and widening the scope beyond individual devices and users, can both be achieved by changing the fundamental way we think about and go about doing interaction design, from a traditional scientific way towards a designerly way.

4. TOWARDS A DESIGNERLY WAY

I will now shift focus from reviewing past mobile interaction design practice and research to discussing how we can facilitate that these continue to be carried out with significance and impact in the future. In this discussion I will draw on design research literature from outside the field of interaction design, and show how a synthesis of these thoughts and concepts can help enrich the way we think about and design mobile interactions today.

I am going to start by discussing the need for doing, and thinking about, interaction design in *a designerly way* rather than in a traditional scientific way. This is not just a matter of wording or branding, but involves some profound changes to the underlying philosophy, approach, reasoning, and focus in interaction design as a profession and research area. For this purpose I will discuss four different perspectives on design and introduce *contextualism* as an alternative, constructivistic, worldview to the, essentially positivistic, ones still pervading large parts of mobile interaction design today. I will also illustrate how contextual views have informed information systems research, and shaped modern architecture and architectural theory in ways that can enrich our thinking about interaction design. Finally, I will revisit the user-centred design (UCD) approach in light of introduced concepts and perspectives, and elaborate on this model towards a contextual approach focusing on the continual convergence of form and context.

4.1. From technical rationality to continual convergence

While mobile interaction design is a relatively young field of study, *design* has been a recognisable topic of research since the early 1960s. The area of "Design Research" promotes "the study of and research into the process of designing in all its many fields" (Design Research Society 1966), and has had a strong presence within the research area of information systems dating back to the early 1980s. Surprisingly, however, given its seemingly obvious relevance, Design Research has only had very small presence in interaction design – and even less in *mobile* interaction design. Notable exceptions include the work of Rogers (2004), Buxton (2007), Moggridge (2007), and Stolterman (2008).

Design Research has in itself undergone an interesting evolution in philosophy, approach, and focus over the last 50 years, and some of the thinking about design that it has spawned can be used to frame and indicate possible ways forward within the design of mobile interactions. In brief, the origins leading to Design Research can be traced back to early 1920s movements of modern design based on values of objectivity and rationality. The aspiration to "scientize" design re-emerged in post-World War II's increased interest in the use of "systematic methods of problem solving, borrowed from computer techniques and management theory, for assessment of design problems and development of design solutions" (Archer 1965), leading to the "design methods movement" of the 1960's where objectivity and rationality began being applied to processes and not just products (Cross 2001, de Figueiredo and da Cunha 2007). The underlying assumption of instrumental or technical rationality was that design, like science, can be dealt with through decomposition and systematic search through possible solutions. Prominent work within this approach to design included Herbert Simon's "The Sciences of the Artificial" (Simon 1969), introducing concepts such as "bounded rationality", and even the early work of Christopher Alexander (1964), which had a flavour of rational methodology – although Alexander himself later declared his complete dissociation with this. In the 1970s, however, there was a backlash against the positivistic scientization of design, responding to the seeming inability of technical rationality to "help society achieve its objectives and solve its problems" (Schön 1983), and fuelled by the social and political movements of the time rejecting the conservative values underlying it (Cross 2001). The rejection of technical rationality led to renewed vigour in design research in the early 1980s, seeking the development of design as a discipline with epistemology and foundation in its own history of practice rather than in science. This is largely captured in Donald Schön's "The Reflective Practitioner" (1983), which "explicitly challenged the positivist doctrine underlying much of the design science movement" (Cross 2001) but also partly in Nigel Cross' quest for a "designerly" way of knowing, thinking and acting (Cross 1982).

One of Schön's central messages is that from a technical rationality perspective, professional practice, such as design, is a process of problem *solving,* but that in real world practice problems are not given but need to be constructed. This requires a problem *setting* approach, in which we name what we will attend to and frame its context (Schön 1983 p. 40). In a similar way, Cross argues that there are designerly ways of knowing, thinking and acting, fundamentally different from the generally recognised scientific and scholarly ways, and that these are more about defining the limits of the problem and suggesting the nature of its possible solution, than they are about exhaustive systematic analysis (Cross 1982). This thinking is echoed in a seminal paper by Giovan Lanzara (1983) who outlines three different views on design as 1) functional analysis, 2) problem-solving, and 3) problem-setting, and argues that understanding the underlying views on design can help explain the problems that constrain it. The first two of these correspond to the technical rationality approach while the third corresponds to Schön's concepts of *reflection-in* and *-on action,* and partly Cross's *designerly way* of knowing, thinking and acting. In design as *functional analysis,* design is a process of systematically breaking down the problem through rational analysis and thereby revealing the structure of the one optimal solution. In design as *problem-solving,* design is about finding solutions to problems. In doing this the designer can learn how to structure his search and how to proceed to the next step from the context. The solution only needs to be "satisficing" (Simon 1969), and it is known that it is just one of many possibilities. In design as *problem-setting,* design is a process of collective inquiry and search taking place through transactions and conversations among several actors with mixed interests in the problem at hand. What needs to be created is what the problem solving view takes for granted, and problem representations are not context-free but largely context-sensitive (Lanzara 1983). Design, thereby, becomes a reflective conversation with the materials of the situation (Schön 1992, Winograd 1996 chapter 9).

Design as dealing with emergent goals

Extending on the thoughts of Lanzara (1983) and Schön (1983, 1992), de Figueiredo and da Cunha (2007) and Gasson (2006) discuss a fourth view to represent a contemporary understanding of design processes. In design as *emergent evolutionary learning,* design is the continuous convergence of problem-understanding and solution-proposition through a cyclical process of learning about a situation and responding through design that is deliberately short-term and partial. This view resonates strongly with Schön's later notion of design as a continual reflective conversation with materials (Schön 1992, Winograd 1996) but emphasises to a larger degree than Schön the dynamic and emergent nature of the *context and goals* of a design and not only of the design itself. The process of design is still rooted in a process of collective inquiry and search, but the notion that structure is inherent in a situation, as assumed in design as problem-solving, is explicitly rejected. Instead, contexts are seen as inherently dynamic and evolving, hence requiring design to be a continuously evolving process too, with an ongoing focus on *both* changing form and changing context. Not only are the problems unclear at the start of the process, so are the *goals* of the design (de Figueiredo and da Cunha 2007). However, as partial solutions are explored, an understanding of the problems, and appropriate design goals, emerges (Gasson 2006).

The view on design as dealing with emergent goals captures, to a large degree, the essence of Christopher Alexander's most recent work on "The Nature of Order" (2002-05). Here, Alexander goes a step beyond the participatory and reflective approach outlined in "A Pattern Language" (Alexander et al. 1977) and "The Timeless Way of Building" (Alexander 1979), and promotes a holistic process of design unfolding where both the context and the form are continuously evolving through a step-by-step morphogenetic process of *wholeness-extending transformations*, that are each intentionally short and open ended. The central contribution of this work of Alexander's in relation to Simon (1969) and Schön (1983) is, in my opinion, that it helps transcend the whole notion of a "problem", and instead makes us focus on the "desiderata", or "that-which-is-desired" (Nelson and Stolterman 2003) in the broader facilitation of human life in all its aspects and richness. Rather than setting and solving problems, Alexander's wholeness-extending design view allows us to treat design as a matter of creating new desired practice, and of dealing with the sometimes ill-defined goals emerging from the design process itself. It may set and solve a problem, but it may also just enable humans to do activities that they couldn't do before – on smaller or larger scales of importance for their life – like streaming a movie or photo from their mobile to their TV, or seeing on their mobile what friends and family on a different continent are up to at the moment. These are interactions that are hard to describe well as "solutions to problems" but easy to describe as meaningful and desirable extensions of a wholeness that was already there.

Alexander's notion of wholeness extending transformations and the view on design as emergent evolutionary learning reflect an evolution of the design discipline over the last three decades, in correspondence with a larger shift in society's world view from one of positivism towards one of constructivism (de Figueiredo and da Cunha 2007, Alvarez and Kilbourne 2002). This evolution of design and shift in world view is of particular interest for contemporary research in the design of mobile interactions because it has parallels with the shift in focus beyond an individual user's interaction with an individual device for a well defined activity, towards the wholeness of mobile technology user experience in indefinable and unbounded activities within the larger, and dynamic, contexts of emergent everyday life.

Looking critically at the predominant design approaches within the area of human-computer interaction, discussed in section 2.3, one could argue that these all fall within Lanzara's three design views, and none of them within the emerging holistic perspective.

User-centred design in the form of usability engineering and Contextual Design are essentially about problem-solving as descried by Lanzara (1983). It is about developing simplified models of the real world, and thereby "bounding" the problem until it becomes sufficiently well-defined to be resolved. Solving the problem is then done by evaluating alternative solutions until one that fulfils a set of criteria is discovered. Hence, design is a rational search process, and solutions are never considered universally optimal, but merely "satisficing" (Simon 1969). The participatory design methodology is different from the usability engineering and Contextual Design approaches to user-centred design in that it particularly addresses the problem-setting aspect of design, or at least aims to do so. Corresponding to Lanzara's (1983) collective inquiry through transactions and conversations among several actors with mixed interests, Participatory Design assumes that there are different perspectives on what the problem is and how it should be solved (Bratteteig 2007), and seeks to uncover these through active involvement of users and other stakeholders in the design process. Hence, design is a reflective conversation with the situation (Schön 1983, 1992), and solutions are unique and appropriate. They are shaped by the shaping of the situation.

Technology-centred design can to some degree be described as functional-analysis in that complexity is reduced by applying scientific reductionism to the (technical) problem at hand. However, in light of Simon's (1969) notions of bounded rationality, rejecting the idea of one rationally optimal solution, it is fairer to describe most technology-centred design as matters of problem-solving. Like usability engineering it is about reducing, or "bounding", the problem by taking ill-structured problems and reducing them into well-structured ones through inductive abstraction, rather than rational decomposition, and then exploring a range of possible solutions essentially through a process of trial-and-error. Unlike usability engineering, however, the initial ill-defined problems are not related to the use of technology, but to the functioning of it.

Design as continual convergence

What is then really missing from the current palette of design methodologies in interaction design is approaches corresponding to the view on design as wholeness-extending (Alexander 2002-05) or emergent evolutionary learning. Such "post–Simon" and "post–Schön" design approaches would take a step further and explicitly subscribe to the view that most activities are unbounded and situated in dynamic contexts, and that the relationship between context and form is therefore a continually changing one requiring that design is inherently cyclic, able to deal with emergent and changing goals, and about construction of context as well as form. They would retain the concepts of "satisficing" design from Simon (1969), though viewed as a much more ephemeral quality, and design as reflection-in-action from Schön (1983), but with added emphasis on designing-through-doing and explicit cyclic exploration of partial solutions as ways of converging requirements and solutions (Alexander 2002-05, Turner 1987). Rather than a rational *scientized* way of interaction design, this would truly be a *designerly* way: rhetorical, exploratory, emergent, opportunistic, abductive, reflective, ambiguous, and risky (Cross 1999). True to its own epistemology and practice, and justified by the efficacy of its results rather than the rigor of its methods (Archer 1992).

Instead of emergent evolutionary learning I will express such a design approach as one of continual convergence of form and context, and claim that apart from also being increasingly problem-setting, rather than merely problem-solving, mobile interaction design should embrace a designerly way of dealing with dynamic contexts and continuously emergent goals and forms. In my use of the term "form", as discussed earlier, like Alexander (1964) I refer to the unity of shape, look, function and content that is formed through design. To extend on the line of thought by Greenbaum and Mathiassen (1990), there is nothing wrong with setting and solving problems when designing mobile interactions. But a focus on the continual convergence of form and context helps shift emphasis towards the larger organic wholes, or digital ecosystems if you will, that interactive mobile systems and services have become a part of.

4.2. World views, root metaphors, and modes of inference

The four views on design as functional analysis, problem-solving, problem-setting, and continual convergence of form and context described above, and the shift from technical rationality to reflective practice, or even artistry, can be further conceptualised and understood by applying the theoretical lens of philosopher Stephen Pepper (1942).

According to Pepper (1942) there are four distinct world views, or root metaphors, through which we can understand the world: *formism, mechanism, organicism,* and *contextualism.* The root metaphor of formism is similarity. Taking a formist view, we seek to understand what the world *is like* by indentifying similarities and differences between things and placing them into meaningful categories. Formist analysis describes either the similarity or differences between two objects, or describes the form that an object exemplifies. The root metaphor of mechanism is the machine. Taking a mechanist view we seek to understand *how things work.* It is assumed that the whole is equal to the sum of the parts, and we try to understand how it works by decomposing complexity into individual parts and looking for cause and effect. Organicism's root metaphor is organic development. The organic perspective is concerned with the coherence between the parts and the whole and tries to understand *how it develops.* Unlike mechanism, organicism does not consider the whole to be simply the sum of the parts, but views the whole as primary and individual parts, only meaningful in relation to this whole. Instability is an inherent characteristic of an organic system. Thus change is given, and it is stability that needs to be explained. Contextualism, as the last of the four, is based on the root metaphor of act-in-context. Taking a contextual view we are concerned with seeing the world in its complexity of context and the need to continuously adapt to its unpredictability and chance happenings, and seek to understand *how this is happening.* The contextual view sees the world not as forms or machines but as ongoing acts that are inseparable from their history, current context, and threads into the future. Like organicism it views the world holistically rather than as a sum of individual parts, but it is more pre-occupied with the active present. Looking across these four world views, formism and mechanism can be described as analytic types of theory (i.e. reductionistic) whereas organicism and contextualism are synthetic (i.e. holistic) types of theory (Pepper 1942). Formism focuses on the concrete, mechanism on laws and principles, organicism on relationships, and contextualism on the contexts in which phenomena occur.

Using Pepper's terms, the shift from viewing design as a matter of functional analysis and problem solving towards a matter of problem setting and continual convergence of form and context is a shift from *formist* and *mechanist* visions of the world to a world view of *organicism* and *contextualism* (de Figueiredo and da Cunha 2007 p. 66). It is a shift from a reductionistic to a holistic paradigm of design, echoing design research thinking subscribing to the view that the process of solving a design problem is identical with the process of understanding its nature. Similar to dealing with "wicked problems", understanding and resolution of a problem are concomitant (Rittel and Webber 1973 p. 161-162) and the designer's ideas for solving a problem influences what information is needed to understand it (de Figueiredo and da Cunha 2007 p. 66). Where formism and mechanism, like science, seek to isolative phenomena from the complex situations they are embedded in and extract generalizable principles (Archer 1992), organicism and contextualism, like design, seek to *embrace* complexity and conceive non-universal and "ultimate particular" results (Stolterman 2008, Nelson and Stolterman 2003), purposely embodying a selection of values determined by their context. Applying organicist or contextualist visions to the activity of design, it becomes a matter of viewing the world holistically, looking at the coherence between parts, seeing it in its complexity, and adapting to its unpredictability. Organic and contextual design are perpetual convergences of solutions and problems – continual convergence of form and context.

Pepper (1942) emphasises that there is a strong tendency to combine contextualist and organicist views, but that there are still some notable differences. Whereas organicism is related to the philosophy of idealism, and explicitly *rejects* reductionism (i.e. formism and mechanism), contextualism is closely related to the philosophy of *pragmatism*, in which positivist principles are accepted, if pragmatically applicable (de Figueiredo and da Cunha 2007). In terms of practice, this means that most non-reductionist design tends to be better characterized as informed by an underlying contextual rather than organicist thinking simply because it typically involves a less rigid – and more pragmatic – view on which principles, techniques, and theories are "allowed" in the design process, as long as it confirms to an overall focus on the contexts in which the design is situated. In pursuit of interdisciplinarity in mobile interaction design, which obviously requires embracing principles, techniques, and theories from other disciplinary areas including ones grounded in positivism, this renders the contextualist (pragmatist) world view most useful.

Design as abductive thinking

When taking a contextual approach to design as proposed here, we also implicitly subscribe to the philosophical worldview of pragmatism, and, at least the essence of, the thoughts of its founding thinkers, such as William James, Charles Sanders Peirce, John Dewey and George Herbert Mead. Pragmatism is a philosophical tradition concerned with the interplay between theory and practice. Of particular interest for design research, Peirce's work on the logic of science explores the issue of idea generation, and promotes a certain mode of inference for explaining this. According to Peirce, new ideas come into being through "logical leaps of the mind", which reflects a third mode of logical reasoning, different from the more conventional forms of deductive and inductive logic. He named this third form of reasoning *abductive* logic. In abductive reasoning, rather than seeking the effect given the cause and rule, or seeking the rule given the cause and effect, we know the rule and the effect but are looking for the *cause*. Through abductive reasoning we create hypotheses through “inference to the best explanation” (Harman 1965) or guided by our “guessing instinct” (Peirce 1931-58). Peirce came to this viewpoint through a fascination with the origin of new ideas, and the observation that they arose when thinkers encountered data that didn’t fit with established models (Martin 2009). The first step in reasoning from here is not observation but *wondering*, and then imagining what could possibly be true. During this process, new ideas emerge as we collect, combine and organise our thoughts in different ways.

Looking at the activity of design *in a designerly way*, as discussed above, it is clear that we are dealing with Peirce’s world of abduction: actively seeking new data or signs of effect, challenging accepted explanations or rules, and inferring possible new worlds or causes (Martin 2009). Design *is* abductive (Cross 1999, 2011). It is not about predicting an effect or composing a rule. It is about suggesting possible causes (design) that will create an effect given what we know. Rather than trying to deduct “what is”, or induct “why something is”, designers are seeking to propose “what might be” (Martin 2009).

As discussed by Martin (2009), a fair critique of abductive thinking in design is, of course, that it does not guarantee success but might possibly lead to poor results – which is why it often scares management and others thinking in a non-designerly way. Therefore, abductive thinking in design should not exclude other forms of reasoning, and should not be used as an alibi for basing everything on intuition and guesswork. But at the same time, as pointed out by Cross, design *is* risky (Cross 1999, 2011), and *not* taking risks in design is more likely a guarantee of failure.

The discussion above is summed up in table 2, outlining the approximate alignments between views on design, interaction design approaches, epistemological positions, ways of thinking and acting, philosophical world views and modes of inference. Table 2 illustrates that contextual interaction design is highly compatible with Schön’s notion of reflective design. The key differences lie in the distinction between Lanzara’s (1983) category of design as *problem setting* and Alexander’s notion of design as *wholeness extending.* This leads to a difference in epistemological position between seeing design as essentially about *thinking*, or about *creation*. The former focuses on the *reflective practice* that happens in design, while the latter focuses on the *continual convergence.*

Table 2. Summary of views on design and associated underlying paradigms and world views.

Design viewed as	Associated with	Interaction Design approach	Epistemological position	Way of thinking and acting	Philosophical world view	Mode of Inference
Functional analysis	Taylor (1947)	Technology-centred	Technical rationality	Scientific and reductionistic	Formism and mechanism	Deductive and inductive
Problem solving	Simon (1969)	Technology-centred User-centred				
Problem setting	Schön (1983)	Participatory	Reflective practice	Designerly and holistic	Contextualism and organicism	Inductive and abductive
Wholeness extending	Alexander (2002-05)	Contextual	Continual convergence			

With these conceptual foundations in place, and the suggestion to take a contextual approach to designing mobile interactions driven by abductive "leaps of the mind" and designerly ways of knowing, thinking and acting, I will now take a brief look at some examples of contextualism in related areas of research and design.

4.3. Contextualism in research and design

The contextual world view has had an explicit presence in the disciplines of information systems and architecture for several decades. In information systems research, Pettigrew (1985, 1987, 1990) has promoted the use of contextualist theory and method for the study of organizational change involving a focus on *content, context* and *process*. Like design, Pettigrew argues that *change* is a phenomenon that is ill-understood in isolation from its context, and that explanations of it are bound to be holistic and multifaceted. Therefore, rather than limiting studies of change to single events or discrete episodes separate from their past, he seek to capture the holistic and dynamic nature of this phenomenon by encouraging research that is contextual in character (Pettigrew 1990). This involves analysis that is both vertical and horizontal – looking at the phenomenon at multiple levels and in perspective of past, present and future time. As exemplified in the work of Mathiassen et al. (2012), Pettigrew's contextualist inquiry framework is a useful approach to the study of change that involves the design of new technology. Framing such study in Pettigrew's terminology, the *content* of change is the transformation from one technology to another (Mathiassen 2012), the *context* of change is the organizational setting on multiple levels including its history, and the *process* of change is the way by which the new technology is developed and deployed. Thinking about this from a design perspective, it is interesting to note that the concepts of context and process are similar, but what is considered the *content* of IT-enabled change in Pettigrew's framework equals, or at least has a large overlap with, what is considered the *form* of interaction design, in Alexander's use of the term. They are both concerned with the introduction of new technology and its influence on people and use – and herein lies exactly the link between a contextual approach to information systems research and a contextual approach to interaction design, as illustrated in figure 14. Applying Pettigrew's (1987) contextualist framework to interaction design emphasises the importance of developing interaction design *processes* or approaches that can inform and stimulate the interplay between form and context. The design of mobile interactions as a continual convergence of form and context is an example of such an approach.

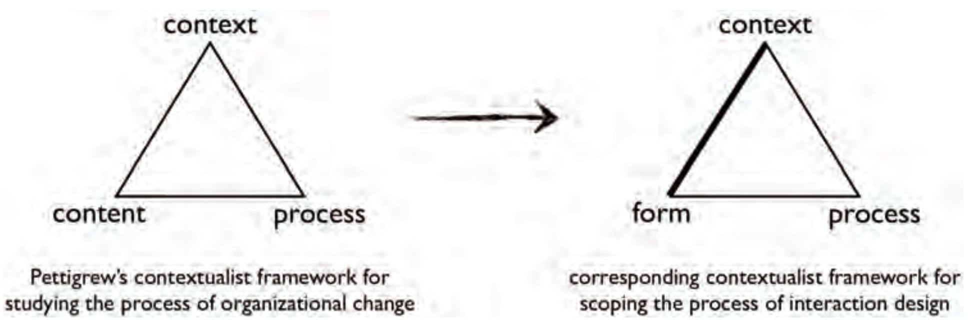


Figure 14. Towards a contextualist framework for interaction design, based on Pettigrew (1987).

Another key insight for interaction design to be drawn from Pettigrew's contextualist framework is the notion of vertical and horizontal analysis. Vertical analysis as an acknowledgement of the contextually multi-layered nature of any design challenge spanning, at least, from detailed human-computer interaction, over user experience, to wider use in context of work or life. Horizontal analysis as a recognition that, like change, design is always situated in a historical context. It has a past, a present and a future, and in doing interaction design we need to have a sensitivity for the fact that new designs do not appear in a vacuum, but almost certainly replace or add to something that was already there. By doing that, design inherently changes its own context and thereby the preconditions for future design.

Contextual architeture

As Schön points out, architecture is "perhaps the oldest recognized design profession and, as such, functions as a prototype for design in other professions. If there is a fundamental process underlying the differences among design professions, it is in architecture that we are most likely to find it" (Schön 1983 p. 77). Some movements within modern architecture particularly emphasize the importance of matching buildings to their surroundings. This design philosophy, known as "contextual architecture" (Brolin 1980, Ray 1980, Shane 1976), has given rise to several highly acclaimed buildings around the world praised by their inhabitants for the way they fit naturally with their surroundings. Apart from having a notable effect on its *outcomes*, working within this design philosophy also has some profound impacts on the *process* of design. Architects working closely with the context of their buildings spend significant amounts of time developing and assessing their design *on the building site* rather than at the drawing board in their studio. As an example, it is a well-known fact that the Danish architect Jørn Utzon, who is probably best known for the Sydney Opera House, spent considerable time on building allotments exploring their contextual properties before and during the development of his building designs. In a rare interview he even described how he would sometimes map out the possible location of walls and windows by placing lines of small rocks on the ground, and then walk around imagining the view of the surrounding environment from these as yet un-built rooms. In an account of the works of Alex Popov (a Sydney-based contextual architect and former associate of Utzon) it is described how the result of buildings created with such sensitivity to the way they engage with their surrounding environment is that they do not just fit their context well, they themselves become part of that evolving context (McGillick and Carlstrom 2002).

In architectural theory, contextual architecture is described as a matter of pursuing the notion of *genius loci* (the protective spirit of a place in classical Roman religion) by responding to the topographical, geographical, social and cultural context of a building site (McGillick and Carlstrom 2002). This concept is most notably explored by Norwegian architect and theorist Christian Norberg-Schulz in his phenomenology of architecture (Norberg-Schulz 1980) arguing that genius loci – or sensitivity to context – has profound implications for place making. Echoing this line of thought, Christopher Alexander argues that architecture exhibiting a quality of "timelessness" always evolves through a series of "wholeness-extending transformations" (Alexander 2007) in which the designer has not only focused on the creation of new form but done this with deep understanding of and respect for the existing context. Alexander's Pattern Language (Alexander et al. 1977) contains a collection of form-context pairs composed to help evoke the readers' imagination of a future design in context – enabling such wholeness preserving contextual architecture in practice. Such quality of design is also described as "ensoulment" by Nelson and Stolterman (2003 p. 285) who writes that "ensoulment is about wholeness and composition, as well as value and meaning" and that "to ensoul a design – in a way that attracts attention and appreciation – demands a respect for the materials, the structure, the shape and its social dimensions" (i.e. its context).

Over the course of more than four decades of empirical work, Christopher Alexander's view on design has evolved considerably towards a holistic and dynamic perspective of the world, resembling the vertical progression outlined in table 2. From the borderline technical-rationality thinking about the interplay between form and context expressed in "Notes on the Synthesis of Form" (1964), over the participatory and reflective approach outlined in "A Pattern Language" (1977) and "The Timeless Way of Building" (1979), to the continual transformation of wholeness expressed in "The Nature of Order" (2002-05). Using the terms of Pepper (1942), the bulk of this work is an exemplar of contextual thinking – probably even bordering on organicism in its latest propositions. It reflects a designerly way of thinking and acting dominated by abductive reasoning, deals with the setting of the problem and not just its solution, and it emphasises that design is about the continuous construction of wholes that amount to more than the sum of their parts. This makes Alexander's thinking particularly relevant for a contextual perspective on designing mobile interactions because it resonates with the emerging need for a perspective transcending user- or technology- centeredness and capable of informing the design of digital ecosystems and holistic user experiences rather than just single devices, systems and interactions.

The nature of order

Alexander's theory about the nature of order (Alexander 2002-05) is a complex piece of design philosophy. In this thinking, design is a matter of pursuing a quality that creates a deep subjective feeling of connectedness in people in its presence, but that we do not have an established name for. Alexander himself calls this quality "life", "wholeness" or "living structures", and argue that the degree of life in things is an objectively observable quality that can be measured empirically. He also argues that the degree of life in things is correlated with the repeated appearance of fifteen empirically identified properties[3]. These are seen in human-made design with a holistic quality, but also *in natural systems*, which, Alexander speculates, may be why people respond emotionally and cognitively positive to them when encountered in artefacts. The feeling of life in a natural thing or in a design stems from an experience of *its whole* and not simply from the sum of its parts. In fact, according to Alexander, it is the whole that defines the parts and give them meaning, and not the other way around. "The flower is not *made* from petals. The petals are made from their role and position in the flower" (Alexander 2002b p. 87). Instead of individual "parts", wholes may consist of smaller entities, at different scales, each with their own

3) Levels of scale, strong centers, boundaries, alternating repetition, positive space, good shape, local symmetries, deep interlock and ambiguity, contrast, gradients, roughness, echoes, the void, simplicity and inner calm, and not-separateness.

localised quality of wholeness. Such sub-wholes are denoted as “living centres”, which can be described as interrelated focal points within the larger whole, each reflecting “an organized zone of space, which because of its internal coherence, and because of its relation to its context, *exhibits centeredness*” (Alexander 2002b p. 84).

The value of the concept of living centers is that it captures the main features that make a difference for our experience of the world, and contributes to its wholeness. In terms of design it thereby becomes a possible means of navigating the challenge of achieving wholeness by providing cues about focal points within it. What is important to notice here, however, is that following the holistic mindset, the centers that make up a given wholeness do not exist independently, but only become centers as a result of the configuration of the whole. Hence, one cannot simply “break down” an overall design challenge into design of individual centers, but have to maintain a simultaneous focus on the whole as well as the parts.

Alexander’s perspective on good design as “living” wholes of form and context has some profound implications for the *process* of design. In terms of process, Alexander emphasises that new design with the quality of wholeness or life never just appears out of nowhere but always *evolves* from a previous state of wholeness, which initial qualities it is able to maintain or expand. Methodologically, this is reflected in the principle of “structure-preserving” (Alexander 2002b) or “wholeness-extending” transformations (Alexander 2007), through which living form-context ensembles gradually evolve, or “unfold”, over time (figure 15). The unfolding of wholes through wholeness-extending transformations builds on the fundamental view of Alexander’s that *future* design wholeness is already *latently present* in current wholeness, and that designing therefore has to be a process of step-by-step adaptation of form and context towards increasing quality and complexity, rather than a matter of defining a desired end-state up front, and then setting out to produce this efficiently and with little or no change. Wholeness and life is, according to Alexander, simply not something that can just be defined or specified and then built. Like generated structures in nature it inherently has to emerge from a growth process of modification and adaptation that happens gradually in response to feedback about “the extent to which an emerging structure supports and embellishes the whole” (Alexander 2002b p. 230). Or as more pragmatically expressed by Moggridge (2007), in design “whatever you come up with will automatically build on the past”, and hence “if you are designing a new version of something that already exists, “state-of-the-art” is the most useful starting point” (Moggridge 2007 p. 728).

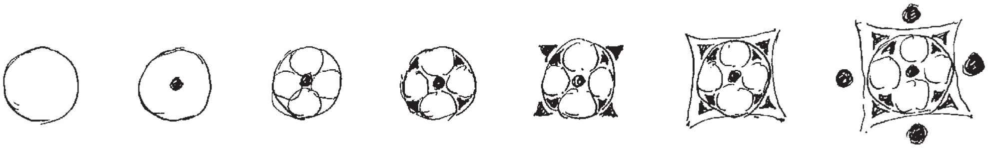


Figure 15. Six wholeness-extending transformations (left to right) (Alexander 2002b p. 52)

While Alexander provides lengthy and detailed discussions about the properties of holistic and living processes and incites a process *"which can support the continuous creation of an emerging living structure in the world"* (Alexander 2002b p. 508), he does not himself present a straightforward operational model or methodology of what such process may look like in practice. Hence, the question remains about how exactly we may procedurally go about achieving wholeness in design?

Overall, Alexander states that the types of processes that are capable of intensifying wholeness are the ones that place more emphasis on the context, and, in doing so, encourage the use of wholeness-extending transformations, and the creation of living centers. He also states that such processes must be "morphogenetic", or "architectural", meaning that they *create coherent form in the world* and explicitly emphasise this form-creating aspect of the process – that is its "designerly" nature, to use Cross's (1982) term. Thirdly, he highlights that in order to create unpredictable (i.e. non-trivial) or unexpected outcomes (Nelson and Stolterman 2003), the design process must be open-ended and itself partly unpredictable in order to truly accommodate for unforeseen adaptation and unfoldings of form and context. According to Alexander, such open-ended, unpredictable, morphogenetic processes, may only be achieved through step-by-step adaptation with appropriate feedback mechanisms for continually assessing the outcome of the process so far, and informing its immediate future direction. *"The process must go gradually, in a way that allows assessments, corrections, and improvements to be made about the degree of life which occurs throughout the structure, at all scales and at all levels. This process must occur continually throughout the conception, design and construction"* (Alexander 2002b p. 237). Hence, the designer needs to be able to shift focus and technique *as needed* whenever in the process – go back to the drawing board, study an aspect of the design in more depth, build and test a prototype, analyse data from a new perspective – instead of being confined to a predefined sequence of activities. A process like this is inherently contextual, iterative and multidirectional. It views the world as inseparable from its history, current context, and threads into the future. It unfolds stepwise but can at any point go in any direction, and it has no predefined starting or ending points.

4.4. Elaborating on user-centred design

In light of these perspectives on design, I will now turn my attention back to the prevailing user-centred design (UCD) approach to interaction design, and respond to some of the shortcomings of this approach discussed in contemporary interaction design research.

As described earlier, user-centred design typically follows a cycle consisting of four stages in which we study, design, build and evaluate technology, as depicted in figure 16. Other labels exist, but essentially they denote the same activities (Harper et al. 2008).

Figure 16. The traditional, four-stage UCD model (Harper et al. 2008)

One of the shortcomings of the UCD model is that it is almost too simplistic and high level. It does not say much about the outcomes from each stage, what they are, and how they differ. It is also unclear how the iterative nature of the model actually works: how different iterations can take different shapes and have different outcomes, and how continuous iteration is cumulative and going forwards towards better outcomes, rather than just going around in a circle. Looking at support for how to iterate through the stages it is also unclear why the activity of "evaluating" is followed by the activity of "studying". If working truly iteratively by designing on the basis of user studies and then evaluating prototype solutions by putting these back in the hands of prospective users, then aren't evaluating and studying in fact two variations of the same type of empirical activity? Finally, it can be questioned if a predefined cyclic sequence really depicts the quintessence of well-functioning interaction design processes, or if such processes are in fact, or ought to be, far more elastic and irregular?

As a response to some of these shortcomings, in the recent Microsoft research report "Being Human: Human-Computer Interaction in the Year 2020" (Harper et al. 2008) it is proposed that the traditional four-stage UCD model is extended with an additional stage in order to better accommodate for "third wave" (Bødker 2006) HCI research that focuses on human values and shaping society's new relationships with emerging ecosystems of computer technologies (figure 17). The additional stage is labelled "understanding" and is placed into the model as an initial activity of conceptual analysis to "focus on human values and to pinpoint those that we wish to design for and to research" (Harper et al. 2008 p. 59). It is a stage meant to involve specifying what kinds of people are the focus of a particular project, and understanding their domains of activity, cultures etc. which will in turn "either point to some fundamental research which needs to be carried out in Stage 2, or will provide guidance toward relevant research which has already been carried out" (Harper et al. 2008 p. 59). While it is noted that understanding a problem is traditionally a part of studying, it is proposed that "it be elevated to become a more explicit process, where the various human values at play are thought through and the trade-offs are examined in a systematic way" (Harper et al. 2008 p. 58).

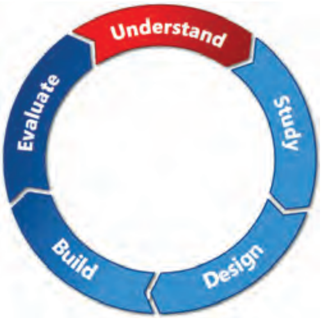


Figure 17. The extended, five-stage UCD model (Harper et al. 2008)

The motivation behind this extended approach to HCI research and design is sound. As we are moving towards the third wave of HCI research, understanding the contextual richness and width of the people, activities, cultures etc. that we are designing for is hugely important. Hence, at first glance, I myself embraced and promoted the inclusion of understanding as an additional step in the UCD model. However, after having thought about it for a while, the proposed five-stage model is, in my opinion, actually a bit problematic and raises more confusion to me than clarity.

One of the problems is that it separates further the already overlapping activities of evaluating and studying with a stage that essentially doesn't fit there. How can you really understand something before you study it? Another problem is that the proposed new stage to "understand" does not type-match the other four stages in the model. Studying, designing, building and evaluating are all *activities* – something you can say to be doing. Understanding is not. It is the *outcome* or *purpose* of an activity. Thereby the proposed extended model now confuses the "how" and the "what" of the process of interaction design. While it might make perfect sense to suggest that the first stage of a design or research project should be to "understand" something, the question that needs to be raised must necessarily be *how* this understanding is obtained? My suggested answer to this is simple: *study and analyse.* If we wish to understand the broader contexts of an interaction design challenge better, which I agree we should, then it is a matter of specifying *what* should be studied and analysed, and *how* it can be done in order to obtain the desired insight.

In response to this, and informed by the discussion of design research, I suggest three alternative changes to the traditional UCD approach:

1. Separating and redefining activities and outcomes
2. Shifting the gravity point away from user-centeredness
3. Making the process flexible and unpredictable

Separating and redefining activities and outcomes

My first suggestion for change is about activities and outcomes. In the alternative modification of the traditional UCD model depicted in figure 18, three key factors are changed. Firstly, the activities and the outcomes are separated. Activities are depicted as arrows in a circle while the outcomes are depicted outside the circle towards the top or bottom. This explicitly highlights the dual-purpose of interaction design research and practice: it is about creating *understanding* and *artefacts.* Understanding is the result of the activities of studying and analysing, and artefacts (i.e. interactive systems) are the result of the activities of designing and building. Where understanding constitutes the foundation on which artefacts are designed and built, conversely artefacts, whether they are the designer's prototypes or already existing ones in the use context, constitute the foundation on which understanding is created through studying and analysing.

Secondly, the activities of evaluating and studying have been merged into one, reflecting the view that evaluating is in fact a *type* of studying and that differentiating between the two creates an unclear intersection between them when looking at the overall iterative process. Thirdly, the activity of *analysing* has been added as an explicit stage between studying and designing. This reflects the view promoted by Harper et al. (2008) that we do indeed need to elevate the importance of developing theoretical and conceptual understanding of the problem at hand in all of its richness and detail, including issues such as human values, context of use, and user experience. However, contrary to Harper et al.'s (2008) extended UCD model, this is reconceptualised here as an explicit theoretical activity of *analysis* closely related, but subsequent, to the empirical activity of studying. This activity *leads to* understanding.

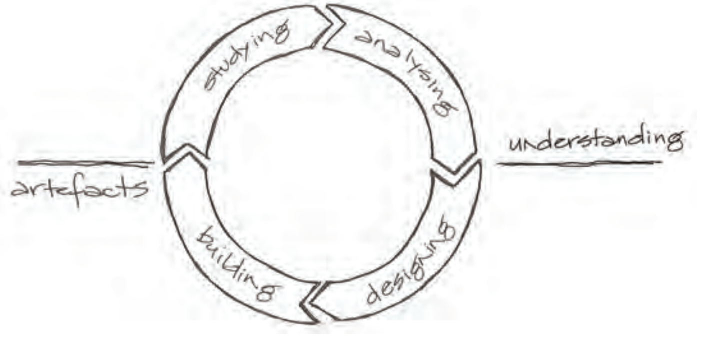


Figure 18. Revised and extended four-stage UCD model. Understanding and artefacts represent the two primary types of outcomes produced during the top and bottom two activities

Shifting the gravity point

My second suggestion for change is about focus. In response to the ongoing debate about UCD being inadequate for informing the creation of novel and innovative interactive systems due to its built-in orbit around *users,* my second suggestion is to change this fundamental point of gravity towards one that better captures the essence of the emergent challenges at hand. In his 2005 Interactions magazine column, "Human-Centred Design Considered Harmful", Donald Norman (2005) already alluded to such fundamental change in gravity point by proposing *Activity*-Centred Design (ACD) as an alternative to UCD. Activity-Centred Design doesn't focus on user goals or preferences but instead more broadly on what they are doing (Saffer 2007). Although this suggestion is a step in the right direction, making "activity" the centre of orbit still doesn't quite suffice - at least not in the case of *mobile* interaction design. In order to better fit the challenges of designing mobile interactions, the gravity point, or unit of analysis, really needs to be shifted towards one with a wide enough scope to encompass the more extensive phenomena of *contextual user experiences* in a holistic way. Activity is a part of this, but equally so are other factors such as settings, people, artefacts, technologies, time, and more importantly the contextual whole that is made up by all of these. A similar point of view is made by Bill Moggridge (2007) in a critical self-reflection on the scope of the people centred prototyping approach used at IDEO, when he asks: "Is this focus on people and prototypes enough? Can we rely on just those two simple strategies to create excellent designs? I'm afraid not, as the constraints will come from the *full context* of the design problem, *not just the people.* (Moggridge 2007 p. 725, italics added). He continues, "you will need to understand as much as possible about everything that will affect the solution (…) find out as much as [you] can about the *context*" (Moggridge 2007 p. 726, italics added).

In response to this, and echoing Alexander's (2002-05) views on design as a continual interplay between form and context discussed earlier, I suggest that a better point of gravity for mobile interaction design would be *the ensemble, symbiosis, or convergence of form and context.* This is illustrated in figure 19.

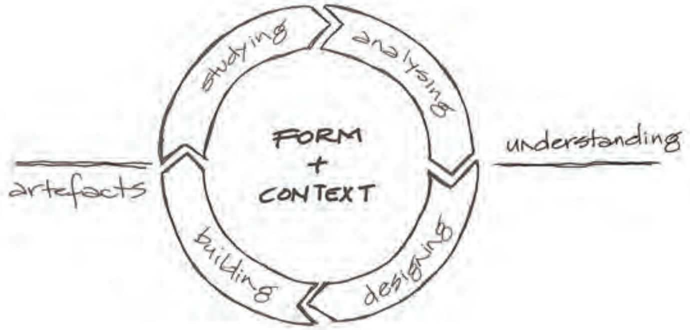


Figure 19. Shifting the gravity point to *form-context* syntheses rather than *users*

Shifting the gravity point of the underlying process model from users to form-context convergence echoes the need for emerging interdisciplinary research fields to change the unit of analysis towards one that transcends the individual contributing disciplines as discussed previously. As discussed in section 3, the methodological foundations and contributing disciplines of mobile interaction design are not yet integrated in an *inter*-disciplinary way, where researchers from distinct disciplines generate novel concepts and integrate different levels of explanation in joint pursuit of a more profound research problem impossible to grasp from one perspective alone. This is evidenced in the observation that current mobile interaction design is predominantly either technology- or user-centred, and hence to a large extent still resonate the classic disciplinary separation between science and engineering on one side, and art and design on the other. In relation to this observation it could be argued that the user-centred design approach, although originally developed to facilitate interdisciplinarity within systems development, in fact works *against* increased disciplinary integration, as much as purely engineering driven systems development, by strongly promoting one particular perspective rather than a holistic one in which technology and users are considered and weighted as equally important. If interdisciplinarity is to take place within mobile interaction design, this is not a viable course. Instead, what is needed is a paradigmatic shift stimulating new ways of conceptualizing and working by framing new issues, research questions and challenges *beyond* those of the individual disciplines involved. *Form-context convergence* is an interesting candidate for such higher-level unit of analysis.

Using this conceptual optic on the revised approach in figure 19, the activities of studying, analysing, designing and building are all about considering the ensemble of a particular form (an interactive mobile system) in relation to its context (users, technology, settings, activities etc.) from different perspectives, with the purpose of producing either understanding or artefacts depending on whether you are in the top or the bottom part of the circle. Consequently, the interaction design process becomes neither user- nor technology-centred but instead transcends and continuously includes both of these viewpoints within a broader perspective on the contextual whole in which users and technologies interact.

Putting form-context convergence in the centre also respond to the concerns raised by, among others, Rasmussen (2007) that the original intentions of human-centeredness, as envisioned in the 1970s, have in UCD degenerated to being a limited paradigm of creating user friendly computer interfaces rather than focusing on the broader societal issues facing human kind through "socially useful production" and "human machine symbiosis". While some of the thoughts behind human-centeredness can still be still traced in the user-centred design literature, the focus has become predominantly on how people interact with computers rather than on "how the technology can be shaped to support enrichment of human skills and socially useful products" (Rasmussen 2007 p. 475). According to Rasmussen (2007) and Gasson (2003) a similar tendency can also be observed within the discipline of Interaction Design when the discourse starts with the concept of computer-based technology and when designers "ignore the context of design as systems situated in physically and socially constituted environments" (Rasmussen 2007 p. 476). Rather than producing socially and holistically useful products, this results in incremental or "problem-closure" development of products framed by relatively limited tasks in isolation from the social world that surrounds them (Gasson 2003 p. 36). In reaction to this, Rasmussen (2007) proposes a "human-context centred" approach revitalising the original intentions of the human-centred tradition and promoting the fundamental view that "although human beings are important creatures in the world, they are still a part of a much larger context of natural and social relationships, in which they should try to act and interact in a sustainable manner" (Rasmussen 2007 p. 478). At the core of this proposal is the principle of dialectical thinking in order to overcome the weaknesses of differentiating between technology- and user-centred approaches. Rather than making such clean-cut distinction, or finding an optimal balance between the two, a possible *fusion* of the opposing interests and forces is sought by transcending to the higher-level unity of form-context convergence.

Making the process flexible and unpredictable

My third suggestion for change is about process. In light of the view that designerly processes of creation unfold stepwise through unpredictable sequences of assessment, corrections and improvements of form-context ensembles with respect to their larger whole, my third suggestion is to discard the view that interaction design should follow a predefined cyclical sequence of activities, like the one depicted in the traditional UCD model. Inspired by Alexander (2002-05) I instead propose a web process model that explicitly allows unpredictable, less orderly and more complex sequences of design, allowing the designer to shift focus and techniques as deemed necessary on basis of the continual consideration of outcomes so far. This is illustrated in figure 20.

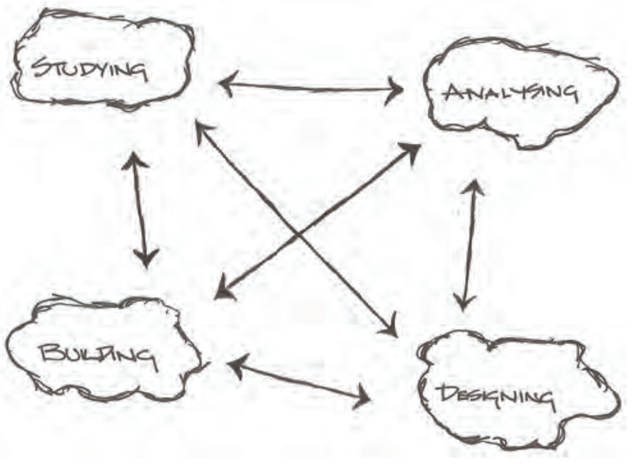


Figure 20. Replacing the predefined cyclical sequence of activities with a flexible web process model that allows unpredictable, less orderly and more complex sequences

The key difference between this and the traditional cyclical UCD approach is that it encourages a much more flexible and pragmatic way of dealing with the design challenge at hand. Rather than being stuck in a particular activity until it is “finished”, and then having to wait for a whole iteration before attending to this activity again, a web, or “stepping stone” (Stokholm 2008, 2010), approach allows the designer to jump freely and frequently between activities, thereby responding better to the emergent needs of the process, and letting the evolving form-context symbiosis control the process rather than the other way around. Clockwise circular sequences are of course not prevented, but they only happen when considered appropriate.

By taking a more flexible and unpredictable process like the one depicted in figure 20 as the methodological foundation for design, rather than a predefined cyclic one as prescribed by UCD, one would in fact support what often happens in real world interaction design practice, as described by Bill Moggridge (2007 p. 649-650, p. 729-730). One would recognise that the most successful design processes are usually out of sequence, apparently unstructured, and sometimes almost seem random – like the ball bouncing in unexpected directions inside a pinball machine, to use Moggridge’s own analogy – and acknowledge the knowledge amongst experienced designers, that the fastest way to achieve a successful design is to use the tools and techniques of the trade “quickly and repeated frequently, but usually not in the same order” (Moggridge 2007 p. 729).

5. THE CONTINUAL CONVERGENCE OF FORM AND CONTEXT

Having arrived at a higher level of insight into designerly ways of thinking and working, and having used this to respond to shortcomings of the prevailing user-centred approach, what then could an alternative, designerly and contextually oriented, approach to designing mobile interactions look like? In this section I will sum up my position on this based on the discussions and synthesis of thoughts and concepts in section 4.

The view emerging from my research is that as we move towards a society of increasingly widespread pervasive computing and digital ecologies, the design of mobile interactions accordingly needs to transcend existing approaches and develop a holistic and contextual view on technology design. My approach to doing this is to use the ensemble of form and context as our central unit of analysis and embrace a designerly way of achieving convergence between form and context through a *contextually grounded, wholeness sensitive, and continually unfolding process of design*. I describe such an approach to designing mobile interactions as one of *continual convergence of form and context*. This approach describes the dual-purpose unfolding of mobile interaction design research and practice, and ties together *empirical, creative, technical* and *theoretical* types of work and thinking that takes place within the design activities of *studying, analysing, designing* and *building* interactive mobile systems. It is neither user- nor technology-centred, but instead enables and encourages truly interdisciplinary research and design at the intersection between technology and liberal arts by transcending these two viewpoints.

My proposed approach to designing mobile interactions can be characterized and described by the following seven principles:

1. Emergence and unpredictability
2. Form and context unity
3. Form and context convergence
4. Oscillation between understanding and artefacts
5. Oscillation between concrete and abstract
6. Four types of design activity
7. Four types of design ripples

These principles are illustrated and described in the following.

5.1. Emergence and unpredictability

In designing mobile interactions the emergence of new artefacts and understanding happens by continually stepping between the four activities of studying, analysing, designing and building. The sequence of activities is not predefined cyclical but flexible and multidirectional, meaning that the designer can jump freely and frequently between activities as needed in response to the continual consideration of outcomes and emergent needs, challenges and opportunities. This allows the process of form-context convergence to be stepwise but unpredictable (Alexander 2002-05) and less orderly (Moggridge 2007), yet continual and cumulative. This is depicted in figure 21 where the upward spiral illustrates the cumulative emergence of artefacts and understanding as the design process unfolds.

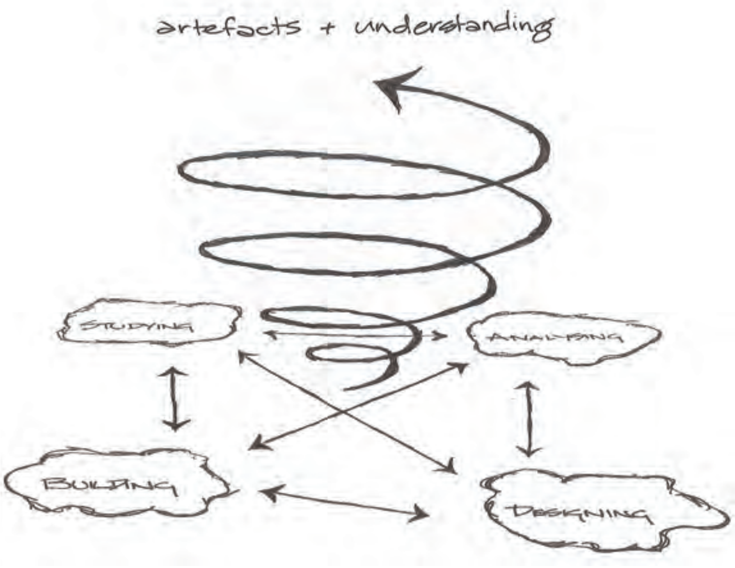


Figure 21. An unpredictable process towards cumulative emergence of artefacts and understanding

Departing further from the circular UCD models, this approach emphasises the evolving character of iterative research and design where each step through a particular activity is purposely different from the last time in terms of focus, scope and type of outcome. For example, the empirical activity of "studying" is different depending on whether it is an early activity of inquiring into the context of a design challenge, or if it is a later activity of inquiring into the user experience of a new interactive prototype system. Similarly, theoretical research involves different levels of analytical abstraction depending on how many steps we have been through, and leads to increasingly higher levels of understanding from each oscillation between the concrete and the abstract. As the process unfolds over time, designs are increasingly detailed and refined, and the artefacts we produce are created with increasing levels of fidelity and completeness.

This view on the interaction design process obviously takes some of its inspiration from Boehm's (1988) spiral model for systems development iterating through particular phases, emphasising that each iteration is different from, and builds on, the previous one. But it is different from Boehm's model in the sense that it does not to the same degree dictate a spiralling sequence of specific activities in detail, and that it traces the processes through very different types of stages. Also, whereas Boehm's model is a tool for managing a software development process through careful planning and risk management, the continually cumulative aspect of designing mobile interactions serves the purpose of illustrating that each step of studying, analysing, designing and building creates additional empirical, theoretical, creative, and technical insight and value in relation to the design that unfolds from the process.

5.2. Form and context unity

The basic unit of analysis in my view on the design of mobile interactions is the unity of form and context. Putting the unity of form and context in the centre echoes the contextual view that the world is inherently dynamic, that the relationship between context and form is a continually changing one, and that design is therefore about the construction of both of these in concert. This view essentially transcends user- and technology-centeredness. The synthesis of form and context is illustrated in figure 22 where these are depicted as an intertwined unity, creating a whole that is bigger than the sum of its individual parts.

Figure 22. Form and context as an intertwined unity

In the design of mobile interactions the unity of form and context provides a common reference point for the activities of studying, analysing, designing and building. Echoing the early work of Alexander (1964), "form" does not just mean shape, but is the unity of *shape, look, function* and *content* – i.e. the interactive system artefacts we design. Context is what defines or frames the situation in which these interactive systems or forms are deployed. Elaborating on Alexander's work, the design of mobile interactions is about actively designing not only the form but also the context – i.e. designing new use situations and practices.

5.3. Form and context convergence

The unpredictable and continual process of emergence and the unity of form and context are tied together in figure 23, providing an overall picture of my view on the design of mobile interactions. In this view on interaction design the unity of form and context is continually explored and refined by stepping between the four activities of *studying, analysing, designing and building* in an unpredictable order gradually leading to the emergence of new artefacts and understanding. Throughout this process, the designer can step from any of the four corners to any of the others, meaning that the sequence of activities is not fixed but flexible. The result of the process is the gradual and unpredictable emergence of form-context convergence.

Each of the four activities contribute to the unfolding of design and gradual convergence of form and context through "ripples" of *assessment, abstraction, exploration and synthesis* towards the specific ensemble of form and context in the centre. In stepping freely between activities, the interaction design process oscillates between producing *understanding* and *artefacts*, and between working with the *concrete* and with the *abstract.*

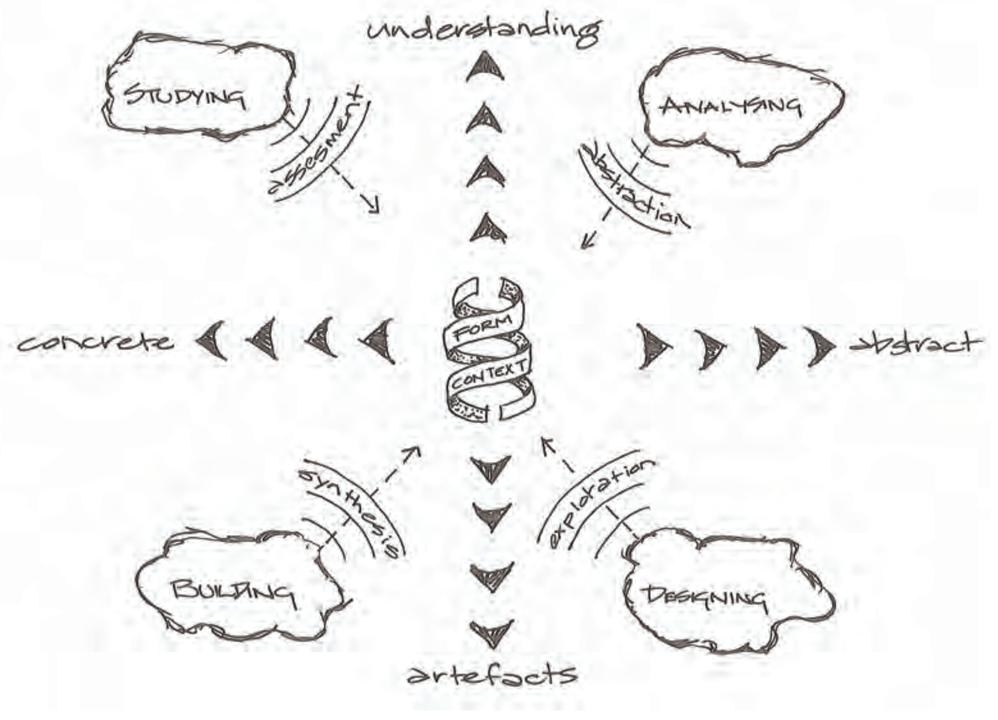


Figure 23. Form and context convergence

The proposed model above encourages a more flexible and pragmatic way of designing mobile interactions, while at the same time defining a shared focus for all activities and describing how they each contribute to the unfolding of the whole.

5.4. Between understanding and artefacts

The distinction between producing understanding and artefacts reflects the contextual view that design is a matter of constructing both problem-understandings and solution-propositions through a cyclical process of learning about a situation, and responding to it through suggestions for design. This is illustrated in figure 24 depicting a design process with oscillations of changing frequencies and amplitude, reflecting the continual and unpredictable sequence of activities, and a tendency towards increasingly mature understanding and artefacts.

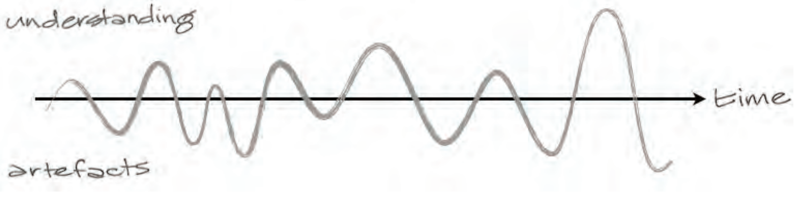


Figure 24. Oscillating between producing understanding and artefacts

In designing mobile interactions, understanding results from the activities of studying and analysing, and artefacts results from designing and building. Understanding is the foundation on which artefacts are designed and built, and artefacts are vehicles for creating understanding through study and analysis. Artefacts cover the range of tangible design products emerging from the process such as sketches, models, mock-ups, simulations, prototypes, and functional systems. On the opposite side, understanding covers the range of less tangible products from the process such as empirical data, personas, scenarios, models, concepts, frameworks, and theory.

5.5. Between concrete and abstract

Introducing a distinction into the form-context convergence model in figure 23 between working with the concrete and working with the abstract is inspired by related discussions of the process of interleaved research and design by, for example, Dubberly et al. (2008), IDEO (2009), Mendel & Yeager (2010) and Dubberly & Evenson (2011). In these discussions it is suggested that the two major phases of a design process of analysis and synthesis (respectively leading to understanding and to artefacts) both involve an orthogonal continuum between "the concrete work we inhabit or could inhabit" and "abstractions and models of what is or what could be, which we imagine and share with others" (Dubberly et al. 2008). It is by elevating our understanding from a concrete to an abstract level before attempting to move from understanding towards new form (i.e. interactive systems), and by exploring new form opportunities in the abstract before implementing and assessing interactive system artefacts in the concrete, that we are able to talk about a process as one of *design* as opposed to one of simple, or *unselfconscious* (Alexander 1964), form making. Bridging between understanding and artefacts on an abstract rather than a concrete level is also what facilitates the conception of solutions beyond incremental improvements to misfit through unreflective reaction. During the design process we continually shift between these two ends of the orthogonal continuum. This is depicted in figure 25, illustrating the same pattern of oscillations with changing frequencies and amplitude caused by the unpredictable sequence of activities.

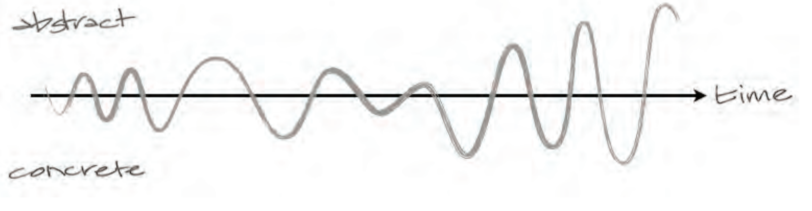


Figure 25. Oscillating between working with the abstract and the concrete

In designing mobile interactions, empirical and technical research takes place in the concrete end of the continuum, dealing with studying and building "what is" and "what could be". In the other end of the continuum, theoretical and creative research takes place in the abstract, analysing and designing what is and what could be. Introducing the concrete-abstract continuum into the model emphasises important differences between studying/analysing and designing/building, and the outcome from these activities in form of ripples of *abstraction* and *synthesis*, as discussed more in section 5.7.

5.6. Four types of design activity

The two orthogonal distinctions between creating understanding or artefacts and between working with the concrete or the abstract define a space involving four distinct types of work: empirical, theoretical, creative, and technical, and illustrate that the design of mobile interactions involves all of these. This is outlined in table 3.

Table 3. Four types of design activity

	Concrete	**Abstract**
Understanding	**Empirical** Studying: working with practice	**Theoretical** Analysing: working with concepts
Artefacts	**Technical** Building: working with prototypes	**Creative** Designing: working with opportunities

In the top left quadrant, mobile interaction design is *empirical*. Empirical work embraces the fundamental concept of modern science that insight must be based on observable evidence – or empirical data. When designing mobile interactions, being in the empirical quadrant means that we are working with the practice or actuality in which our products and solutions are supposed to fit and be used. This work is done through studies of real-world practice using empirical methods such as observations, probes and experimentation, and covers the empirical study of people, technology and context.

In the top right quadrant, mobile interaction design is *theoretical*. Theoretical work seeks to explain empirical phenomena in a consistent way, enabling us to understand and predict a given subject matter. When designing mobile interactions, being in the theoretical quadrant means that we are working with theoretical models and descriptions of our subject matter. This is done through analysing empirical data using theoretical frameworks and concepts derived either from previous research or produced through grounded theory or analysis.

Moving to the bottom right quadrant, mobile interaction design becomes *creative*. Creative work is a process by which a person creates something novel that is of value for other people, society, etc. When designing mobile interactions, being in the creative quadrant means that we are working with new opportunities for mobile computing inspired and informed by our empirical and theoretical insight and our knowledge about the potentials of technology. This is done by conceiving and refining original design ideas and solutions through an iterative process of designing potential artefacts and products making use of flexible and incomplete design instantiations such as sketches, models, and mock-ups, that are purposely suggestive, explorative, and even provocative, rather than descriptive, delimited, and definitive.

Finally, in the bottom left quadrant, mobile interaction design is *technical*. Technical work in this relation seeks to provide concrete instantiations of design propositions or solutions in response to opportunities, problems, challenges or needs. When designing mobile interactions, being in the technical quadrant means that we are working with prototypes of our proposed design ideas. This is done through the building of artefacts such as simulations, prototypes, or functional systems, investigating technical feasibility, and subsequently enabling empirical investigations of use quality.

5.7. Four types of design ripples

The distinction between four types of interaction design work naturally leads on to four types of design *ripples* emerging from these. These ripples, or pulses, describe what happens in the oscillations between the concrete and the abstract, and between artefacts and understanding. In designing mobile interactions they can be described as pulses of *abstraction, exploration, synthesis, and assessment,* happening respectively when analysing, designing, building and studying (figure 26).

Figure 26. Ripples in the design of mobile interactions

Abstraction happens when progressing towards abstract theoretical frameworks and concepts from, for example, concrete empirical data. It is about learning more from the world around us. In designing mobile interactions the purpose of this transition is to elevate the level of understanding about context beyond concrete observations and descriptive accounts, and distil insights that explain the observed phenomena. This is done by analysing our data, filtering it, prioritizing it, and organizing it, and in doing so moving further and further away from the concrete, and from artefacts.

Exploration happens when moving either from analysis or building to design. While going back to the drawing board informed by experiences with a prototype system is not particularly difficult, bridging the gap between understanding and artefacts is often perceived as the hard bit of the design process. This is because rather than elaborating on a previous stage with the same type of aim, it is an orthogonal direction of work with a completely different type of outcome. As described by Dubberly et al. (2008) this transition is what lies "at the heart of designing" – moving from analysis to design, from problem to solution, from the current situation to the future, from research towards prototypes. It is about opening up a design space for future form. In designing mobile interactions the purpose of this transition is to use our understanding of context as foundation for investigating and suggesting what the future might look like, or to change our design in light of new technical insight. This is done by speculating, hypothesising, and imagining possible futures in abstract form, and grounded in data and theory.

Synthesis happens when progressing towards concrete prototypes from, for example, abstract design. It is the opposite transition to abstraction and is about making things real. In designing mobile interactions the purpose of this transition is to manifest our design ideas in concrete artefacts showcasing what the future could actually look like and making it available for us in a tangible form that can be put into context. This is done by synthesizing design ideas and technologies into concrete interactive systems, and bringing them to life through prototype implementations investigating the feasibility of our ideas and concepts. In doing so we move further and further away from the abstract.

Assessment happens when moving either from building or analysing to studying. This is the opposite transition to exploration and is about measuring the quality of our design instantiations or examining our theoretical understanding in order to improve them. Like exploration, assessment may involve an orthogonal direction of work from abstraction and synthesis, and crossing a gap that is hard to bridge. Crossing back from artefacts to understanding is difficult and sometimes overlooked as an important and integrated part of the design process. Instead "evaluation" is often left as an appendix, and the subsequent process sadly shortcut by heading unreflected into re-implementation rather than seeking better understanding in order to subsequently explore radically different design opportunities. In designing mobile interactions the purpose of this transition is to return to the empirical realm of work that laid the grounds for our understanding and artefacts in the first place. This is done by feeding our theories or designs back into their context of origin, or intended context of use, assessing their fitness, and using our newly gained insight as the starting point for further abstraction, exploration or synthesis.

5.8. The contextual approach and my own research

The contextual approach described above can be used as a roadmap for summarizing my own research contributions to the design of mobile interactions, as outlined in section 3.2. In my research I have addressed what happens in the four activities of *studying, analysing, designing* and *building* when taking a contextual view, and what the ripples of abstraction, exploration, synthesis and assessment emerging from these activities may involve and look like, and how they may influence the convergence of form and context in specific interaction design projects. Furthermore, my research has addressed the creation of concrete *artefacts* that are contextually grounded, and the creation of abstract *understanding* about the relationships between interactive mobile systems, users, and context. In concert this research illustrates the elements, principles, and dynamics of the contextual approach, and exemplify the types of outcomes that it aspires to produce.

In the next section I will take a closer look at these individual research contributions, and use them to unfold the different activities and outcomes of the contextual approach. I will describe how the five parts of the thesis underpin the contextual approach to mobile interaction design as a whole, and how each chapter contribute to the individual parts.

6. CONTRIBUTIONS

In this section I will summarize and relate my individual research contributions on the design of mobile interactions. These contributions are built on a foundation of research conducted by myself, and in collaboration with colleagues, at Aalborg University in Denmark and The University of Melbourne in Australia between 2001 and 2012. It has included several mobile computing projects concerned with various user groups and use domains – from officers on large container vessels and nurses at hospitals, to young urban residents out on the town and intimate family members at home. Common for these projects are that they have been contextual in nature, taken a designerly approach, and sought to combine and integrate methods and knowledge from different disciplines. Hence, most of the projects have involved empirical, theoretical, creative and technical research, and shifting between the different activities of designing mobile interactions described in the previous section.

The individual contributions are twenty-two journal and conference articles published between 2003 and 2012 included in this thesis as chapters 2-23 (table 4). The first article describes state-of-the-art in mobile interaction design research in the early 2000s, and constitute the foundation for discussion of this back in section 3. The other twenty-one articles are compiled into five parts, each illustrating a specific aspect of designing mobile interactions that I have worked with. Parts I, II, III address the empirical, theoretical, creative and technical activities of studying/evaluating, analysing, designing and building. Parts IV and V address the two different types of outcomes from interaction design: artefacts and understanding.

Table 4. Research themes, questions and selected contributions

Theme	Research question	Contributions
Challenges and opportunities	What are the challenges and opportunities for mobile interaction design research in terms of approaches and focus?	Chapter 2
<u>Part I</u> Studying and analysing	How can we study, analyse and understand aspects of context relevant for mobile interaction design?	Chapters 3, 4, 5, 6
<u>Part II</u> Designing and building	How can we design and build interactive mobile systems grounded in context?	Chapters 7, 8, 9, 10
<u>Part II</u> Improving evaluation	How can we improve our techniques for studying the user experience of mobile interaction design in context?	Chapters 11, 12, 13, 14
<u>Part IV</u> Artefacts	How can we make use of context in the implementation of concrete interactive mobile systems?	Chapters 15, 16, 17, 18, 19
<u>Part V</u> Understanding	How can we abstractly describe and understand the relationships between interactive mobile systems, users, and context?	Chapters 20, 21, 22, 23

Figure 27 below maps the individual chapter contributions in Parts I-V onto the conceptual space introduced previously to describe the different types of design work and their outcomes.

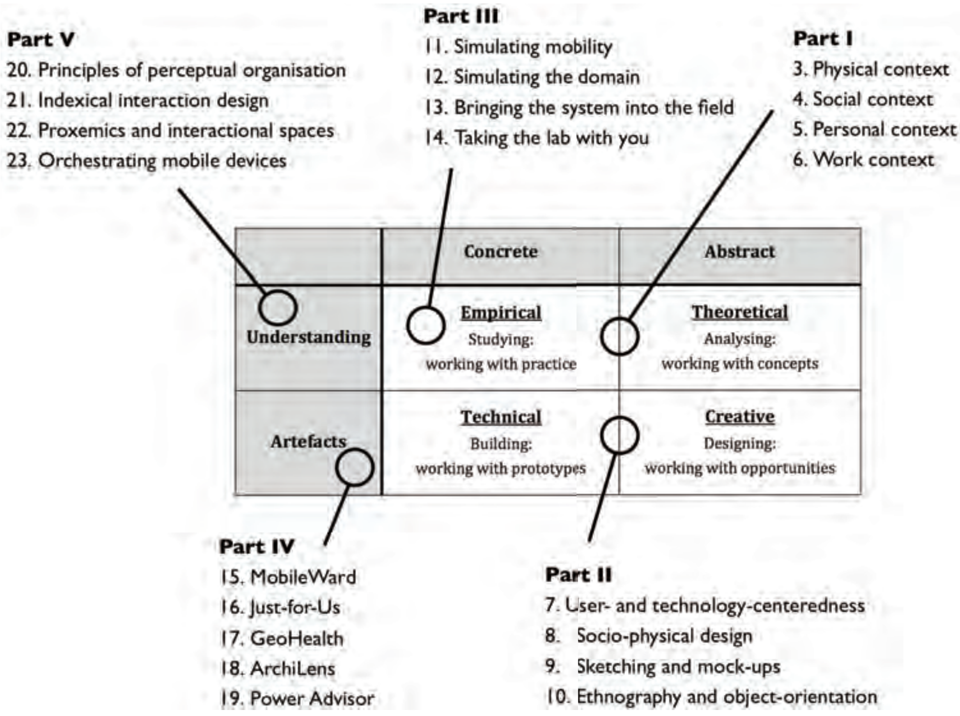


Figure 27. Overview of chapters 3 to 23 in relation to activities in the design of mobile interactions

In the following I will outline the focus of each these five parts, and summarise each of the contributing chapters.

6.1. Part I – Studying and analysing

Part I addresses the question *how can we study, analyse and understand aspects of context relevant for mobile interaction design?* In order to do contextual interaction design we need to understand what context is, for a particular design task. What dimensions of context are important, what role they play, how they interact and influence each other, etc. This is an empirical and theoretical activity where we study form and context in the real world, and elevate our insight from this from a concrete to an abstract level.

Four of my own contributions to this are included in chapters 3-6 and summarized in the following sections. These four chapters each address and explore a particular dimension of context for mobile interaction design: physical, social, personal and work, by specifically investigating four contextual entities: places and surroundings, people and interactions, families and relationships, and tasks and coordination as illustrated in figure 28. The contributions of these works are both methodological and conceptual. Firstly, they offer empirical approaches for studying concrete physical, social, personal and work contexts, and theoretical approaches for analysing them. Secondly, they offer conceptual abstractions of physical, social, personal and work contexts in the form of graphical representations, conceptual frameworks, and structure models derived through analysis of empirical data.

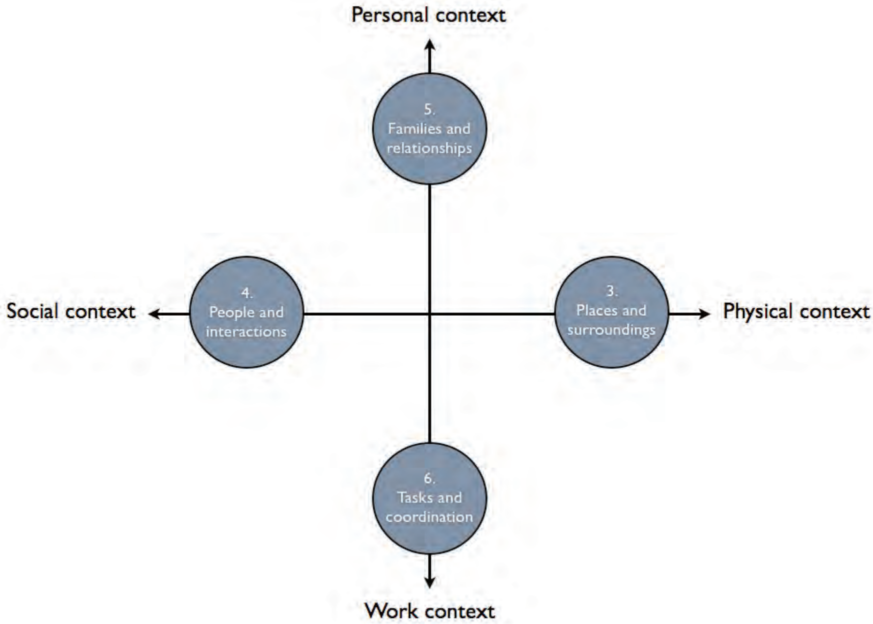


Figure 28. Four dimensions of context investigated in chapters 3-6

In concert the four chapters in Part I illustrate how we can study, analyse and understand different aspects of context relevant for designing mobile interactions. Although context for mobile interaction design is a complex and multi-dimensional concept, it can be broken down into more specific entities or dimensions that can then be studied carefully in detail. Methodologically, there are several potential approaches for studying context empirically. Common for the approaches presented in Part I is that they all involve natural setting research going out into the field and investigating the context of interest first hand. However, as also illustrated in these chapters, this overall approach still leaves room for variety in specific methods and techniques.

Physical context
Chapter 3 investigates physical context. The chapter presents a field study and architectural analysis of a concrete urban environment, Federation Square in Melbourne, Australia. The study was conducted by combining empirical and analytical methods from architecture and city planning developed by Kevin Lynch (1960) and Christopher Alexander et al. (1977) to survey, model and represent prominent architectural and informational aspects of a built environment such as districts, buildings, structures and signage. The result of the study is a descriptive framework summarizing the architectural and informational properties of a built environment. The use of such understanding in mobile interaction design is illustrated through a location-based mobile guide.

Social context
Chapter 4 investigates social context. The chapter presents a field study and sociological analysis of small groups of friends socializing "out on the town" in the same urban area as investigated in chapter 3. The study combined rapid ethnography (Millen 2000), contextual interviews, and grounded analysis (Strauss and Corbin 1990), and was

guided conceptually by McCullough's (2004) typology for situated interactions. This resulted in a conceptual framework encapsulating and describing the interplay between people, activity and places during situated social interactions in an urban environment structured around the three key concepts of knowledge, situation and motivation. The use of such understanding in the design of mobile interactions is illustrated through a context-aware mobile guide representing and adapting to social context.

Personal context
Chapter 5 investigates personal context. The chapter distils eight dimensions of intimacy from the literature, and presents a longitudinal field study of technology-mediated intimacy and strong-tie relationships with six families in Melbourne, Australia. In order to minimize researcher intervention in peoples' private lives, the study deployed the "auto-ethnographic" approach of cultural probes (Gaver et al. 1999) in combination with contextual interviews. The outcome of the study is a thematic understanding of what constitutes intimacy and how interactive systems are appropriated and used within intimate relationships. The use of such insight in mobile interaction design is illustrated through three early design ideas for technologies supporting intimacy.

Work context
Chapter 6 investigates work context. The chapter presents an ethnographic field study of communication on board large container vessels while manoeuvring inside harbour basins (Andersen 2001), and a simulator study of prototype technology in use. The outcome of the study is a detailed understanding of communication and coordination in the work context, based on Winograd and Flores' (1986) model of conversation for action. From this understanding it is possible to model communication flow in distributed work and use this to facilitate persistency in communication. This is demonstrated through the design of a mobile text-based communication system.

6.2. Part II – Designing and building

Part II addresses the question *how can we design and build interactive mobile systems grounded in context?* In order to make use of our empirical and theoretical understanding of context in the exploration of a design space, we need to develop ways of transferring this knowledge into the design activity, develop ways of grounding design ideation and exploration in such input, and support synthesising multitudes of ideas into concrete interactive systems. This activity is creative where we explore opportunities for form-context convergence, and technical where we concretise our ideas into prototypes.

In order to be successful at these it is my belief that interaction designers both need a high degree of free and unrestricted exploration of new design opportunities *as well as* some level of systematisation in order to control and direct the design process within a desired area of focus. Using the terms of Shneiderman (2000) these two requirements reflect two different types of thinking and working within design namely *inspirationalistic creativity* and *structuralistic creativity*. Inspirationalistic creativity is largely intuition-based in concert with preparation and incubation leading to moments of illumination. It promotes techniques such as brainstorming, free association, lateral thinking and divergence to support idea generation (Luther and Diakopoulos 2007). Although the outcome of inspirationalistic creativity is not necessarily art, much art is driven by inspirationalistic creativity. In contrast, structuralistic creativity is driven by more systematic approaches and highlights the importance of methodical exploration of possible solutions by systematically breaking down the challenge and structuring the response accordingly. It promotes techniques such as flow charts, decision trees, and other structured diagrams (Shneiderman 2000), and would also include software engineering modelling such as object-oriented analysis. Although the outcome of structuralistic creativity is not necessarily science, much science is driven by structuralistic creativity. As inspirationalistic and structuralistic creativity are often in conflict with each other, aspiring to support and maintain both modes of thinking and working when designing mobile interactions we are faced with a challenge of devising methods and techniques that are interdisciplinary and on the intersection of arts and science.

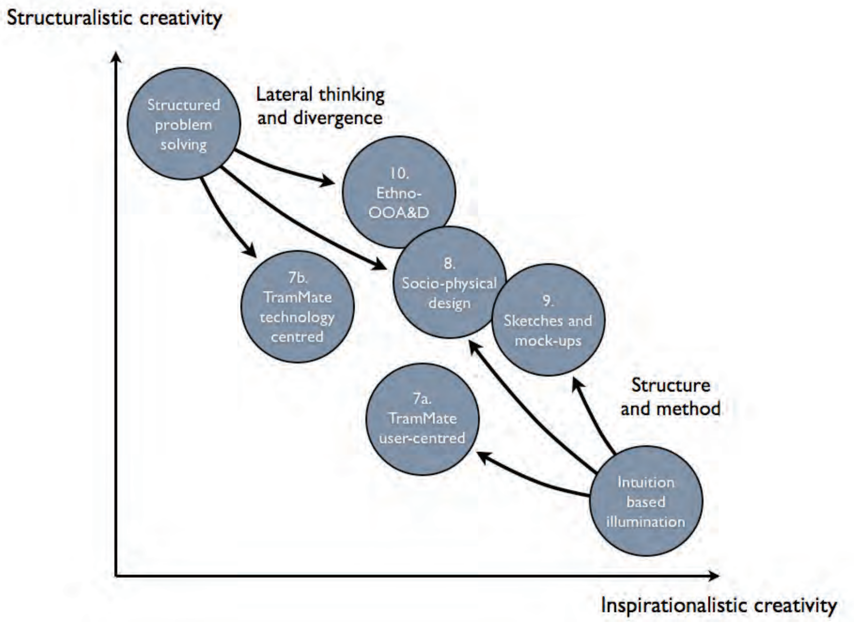


Figure 29. Achieving structuralistic and inspirationalistic creativity in mobile interaction design

Four of my own contributions on this topic are included in chapters 7-10. Figure 29 maps these four chapters into a space between inspirationalistic and structuralistic creativity, illustrating how they each represent a shift from pure structured problem solving or intuition based illumination towards a hybrid approach to designing mobile interactions involving both types of creativity. Chapters 7-10 address and discuss possible ways of supporting the creative process of moving from abstract understanding towards concrete artefacts by using *form and context ensembles* as the central point of gravity for design exploration. The contributions of these works are primarily methodological illustrated through case studies of mobile interaction design. Together they offer a palette of techniques for exploring a design space in an empirically and theoretically informed way. This is done through the use of various representations of context and form, such as abstract models, sketches, paper prototypes, and mock-ups in varying levels of fidelity, in combination with stepwise descriptions for the use of these representations in the continual development and refinement of design from initial activities of broad ideation and exploration, to the later activities of detailed explanation and specification.

Together the four chapters in Part II illustrate how we can support contextual design of mobile interactions by combining inspirationalistic creativity and structuralistic creativity. Although this is a challenging task with opposing interests and ways of thinking and working, there are several ways where a creative and contextual process of designing mobile interactions can include a high degree of free and unrestricted exploration of new design opportunities as well as some level of systematization.

User- and technology-centeredness

Chapter 7 discusses user- and technology-centeredness. In this chapter we ground the design of mobile interactions in our experiences with processes that are either user- or technology-centred. The chapter provides a meta-commentary on the relative values and limitations of inspirationalistic and structuralistic creativity for addressing the same mobile interaction design brief as a part of the "TramMate" project in Melbourne, Australia (Kjeldskov et al. 2003). The chapter concludes that neither of the two approaches was superior, but that in concert they appeared to have potentials for complementing each other very well within a broader approach: the user-centred approach grounding design in the contextual complexity of current practice, and the technology-centred approach providing a counterpoint grounded in the context of future technical possibilities.

Socio-physical design

Chapter 8 explores a socio-physical approach to design. In this chapter we propose and illustrate an inter-disciplinary approach to designing mobile interactions combining rapid ethnography, architectural analysis, design sketching and paper prototyping. The socio-physical approach combines inspirationalistic and structuralistic creativity by, on one hand, informing unrestricted idea generation through inspirational and ambiguous representations of context and, on the other hand, offering a stepwise and structured process for translating the understanding of socio-physical context into specific design ideas. The approach was explored through a mobile interaction design case study following on from the empirical work presented in chapters 3 and 4 and leading to the design and implementation of an interactive mobile prototype system.

Sketches and mock-ups

Chapter 9 illustrates the use of sketches and mock-ups. In this chapter we revisit the study of intimacy from chapter 5, and describe how our understanding of personal context was used to inform mobile interaction design. The chapter discusses the use of interactive technologies to support intimacy and describes our activities of ideation and interaction design. These took place through a semi-structured process of sketching and mock-up development in collaboration with study participants, adding a level of structure and method to the design team's mainly inspirationalistic type of creativity. The diversity in scope and form of designs produced illustrate how sketches and mock-ups are a highly flexible means for exploring and communicating ideas in a heterogenous design team.

Ethnography and object-orientation
Chapter 10 explores the combination of ethnography and object-orientation. In this chapter we propose and describe an inter-disciplinary combination of methods for supporting the design and development of interactive mobile computer systems. This combines inspirationalistic and structuralistic creativity by integrating an open-ended method with a structured one. The chapter presents two case studies of mobile system development, following on from the work presented in chapter 6. The two studies show that ethnographic data is highly valuable for developing object-oriented models by providing contextual richness. In return, object-oriented analysis is a valuable way of working with ethnographic field data by providing structure. Combining the two we were able to strongly inform system design with our ethnographic field studies.

6.3. Part III – Improving evaluation

Part III addresses the question *how can we improve our techniques for studying the user experience of mobile interaction design in context?* In order to realistically assess the quality of mobile interaction designs we need to systematically study prototypes under conditions that are appropriately representative of the future use context. This is an empirical activity where we study concrete artefacts in actual use, with the purpose of understanding the interplay between form and context.

The challenge of assessing *mobile* interaction design as opposed to non-mobile has led to a long lasting, and sometimes heated, discussion of what methods and techniques are appropriate. Most notably, this discussion has been contrasting field-based approaches with lab-based ones, polarising the research field into two distinct camps of thought, one taking an ethnographic research approach, the other taking a usability engineering one. Ethnographic field studies are characterized by taking place in "the real world" with researchers spending considerable amounts of time in the actual context of their study. Data is typically gathered through observations and interviews, and the studied phenomena (i.e. interactive mobile systems in use) are placed in a social and cultural context. The major advantages of ethnographies are the gathering of large amount of rich and grounded data, and a high level of *ecological validity* (methods, materials and settings resembling the real-life situation being investigated). The major disadvantages are unknown biases, unknown external validity/generalizability and, typically, lack of control. In contrast, usability evaluations traditionally take place in controlled environments created for the purpose of research, and data is gathered through video recording, logging, questionnaires, and interviews. The major advantages of lab based usability evaluations are the ability to focus in detail on specific phenomena of interest, and large experimental *control* before and during the study. Usability evaluations are also highly replicable and allow high quality data collection. The major disadvantages are the limited relations to the real world, unknown external validity and, typically, low level of ecological validity.

For my part, I find myself in-between these two camps of thought – my position being that natural setting research is an essential approach in most designs of mobile interactions, with its advantage of high ecological validity. At the same time, however, I also believe that some aspects of mobile interaction design can advantageously be assessed through artificial setting research, with the benefit of high levels of control. Hence, in order to shed light on the topic, in 2003 we carried out a comparative study of the added value of evaluating the usability of context-aware interactive mobile system in the field (Kjeldskov et al. 2004a). To our surprise, we found that *if looking for usability problems* the added value of going in to the field was relatively small, while the added complexity and time spent was considerably high. This finding led us to ask the question if field studies were really "worth the hassle", leading to widespread debate in the research field and numerous follow-up studies and articles[4]. In hindsight, judging from the subsequent, sometimes rather simplistic and black-or-white, debate we probably asked the wrong question. Rather than "if" it is worth the hassle the question is really "when" and "how" it can be advantageous to study mobile interaction design in the field, and what we should be looking for out there, beyond usability problems.

This question remains largely unanswered in a balanced and unbiased way that does not simply restate existing disciplinary doctrines. Even better, however, the question we really ought to be asking ourselves is how the challenges and advantages of field and lab-based evaluation studies can be overcome and combined through new hybrid methods and techniques integrating elements of the two? As research control and the ecological validity of a study are often, fundamentally, in conflict with each other, aspiring to support them both presents us with a challenge to develop methods and techniques that are essentially interdisciplinary on the intersection between social and technical science.

Four of my own contributions on this topic are included in chapters 11-14.

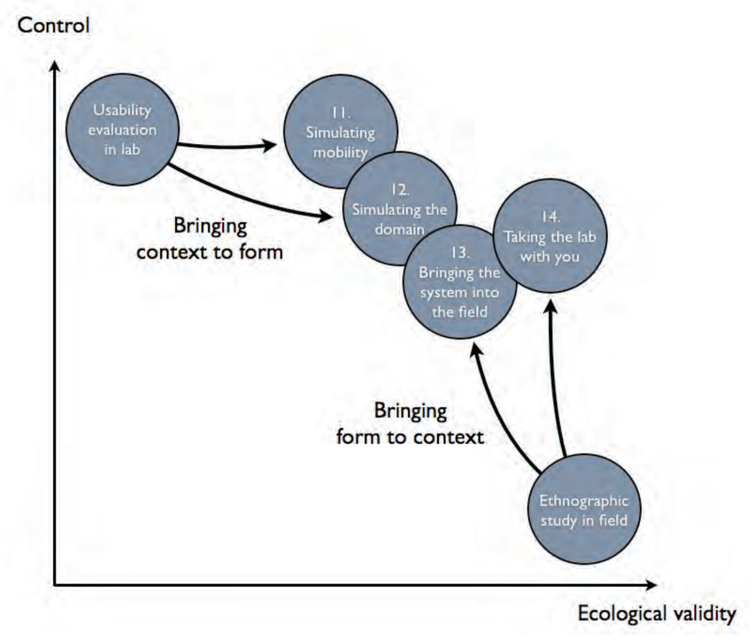


Figure 30. Achieving control and ecological validity when evaluating interactive mobile systems

4) See, for example, the articles "Its worth the Hassle" by Nielsen et al. (2006) and "Why its worth the hassle" by Rogers et al. (2007).

Figure 30 maps these four chapters into a space between control and ecological validity, illustrating how they each represent a shift from pure ethnographic studies in the field or pure usability evaluations in the lab towards a hybrid approach aspiring for both methodological qualities. This is done by combining elements from usability evaluation methods with elements of ethnographic field study methods – simulating context when evaluating in controlled environments, and guiding focus and improving data collection when evaluating in the real world. Collectively, the four chapters in section III illustrate different ways of evaluating mobile interaction design with both ecological validity and control. The described evaluation methods show how elements of context can be simulated in controlled artificial setting research environments, and how means for guiding focus and enabling high quality data collection can be applied to natural setting research environments. The methods also illustrate how new hybrid approaches can be developed by integrating elements from different disciplinary traditions.

Simulating mobility
Chapter 11 investigates how mobility can be simulated in a controlled setting. In this chapter we aim to increase the ecological validity of laboratory evaluations by simulating the user being physically mobile during use. The chapter presents and evaluates five lab techniques involving various aspects of physical motion combined with either needs for navigation in physical space or division of attention, using the case of walking down a pedestrian street as base line reference. The findings from the study show that each of the proposed techniques had similarities to the real world reference, but that none of them were completely identical to the field condition. The best simulation of mobility in the real world context was obtained when using a treadmill running at varying speed.

Simulating the domain
Chapter 12 investigates how use domains can be simulated in controlled settings. In this chapter we aim to increase the ecological validity of laboratory evaluations by simulating the use domains of interactive mobile systems. The chapter presents two case studies of mobile interaction design evaluations in controlled high-fidelity simulations of the real world. Findings show that simulating the domain provided better results than the traditional non-contextual lab approach, but that the field evaluation provided additional insight into real world use. The study concludes that although not as ecologically valid as a field study, it is possible to obtain a higher level of ecological validity by simulating significant elements of a use domain in controlled laboratory settings.

Bringing the system into the field
Chapter 13 investigates how a structured evaluation can be carried out in the field. In this chapter we aim to facilitate increased control in field evaluations without compromising its ecological validity. The chapter presents a multi-method evaluation of an interactive mobile system with the purpose of investigating how a field-based evaluation guided by tasks performs against other evaluation methods. Findings show that each of the methods had its own benefits. In relation to context, however, the field evaluation was able to uniquely highlight a number of issues of real world use. The study concludes that a field study, with a high level of ecological validity, can advantageously be combined with techniques from lab-based evaluations if more control and replicability is needed.

Taking the lab with you
Chapter 14 investigates how data collection can be improved in the field. In this chapter we aim at facilitating increased control and better data collection in field studies by exploring the use of small wireless cameras attached to users and their mobile devices. The chapter presents the development and use of a "field-laboratory" over four years of evaluating mobile interaction design in the field. It describes the current setup and explains the rationales for key decisions on technology and form factors. The study shows that it is possible to collect ecologically valid field data about mobile interaction design in use in a quality matching stationary usability laboratories by means of a field-lab. It also shows that field-labs can be made small, lightweight, and operational for hours.

6.4. Part IV – Artefacts

Part IV addresses the question *how can we make use of context in the implementation of concrete interactive mobile systems?* In order to create actual interaction design artefacts, we need to know what is technologically possible now and in the near future, and we need to know how emerging technologies can be used for pushing this frontier further. This requires building concrete interactive mobile systems exploring the feasibility of abstract design ideas in practice.

The technical dimension of designing mobile interactions contains several specific topic areas with their own challenges and focus, such as input and output devices, advanced graphic interfaces, mobile web browsing, wireless network connectivity, and context-awareness. From a mobile interaction design perspective these are all important technical topic areas, which if advanced further could potentially facilitate the creation of new and better design solutions. The technical focus of my own research in designing mobile interactions falls within the area of "context-awareness". Within this area there has been a lot of technical research into possible ways of making mobile systems capable of sensing and automatically adapting to their context, and using this capability for streamlining user interaction. Context-aware mobile systems can be differentiated technically in several ways. Chen and Kotz (2000) distinguish between *active* and *passive* context-awareness describing whether the system 1) actively takes initiative to push context-based content to the user, or 2) passively awaits the user to perform an act of information pull. Another distinction can be derived from Oulasvirta et al. (2005) who argue that the debate about context consists of two disparate camps of thought: Realism or Constructivism. As discussed in (Kjeldskov et al. 2012b), the philosophical difference between these two results in two distinct types of context-aware systems that either 1) automatically adapt their behaviour to the context as perceived by the computer system, or 2) mediate context information for the user to interpret and make use of. We label these two classes *adaptive* and *mediated* context-awareness.

Five of my own contributions to the building of concrete interactive mobile systems are included in chapters 15-19. These contributions present our experiences with the technical implementation of five prototype systems, which have all been studied in use in the field. Figure 31 map these five systems into a two-dimensional space defined by the four different characteristics of context-aware systems described above: active vs. passive and adaptive vs. mediated.

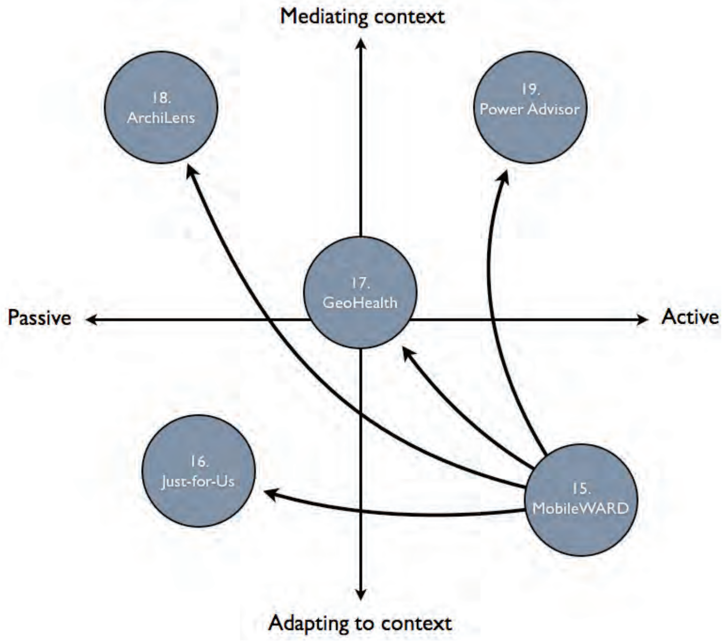


Figure 31. Five interactive prototypes mapped out in terms of their contextual characteristics

The sequential order of the systems represents the order in which they were built. This sequence also illustrates a development in thinking about contextual systems from traditional adaptive and active context-awareness in MobileWARD, towards other types of interactive mobile systems that are partly or entirely passive, and partly or entirely mediated.

Collectively, the five chapters in section IV illustrate different ways of making explicit use of context in the creation of concrete artefacts through the exploration of emerging technologies as they have become available on commercial mobile devices. The described system implementations illustrate how the feasibility of mobile interaction design ideas can be investigated through the technical construction of functional interactive prototype systems, which can then be used as research vehicles for empirical studies in realistic use contexts. The selection of systems discussed in section IV also illustrates how an interaction design space can be defined technically for a specific class of interactive mobile systems, in this case context-aware systems being active/passive and adaptive/mediated, and how such a technical interaction design space can be used to guide the design focus towards deliberately exploring particular technical properties.

MobileWARD

Chapter 15 presents MobileWARD. In this chapter we explore the construction of an interactive context-aware mobile prototype system for the healthcare domain. The system deploys a combination of active and adaptive context-awareness by automatically pushing information and functionality to the user filtered on the basis of their location, current work activities and people nearby. The chapter presents the details of the prototype system and results from an empirical study of its use. It concludes that context-aware mobile information systems hold potential value within the healthcare domain as a component of a ubiquitous computing environment supporting the mobile and distributed nature of work activities in this domain. However, the implementation of interaction designs for such systems is highly complex and must be carefully thought out and evaluated in order to ensure a good fit between systems, users, and their context.

Just-for-Us
Chapter 16 presents Just-for-Us. In this chapter we explore the construction of an interactive context-aware mobile prototype system for socialising in the city. The system primarily deploys passive adaptive context-awareness by filtering user requested content and functionality on the basis of their context. Just-for-Us was an early attempt at making mobile context-aware systems web-based. The aim of this technical approach was to explore system alternative architectures that would allow such mobile systems to benefit from the fast paced development of new programming facilities for the mobile web. The chapter concludes that the web-based approach holds interesting potentials for the creation of highly dynamic and graphical mobile context-aware applications, but that mobile web browsers and programming environments of the time lacked a series of capabilities for handling dynamic exchange of information between clients and servers.

GeoHealth
Chapter 17 presents GeoHealth. In this chapter we explore the construction of a web-based location-based service for home healthcare workers distributed over a large geographical area. The system combines active and passive context-awareness depending on importance of information, and combines adaptation and mediation of context through a mesh-up of information from various sources on an interactive map. Extending directly from the work presented in chapter 16, the GeoHealth system explores the powers of Web 2.0 technologies in combination with GPS positioning, Google Maps and the Mobile Internet. The chapter presents the prototype system in detail followed by results from an empirical study of its use. It shows that mobile location-based services built around Web 2.0 technologies and interactive maps have unexploited potentials for mobile interaction design within the domain of home healthcare.

ArchiLens
Chapter 18 presents ArchiLens. In this chapter we explore the construction of a mobile augmented reality system for architectural visualization. The system deploys passive context-awareness and mediates contextual information by allowing the user to explore the visual and spatial characteristics of their future house in context. This is done by overlaying the 3D scene onto the live images from the phone's built-in camera. The ArchiLens system was implemented as an application for the Android operating system, making use of its powerful graphics engine mainly designed for interactive 3D games. The chapter presents the prototype in detail followed by results from a study of use with 40 participants. It show that the 3D capabilities of modern mobile devices in combination with their contextual sensors and built-in high quality video cameras have strong potentials as a platform for the design of mobile interactions.

Power Advisor
Chapter 19 presents Power Advisor. In this chapter we explore the construction of an interactive mobile system to promote sustainability by allowing people to monitor their domestic electricity consumption and adjust usage behaviour accordingly. The system deploys active and mediated context-awareness by pushing information about the household's electricity consumption to the user as a resource for interpretation and exploration. This information is collected wirelessly from a "smart" power meter unit. The Power Advisor system was implemented as a mobile web application allowing it to be used on Android and iOS enabled devices. The chapter presents the prototype system in detail followed by results from a study of use in 10 households over a period of 7 weeks. The findings provide insight into peoples awareness of electricity consumption in their home and how this may be influenced through interaction design of mobile systems.

6.5. Part V – Understanding

Part V addresses the question *how can we abstractly describe and understand the relationships between interactive mobile systems, users and context?* In order to facilitate research progress beyond small incremental steps from one design to the next, we need to develop a cumulative body of knowledge that can help explain, theoretically, the relationship between people, technology, and their context. This requires research elevating our understanding from a concrete to an abstract level.

With form-context convergence as the central unit of analysis, the theoretical approach to this phase of research is inherently one of holism rather than of reductionism. Whereas reductionism believes that complex systems can be understood by breaking them down into fundamental parts, which can then be studied in detail individually, holism believes the opposite, that the whole is larger than the sum of the parts and therefore can not be understood through explanation of its components alone but has to be considered in its totality. This can be done by investigating the whole through theoretical and conceptual lenses that focus on explaining the relations and interactions between elements in a system rather than on their individual mechanics.

Four of my own contributions on this topic of understanding the relationships between interactive mobile systems, users, and context, are included in chapters 20-23. When trying to understand the contextual user experience of a mobile interaction design, the central question is how people create meaning from such artefacts in context. Figure 32 illustrates how these four chapters each take an individual approach to the issue of sense making in context on different levels of abstraction – the way we perceive the world by identifying meaningful patterns and wholes, the way we interpret the world by assigning meaning to signs, the way we use our joint embodied presence in the world to create shared meanings, and the way we organise and orchestrate the world around us in order to create meaningful systems of systems.

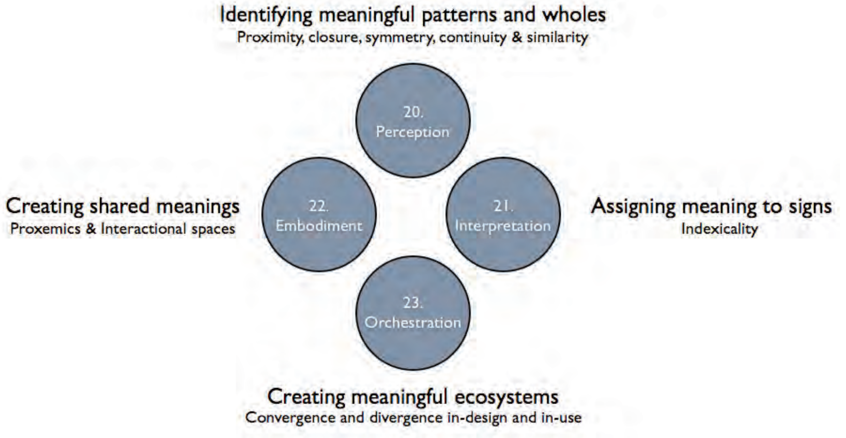


Figure 32. Four approaches to understanding contextual mobile user experience

Principles of perceptual organisation
Chapter 20 presents and discusses five principles that can be applied for explaining how people identify meaningful patterns and wholes from ensembles of mobile systems and their context. The discussion is informed by a field study of mobile user experience at Federation Square in Melbourne, Australia, and takes its theoretical inspiration from the discipline of Gestalt psychology. Based on a theoretical analysis of our empirical findings, we argue that the user experience of location-based mobile interaction designs in context can be described and understood through Gestalt theory's five principles of perceptual organisation: *proximity, closure, symmetry, continuity,* and *similarity*. Specifically, we argue that these principles assist us in explaining how people create meaningful wholes from incomplete and fragmented information on mobile devices. They do so by "drawing from a larger canvas" to which both mobile devices and their context are contributing. Consequently, as mobile interaction designers we need to design for this "larger canvas" rather than merely for the "smaller canvas" of the mobile device in isolation.

Indexical interaction design
Chapter 21 discusses how people interpret information representations on mobile devices in context by assigning meaning to indexical signs. The discussion is informed by three field studies of mobile user experience in Denmark and Australia, and is grounded theoretically in the discipline of semiotics. Based on our empirical and analytical work, we argue that the relationship between users, user interface representations, and context can be described and understood through the semiotic concept of *indexicality.* In relation to mobile interaction design, we argue that information in an interactive mobile system can be understood as a special type of indexical sign, where meaning is created through interpretation of the ensemble of system and context. We also argue that increasing the level of indexicality in an information representation by locating it in time and space results in a reduction of symbolic and iconic representations required to communicate a specific piece of information. Of particular importance for the design of mobile interactions, this allows a reduction of explicit information presented to the user.

Proxemics and interactional spaces

Chapter 22 discusses how people create shared meanings through embodiment in shared physical spaces, and how an understanding of embodiment and proxemics can be used to guide the design of interactional spaces or digital artefact ecologies. The discussion is informed by a theoretical analysis of human interaction in shared physical spaces drawing on the philosophical foundations of Husserl, Heidegger and Merleau-Ponty. The key point is that coordinated action, meaning-making and intersubjective understanding are shaped, in part, from our embodied actions in space and the availability to others of these actions, for example, the way we move, point, touch and gesture in relation to objects and other people. In respect to understanding mobile device user experiences we argue that this is profoundly influenced by spatial factors such as proxemics and the physical design of the interactional spaces in which they are used. We exemplify this through the "blended interaction space" prototype including an ecosystem of interactive surfaces and devices and facilitating various forms of "proxemic interactions" (Greenberg et al. 2011).

Orchestrating mobile devices

Chapter 23 discusses how people create and orchestrate meaningful digital ecosystems of interactive mobile systems and devices. The discussion is informed by a cultural probe study of early adopters of mobile devices in Melbourne, Australia, and takes its offset in the debate about convergence versus divergence as principles for mobile interaction design (as discussed in section 2.1). The chapter presents three seemingly irreconcilable perspectives on the relationship between functionality and user experience drawn from the literature, and argues that these are, in fact, complementary views, when observed in a broader perspective. The key point in this argument is the observation that convergence and divergence are not just principles of design, but also principles of orchestration in use.

7. CONCLUSIONS

In this final section I briefly summarize the main conclusions of my work. The conclusions are expressed as eight lessons about the design of mobile interactions as a matter of continual convergence of form and context, based on my experiences from taking this view and approach in my own research over the last decade.

Table 5. Key lessons about the design of mobile interactions

Lesson 1:	Mobile interaction design has become a discipline at the intersection between technology and liberal arts where the best results yield from combining the two. Doing this well requires approaches that transcend technology- and user-centeredness.
Lesson 2:	Form-context unity is a central concept in all phases of designing mobile interactions. This makes it a suitable higher-level unit of analysis for interdisciplinary research and design transcending focus beyond technology- or user-centeredness.
Lesson 3:	Contextual interaction design guides the continual convergence of form and context through a designerly process of shifting freely between empirical, theoretical, creative and technical work, oscillating between understanding and artefacts, concrete and abstract.
Lesson 4:	In studying and analysing the foundations for designing mobile interactions, a contextual approach defines a broad perspective on the interplay between both users, technology, and advocates the use of theories, concepts and frameworks from different disciplines.
Lesson 5:	In designing and building interactive mobile systems, a contextual approach supports open-minded, yet structured, exploration of design opportunities and prototypes by facilitating both inspirationalistic and structuralistic creativity.
Lesson 6:	In studying the user experience of mobile interaction design, a contextual approach adds improved techniques that facilitate both control and ecological validity to the palette of methods for evaluating interactive mobile systems.
Lesson 7:	In building interactive mobile system artefacts, a contextual approach promotes convergence of form and context that make close relationships between content, functionality, interaction, system behaviour and the surrounding context.
Lesson 8:	In understanding the user experience of mobile interaction design, a contextual approach promotes a holistic perspective on the interplay between people, interactive systems, and context, seeing it as a whole that is larger than the sum of its parts.

Lesson 1: transcending technology- and user-centeredness
Mobile interaction design has become a discipline at the intersection between technology and liberal arts where the best results yield from combining the two. Doing this well requires approaches that transcend technology- and user-centeredness. Although current mobile interaction design is multi-methodological and involves multiple disciplines, there is still an assumption that users and technology can advantageously be studied separately. In contrast, taking a contextual approach to interaction design means that focus is explicitly broadened to the higher-level unity of form and context. Combining and integrating methods and techniques from different disciplines into new and hybrid ones, with a transcendent unit of analysis, allows us to maintain this broader focus throughout all of the different activities of the interaction design process.

Lesson 2: form-context unity
Form-context unity is a central concept in all phases of designing mobile interactions. This makes it a suitable higher-level unit of analysis for interdisciplinary research and design transcending focus beyond technology- or user-centeredness. Viewing the design of mobile interactions as continual convergence of form and context facilitates a paradigmatic shift stimulating new ways of conceptualizing and working by framing new issues, research questions and challenges beyond those of the disciplines involved with mobile interaction design individually. As a central unit of analysis, form-context unity provides a wide enough scope to encompass the more extensive phenomena of mobile interaction design user experience in a holistic way. From a contextual perspective the activities of studying, analysing, designing and building are all about considering the ensemble of a particular form in relation to its context: an interactive mobile system in relation to users, technology, settings, activities, etc. Consequently, the interaction design process becomes neither user- nor technology-centred but instead continuously includes both of these viewpoints within a broader perspective.

Lesson 3: a designerly way
Contextual approaches to interaction design guides the continual convergence of form and context through an unpredictable process of shifting freely between empirical, theoretical, creative and technical work, oscillating between producing understanding and artefacts, and between working in the concrete and in the abstract. The process of stepping freely between the activities of studying, analysing, designing and building creates what can be described as ripples of abstraction, exploration, synthesis, and assessment towards the unity of form and context, which is always in focus and continually evolving. In this way, viewing the design of mobile interactions as continual convergence of form and context promotes a designerly way of thinking and working where new knowledge, artefacts, and contexts emerge from a series of intentionally short and open ended steps. It allows us to do mobile interaction design that is intentionally rhetorical, exploratory, emergent, opportunistic, abductive, reflective, ambiguous, and risky, and it allows us to reach beyond interaction design as a matter of problem setting and solving, and treat it also as a matter of creating entirely new practices – enabling humans to do things in their lives that they couldn't do before.

Lesson 4: studying and analysing
In studying the background and foundations for designing mobile interactions, viewing interaction design as continual convergence of form and context defines a broader perspective than, for example, a user- or technology-oriented approach, and directs the focus of attention towards understanding dimensions of context of particular importance or relevance to the specific case. In analysing empirical data, the design of mobile interactions retains this broader perspective and advocates the use of theories, concepts and frameworks from different disciplines for elevating empirically grounded understanding from concrete to abstract. Physical, social, personal, and work context are examples of such dimensions of context, which I have specifically investigated in my work presented here. Taking a contextual approach during empirical and theoretical activities of interaction design means that both users and technology, as well as other contextual factors, are of importance, but that the central unit of analysis transcends these for a holistic perspective on the form-context ensemble to which they all contribute.

Lesson 5: designing and building
In designing and building interactive mobile systems, the contextual approach supports the process of moving from abstract understanding to concrete artefacts by encouraging and supporting both inspirationalistic and structuralistic creativity in the exploration of design opportunities and prototypes. This is done through the use of various representations of context and form, such as abstract models, sketches, paper prototypes, mock-ups, and prototypes on different levels of fidelity combined with stepwise descriptions for their use in the gradual development and refinement of design, from the initial broader activities of ideation to the later narrower activities of specification. Using this mix of representations of form and context adds lateral thinking and divergence to structured problem solving, and structure and method to intuition based illumination. Combining technology-centeredness with user-centeredness, architectural analysis with sociological analysis, sketching with mock-ups and prototypes, and ethnography with object-orientation, are examples of specific design methods, which we have investigated in the work presented here. Taking a contextual approach as exemplified by these methods and techniques means that the creative and technical activities gain and maintain a holistic perspective on the design challenge at hand, transcending that of user- or technology-centeredness.

Lesson 6: improving evaluation
In studying the user experience of mobile interaction design, the contextual approach adds improved techniques for systematic assessments under realistic conditions representative of the future use situations and settings to the palette of evaluation methods for interactive mobile systems. This can be done through hybrid methods and techniques that inherit the qualities of control from traditional laboratory-based usability evaluations, and the qualities of ecological validity of ethnographic field studies by, for example, simulating context in artificial settings, or guiding focus and data collection in natural settings. Simulating mobility or the use domain, bringing the system in to the field, and designing a lab that you can take with you, are specific examples of such hybrid methods and techniques, which we have investigated in the work presented here. Taking a contextual approach to the assessment of mobile interaction design puts explicit

emphasis on the importance of making empirical studies as realistic as possible in terms of the settings and situations they are carried out in, while also ensuring that detailed, comparable and generalizeable data is collected. As demonstrated in the empirical studies discussed, it is possible to design evaluations of mobile interaction design that achieve both ecological validity and experimental control.

Lesson 7: artefacts

In building interactive mobile system artefacts, the contextual approach promotes convergence of form and context that explores the use of established and emerging technologies for making close relationships between content, functionality, interaction, system behaviour and the surrounding context. This can be done by building concrete interactive mobile system prototypes that implements different ways of relating their content and functionality to their context, ways of facilitating user interaction with such content and functionality, and ways of making content and functionality respond to contextual changes. Building prototypes that mediate or adapt to context, and do so passively or actively, are examples of such systematic synthesis of new technologies for the design of mobile interactions, which we have investigated in the work presented here. Taking a contextual approach to the implementation of such artefacts encourages the technical construction of a certain type of mobile interaction design where there is an inherently close and explicit relationship between interactive systems, users and their surrounding context. As demonstrated with the interactive prototype systems discussed, this is not simply a matter of making them respond automatically to context, but offers a much wider range of new and interesting opportunities for mobile interaction design.

Lesson 8: understanding

In understanding the user experience of mobile interaction design, the contextual approach promotes a holistic perspective on the interplay between people, interactive systems, and context, seeing it as a whole that is larger than the sum of its individual parts. From a contextual perspective, the user experience of interactive mobile systems is a complex system that cannot be understood well through reductionism alone but also has to be considered in its totality. This can be done by scrutinising the phenomenon through theoretical and conceptual lenses that particularly focus on explaining the relations and interactions between elements in an interactive system rather than on explaining the individual mechanics of these elements. When trying to understand, abstractly, the user experience of a mobile interaction design, the central question is how people create meaning from such artefacts in context. Perceiving meaningful patterns and wholes, interpreting meaning of signs, creating shared meaning through our embodiment in the world, and orchestrating mobile devices by creating meaningful ecosystems are examples of specific holistic perspectives, which we have investigated in the work presented here. Taking a contextual approach to understanding the user experience of mobile interaction design means that the analytical outcomes transcend specific technologies and users, facilitating further research and design at a pace beyond incremental steps from one design to the next.

8. EPILOGUE

Nearing the end of my position summary, there remain some questions unaddressed. I have presented a contextual approach to designing contemporary interactive mobile computer systems, that promotes a designerly way of achieving convergence between form and context through a wholeness sensitive and continually unfolding process of design. The value of this approach and way of thinking has been discussed in relation to other design approaches and illustrated through my contributing research publications. But what are the potential challenges and limitations of such a contextual approach and holistic view? And where do we go from here?

Challenges for a contextual approach

One potential challenge in taking a contextual approach to the interaction design process, rather than a traditional user- or technology-centred one, is that the concept of *context*, and *form-context ensembles,* may appear too abstract and difficult to grasp. In contrast, *users* are entities that we can relatively easy define, identify, study, and simply go and talk to if we don't quite understand them. Similarly, *technology* is a relatively tangible thing in interaction design that we can often simply look at, touch, and try out. The intangibility of entities like context and wholeness make them more difficult to deal with by comparison. If a contextual approach to interaction design is to succeed, we will need to explore it further, and to develop concepts, techniques, and best practices that make it accessible for interaction designers and interaction design researchers to embrace and hence practice contextual and holistic thinking in their work. My own contributions to this included in this thesis are steps in that direction. But they are in no way complete. My hope is that this is something others will find interesting to pursue, and that researchers and practitioners currently grounded in user- or technology-centred design will not see the views that I have presented here as a disparagement against the quality or importance of their work. After all, taking a contextual approach is not a matter of throwing away the legacy of user- or technology-centred design at all. It is a matter of trying to include both of these view-points, equally, and within a broader scope that enables us to transcend them.

Downsides of holism in interaction design

In terms of the downsides of a holistic view on interaction design, an obvious concern is that by focussing on the whole, you might erroneously neglect or ignore important details of the parts. This is a valid concern, and one that is important to keep clear in mind when taking a holistic stance. The kind of holistic thinking that I have promoted in this thesis falls within what Edmonds (1999) calls *pragmatic holism.* Rather than the all-embracing view of a system in *experiential* holism, this is the kind of *non-linear* holism that refers to the phenomenon of emergence in that "when A and B are combined, the resulting C has more properties than what each of the components bring" (Raman 2005). While reductionism may struggle to deal with such phenomena of emergence in complex systems and how individual elements can converge into something very different, holism on the other hand may struggle to deal with identifying and explaining what "lies beneath" a larger whole in a way that enables us to understand it, and possibly reproduce it.

Essentially these concerns put us in the middle of the highly polarised debate about whether reductionism or holism is the better approach for viewing and dealing with the world (see, for example, Edmonds 1999, Raman 2005). I won't go into this debate here, but in my opinion neither of the two are, in their extreme forms, very useful positions to hold as interaction designers or interaction design researchers. What is needed in interaction design are world views that are less dogmatic and more pragmatic, seeking *useful* accounts, models, and understandings of the phenomena in the world that we are interested in designing and designing for. Both reductionism and holism are legitimate and have value in this respect. But they provide us with very different pictures and understanding of the same phenomenon. As described by Raman (2005), like a microscope and a telescope, reductionism and holism are two powerful instruments to explore the world. "Each is relevant and important in its own context. The more we focus on one, the more the other becomes blurred. Thus, reductionism and holism are complementary in the Bohr sense of the term" (Raman 2005), and to get a full picture, we need them both. In the words of Herbert Simon "in the face of complexity, an in-principle reductionist may be at the same time a pragmatic holist" (Simon 1962). Conversely an in-principle holist may at times need to apply reductionist principles when pragmatically useful and not compromising the overall view of the whole.

In mobile interaction design the reductionist view is already strongly present, but the complementary holistic view is not.

Towards digital ecology

The final thing I wish to touch upon are the notions of digital ecosystems and artefact ecologies. As I have discussed in section 2.1 the currently emerging trend within mobile computing is the creation of digital ecosystems where interactive mobile systems and devices are viewed less in isolation and more as parts of larger use contexts or artefact ecologies (see, for example, Jung et al. 2008, Bødker and Klokmose 2011). In my opinion this is an avenue for further research that is particularly interesting, and one that I look forward to engaging with more deeply. As a starting point for this, I believe that the contextual approach on designing mobile interactions presented in this thesis holds potentials for designing digital ecosystems and artefact ecologies. The reason for this is that it already inolves designing for the whole and has a build-in sensitivity for the continual emergence and convergence of form and context that characterises such ecosystems and ecologies. What is still needed though is the further development of a theoretical and conceptual lens through which we can view, address and describe this emerging phenomenon in a way that informs and inspires design and further thinking. This work may find inspiration and traction in some of the conceptually stronger and less technical literature on ubiquitous and pervasive computing that has started to appear in recent years, such as Adam Greenfield's book "Everyware" (2006).

As a way of encapsulating and labeling this work, I suggest using and developing the term *digital ecology.* Ecology is the study of elements making up an ecosystem, and is very generally about understanding the interactions between organisms and their environment. It is inherently holistic and has an interdisciplinary nature, and it is not synonymous with "the environment" or with "environmentalism". Nor is ecological thinking limited to the discipline of biology. For example, "industrial ecology" studies material and energy flows through networks of industrial processes, and "human ecology" is as interdisciplinary area of research that provides a framework for understanding and researching human social interaction. In a similar fashion, I believe "digital ecology" may be a useful way of describing the study of elements making up digital ecosystems and the holistic understanding of interactions between these elements and their environment[5]. By digital ecology I thereby mean the study of interrelated digital systems (e.g. mobile and pervasive computing) and the processes by which these systems work and interact, and are conceived, emerge, converge, and evolve. It is about understanding the functioning, use and experience of digital ecosystems and digital artefact ecologies around us, and the design processes that create and advance them.

ACKNOWLEDGEMENTS

This work was funded in part by the Danish Technical Research Council, grants No. 26-03-0341 and 26-04-0026, several project and travel grants from The Obel Family Foundation, Aalborg University's Faculty of Engineering and Science, and the Department of Computer Science. I thank all of my co-authors on the included research publications, in particular my primary collaborators and co-authors over the last decade: Jeni Paay, Mikael B. Skov, Steve Howard, Jan Stage, Frank Vetere, Connor Graham, Kenton O'Hara, and Sonja Pedell. Very special thanks to Lars Mathiassen and Steve Howard for feedback and discussions on early drafts of the thesis, to Erik Frøkjær for ongoing fruitful discussions of my work, and to members of the Information Systems group/Centre for Socio-Interactive Design for constructive comments about how to strengthen it. I also thank Ellen Christiansen for introducing me to the works of Christopher Alexander. Finally, I wish to thank my Ph.D. and many Masters students in Human-Computer Interaction over the years, who have contributed to my research through prototype experiments, user experience studies and discussions on mobile interaction design, in particular Dimitrios Raptis, Henrik Sørensen, Rahuvaran Pathmanathan, Claus M. Christensen, Klaus Kjeldsen, Niels Husted, Jacob Nørskov, Kenneth Pedersen, Glen W. Nielsen, Søren Thorup, Michael Vestergaard, Rene Vutborg, Liv Stahl Madsen, Simon Lind Damgaard, Michael Bønnerup, Søren H. Møller, Thulasika Rasenthiran, Anders Christensen, Eva Andersen, and Lars Hedegaard.

5) The term "digital ecology" has elsewhere been used to describe the fusion of virtual and real life forms, or the mix of digital code and environmentalism. These are not related to my suggested use of the term.

REFERENCES

Abowd, G. D. and Mynatt E. D. (2000) Charting Past, Present and Future Research in Ubiquitous Computing. *ACM Transactions on Computer-Human Interaction, 7*(1), 29-58.

Admob (2009) *March 2009 Metrics Report.* Retrieved January 2011, from: http://metrics.admob.com/

Agre, P. (2001). Changing Places: Contexts of Awareness in Computing. *Human-Computer Interaction 16*(2), 177-192.

Agre, P. (1997) Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI. In G. Bowker, S. L. Star, W. Turner and L. Gasser (Eds.), *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide* (pp. 131-158). New Jersey: Lawrence Erlbaum Associates Publishers.

Alexander, C. (2007) Empirical Findings from The Nature of Order. *Environmental & Architectural Phenomenology Newsletter*, winter 2007. Retrieved January, 2011, from http://www.arch.ksu.edu/seamon/Alexander_Nature%20of%20Order.htm

Alexander, C. (2005) *The Nature of Order. Vol. 3: A Vision of A Living World.* Berkeley, California: CES Publishing.

Alexander, C. (2004) *The Nature of Order. Vol. 4: The Luminous Ground.* Berkeley, California: CES Publishing.

Alexander, C. (2002b) *The Nature of Order. Vol. 2: The Process of Creating Life.* Berkeley, California: CES Publishing.

Alexander, C. (2002a) *The Nature of Order. Vol. 1: The Phenomenon of Life.* Berkeley, California: CES Publishing.

Alexander, C. (1979) *The timeless Way of Building.* New York: Oxford University Press.

Alexander, C., Ishikawa, S., Silverstein, M. with Jacobson, M., Fiksdahl-King, I. and Angel, S. (1977) *A Pattern Language: Towns, Buildings, Constructions.* New York: Oxford University Press.

Alexander, C. (1964) *Notes on the Synthesis of Form.* Cambridge, MA: Harvard University Press.

Alvarez, I. and Kilbourn, B. (2002) Mapping the information society literature: topics, perspectives and root metaphors. *First Monday, 7*(1). Retrieved January, 2011, from: http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/922/844

Andersen, P. B. (2001) Maritime Work and Communication. *Australian Journal of Information Systems. 8*(2), 83-102.

Aoki, P. M., Honicky, R. J., Mainwaring, A., Myers, C., Paulos, E., Subramanian, S. and Woodruff, A. (2009) A vehicle for research: using street sweepers to explore the landscape of environmental community action. In *Proceedings of CHI 2009*, Boston, MA (pp. 375-384). New York: ACM.

Archer, B. (1992) The nature of research in design and design education. In B. Archer, K. Baynes and P. Roberts (Eds.), *The Nature of Research into Design and Technology Education: Design Curriculum Matters.* Loughborough, UK: Loughborough University.

Archer, B. (1965) *Systematic Method for Designers.* London: Council of Industrial Design.

Atkinson, P. (2005) Man in a Briefcase - The Social Construction of the Laptop Computer and the Emergence of a Type Form. *Journal of Design History, 18*(2), 191-205.

Augsburg, T. (2005) *Becoming Interdisciplinary: An Introduction to Interdisciplinary Studies.* Kendall Hunt Publishing.

Bagnara, S. and Smith, G. S. (Eds.) (2006) *Theories and Practice in Interaction Design.* London: Lawrence Erlbaum Associates Publishers.

Ballard, B. (2007) *Designing the Mobile User Experience*. Padstow: John Wiley and Sons Ltd.

Bannon, L. (1992) Interdisciplinarity or interdisciplinary theory in CSCW? In *Workshop proceedings of CSCW 1992*. Workshop on Interdisciplinary Theory for CSCW Design, Toronto, Canada.

Bardram, J. E. (2009) Activity-based computing for medical work in hospitals. *ACM Transactions on Computer-Human Interaction, 16*(2), 1-36.

Barkhuus, L. and Dey, A. (2003) Is Context-Aware Computing Taking Control away from the User? Three Levels of Interactivity Examined. In *Proceedings of UbiComp 2003,* LNCS (pp. 149 – 156). Berlin: Springer-Verlag.

Basili, V. R., Selby, R. W. and Hutchins, D. H. (1986) Experimentation in software engineering. *IEEE Transactions on Software Engineering, SE-12* (1986), 733-743.

Benbasat, I. (1985) An analysis of research methodologies. In F. W. MacFarlan (Ed.) *The Information System Research Challenge* (pp. 47-85). Boston: Harvard Business School Press.

Benford S., Giannacji G., Koleva B. and Rodden T. (2009) From Interaction to Trajectories: Designing Coherent Journeys Through User Experiences. In *Proceedings of CHI 2009*, Boston, MA (pp. 709-718). New York: ACM.

Benyon, D., Turner, P. and Turner, S. (2005) *Designing Interactive Systems.* Harlow: Addison-Wesley.

Bergman E. (Ed.) (2000) *Information Appliances and Beyond.* San Francisco: Morgan Kaufmann Publishers.

Bergman E. and Haitani R. (2000) Designing the PalmPilot: a Conversation with Rob Haitani. In E. Bergman (Ed.), *Information Appliances and Beyond*. San Francisco: Morgan Kaufmann Publishers.

Betiol, A. H. and de Abreu Cybis, W. (2005) Usability Testing of Mobile Devices: A Comparison of Three Approaches. In *Proceedings of INTERACT 2005*, LNCS (pp. 470-481). Berlin: Springer-Verlag.

Beyer, H. and Holtzblatt, K. (1998) *Contextual design: Defining customer-centred systems.* San Francisco: Morgan Kaufmann Publishers.

Blevis, E. and Stolterman, E. (2009) Transcending Disciplinary Boundaries in Interaction Design. *Interactions, 16*(5), 48-51.

Blom, J., Chipchase, J. and Lehikoinen, J. (2005) Contextual and cultural challenges for user mobility research. *Communications of the ACM, 48*(7), 37-41.

Boehm, B. W. (1988) A Spiral Model of Software Development and Enhancement. *IEEE Computer, 21*(5), 61-72.

Bondo, J., Barnard, D., Burcaw, D., Novikoff, T., Kemper, C., Parrish, C., Peters, K., Siebert, J. and Wilson, E. (2009) *iPhone User Interface Design Projects*. Apress.

Bradley, N. A. and Dunlop, M. D. (2002) Understanding Contextual Interactions to Design Navigational Context-Aware Applications. In *Proceedings of Mobile HCI 2002* (pp. 349-353). Berlin: Springer Verlag.

Bratteteig, T. (2007) Design Research in Informatics: A response to Livari. *Scandinavian Journal of Information Systems, 19*(2), 65-75.

Brewster, S. (2002) Overcoming the Lack of Screen Space on Mobile Computers. *Personal and Ubiquitous Computing, 6,* 188-205.

Brolin, B. C. (1980) *Architecture in context: fitting new buildings with old*. New York: Van Nostrand Reinhold.

Brown, B. and Randell, R. (2004) Building a Context-Sensitive Telephone: Some Hopes and Pitfalls for Context Sensitive Computing. *Computer-Supported Cooperative Work, 13*(3-4), 329-345.

Brown, B. A., Sellen, A. J. and O'Hara, K. (2000) A diary study of information capture in working life. In *Proceedings of CHI 2000*, The Hague, The Netherlands (pp. 438-445). New York: ACM.

Buxton, B. (2007) *Sketching User Experiences: getting the design right and the right design.* San Francisco: Morgan Kaufmann Publishers.

Buxton, W. (2001) Less is More (More or Less): Uncommon Sense and the Design of Computers. In P. Denning (Ed.), *The Invisible Future: The seamless integration of technology in everyday life* (pp. 145-179). New York: McGraw Hill.

Bødker, S. and Klokmose, C. (2011) The Human-Artifact Model: An Activity Theoretical Approach to Artifact Ecologies. *Human-Computer Interaction, 26*(4), 315-371.

Bødker, S. (2006) When Second Wave HCI meets Third Wave Challenges. In *Proceedings of NordiCHI 2006* (pp. 1-8). New York: ACM.

Bødker, S. (1996) Creating conditions for participation: Conflicts and resources in systems design. *Human Computer Interaction, 11*(3), 215-236.

Bødker, S., Ehn, P., Kammersgaard, J., Kyng, M. and Sundblad, Y. (1987) A Utopian experience: On design of Powerful Computer-based tools for skilled graphic workers. In G. Bjerknes, P. Ehn and M. Kyng. (Eds.), *Computers and democracy: A Scandinavian challenge* (pp. 251–278). Aldershot, UK: Avebury.

Chalmers, M. (2004) A Historical View of Context. *Computer Supported Cooperative Work, 13*, 223-247.

Chen, B. X. (2010) *What the iPad means for the future of computing*. Retrieved January, 2011, from: http://www.wired.com

Chen, G. and Kotz, D. (2000) *A survey of context-aware mobile computing research*. (Paper TR2000-381). Department of Computer Science, Darthmouth College.

Cheverst, K., Davies, N., Mitchell, K. and Efstratiou, C. (2001) Using Context as a Crystal Ball: rewards and Pitfalls. *Personal and Ubiquitous Computing, 5*(1), 8-11.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C. (2000) Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. In *Proceedings of CHI 2000*, The Hague, Amsterdam (pp. 17-24). New York: ACM.

Crabtree, B. and Rhodes, B. (1998) Wearable Computing and the Remembrance Agent. *BT Technology Journal, 16*(3), 118-124.

Cross, N. (2011) *Design Thinking.* Oxford: Berg Publishers.

Cross, N. (2001) Designerly ways of knowing: design discipline versus design science. *Design Issues, 17*(3), 49-55.

Cross, N. (1999) Natural intelligence in design. *Design Studies, 20*, 25-39.

Cross, N. (1982) Designerly ways of knowing. *Design Studies, 3*(4), 221-227.

Dahlbom, B. and Mathiassen, L. (1993) *Computers in context: the philosophy and practice of systems design.* Malden, MA: Blackwell Publishers Inc.

Danis, C. and Karat, J. (1995) Technology-Driven Design of Speech Recognition Systems. In *Proceedings of the 1st conference on Designing interactive systems: processes, practices, methods, techniques* (DIS '95), New York, USA (pp. 17-24). New York: ACM.

de Figueiredo, A. D. and da Cunha, P. R. (2007) Action Research and Design in Information Systems: Two Faces of a Single Coin. In N. Kock (Ed.), *Information Systems Action Research: An Applied View of Emerging Concepts and Methods* (pp. 61-95). New York: Springer.

de Sá, M. and Carrico, L. (2011) Designing and Evaluating Mobile Interaction: Challenges and Trends. *Foundations and Trends in Human Computer Interaction, 4*(3), 175-243.

de Sá, M. and Carrico, L. (2009) A Mobile Tool for In-Situ Prototyping. In *Proceedings of MobileHCI 2009*, Bonn, Germany (article 20). New York: ACM.

Design Research Society (1966) *First statement of rules, inaugural meeting of the society, 1966*. Retrieved January, 2011, from: http://www.designresearchsociety.org/

Dey, A. K. (2001) Understanding and Using Context. *Personal and Ubiquitous Computing, 5*(1), 4-7.

Dey, A. K. and Gregory, D. A (2000) Towards a better understanding of context and context- awareness. In *Proceedings of Workshop on the What, Who, Where, When and How of Context-Awareness, CHI 2000*. Retrieved from: ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99-22.pdf

Dix, A., Finlay, J., Abowd, G. and Beale, R. (2004) *Human-Computer Interaction* (3rd ed.). London: Prentice Hall Europe.

Dix, A., Rodden, T., Davies, N., Trevor, J., Friday, A. and Palfreyman, K. (2000) Exploiting Space and Location as a Design Framework for Interactive Mobile Systems. *ACM Transactions on Computer-Human Interaction 7*(3), 285-321.

Dourish, P. (2004) What we talk about when we talk about context. *Personal and Ubiquitous Computing, 8*(1), 19-30.

Dourish, P. (2001a) *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge: The MIT Press.

Dourish, P. (2001b) Seeking a Foundation for Context-Aware Computing. *Human-Computer Interaction, 16*(2), 229-241.

Dreyfuss, H. (1955) *Designing for People* (2003 ed.). New York: Allworth Press.

Dubberly, H. and Evenson, S. (2011) Design as Learning – or "Knowledge Creation" – the SECI Model. *Interactions, 18*(1), 75-79.

Dubberly, H., Evenson, S. and Robinson, R. (2008) On modeling: The analysis-synthesis bridge model. *Interactions, 15*(2), 57-61.

Edmonds, B. (1999) Pragmatic Holism (or pragmatic reductionism). *Foundations of Science, 4*(1), 57-82.

Edwards, W. K. (2005) Putting Computing in Context: An Infrastructure to Support Extensible Context-Enhanced Collaborative Applications. *ACM Transactions On Computer-Human Interaction, 12*(4), 446-474.

Ehn, P. and Kyng, M. (1991) Cardboard Computers: Mocking-it-up or Hands-on the Future. In J. Greenbaum and M. Kyng (Eds.), *Design at Work* (pp. 169–196). Hillsdale, New Jersey: Laurence Erlbaum Associates Publishers.

Ehn, P. and Kyng, M. (1987) The Collective Resource Approach to Systems Design. In G. Bjerknes, P. Ehn and M. Kyng (Eds.), *Computers and democracy: A Scandinavian challenge* (pp. 251–278). Aldershot, UK: Avebury.

Erickson, T. (2006) Five Lenses: Towards a Toolkit for Interaction Design. In S. Bagnara and G. S. Smith (Eds.), *Theories and Practice in Interaction Design*. London: Lawrence Erlbaum Associates Publishers.

Einstein, A. (1931) *Cosmic Religion: with other Opinions and Aphorisms.* New York: Covici-Friede.

Fling, B (2009) *Mobile Design and Development: Practical Concepts and Techniques for Creating Mobile Sites and Web Apps.* O'Reilly Media.

Fortunati, L. (2001) The Mobile Phone: An Identity on the move. *Personal and Ubiquitous Computing, 5*(2), 85-98.

Frederick, G. and Lal, R. (2010) *Beginning Smartphone Web Development: Building Javascript, CSS, HTML and Ajax-Based Applications for iPhone, Android, Palm Pre, Blackberry, Windows Mobile and Nokia S60.* Apress.

Gasson, S. (2006) *Emergence in Organizational 'Problem-solving: Theories of Social Cognition*. Retrieved Sept, 2011, from: http://www.ischool.drexel.edu/faculty/sgasson/papers/probSolv.html

Gasson, S. (2003) Human-Centered Vs. User-Centered Approaches to Information System Design. *Journal of Information Technology Theory and Application, 5*(2), 29-46.

Gaver, B., Dunne, T. and Pacenti, E. (1999) Design: Cultural Probes. *Interactions, 6*(1), 21-29.

Graham, R. and Carter, C. (1999) Comparison of Speech Input and Manual Control of In-Car Devices while on-the-move. In *Proceedings of the Second Workshop on Human Computer Interaction with Mobile Devices (Mobile HCI 1999*), Edinburgh, Scotland.

Green, N., Harper, R. and Cooper, G. (2001) Configuring the Mobile User: Sociological and Industry Views. *Personal and Ubiquitous Computing, 5*(2), 146-156.

Greenbaum, J. and Mathiassen, L. (1990) Zen and the Art of Teaching Systems Development. *ACM Computers and Society, 20*(1), 26-30.

Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R. and Wang, M. (2011) Proxemic interactions: the new ubicomp? *Interactions, 18*(1), 42-50.

Greenberg, S. (2001) Context as a dynamic construct. *Human-Computer Interaction, 16*(2), 257-268.

Greenfield, A. (2006) *Everyware: the dawning age of ubiquitous computing.* Berkeley: New Riders.

Gye, L. (2007) Picture This: the Impact of Mobile Camera Phones on Personal Photographic Practices- Continuum. *Journal of Media and Cultural Studies, 21*(2), 279-288.

Hagen, P., Robertson, T., Kan, M. and Sadler, K. (2005) Emerging research methods for understanding mobile technology use. In *Proceedings of OZCHI 2005*, Canberra, Australia (pp. 1-10). New York: ACM.

Harman, G. H. (1965) The Inference to the Best Explanation. *Philosophical Review, 74* (1), 88-95.

Harper, R., Rodden, T., Rogers, Y. and Sellen, A. (2008) *Being Human: Human-Computer Interaction in the year 2020.* Cambridge: Microsoft Research.

Harrison, S., Tatar, D. and Sengers, P. (2007) The three paradigms of HCI. In *Proceedings of CHI'07, alt.chi*. New York: ACM.

Helal, A., Haskell, B., Carter, J. L., Brice, R., Woelk, D. and Rusinkiewich, M. (1999) *Any time, anywhere computing: mobile computing concepts and technology.* Boston, MA: Kluwer Academic Publishers.

Hinckley, K., Pierce, J., Sinclair, M. and Horvitz, E. (2000) Sensing techniques for mobile interaction. In *Proceedings of UIST 2000* (pp. 91-100). New York: ACM.

Hinckley, K. and Horvitz, E. (2001) Toward More Sensitive Mobile Phones. In *Proceedings of UIST 2001*, Orlando, Florida, USA (pp. 191-192). New York: ACM.

Hinckley, K., Pierce, J., Horvitz, E. and Sinclair, M. (2005) Foreground and background interaction with sensor-enhanced mobile devices. *ACM Transactions on Computer-Human Interaction, 12*(1), 31-52.

Hosbond, J. H. (2005) Mobile Systems Development: Challenges, Implications and Issues. In *Proceedings of MOBIS 2005*, Leeds, UK, IFIP TC8.

Hutchins, E. (1995) *Cognition in the Wild.* Cambridge, MA: The MIT Press.

Høegh, R. T., Kjeldskov, J., Skov, M. B. and Stage, J. (2008) A Field Laboratory for Evaluating In Situ. In J. Lumsden (Ed.), *Handbook of Research on User Interface Design and Evaluation for Mobile Technology* (pp. 982-996). PA, USA: Idea Group Inc (IGI).

Höök, K. (2012) A Cry for More Tech at CHI!. *Interactions, 19*(2), 10-11.

IDEO (2009) Human Centered Design Toolkit, 2nd Edition. Retrieved January, 2011, from: http://www.ideo.com/images/uploads/hcd_toolkit/IDEO_HCD_ToolKit_Complete_for_Download.pdf

Jameson, A. (2001) Modeling both the Context and the User. *Personal and Ubiquitous Computing, 5*(2), 29-33.

Jobs, S. (2011) *Apple Special Event*. San Francisco, 2 March, 2011. Retrieved March, 2011, from: http://itunes.apple.com/us/podcast/apple-keynotes/id275834665

Jobs, S. (2010) *Apple Announces iPad*. San Francisco, 27 January, 2010. Retrieved February, 2010, from: http://itunes.apple.com/us/podcast/apple-keynotes/id275834665

Johnson, P. (1998) Usability and Mobility: Interactions on the move. In *Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices*, Glasgow, Scotland (GIST Technical Report G98-1).

Jones, M. and Marsden, G. (2006) *Mobile Interaction Design.* Glasgow: John Wiley and Sons, Ltd.

Jones, Q., Grandhi, S. A., Terveen, L. and Whittaker, S. (2004) People-to-People-to-Geographical-Places: The P3 Framework for Location-Based Community Systems. *Computer Supported Cooperative Work, 13*, 249-282.

Jung, H., Stolterman, E., Ryan, W., Thompson, T. and Siegel, M. (2008) Toward a Framework for Ecologies of Artifacts: How Are Digital Artifacts Interconnected within a Personal Life? *In Proceedings of the 5th Nordic Conference on Human-Computer Interaction, NordiCHI 2008*, Lund, Sweden (pp 201-210). New York: ACM.

Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T. and Kankainen, A. (2005) Usability Testing of Mobile Applications: A Comparison between Laboratory and Field Testing. *Journal of Usability Studies, 1*(1), 4-17.

Karapanos, E., Zimmerman, J., Forlizzi, J. and Martens, J. (2009) User Experience Over Time: An Initial Framework. In *Proceedings of CHI 2009*, Boston, MA, USA (pp. 729-738). New York: ACM.

Kay, Alan (1972). A Personal Computer for Children of All Ages. In *Proceedings of ACM National Conference*. New York: ACM.

Kensing, F. and Blomberg, J. (1998) Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work, 7*(1), 167–185.

Kindberg, T., Spasojevic M., Fleck R. and Sellen A. (2005) The Ubiquitous Camera: An In-Depth Study of Camera Phone Use. *IEEE Pervasive Computing, 4*(2), 42-50.

Kjeldskov, J. and Paay, J. (2012) A longitudinal review of mobile HCI research methods. In *Proceedings of Mobile HCI 2012,* San Francisco, USA (pp. 69-78). New York: ACM.

Kjeldskov, J., Skov, M. B., Nielsen, G. W., Thorup, S. and Vestergaard, M. (2012b) Digital Urban Ambience: Mediating Context on Mobile Devices in the City. *Journal of Pervasive and Mobile Computing* (in press).

Kjeldskov, J., Skov, M. B., Paay, J. and Pathmanathan, R. (2012a) Using Mobile Phones to Support Sustainability: A Field Study of Residential Electricity Consumption. In *Proceedings of CHI 2012*, Austin, Texas, USA. New York: ACM.

Kjeldskov, J. and Stage, J. (2012) Combining ethnography and object-orientation: contextual richness and abstract models for mobile interaction design. *International Journal of Human-Computer Studies, 70*(3), 197–217.

Kjeldskov, J. and Paay, J. (2010) Indexicality: understanding mobile human-computer interaction in context. *ACM Transactions on Computer-Human Interaction (TOCHI). 17*(4)

Kjeldskov, J., Christensen, C. M. and Rasmussen, K. K. (2010) GeoHealth: a location-based service for home healthcare workers. *Journal of Location-Based Services, 4*(1), 3-27.

Kjeldskov, J. and Skov, M. B. (2007b) Exploring Context-Awareness for Ubiquitous Computing in the Healthcare Domain. *Personal and Ubiquitous Computing, 11*(7), 549-562.

Kjeldskov, J. and Skov, M. B. (2007a) Studying Usability in Sitro: Simulating Real World Phenomena in Controlled Environments. *International Journal of Human-Computer Interaction, 22*(1), 7-37.

Kjeldskov J. and Stage J. (2006) Exploring "Canned Communication" for Coordinating Distributed Mobile Work Activities. *Interacting with Computers, 2006*(18), 1310-1335.

Kjeldskov, J. and Paay, J. (2006) Public Pervasive Computing in the City: Making the Invisible Visible. *IEEE Computer, 39*(9), 60-65.

Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S. and Davies, J. (2005) Evaluating the Usability of a Mobile Guide: The influence of Location, Participants and Resources. *Behaviour and Information Technology, 24*(1), 51-65.

Kjeldskov, J. and Paay, J. (2005) Just-for-Us: A Context-Aware Mobile Information System Facilitating Sociality. In *Proceedings of Mobile HCI 2005*, Salzburg, Austria (pp. 23-30). New York: ACM.

Kjeldskov, J., Gibbs, M., Vetere, F., Howard, S., Pedell, S., Mecoles, K. and Bunyan, M. (2004b) Using Cultural Probes to Explore Mediated Intimacy. Australasian Journal of Information Systems, 12(1), 102-115

Kjeldskov, J., Skov, M. B., Als, B. S. and Høegh, R. T. (2004a) Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In *Proceedings of MobileHCI 2004*, Glasgow, Scotland, LNCS (pp. 61-73). Berlin: Springer-Verlag.

Kjeldskov, J. and Howard, S. (2004) Envisioning Mobile Information Services: Combining User- and Technology-Centered Design. In *Proceedings of the 6th Asia-Pacific Conference on Human-Computer Interaction (APCHI 2004*), Rotorua, New Zealand, LNCS (pp. 180-190). Berlin: Springer-Verlag.

Kjeldskov, J. and Stage, J. (2004) New Techniques for Usability Evaluation of Mobile Systems. *International Journal of Human-Computer Studies, 60*(2004), 599-620.

Kjeldskov, J. (2003) *Human-Computer Interaction Design for Emerging Technologies: Virtual Reality, Augmented Reality and Mobile Computer Systems.* Ph.D. Thesis, Department of Computer Science, Aalborg University, Denmark. ISSN 1601-0590 (no 23).

Kjeldskov, J. and Graham, C. (2003) A Review of MobileHCI Research Methods. In *Proceedings of the 5th International Mobile HCI 2003 conference*, Udine, Italy, LNCS (pp. 317-335). Berlin: Springer-Verlag.

Kjeldskov J., Howard S., Murphy J., Carroll J., Vetere F. and Graham C. (2003) Designing TramMate - a context aware mobile system supporting use of public transportation. In *Proceedings of the 1st Conference on Designing User Experiences (DUX 2003)*, San Francisco, CA, USA (pp. 1-4). New York: ACM.

Kjeldskov, J., Skov, M. B. and Stage, J. (2002) *Usability Evaluation of the Autolocate WAP-Service* (in Danish). (HCI Lab Technical Report no. 2002/1, January 2002). Department of Computer Science, Aalborg University, Denmark.

Kostakos, V., Nicolai, T., Yoneki, E., O'Neill, E., Kenn, H. and Crowcroft, J. (2009) Understanding and measuring the urban pervasive infrastructure. *Personal Ubiquitous Computing, 13*(5), 355-364.

Krogstie, J., Lyytinen, K., Opdahl, A.L., Pernici, B., Siau, K. and Smoland K. (2004) Research areas and challenges for mobile information systems. *International Journal of Mobile Communications, 2*(3), 220-234.

Köhler, W. (1947) *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. New York: Liveright Publishing Corporation.

Lanzara, G. F. (1983) The design process: frames, metaphors and games. In U. Briefs, C. Ciborra, L. Schneider (Eds.), *Systems Design For, With and By The Users*. North-Holland Publishing Company.

Lauesen, S. (2005) *User Interface Design: A software engineering perspective*. Harlow: Addison-Wesley.

Laurel, B. (Ed.) (1990) *The Art of Human-Computer Interface Design.* Reading, MA: Addison-Wesley.

Lewis, I. M. (1985) *Social Anthropology in Perspective*. Cambridge University Press.

Lindholm, C., Keinonen, T. and Kiljander, H. (2003) *Mobile Usability: How Nokia Changed the Face of the Mobile Phone.* New York: McGraw-Hill.

Lindley, S. E., Harper, R., Randall, D., Glancy, M. and Smyth, N. (2009) Fixed in Time and "Time in Motion": Mobility of Vision through a SenseCam Lens. In *Proceedings of MobileHCI 2009*, Bonn, Germany (article 2). New York: ACM.

Ling, R. (2001) We Release Them Little by Little: Maturation and Gender Identity as Seen in the Use of Mobile Telephony. *Personal and Ubiquitous Computing, 5*(2), 123-136.

Little, L. and Briggs, P. (2009) Private whispers/public eyes: Is receiving highly personal information in a public place stressful? *Interacting with Computers, 21*(4), 316-322.

Luff, P. and Heath, C. (1998) Mobility in Collaboration. In *Proceedings of CSCW'98,* Seattle, USA (pp. 305-314). New York: ACM.

Luther, K. and Diakopoulos, N. (2007) Distributed Creativity. In *Proceedings of Creativity and Cognition Workshop on Supporting Creative Acts Beyond Dissemination,* Washington, DC.

Lynch, K. (1960) *The Image of the City.* Cambridge, MA: The MIT Press.

Lyons, K. and Starner, T. (2001) Mobile Capture for Wearable Computer Usability Testing. In *Proceedings of the 5th IEEE International Symposium on Wearable Computers*, Zurich, Switzerland. IEEE Press.

Martin, R. (2009) *The Design of Business: Why Design Thinking is the Next Competitive Advantage.* Harvard Business Press.

Mathiassen, L. (2012) Development of IT-enabled Chronic Care Management: A Contextualist Framework. *Information Systems Journal* (forthcoming).

McGillick, P. and Carlstrom, K. (2002) *Alex Popov: Buildings and Projects*. Axel Menges.

Meggs, P. B. and Purvis, A. W. (2005) *Meggs' History of Graphic Design* (4th ed.). John Wiley and Sons.

McCullough, M. (2004) *Digital Ground - Architecture, Pervasive Computing, and Environmental Knowing.* Cambridge, MA: The MIT Press.

Mendel, J. and Yeager, J. (2010) Knowledge Visualization in Design Practice: Exploring the Power of Knowledge Visualization in Problem Solving. *Parsons Journal for Information Mapping, 2*(3), 1-4.

Mikkonen, M., Vayrynen, S., Ikonen, V. and Heikkila, O. (2002) User and Concept Studies as Tools in Developing Mobile Communication Services. *Personal and Ubiquitous Computing, 2002*(6), 113-124.

Millen, D. R. (2000): Rapid ethnography: time deepening strategies for HCI field research. In *Proceedings of DIS 2000*, New York, USA (pp. 280-286). New York: ACM.

Miller, F. P., Vandome, A. F. and McBrewster, J. (2010) *Digital Ecosystem*. Alphascript Publishing.

Moggridge, B. (2007) *Designing Interactions*. Cambridge, MA: The MIT Press.

Murphy, J., Kjeldskov, J., Howard, S., Shanks, G. and Hartnell-Young, E. (2005) The Converged Appliance: "I Love it… But I Hate it". In *Proceedings of OzCHI 2005*, Canberra, Australia (pp. 1-10). New York: ACM.

Myers, M. D. (1997) Qualitative Research in Information Systems. *MIS Quarterly, 21*(2), 241-242.

Nelson, H. G. and Stolterman, E. (2003) *The design way – intentional change in an unpredictable world*. Englewood Cliffs, New Jersey: Educational Technology Publications.

Nielsen, C. M., Overgaard, M., Pedersen, M. B., Stage, J. and Stenild, S. (2006) It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. In *Proceedings of NordiCHI 2006* (pp. 272-280). New York: ACM.

Nielsen, J. (2000b) *WAP Field Study Findings*. Alertbox. Retrieved December, 2000, from: http://www.useit.com/alertbox/20001210.html

Nielsen, J. (1993) *Usability Engineering*. Boston: Academic Press.

Nielsen, J. and Molich, R. (1990) Heuristic evaluation of user interfaces. In *Proceedings of CHI 1990*, Seattle, WA (pp. 249-256). New York: ACM.

Nonaka, I. and Toyama, R. (2002) A firm as a dialectical being: towards a dynamic theory of a firm. *Industrial and Corporate Change, 11*(5), 995-1009.

Norberg-Schulz, C. (1980) *Genius Loci: Towards a Phenomenology of Architecture.* New York: Rizzoli.

Norman, D. A. (2010). Technology First, Needs Last: The Research-Product Gulf. *Interactions, 17*(2), 38-42.

Norman, D. A. (2005) Human-Centred Design Considered Harmful. *Interactions, 12*(4), 14-19.

Norman, D. A. (1998). *The Invisible Computer - Why Good Products Can Fail, the Personal Computer Is So Complex and Information Appliances Are the Solution*. Cambridge, MA: The MIT Press.

Norman, D. A. (1990) Four (more) issues for cognitive science. (Cognitive Science Technical Report No. 9001). Department of Cognitive Science, University of California, San Diego.

Norman, D. A. and Draper, S. W. (Eds.) (1986) *User Centred System Design: New Perspectives on Human-Computer Interaction.* Hillsdale: Lawrence Erlbaum Associates Publishers.

O'Hara, K., Kjeldskov, J. and Paay, J. (2011) Blended Interaction Spaces for Distributed Team Collaboration. *ACM Transactions on Computer-Human Interaction, 18*(1), article 3.

Olofsson, E. and Sjölén, K. (2005) *Design Sketching. Including an extensive collection of inspiring sketches by 24 students at the Umeå Institute of Design* (2nd ed.). Klippan: Ljungbergs Tryckeri.

Oulasvirta, A. (2009) Field Experiments in HCI: Promises and Challenges. In P. Saariluoma and H. Isomaki (Eds.), *Future Interaction Design II*. Berlin: Springer-Verlag.

Oulasvirta, A. and Nyyssonen, T. (2009). Flexible hardware configurations for studying mobile usability. *Journal of Usability Studies*, *4*(2), 93-105.

Oulasvirta, A., Tamminen, S. and Höök, K. (2005) Comparing two approaches to context: realism and constructivism. In *Proceedings of the 4th Decennial Conference on Critical Computing (CC'05)*, Aarhus, Denmark (pp. 195-198). New York: ACM.

Paay, J., Kjeldskov, J., Howard S. and Dave, B. (2009b) Out on the town: a socio-physical approach to the design of a context aware urban guide. *Transactions on Computer-Human Interaction, 16*(2), 7-34.

Paay, J., Sterling, L., Vetere, F., Howard, S. and Boettcher, A. (2009a) Engineering the social: The role of shared artifacts. *International Journal of Human-Computer Studies, 67*(5), 437-454.

Paay, J., (2008) From ethnography to interface design. In J. Lumsden (Ed.), *Handbook of Research on User Interface Design and Evaluation for Mobile Technology* (pp. 1-15). PA, USA: Idea Group Inc (IGI).

Paay, J. and Kjeldskov, J. (2008b) Understanding the user experience of location based services: five principles of perceptual organization applied. *Journal of Location-Based Services, 2*(4), 267-286.

Paay, J. and Kjeldskov, J. (2008a) Situated Social Interactions: a Case Study of Public Places in the City. *Computer-Supported Cooperative Work, 17*(2-3), 275-290.

Paay, J. and Kjeldskov, J. (2005) Understanding and Modelling the Built Environment for Mobile Guide Interface Design. *Behaviour and Information Technology, 24*(1), 21-35.

Palen, L., Salzman, M. and Youngs, E. (2000) Going Wireless: Behavior & Practice of New Mobile Phone Users. In *Proceedings of CSCW 2000*, Philadelphia, PA, USA (pp. 201-210). New York: ACM.

Pascoe, J., Ryan, N. and Morse, D. (2000) Using While Moving: HCI Issues in Fieldwork Environments. *Transactions on Computer-Human Interaction, 7*(3), 417-437.

Peirce, C. S. (1931-58) Collected Writings (8 Vols.). In C. Hartshorne, P. Weiss, and A. Burks (Eds.), *Collected Papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.

Pepper, S. C. (1942) *World Hypothesis: A Study in Evidence*. Berkeley: University of California Press.

Perry, M., O'Hara, K., Sellen, A., Brown, B. and Harper, R. (2001) Dealing with mobility: understanding access anytime, anywhere. *ACM Transactions on Computer-Human Interaction, 8*(4), 323-347.

Pettigrew, A. M. (1990) Longitudinal field research on change: theory and practice. *Organization Science, 1*(3), 267-292.

Pettigrew, A. M. (1987) Context and Action in the Transformation of the Firm. *Journal of Management Studies, 24*(6), 649-670.

Pettigrew, A. M. (1985) Contextualist Research: a Natural Way to Link Theory and Practice. In E. E. Lawler (Ed.), *Doing Research That Is Useful in Theory and Practice*. San Francisco: Jossey Bass.

Preece, J., Rogers, Y. and Sharp H. (2002) *Interaction Design: beyond human-computer interaction.* New York: John Wiley and Sons Ltd.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey, T. (1994). *Human-Computer Interaction*. Workingham: Addison-Wesley.

Preece, J., Benyon, D., Davies, G., Keller, L. and Rogers, Y. (1993) *A Guide to Usability: Human Factors in Computing.* Harlow: Addison-Wesley.

Raman, V. V. (2005) Scientific Reductionism and Holism: Two Sides of the Perception of Reality. *Theology and Science, 3*(3), 250-253.

Ramsay, M. and Nielsen, J. (2000a) *WAP Usability: Déjà Vu: 1994 All Over Again*. (Nielsen Norman Group Report, December 2000). Retrieved from: http://www.nngroup.com/reports/wap/

Rantanen, J., Impio, J., Karinsalo, T., Reho, A., Tasanen, M. and Vanhala, J. (2002) Smart Clothing Prototype for the Arctic Environment. *Personal and Ubiquitous Computing, 2002*(6), 3-16.

Rapoport, R. N. (1970) Three Dilemmas in Action Research. *Human Relations, 23*(4), 499-513.

Raskin, J. (2000) *The Humane Interface: New Directions for Interactive Systems*. Boston: Addison-Wesley.

Rasmussen, L. B. (2007) From human-centred to human-context centred approach: looking back over 'the hills', what has been gained and lost? *AI & Society, 2007*(21), 471-495.

Ray, K. (1980) *Contextual Architecture: Responding to existing style*. New York: McGraw-Hill.

Reichl, P., Frohlich P., Baillie L., Schatz R. and Dantcheva A. (2007) The LiLiPUT Prototype: A Wearable Lab Environment for User Tests of Mobile Telecommunication Applications. *CHI 2007 Extended Abstracts*, San Jose, California, USA (pp. 1833-1838). New York: ACM.

Rittel H. W. J. and Webber M. M. (1973) Dilemmas in a General Theory of Planning. *Policy Sciences, 4*(2), 155-196.

Robertson, T. (1997) Cooperative Work and Lived Cognition: A Taxonomy of Embodied Actions. In *Proceedings of ECSCW 1997*, (pp. 205-220).

Rodden, T., Cheverst, K., Davies, N. and Dix, A. (1998) Exploiting Context in HCI Design for Mobile Systems. In *Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices*, Glasgow, Scotland (GIST Technical Report G98-1).

Rogers, Y., Sharp, H. and Preece, J. (2011) *Interaction Design: beyond human-computer interaction* (3rd ed.). New York: John Wiley and Sons Ltd.

Rogers, Y., Connelly, K., Tedesco, L., Hazlewood, W., Kurtz, A., Hall, R. E., Hursey, J. and Toscos, T. (2007) Why it's worth the hassle: the value of in-situ studies when designing Ubicomp. In *Proceedings UbiComp 2007*, LNCS (pp. 336-353). Berlin: Springer-Verlag.

Rogers, Y., Scaife M. and Rizzo, A. (2005) Interdisciplinarity: an Emergent or Engineered Process. In S. J. Derry, C. D. Schunn and M. A. Gernsbacher (Eds.), *Interdisciplinary Collaboration*. Mahwah, New Jersey: LEA.

Rogers, Y. (2004) New theoretical approaches for human-computer interaction. *Annual review of Information Science and Technology (ARIST), 38,* 87-143.

Rogers, Y., Scaife, M., Harris, E., Phelps, T., Price, S., Smith, H., Muller, H., Randell, C., Moss, A., Taylor, I., Stanton, D., O'Malley, C., Corke, G. and Gabrielli, S. (2002). Things aren't what they seem to be: innovation through technology inspiration. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques (DIS 2002)(pp.* 373-378). New York: ACM.

Rosson, M. B. and Carroll, J. M. (2001) *Usability engineering: scenario-based development of human-computer interaction.* San Francisco, CA: Morgan Kaufmann Publishers.

Rowland, D., Flintham, M., Oppermann, L. Marshall, J., Chamberlain, A., Koleva, B., Benford, S. and Peres, C. (2009) Ubikequitous Computing: Designing Interactive Experiences for Cyclists. In *Proceedings of MobileHCI 2009*, Bonn, Germany (article 21). New York: ACM.

Rubin, J. (1994) *Handbook of Usability Testing: How to plan, design, and conduct effective tests.* New York: John Wiley and Sons Ltd.

Saffer, D. (2007) *Designing for Interaction: Creating Smart Applications and Clever Devices*. Berkeley, CA: AIGA Design Press/New Riders.

Schmidt, A., Aidoo, K. A., Takaluoma, A., Tuomela, U., Van Laerhoven, K. and Van de Velde, W. (1999b) Advanced Interaction in Context. In *Proceedings of HUC 1999* (pp. 89-101). London: Springer-Verlag.

Schmidt, A., Beigl, M. and Gellersen, H. (1999a) There is more to Context than Location. *Computers and Graphics Journal, 23*(6), 893-902.

Schilit, B. and Theimer, M. (1994) Disseminating active map information to mobile hosts. *IEEE Network, 8*(5), 22-32.

Schön, D. A. (1983) *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Schön, D. A. (1992) Designing as reflective conversation with the materials of a design situation. *Research and Engineering Design, 3*(3), 131-147.

Shane, G. (1976) Contextualism. *Architectural Design, 46*(11), 676-679.

Sharp, H., Rogers, Y. and Preece, J. (2007) *Interaction Design: beyond human-computer interaction* (2nd ed.). Barcelona: John Wiley and Sons Ltd.

Sharples, M., Corlett, D. and Westmancott, O. (2002) The Design and Implementation of a Mobile Learning Resource. *Personal and Ubiquitous Computing, 2002*(6), 220-234.

Shneiderman, B. (2000) Creating creativity: user interfaces for supporting innovation. *ACM Transactions on Computer-Human Interaction, 7*(1), 114-138.

Shneiderman, B. (1998) *Designing the User interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Workingham: Addison-Wesley.

Simon, H. A. (1969) *The Sciences of the Artificial*. Cambridge, MA: The MIT Press.

Simon, H. A. (1962) The Architecture of Complexity. *Proceedings of the American Philosophical Society, 106*(6), 467-482.

Skov, M. B., Kjeldskov, J., Paay, J., Husted, N., Nørskov, J. and Pedersen, K. (2012) Designing on-site: Facilitating Participatory Contextual Architecture with Mobile Phones. *Journal of Pervasive and Mobile Computing (in press).*

Smith, G. C. (2007) What in Interaction Design? In B. Moggridge (2007) *Designing Interactions*. Cambridge, MA: The MIT Press.

Snyder, C. (2003) *Paper Prototyping: the fast and easy way to design and refine user interfaces*. Amsterdam: Morgan Kaufmann Publishers.

Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C. and DeAngelo, T. (1999) *Web Site Usability: A designers guide.* San Francisco: Morgan Kaufmann Publishers.

Star, S. L. and Griesemer, J. R. (1989) Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*, 387-420.

Stokholm, M. (2010) *Stepping Stones I - en model for integrering af proces elementer*. Unpublished working paper.

Stokholm, M. (2008) A Holistic Approach to Interdisciplinary Innovation Supported by a Simple Tool. In *Proceedings of the 9th International Conference on Human Factors in Organizational Design and Management* (pp. 143-149). Santa Monica: IEA Press.

Stolterman, E. (2008) The nature of design practice and implications for interaction design research. *International Journal of Design, 2*(1), 55-65.

Strauss, A. and Corbin, J. (1990) *Basics of Qualitative Research*. Newbury Park, California: Sage Publications.

Tamminen, S., Oulasvirta, A., Toiskallio, K. and Kankainen, A. (2004) Understanding mobile contexts. *Personal and Ubiquitous Computing, 8*(2), 135-143.

The Royal Society (1996) *Interdisciplinarity - Transport and the Environment*. Retrieved May, 2010, from: http://royalsociety.org/WorkArea/DownloadAsset.aspx?id=5954

Tufte E. R. (2001) *The Visual Display of Quantitative Information* (2nd ed.). Cheshire, Connecticut: Graphics Press.

Tufte, E. R. (1997) *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Connecticut: Graphics Press.

Tufte, E. R. (1990) *Envisioning Information.* Cheshire, Connecticut: Graphics Press.

Turner, J. A. (1987) Understanding The Elements Of Systems Design. In R. J. Boland and R. A. Hirschheim (Eds.), *Critical Issues In Information Systems Research*. John Wiley and Sons Ltd.

van den Ende, J. and Dolfsma, W. (2004) Technology-push, demand-pull and the shaping of technological paradigms: Patterns in the development of computing technology. *Journal of Evolutionary Economics, 15*(1), 83-99.

Verganti, R. (2010) Apple's Secret? It Tells Us What We Should Love. Retrieved January, 2011, from: http://blogs.hbr.org/cs/2010/ 01/how_apple_innovates_by_telling.html

Verganti, R. (2009) *Design-driven innovation: Changing the rules of competition by radically innovating what things mean.* Boston: Harvard Business Press.

Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoles, K. and Bunyan, M. (2005) Mediating Intimacy: Designing Technologies to Support Strong-Tie Relationships. In *Proceedings of CHI 2005*, Portland, Oregon, USA (pp. 471-480). New York: ACM.

Viller, S. and Sommerville, I. (2000) Ethnographically informed analysis for software engineers. *International Journal of Human–Computer Studies, 53*, 169-196.

Weilenmann, A. (2001) Negotiating Use: Making Sense of Mobile Technology. *Personal and Ubiquitous Computing, 5*(2), 137-145.

Weiser, M. (1991) The Computer for the 21st Century. *Scientific American, 265*(3), 94-104.

Weiss, S. (2002) *Handheld Usability*. Milan: John Wiley and Sons Ltd.

Wilde, O. (1889) *The Decay of Lying*. Penguin Classics

Winograd, T. (1997) The Design of Interaction (aka. From Computing Machinery to interaction design). In P. Denning and R. Metcalfe (Eds.), *Beyond Calculation: The Next Fifty Years of Computing.* Amsterdam: Springer-Verlag.

Winograd, T. (1996) *Bringing Design to Software*. Reading, MA: Addison-Wesley.

Wynekoop, J. L. and Conger, S. A. (1990) A Review of Computer Aided Software Engineering Research Methods. In *Proceedings of the IFIP TC8 WG 8.2 Working Conference on The Information Systems Research Arena of The 90's*, Copenhagen, Denmark.

Yin, R. K. (1994) *Case Study Research, Design and Methods* (2nd ed.). Newbury Park: Sage Publications.

Zuberec, S. (2000) The interaction design of Microsoft Windows CE. In E. Bergman (Ed.), *Information Appliances and Beyond*. San Francisco: Morgan Kaufmann Publishers.

Challenges and opportunities

Chapter 2. A review of mobile interaction design research

Chapter 2

A review of mobile interaction design research

Jesper Kjeldskov and Connor Graham

Abstract. This paper examines and reviews research methods applied within the field of mobile human-computer interaction. The purpose is to provide a snapshot of current practice for studying mobile HCI to identify shortcomings in the way research is conducted and to propose opportunities for future approaches. 102 publications on mobile human-computer interaction research were categorized in a matrix relating their research methods and purpose. The matrix revealed a number of significant trends with a clear bias towards building systems and evaluating them only in laboratory settings, if at all. Also, gaps in the distribution of research approaches and purposes were identified; action research, case studies, field studies and basic research being applied very infrequently. Consequently, we argue that the bias towards building systems and a lack of research for understanding design and use limits the development of cumulative knowledge on mobile human computer interaction. This in turn inhibits future development of the research field as a whole.

1. INTRODUCTION

The study of human computer interaction for mobile devices is a relatively young research field in which commercially successful devices have only been available for less than a decade and leading conferences have only a few years of history. In young research fields there is often a tendency to be highly opportunity and technology driven and to focus primarily on producing *solutions* while reflecting less on methodology. This characterized early computer research and can also be seen in relation to emerging research areas such as virtual and augmented reality. As a research field matures, examining how the research is being conducted and reflecting on the impact of this on the knowledge being produced is necessary in order to be able to understand and influence the future direction of the field. So far, this has not been done consistently within the community

Originally published as Kjeldskov, J. and Graham, C. (2003) A Review of MobileHCI Research Methods. In *Proceedings of Mobile HCI 2003,* Udine, Italy. LNCS (pp. 317-335). Berlin, Springer-Verlag.

of mobile HCI and consequently little knowledge on a methodological level exists about the research field. This analysis and discussion will be borne out by this paper. Inspired by related studies within the field of Information Systems (IS), we aim at evoking more discussion of research methodology in mobile HCI by presenting a snapshot of current research practice within our field, identifying and discussing shortcomings in present research and opportunities for future approaches.

Focus and reflection on research methodology has been a key subject within information system research for decades (see for example Basili et al. 1986; Benbasat 1987; Galliers 1990, Myers 1997, Wynekoop and Conger 1990). Facilitating this discussion, a number of frameworks for describing and categorizing IS-research methods have been developed (see for example Galliers 1990), which could also be relevant in relation to discussions of mobile HCI research. Specifically, we find that the classification of computer-aided software engineering (CASE) research by Wynekoop and Conger (1990) demonstrate a generally usable (and relatively simple) approach to informing the discussion of research methods applied within a given area. Wynekoop and Conger (1990) reviewed and classified 40 IS-research papers in a two-dimensional matrix relating research methods and research purpose, providing a picture of the research field facilitating discussion of current research practice. In this paper we replicate elements from this study by applying its overall approach to the field of mobile HCI. In section 2 and 3 we present the categories of research methods and research purposes used in our classification. In section 4 we describe the conducted review of Mobile HCI research papers and present a matrix describing the resulting classification. Trends highlighted by this matrix are then discussed in section 5 and in sections 6 and 7 we indicate limitations, conclude our study and point out paths for further work.

2. RESEARCH METHODS

Defining and especially differentiating research methods can be a challenge. Definitions are sometimes vague and often different aspects of different methods overlap. As the purpose of this paper is not to discuss definitions of research methods as such, we have chosen to apply eight definitions extracted from Wynekoop and Conger (1990) with supplementary input from general references on research methodology in information systems (Lewis 1985; Myers 1997; Rapoport 1970; Yin 1994). Knowing that these definitions may themselves be objects for disagreement, we refer to Wynekoop and Conger (1990) and Galliers (1990) for further discussion of the definitions.

In this section, we present and review the eight research methods used in our classification of mobile HCI research. For each method, strengths and weaknesses are identified as well as primary uses and possible application in mobile HCI research. This discussion is summarized in table 1. As an overall categorization, we group the eight research methods according to Benbasat's (1985) categories of *natural, artificial* and *environment independent* settings.

Table 1. Summary of research methods, extracted from Wynekoop and Conger (1990)

	Method	Strengths	Weaknesses	Use
Natural setting	Case studies	Natural settings Rich data	Time demanding Limited generalizability	Descriptions, explanations, developing hypothesis
	Field studies	Natural Settings Replicable	Difficult data collection Unknown sample bias	Studying current practice Evaluating new practices
	Action research	First hand experience Applying theory to practice	Ethics, bias, time Unknown generalizability	Generate hypothesis/theory Testing theories/hypothesis
Artificial setting	Laboratory experiments	Control of variables Replicable	Context insensitive No variable manipulation	Controlled experiments Theory/product testing
Environment independent setting	Survey research	Easy, low cost Can reduce sample bias	Context insensitive No variable manipulation	Collecting descriptive data from large samples
	Applied research	The goal is a product which may be evaluated	May need further design to make product general	Product development, testing hypothesis/concepts
	Basic research	No restrictions on solutions Solve new problems	Costly, time demanding May produce no solution	Theory building
	Normative writings	Insight into firsthand experience	Opinions may influence outcome	Descriptions of practice, building frameworks

2.1. Case Studies

Yin (1994) defines a case study as "an empirical enquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident". Thus case studies are often intensive empirical studies of small size entities such as groups, organizations, individuals, systems or tools with the researcher distinct from the phenomena being studied (Wynekoop and Conger 1990). When conducting case studies, data is typically collected by a combination of various qualitative and quantitative means such as observations, interviews and questionnaires etc. with little experimental or statistical control enforced. The data collected is grounded in natural settings, typically very rich and sometimes contradictory or inconsistent, thus often resulting in complicated analysis. Case studies are particularly well suited for research focusing on describing and explaining a specific phenomenon and for developing hypothesis or theory through, for example, applying grounded-theory approaches. However, case studies are very time demanding and generalizing findings can be difficult.

Since mobile HCI is a relatively young research area, case studies could be used to provide rich data explaining phenomena involving mobility or the use of mobile devices in context.

2.2. Field Studies

Generally, field studies are characterized by taking place in "the real world" as opposed to in a laboratory setting. Field studies cover a range of qualitative and quantitative approaches from *ethnographic* studies of phenomena in their social and cultural context inspired by the discipline of social and cultural anthropology (Lewis 1985) to field *experiments* in which a number of independent variables are manipulated (Wynekoop and Conger 1990).

Ethnographic field studies are characterized by researchers spending significant amounts of time in the field and, to some extent, immersing themselves into the environment they study. Typically, data is gathered through observations and/or interviews and the phenomena studied are placed in a social and cultural context. The major advantage of ethnographic field studies is the generation of large amounts of rich and grounded data in relatively short time. The major disadvantages are unknown biases and no guarantee of collected data being representative.

While ethnographic field studies are non-experimental, field experiments are characterized by manipulation of a number of independent variables to observe the influence on dependant variables in a natural setting. The major advantages of field experiments are increased realism and increased control in comparison to ethnographic field studies and support for studying complex situated interactions and processes. Disadvantages include limited control of experiments and complicated data collection compared to, for example, experiments in laboratory settings. Furthermore, as experimental manipulation increases, realism typically decreases.

In relation to mobile HCI research, field studies could be applied for either informing design for or understanding of mobility by ethnographic studies of current practice or for evaluating design or theory by conducting experiments in realistic use settings.

2.3. Action Research

Originating from the social sciences, action research is a well-established research method through which researchers not only apply scientific knowledge to an object of study, but also add to the body of scientific knowledge through that study, thus differentiating action research from applied science or research (Myers 1997). Conducting action research, the researcher participates in the intervention of the activity or phenomenon being studied while at the same time evaluating the results (Wynekoop and Conger 1990). More specifically, Rapoport (1970) defines action research as aiming "to contribute both to the practical concerns of people in an immediate problematic situation and to the goals of social science by joint collaboration within a mutually acceptable ethical framework".

The advantage of action research is the very close relationship between researchers and the phenomena of interest. This facilitates first-hand insight, limits researcher influence on subjects being studied and supports a prosperous way of applying theory to practice and evaluating its outcome. However, action research is very time consuming, and since the researcher takes part in the phenomena studied remaining objective can be difficult. Also, when participating in the intervention of an activity or phenomenon, considerations emerge concerning if it is ethically acceptable for a researcher, for example, to conceal knowledge of particular approaches having better effects than others. Finally, the outcome of this research can be difficult to generalize.

In relation to mobile HCI research, action research could be used for extending field or case studies by researchers participating actively in real world activities involving mobility, introducing different solutions or theories "on-the-fly" as well as evaluating their effects and/or validity.

2.4. Laboratory Experiments

In contrast to field studies, laboratory studies are characterized by taking place in a controlled environment created for the purpose of research. Thus laboratory experiments do not necessarily have to take place in dedicated "laboratories" as such but can be conducted in various controlled environments such as in an office (Tang et al 2001), in a hallway (Bohnenberger et al. 2002) or in a simulator (Kjeldskov and Skov 2003). Laboratory experiments facilitate various types of data being collected using different experimental methods depending on the style of subsequent analysis desired. While traditional quantitative measurements of factors such as error rate and task completion times collected through, for example, cognitive walkthrough methods are suitable for statistical methods of analysis, using more qualitative approaches such as heuristic evaluation or think-aloud protocols during the conduct of experimental tasks also produces results suitable for analysis.

The major advantages of laboratory studies are the opportunity to focus on specific phenomena of interest and a large degree of experimental control in terms of manipulation of variables before and during the experiment through for example assignment of test subjects and exposure to different treatment variables (Wynekoop and Conger 1990). Also, laboratory experiments are typically highly replicable and facilitate good data collection. Disadvantages include limited relation to the real world and an unknown level of generalizability of results outside laboratory settings.

In mobile HCI research, laboratory experiments are suitable for evaluating design ideas, specific products or theories about design and user interaction in controlled environments with little or no interference from the real world.

2.5. Survey research

Surveys usually inform research by providing information from a known sample of people gathered through various systematic techniques such as questionnaires and interviews. Using surveys, data is gathered directly from selected respondents and it is assumed that these are independent of their environment (Wynekoop and Conger 1990). Typically, data from questionnaire surveys is collected without researcher intervention and is analyzed quantitatively while data from interview surveys are analyzed qualitatively.

The major advantages of surveys are that they facilitate large amounts of data to be gathered with relatively little effort, supporting broad generalization of results. Also a high level of control regarding sample subjects makes reduction of bias possible thus increasing validity. However, surveys suffer from providing only snapshots of studied phenomena and rely highly on the subjective views of respondents.

In mobile HCI research, surveys could, for example, facilitate generalizable information being gathered about user needs and requirements for understanding a phenomenon, building theory or developing systems. Also, surveys could be used for gathering data about the user experience of specific products or designs for evaluation purposes.

2.6. Applied Research

According to Wynekoop and Conger (1990) applied research, builds on trial and error on the basis of the researchers capabilities of reasoning through *intuition, experience, deduction* and *induction.* Typically the desired goal or outcome of the research process is known in terms of requirements on some level of abstraction, but methods or techniques for accomplishing this outcome are unknown and thus sought through applying potentially relevant research. The advantages of applied research is that it is very goal directed and (typically) results in some kind of product being produced, which can be evaluated against the initial goals. The major disadvantages of applied research are that initial solutions may be very limited and not generalizable and that appropriate solutions for accomplishing the desired outcome may not be produced at all.

In mobile HCI research, applied research is relevant in relation to design and implementation of systems, interfaces and techniques, which meet certain requirements for performance, user interaction, user satisfaction etc.

2.7. Basic Research

Doing basic research, researchers develop new theories or study well-known problems to which neither specific solutions nor methods for accomplishing solutions are known (Wynekoop and Conger 1990). Like applied research, the approach of basic research is trial and error based relying on the competences of the researcher. The major advantage of basic research is the openness of the research facilitated both in terms of approaches and time, allowing a high level of creativity in the search for methods and solutions. However, basic research, like applied research, can be very time consuming and there is no guarantee of any solution eventually being produced.

In relation to mobile HCI, basic research may be applied to the development of theoretical frameworks for, for example, understanding basic issues of mobility or for identifying new problems and possible solutions related to human-computer interaction while being mobile.

2.8. Normative Writings

In order to include the significant body of so-called “non-research” writings about phenomena of interests in their classification of research methods, Wynekoop and Conger (1990) suggests the category of *normative writings,* covering concept development writings, presentation of “truth” and Benbasat’s (1987) category of “application descriptions”. While concept development writings organize ideas in order to stimulate and indicate directions for future research, such as the case of this paper, normative writings belonging to the “truth” category present ideas, concepts and suggestions, which seem intuitively correct but are not based on theory or research. Application descriptions are defined as “narratives written by practitioners” (Wynekoop and Conger 1990), describing subjective views on a situation and what worked for them in that particular situation. The primary advantage of normative writings is that they require little effort to produce compared to presenting theoretical concepts. Disadvantages include limited theoretical and methodological reflection and limited generalizability.

In mobile HCI, normative writings describing designs and processes that worked well or did not prove successful may be useful for inspiring future research or design.

3. RESEARCH PURPOSE

Research methods as discussed above and research *purpose* are typically closely related but not necessarily determined by one another. Like Wynekoop and Conger (1990) we thus use the second dimension of our matrix for classifying mobile HCI research to describe research purpose. Populating the categories this dimension we borrow the categories and definitions of research purposes originally proposed by Basili et al. (1986) and also used by Wynekoop and Conger (1990). These are briefly defined below.

Understanding is the purpose of research focusing on finding the meaning of studied phenomena through e.g. frameworks or theories developed from collected data.

Engineering is defined as the purpose of research focused towards developing new systems or parts of systems such as e.g. an interaction technique for mobile phones.

Re-engineering describes the purpose of research focusing on improving existing systems by redeveloping them such as e.g. adapting a web browser to a small display.

Evaluating is the purpose of research assessing or validating products, theories or methods e.g. the usability of a specific mobile device design or a theory of interaction.

Describing finally refers to research focusing on defining desirable properties of products e.g. a mobile system.

4. CLASSIFICATION OF MOBILE HCI RESEARCH

In this section we present a classification of selected mobile human-computer interaction research papers in relation to the research methods and purposes discussed above.

A total of 102 conference and journal papers were classified in relation to the described categories of research purpose and research methods applied. These papers constitute all publications related to mobile human-computer interaction between 2000 and 2002 in the following top-level conference proceeding series and journals:

- Conference on Computer-Human Interaction, CHI, ACM
- Conference on Advanced Visual Interfaces, AVI, ACM
- Conference on User Interface Software and Technology, UIST, ACM
- Conference on Computer-Supported Cooperative Work, CSCW, ACM
- Symposium on Mobile Human-Computer Interaction, Mobile HCI
- Symposium on Designing Interactive Systems, DIS, ACM
- Transactions on Computer-Human Interaction, TOCHI, ACM
- Journal of Personal and Ubiquitous Computing, Springer-Verlag

While other conferences and journals exist, presenting interesting research on mobile human-computer interaction, we found that the listed conferences and journals provided a solid and adequately representative base for this study given the number of publications on the topic and the general level of the reviewing processes for these conferences and journals.

The 102 papers specifically focusing on mobile human-computer interaction were identified by thoroughly reading through abstracts (and sometimes introductions) of all publications between 2000 and 2002 in the listed conference proceeding series and journals. A paper was selected for the study if it was in any way related to mobile devices and human-computer interaction. Thus a paper would be omitted if it focused only on mobile network protocol design or did not involve any aspect of mobility of users or systems. All papers were printed, numbered, read through and classified over a period of two weeks by the first author of this paper with particular focus on identifying the purpose of the presented work and the research methods applied in achieving this. The classification is shown in table 2 below.

To ensure consistency, the initial classification of the papers was evaluated by scanning through all the papers a second time on a single day. To ensure validity, the second author of this paper subsequently evaluated the classification by blindly classifying 20 randomly selected papers. As this resulted in a number of disparities, all 102 papers were discussed and classified one by one in collaboration between the two authors.

While in the review presented in Wynekoop and Conger (1990) each paper is only attributed to one research method and purpose, this was not possible with all of the papers on mobile human-computer interaction. Some of the reviewed papers clearly employed more than one research method and had multiple purposes. A common example of this would be papers presenting a system engineered by applying research and subsequently evaluated in a laboratory. Consequently, such papers were given multiple classifications and appear more than once in table 2 above. As a consequence of multiple research methods and purposes in the same paper, aggregate percentages will sometimes amount to more than 100%.

Table 2 shows that 55% of mobile HCI research falls within the applied category (56 of 102 papers). The secondly most used method is laboratory experiments being applied in 31% of the research (32 of 102 papers). 20% of the papers report from field studies and 8% report from survey studies while 7% are normative writings and 6% report from case studies. Only 3 papers were classified as basic research and no entries were found for action research. This distribution shows a clear bias towards environment independent and artificial setting research in the form of applied and laboratory based approaches at the expense of natural setting research focusing on real use and basic and action research generating theory and refining it in practice.

Looking at the research purpose, 51% of mobile HCI research is done for engineering with additional 10% done for re-engineering. Thus in total, 61% of the research reported involves building systems. 41% of the papers involve evaluation, of which 71% is done through laboratory experiments 19% through field experiments and the remaining 10% through surveys. Research for understanding mobile HCI accounts for 18% of the papers, of which 22% reports from the use of surveys and 22% from field studies. Describing different aspects of mobile HCI accounts for 10% of the research, of which 50% are in the form of normative writings. Thus within mobile HCI research there is a clear tendency towards building systems and if evaluating them, doing so in laboratory settings. Understanding and learning from the design and real use of systems is less prioritized, limiting the generation of a cumulative body of knowledge on mobile human-computer interaction.

Table 2. Classification of mobile human-computer interaction research. Numbers refer to indexes in the appendix of reviewed mobile HCI research papers bibliography

Research Method

Research purpose	Case studies	Field studies	Action research	Lab experiment	Survey research	Applied research	Basic research	Normative writings
Understand	10, 11, 51	67, 68, 69, 101		91	14, 25, 53, 72	43	1, 21, 32	16, 20
Engineer	24, 49	3, 65, 71, 85, 94				2, 4, 9, 12, 17, 18, 19, 23, 27, 28, 29, 33, 34, 36, 38, 39, 41, 45, 46, 48, 50, 52, 57, 58, 59, 60, 64, 70, 73, 74, 76, 77, 78, 79, 81, 83, 84, 86, 89, 90, 92, 93, 95, 97, 98		
Re-engineer	22	6, 9, 41, 63, 71, 81, 85, 89				7, 8, 31, 55, 63, 75, 80, 100, 102		
Evaluate				4, 5, 6, 7, 8, 9, 12, 13, 18, 22, 30, 40, 42, 44, 49, 50, 55, 56, 58, 60, 61, 66, 75, 77, 82, 90, 94, 97, 98, 99	26, 33, 64, 93			
Describe		47, 62, 96		47		35		15, 37, 54, 87, 88

Of the 56 papers applying research, 96% do so for the purpose of engineering or re-engineering. Thus in total, these two cells in the matrix account for more than 50% of the mobile HCI research classified. Of the 32 papers in the laboratory experiment category, 94% use this method for evaluation purposes. Of the 20 papers reporting from field studies, 40% use this method for evaluation purpose while 25% use it for engineering. 20% (4 papers) report from field studies for the purpose of understanding. Thus when building systems within mobile HCI research, there is a clear tendency to do so primarily by trial and error and a lesser tendency to do so based on actual user studies. Also, controlled environments are used primarily for product evaluation purposes. Field studies are only applied to inform the design of new systems to a limited extend.

Of the 45 papers reporting applied research with the purpose of engineering systems, only 37% (17 papers) *also* report evaluating the produced solutions. 61% of these evaluations are done through laboratory experiments, 22% through field studies and the remaining 17% by surveys. Of the 9 papers reporting the *re-engineering* of systems, 56% also report evaluations of these systems. 80% of these are through laboratory experiments and 20% through field studies. Thus when building new systems, there is a tendency towards *not* evaluating them while when subsequently *re-building* them, evaluation is more prevalent. When evaluating engineered or re-engineered systems, there is a large bias towards applying laboratory-based approaches.

5. DISCUSSION

Table 2 reveals a lack of focus on real use contexts in relation to engineering and evaluating mobile systems as well as limited construction and use of theory. While field studies *are* being done, natural setting research is not prevalent. One reason for this may be that applied research and laboratory experiments are simply easier to conduct and manage than field studies, case studies and action research. Another reason may be that mobile HCI has strong roots in computer science and human-computer interaction. These fields collectively have a strong bias towards engineering and evaluating systems, with input from fields such as ethnography only recently emerging.

Reflecting further on table 2, a number of features seem to characterize the field of mobile human-computer interaction. Firstly, given the prevalent applied approach to engineering it seems assumed that we already know what to build and which specific problems to overcome such as limited screen real estate, limited means for interaction, dynamic use-contexts and limited network bandwidth. As only a little research actually addresses the question of what is useful and what is perceived problematic from a user-perspective and as a qualitative review of the classified papers reveal that evaluations are often focused on functionality rather than contextual issues, it is difficult to set aside this assumption and identify and face more fundamental challenges to mobile human-computer interaction. Secondly, given the limited focus on real-world studies it seems that the real contexts are not actually important for the mobile system we build and use and that mobile computer systems are a generically applicable solution. The view that building and evaluating systems by trial and error is better than grounding engineering, evaluation and theory in user-based studies weakens research in mobile HCI. Thirdly, given the fact that only few studies are based on a methodological foundation, it seems assumed that methodology matters very little in mobile HCI research. This supposition is problematic as the choice of methods clearly has influence on the results subsequently produced (Myers 1997). From a cognitive psychology perspective, for example, problem solving by applied research is viewed as a rather poor method as it demands huge efforts by researchers that often "translate into poor performance because they require search of a large space of possibilities" (Wynekoop and Conger 1990).

The distribution of research methods and purposes shown in table 2 offers a number of opportunities in the area of mobile HCI. Firstly, the fact that field studies are mostly being used for the purpose of evaluation presents the opportunity to use this method to explore use context and user needs to promote understanding. Field studies could assist with the translation of needs into new designs and the re-engineering of existing designs. Mobility is very difficult to emulate in a laboratory setting, as is the dynamism of changing context. Field studies offer the ideal opportunity for the study of rich real-world use cases. Learning from other disciplines that have struggled with the study of similar "slippery" phenomena, such as ethnography in this regard could provide important insight. The lack of survey and case study research also presents an opportunity. Information Systems uses these approaches widely, with the former research method often being used to collect large amounts of data from, for example, actual end-users of a system. In addition this approach offers a good opportunity to study the use of systems in the hands of a large segment of the population, enabling wider reaching generalizations. Case studies within mobile HCI could increase learning from existing implemented systems within real-world contexts, for example mobile systems and infrastructure within organizations. Such case studies would enable the close scrutiny of pre-defined phenomena in fixed contexts, which could then be used to enrich the collective knowledge in the discipline and to enable key issues to be described and understood. The issues generated could then be used to generate hypotheses to propagate further research. The limited use of action research points to both the lack of a well-established body of theoretical research within the discipline and the unwillingness to implement mobile systems which are uncertain to succeed and take a long time to evaluate and implement. This is perhaps not surprising, given the current cost of such technology and the associated implementation overhead. Nonetheless, this is, again, an opportunity to develop knowledge in the discipline through practice and evaluation. Finally, the lack of basic research means that opportunities exist for the development of theoretical frameworks to promote description and understanding. In addition, the applicability of theories from other disciplines to mobile HCI can be examined through basic and action research.

6. LIMITATIONS

The presented review of research methods has a number of limitations. First of all, the categories of research methods can be criticized for being vague and overlapping. Thus for example, case studies are often done in the field but it is unclear how this method differs from field studies. If a case study were, on the other hand, conduced in a controlled environment, how would it be different from a laboratory experiment? Furthermore, it can be discussed whether the eight categories of methods belong to the same level of abstraction or if some categories could be subordinated others. Combined with the fact that many research papers provide only little information about method, it can be difficult to decide which category a specific paper belongs to. Thus the presented study relies on the researchers comprehension of the categories and ability to make a qualified judgment on the basis of sometimes scarce information. Also, it can, of course, be questioned if the selected papers are representative and to what extent activities within a given area are actually reflected through publications.

7. CONCLUSIONS

In this paper we have examined and reviewed research methods within the field of mobile HCI through classifying 102 research papers. We have identified a number of significant trends in research purpose and methods with a clear bias towards engineering systems using applied approaches and, if evaluating them, doing so in laboratory settings. In addition we have found that research methods examining phenomena in context such as case studies are not widely used. These findings present a number of opportunities for further research suggesting the need for a change of emphasis within mobile HCI.

REFERENCES

Basili, V.R., Selby, R.W. and Hutchins, D.H.: Experimentation in software engineering. IEEE Transactions on Software Engineering, SE-12 (1986) 733-743

Benbasat, I.: An analysis of research methodologies. In MacFarlan, F.W. (Ed.) The Information System Research Challenge. Boston, Harvard Business School Press (1985) 47-85

Benbasat, I., Goldstein, D.R. and Mead, M.: The case research strategy in studies of information systems. MIS Quarterly, Vol. 7(3) (1987) 369-386.

Bohnenberger, T., Jameson, A., Kruger, A. and Butz, A.: Location-Aware Shopping Assistance: Evaluation of a Decision-Theoretic Approach. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

Galliers, R.D.: Choosing Appropriate Information Systems Research Approaches: A Revised Taxonomy. In Proceedings of the IFIP TC8 WG 8.2 Working Conference on The Information Systems Research Arena of The 90's, Copenhagen, Denmark (1990)

Kjeldskov J. and Skov M. B. (2003) Evaluating the Usability of a Mobile Collaborative System: Exploring Two Different Laboratory Approaches. In Proceedings of the 4th International Symposium on Collaborative Technologies and Systems 2003. Orlando, Florida.

Lewis, I.M.: Social Anthropology in Perspective. Cambridge University Press (1985)

Myers, M.D.: Qualitative Research in Information Systems. MIS Quarterly, Vol. 21(2) (1997) 241-242

Rapoport, R.N.: Three Dilemmas in Action Research. Human Relations, Vol. 23(4) (1970) 499-513

Tang, J., Yankelovich, N., Begole, B., Van Kleek, M., Li, F. and Bhalodia, J.: ConNexus to Awarenex: Extending awareness to mobile users. In proceedings of CHI2001, Seattle, WA, USA, ACM (2001)

Wynekoop, J.L. and Conger, S.A.: A Review of Computer Aided Software Engineering Research Methods. In Proceedings of the IFIP TC8 WG 8.2 Working Conference on The Information Systems Research Arena of The 90's, Copenhagen, Denmark (1990)

Yin, R. K.: Case Study Research, Design and Methods, 2nd ed. Newbury Park, Sage Publications (1994)

APPENDIX: REVIEWED MOBILE HCI RESEARCH PAPERS, 2000-2002

1. Abowd, G.D. and Mynatt, E.D.: Charting Past, Present and Future Research in Ubiquitous Computing". ACM Transactions on Computer-Human Interaction, Vol. 7, No. 1 (2000) 29-58
2. Beigl, M.: MemoClip: A Location-Based Remembrance Appliance. Personal and Ubiquitous Computing, Vol. 4. Springer-Verlag, London (2000) 230-233
3. Bertelsen, O. and Nielsen, C.: Augmented Reality as a Design Tool for Mobile Devices. In proceedings of DIS 2000, ACM (2000)
4. Bohnenberger, T., Jameson, A., Kruger, A. and Butz, A.: Location-Aware Shopping Assistance: Evaluation of a Decision-Theoretic Approach. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
5. Brewster, S. and Murray, R.: Presenting Dynamic Information on Mobile Computers. Personal and Ubiquitous Computing, Vol. 4. Springer-Verlag, London (2000) 209-212
6. Brewster, S.: Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 188-205
7. Buyukkokten, O., Garcia-Molina, H., Paepcke, A. and Winograd, T.: Power Browser: Efficient Web Browsing for PDAs. In Proceedings of CHI2000, The Hague, Amsterdam, ACM (2000)

8. Buyukkokten, O., Garcia-Molina, H., Paepcke, A.: Accordion Summarization for End-Game Browsing on PDAs and Cellular Phones. In proceedings of CHI2001, Seattle, WA, USA, ACM (2001)
9. Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C.: Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. In Proceedings of CHI2000, The Hague, Amsterdam, ACM (2000)
10. Cheverst, K., Mitchell, K. and Davies, N.: Investigating Context-aware Information Push vs. Information Pull to Tourists. In Proceedings of Mobile HCI 2001, Lille, France (2001)
11. Cheverst, K., Davies, N., Mitchell, K. and Efstratiou, C.: Using Context as a Crystal Ball: rewards and Pitfalls. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 8-11
12. Chincholle, D., Goldstein, M., Nyberg, M. and Eriksson, M.: Lost or Found? A Usability Evaluation of a Mobile Navigation and Location-Based Service. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
13. Chittaro, L. and Cin, P.D.: Evaluating Interface Design Choices on WAP Phones: Single-choice List Selection and Navigation among Cards. In Proceedings of Mobile HCI 2001, Lille, France (2001)
14. Colbert, M.: A Diary Study of Rendezvousing: Group Size, Time, Pressure and Connectivity. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
15. Constas, I. and Papadopoulos, D.: Interface-Me: Pursuing Sociability Through Personal Devices. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 195-200
16. Coschurba, P., Baumann, J., Kubach, U. and Leonhardi, A.: Metaphors and Context-Aware Information Access. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 16-19
17. Danesh, A., Inkpen, K., Lau, F., Shu, K. and Booth, K.: Geney: Designing a Collaborative Activity for the Palm Handheld Computer. In proceedings of CHI2001, Seattle, WA, USA, ACM (2001)
18. De Bruijn, O,, Spence, R. and Chong, M.Y.: RSVP Browser: Web Browsing on Small Screen Devices. In Proceedings of Mobile HCI 2001, Lille, France (2001)
19. Dey, A.K.:Understanding and Using Context. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 4-7
20. Dix, A., Rodden, T., Davies, N., Trevor, J., Friday, A. and Palfreyman, K.: Exploiting Space and Location as a Design Framework for Interactive Mobile Systems. Transactions on Computer-Human Interaction, Vol. 7, No. 3. ACM (2000) 285-321
21. Dubois, E., Gray, P., and Nigay, L.: ASUR++: A Design Notation for Mobile Mixed Systems. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
22. Ebling, M.R., John, B.E. and Satyanarayanan, S.: The Importance of Transludence in Mobile Computing Systems. ACM Transactions on Computer-Human Interaction, Vol. 9, No. 1 (2002) 42-67
23. Fano, A.: What are a Location's File and Edit Menus? Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 12-15
24. Forlizzi, J. and McCormack, M.: Case Study: User Research to Inform the Design and Development of Integrated Wearable Computers and Web-Based Services. In proceedings of DIS 2000, ACM (2000)
25. Fortunati, L.: The Mobile Phone: An Identity on the move. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 85-98
26. Frohlich, D. and Murphy, R.: The Memory Box. Personal and Ubiquitous Computing, Vol. 4. Springer-Verlag, London (2000) 238-240
27. Geldof, S. and Terken, J.: Talking Wearables Exploit Context. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 62-65
28. Gelgon, M. and Tilhou, K. Automated Multimedia Diaries of Mobile Device Users Need Summarization. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
29. Goldstein, M., Oqvist, G., Bayat-M, M., Ljungstrand, P. and Bjork, S.: Enhancing the Reading Experience: Using Adaptive and Sonified RSVP for Reading on Small Displays. In Proceedings of Mobile HCI 2001, Lille, France (2001)

30. Goldstein, M., Alsio, G. and Werdenhoff, J.: The Media Equation Does Not Always Apply: People are not Polite Towards Small Computers. Personal and Ubiquitous Computing, Vol. 6, Springer-Verlag, London (2002) 87-96
31. Gonzalez-Castano, F.J., Anido-Rifon, L. and Costa-Montenegro, E.: A New Transcoding Technique for PDA Browsers, Based on Context Hierarchy. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
32. Green, N., Harper, R.H.R. and Cooper, G.: Configuring the Mobile User: Sociological and Industry Views. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 146-156
33. Hibino, S. and Mockus, A.: handiMessenger: Awareness-Enhanced Universal Communication for Mobile Users. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
34. Hinckley, K., Pierce, J., Sinclair, M. and Horvitz, E.: Sensing Techniques for Mobile Interaction. In proceedings of UIST2000, San Diego, CA, USA, ACM (2000)
35. Hinckley, K. and Horvitz, E.: Toward More Sensitive Mobile Phones. In proceedings of UIST2001, Orlando, Florida, USA, ACM (2001)
36. Holland, S. and Morse, D.R.: AudioGPS: spatial audio in a minimal attention interface. In Proceedings of Mobile HCI 2001, Lille, France (2001)
37. Holland, S., Morse, D. and Gedenryd, H.: Direct Combination: A New User Interaction Principle for Mobile and Ubiquitous HCI. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
38. Huang, E.M., Terry, M., Mynatt, E., Lyons, K. and Chen, A.: Distributing Event Information by Simulating Word-of-Mouth Exchanges. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
39. Huttenrauch, H. and Norman, M.: PocketCERO – mobile interfaces for service robots. In Proceedings of Mobile HCI 2001, Lille, France (2001)
40. Isokoski, P. and Raisamo, R.: Device Independent Text Input: A Rationale and an Example. In Proceedings of AVI2000, Palermo, Italy, ACM (2000)
41. Izadi, S., Fraser, M., Benford, S., Flintham, M,, Greenhalgh, C., Rodden, T. and Schnadelbach, H.: Citywide: supporting interactive digital experiences across physical space. In Proceedings of Mobile HCI 2001, Lille, France (2001)
42. James, C.L. and Reischel, K.: Text Input for Mobile Devices: Comparing Model Prediction to Actual Performance. In proceedings of CHI2001, Seattle, WA, USA, ACM (2001)
43. Jameson, A.: Modeling both the Context and the User. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 29-33
44. Jones, M., Buchanan, G. and Thimbleby, H.: Sorting Out Searching on Small Screen Devices. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)
45. Kehr, R. and Zeidler, A.: Look, Ma, My Homepage is Mobile! Personal and Ubiquitous Computing, Vol. 4. Springer-Verlag, London (2000) 217-220
46. Kohtake, N., Rekimoto, J. and Anzai, Y.: InfoPoint: A Device that Provides a Uniform User Interface to Allow Appliances to Work Together over a Network. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 264-274
47. Lacucci, G., Kuutti, K. and Ranta, M.: On the Move with a Magic Thing: Role Playing in Concept Design of Mobile Services and Devices. In proceedings of DIS 2000, ACM (2000)
48. Laerhoven, K.V. and Aidoo, K.: Teaching Context to Applications. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 46-49
49. Lamming, M., Eldridge, M., Flynn, M., Jones, C. and Pendlebury, D.: Satchel: Providing Access to Any Document, Any Time, Anywhere. Transactions on Computer-Human Interaction, Vol. 7, No. 3. ACM (2000) 322-352
50. Lehikoinen, J. and Salminen, I.: An Empirical and Theoretical Evaluation of BinScroll: A Rapid Selection Technique for Alphanumeric Lists. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 141-150

51. Licoppe, C. and Heurtin, J.P.: Managing One's Availability to Telephone Communication Through Mobile Phones: A French Case Study of the Development of Dynamics of Mobile Phone Use. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 99-108

52. Lin, J., Laddaga, R. and Naito, H.: Personal Location Agent for Communicating Entities (PLACE). In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

53. Ling, R.: We Release Them Little by Little: Maturation and Gender Identity as Seen in the Use of Mobile Telephony. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 123-136

54. Ljungstrand, P.: Context Awareness and Mobile Phones. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 58-61

55. MacKenzie, I.S., Kober, H., Smith, D., Jones, T. and Skepner, E.: LetterWise: Prefix-Based Disambiguation for Mobile Text Input. In proceedings of UIST2001, Orlando, Florida, USA, ACM (2001)

56. MacKenzie, S.: KSPC (Keystrokes per Character) as a Characteristic of Text Entry Techniques. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

57. Mantyjarvi, J. and Seppanen, T.: Adapting Applications in Mobile Terminals Using Fuzzy Context Information. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

58. Marsden, G., Thimbleby, H., Jones, M. and Gillary, P.: Data Structures in the Design of Interfaces. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 132-140

59. Matsushita, N., Ayatsuka, Y. and Rekimoto, J.: Dual Touch: A Two-Handed Interface for Pen-Based PDAs. In proceedings of UIST2000, San Diego, CA, USA, ACM (2000)

60. Mayol, W.W., Tordoff, B.J. and Murray, D.W.: Wearable Visual Robots. Personal and Ubiquitous Computing (2002), Vol. 6. Springer-Verlag, London (2002) 37-48

61. Mizobuchi, S., Mori, K., Ren, X. and Michiaki, Y.: An Empirical Study of the Minimum Required Size and the Minimum Number of Targets for Pen Input on the Small Display. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

62. Mikkonen, M., Vayrynen, S., Ikonen, V. and Heikkila, O.: User and Concept Studies as Tools in Developing Mobile Communication Services for the Elderly. Personal and Ubiquitous Computing (2002), Vol. 6. Springer-Verlag, London (2002) 113-124

63. Milewski, A. and Smith, T.M.: Providing Presence Cues to Telephone Users. In Proceedings of CSCW2000, Philadelphia, PA, USA, ACM (2000)

64. Nakanishi, Y., Tsuji, T., Ohyama, M. and Hakozaki, K.: Context Aware Messaging Service: A Dynamical Messaging Delivery using Location Information and Schedule Information. Personal and Ubiquitous Computing (2000), Vol. 4. Springer-Verlag, London (2000) 221-224

65. Nigay, L., Salembier, P., Marchand, T., Renevier, P. and Pasqualetti, L.: Mobile and Collaborative Augmented Reality: A Scenario Based Design Approach. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2000)

66. Oquist, G. and Goldstein, M.: Towards an Improved Readability on Mobile Devices: Evaluating Adaptive Rapid Serial Visual Presentation. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

67. Palen, L., Salzman, M. and Youngs, E.: Going Wireless: Behavior & Practice of New Mobile Phone Users. In Proceedings of CSCW2000, Philadelphia, PA, USA, ACM (2000)

68. Palen, L., Salzman, M. and Youngs, E.: Discovery and integration of Mobile Communications in Everyday Life. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 109-122

69. Palen, L. and Salzman, M.: Beyond the Handset: Designing for Wireless Communication Devices. Transactions on Computer-Human Interaction, Vol. 9, No. 2. ACM (2002) 125-151

70. Partridge, K., Chatterjee, S. and Want, R.: TiltType: Accelerometer-Supported Text Entry for Very Small Devices. In proceedings of UIST2002, Paris, France, ACM (2002)

71. Pascoe, J., Ryan, N. and Morse, D.: Using While Moving: HCI Issues in Fieldwork Environments. Transactions on Computer-Human Interaction, Vol. 7, No. 3. ACM (2000) 417-437

72. Perry, M., O'Hara, K., Sellen, A., Brown, B. and Harper, R.: Dealing with Mobility: Understanding Access Anytime, Anywhere. Transactions on Computer-Human Interaction, Vol. 8, No. 4. ACM (2001) 323-347

73. Petrelli, D., Not, E., Zancanaro, M., Strapparava, C. and Stock, O.: Modeling and Adapting to Context. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 20-24

74. Pham, T., Schneider, G., Goose, S. and Pizano, A.: Composite Device Computing Environment: A Framework for Situated Interaction Using Small Screen Devices. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 25-28

75. Pirhonen, A., Brewster, S. and Holguin, C.: Gestural and Audio Methaphors as a Means of Control for Mobile Devices. In Proceedings of CHI2002, Minneapolis, Minnesota, USA, ACM (2002)

76. Pospischil, G., Umlauft, M. and Michlmayr, E.: Designing LoL@, a Mobile Tourist Guide for UMTS. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

77. Poupyrev, I., Maruyama, S. and Rekimoto, J.: Ambient Touch: Designing Tactile Interfaces for Handheld Devices. In proceedings of UIST2002, Paris, France, ACM (2002)

78. Raghunath, M.T., Narayanaswami, C.: User Interfaces for Application s on a Wrist Watch. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 17-30

79. Randell, C. and Muller, H.: The Shopping Jacket: Wearable Computing for the Consumer. Personal and Ubiquitous Computing, Vol. 4. Springer-Verlag, London (2000) 241-244

80. Randell, C. and Muller, H.L.: The Well Mannered Wearable Computer. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 31-36

81. Rantanen, J., Impio, J., Karinsalo, T., Reho, A., Tasanen, M. and Vanhala, J.: Smart Clothing Prototype for the Artic Environment. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 3-16

82. Ren, X., Moriya, S.: Improved Selection Performance on Pen-Based Systems: A Study of Pen-Based Interaction for Selection tasks. Transactions on Computer-Human Interaction, Vol. 7, No. 3. ACM (2000) 384-416

83. Rist, T., Brandmeier, P., Herzog, G. and Andre, E.: Getting the Mobile Users in: Three Systems that Support Collaboration in an Environment with Heterogeneous Communication Devices. In Proceedings of AVI2000, Palermo, Italy, ACM (2000)

84. Rist, T., and Brandmeier, P.: Customizing Graphics for Tiny Displays of Mobile Devices. In Proceedings of Mobile HCI 2001, Lille, France (2001)

85. Ross, D.A. and Blasch, B.B.: Development of a wearable Computer Orientation System. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 49-63

86. Roth, J. and Unger, C.: Using Handheld Devices in Synchronous Collaborative Scenarios. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 243-252

87. Roth, J.: Patterns of Mobile Interaction. In Proceedings of Mobile HCI 2001, Lille, France (2001)

88. Ruuska-Kalliokulja, S., Schneider-Hufschmidt, M., Vaananen-Vainio-Mattila, K. and Von Niman, B.: Shaping the future of Mobile Devices – Results of the CHI2000 Workshop on Future Mobile Device User Interfaces. In Proceedings of Mobile HCI 2001, Lille, France (2001)

89. Sawhney, N. and Schmandt, C.: Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments". Transactions on Computer-Human Interaction, Vol. 7, No. 3. ACM (2000) 353-383

90. Sazawal, V., Want, R. and Boriello, G.: The Unigesture Approach: One-Handed Text Entry for Small Devices. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, Berlin (2002)

91. Schenkman, B.N.: Perceived Similarities and Preferences for Consumer Electronics Products. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 125-131

92. Schmidt, A., Takaluoma, A. and Mantyjarvi, K.: Context-Aware Telephony Over WAP. Personal and Ubiquitous Computing, Vol. 4. Springer-Verlag, London (2000) 225-229

93. Schmidt, A., Stuhr, T. and Gellersen, H.: Context-Phonebook – Extending Mobile Phone Applications with Context. In Proceedings of Mobile HCI 2001, Lille, France (2001)

94. Sharples, M., Corlett, D. and Westmancott, O.: The Design and Implementation of a Mobile Learning Resource. Personal and Ubiquitous Computing (2002) Vol. 6. Springer-Verlag, London (2002) 220-234

95. Silfverberg, M., Mackanzie, I.S. and Korhonen, P.: Predicting Text Entry Speed on Mobile Phones. In Proceedings of CHI2000, The Hague, Amsterdam, ACM (2000)

96. Strom, G.: Mobile Devices as Props in Daily Role Playing. In Proceedings of Mobile HCI 2001, Lille, France (2001)

97. Swindells, C., Inkpen, K.M., Dill, J.C. and Tory, M.: That one There! Pointing to establish device identity. In proceedings of UIST2002, Paris, France, ACM (2002)

98. Tang, J., Yankelovich, N., Begole, B., Van Kleek, M., Li, F. and Bhalodia, J.: ConNexus to Awarenex: Extending awareness to mobile users. In proceedings of CHI2001, Seattle, WA, USA, ACM (2001)

99. Thomas, B., Grimmer, K., Zucco, J. and Milanese, S.: Where Does the Mouse Go? An Investigation into the Placement of a Body-Attached TouchPad Mouse for Wearable Computers. Personal and Ubiquitous Computing, Vol. 6. Springer-Verlag, London (2002) 97-112

100. Trevor, J., Hilbert, D.M., Schilit, B.N. and Koh, T.K.: From Desktop to Phonetop: A UI For Web Interaction On Very Small Devices. In proceedings of UIST2001, Orlando, Florida, USA, ACM (2001)

101. Weilenmann, A.: Negotiating Use: Making Sense of Mobile Technology. Personal and Ubiquitous Computing, Vol. 5. Springer-Verlag, London (2001) 137-145

102. Wobbrock, J.O., Forlizzi, J., Hudson S.E. and Myers, B.A.: WebThumb: Interaction Techniques for Small-Screen Browsers. In proceedings of UIST2002, Paris, France, ACM (2002)

Part I

Studying and analysing

STUDYING AND ANALYSING

Part I addresses the question *how can we study, analyse and understand aspects of context relevant for mobile interaction design?* In order to do contextual interaction design we need to understand what context is, for a particular design task. What dimensions of context are important, what role they play, how they interact and influence each other, etc. Four of my own contributions to this are included in chapters 3-6. These chapters each address and explore a particular dimension of context for mobile interaction design: physical, social, personal and work, by specifically investigating four contextual entities: places and surroundings, people and interactions, families and relationships, and tasks and coordination.

Physical context

Chapter 3 investigates physical context. The chapter presents a field study and architectural analysis of a concrete urban environment, Federation Square in Melbourne, Australia. The study was conducted by combining empirical and analytical methods from architecture and city planning to survey, model and represent prominent architectural and informational aspects of a built environment such as districts, buildings, structures and signage. The result of the study is a descriptive framework summarizing the architectural and informational properties of a built environment. The use of such understanding in mobile interaction design is illustrated through a location-based mobile guide.

Social context

Chapter 4 investigates social context. The chapter presents a field study and sociological analysis of small groups of friends socializing “out on the town” in the same urban area as investigated in chapter 3. The study combined rapid ethnography, contextual interviews, and grounded analysis, and was guided conceptually by a typology for situated interactions. This resulted in a conceptual framework encapsulating and describing the interplay between people, activity and places during situated social interactions in an urban environment structured around the three key concepts of knowledge, situation and motivation. The use of such understanding in the design of mobile interactions is illustrated through a context-aware mobile guide representing and adapting to social context.

Personal context

Chapter 5 investigates personal context. The chapter distils eight dimensions of intimacy from the literature, and presents a longitudinal field study of technology-mediated intimacy and strong-tie relationships with six families in Melbourne, Australia. In order to minimize researcher intervention in peoples’ private lives, the study deployed the “auto-ethnographic” approach of cultural probes in combination with contextual interviews. The outcome of the study is a thematic understanding of what constitutes intimacy and how interactive systems are appropriated and used within intimate relationships. The use of such insight in mobile interaction design is illustrated through three early design ideas for technologies supporting intimacy.

Work context
Chapter 6 investigates work context. The chapter presents an ethnographic field study of communication on board large container vessels while manoeuvring inside harbour basins, and a simulator study of prototype technology in use. The outcome of the study is a detailed understanding of communication and coordination in the work context, based on a model of conversation for action. From this understanding it is possible to model communication flow in distributed work and use this to facilitate persistency in communication. This is demonstrated through the design of a mobile text-based communication system.

Chapter 3

Physical context

Jeni Paay and Jesper Kjeldskov

Abstract. The research presented in this paper aims to inform interface design for mobile guides by understanding and modelling the built environments in which the guide will be used. This is important because research into the use of mobile guides has shown that people have a strong ability to make sense of the physical space in which they are situated and make use of this when using mobile guides. Based on a field study and architectural analysis of the recently built Federation Square in Melbourne, Australia, we present a descriptive framework, MIRANDA, which provides a summarized abstraction of the fundamental architectural and informational features of a built environment. The use of this descriptive framework in HCI design for mobile guides is exemplified through the design of a mobile guide system for Federation Square that was informed by the identified architectural characteristics. On the basis of the field study and example design, we argue that mobile guides interface design can benefit from making use of 'knowledge-in-the-world' by streamlining and indexing information and functionality to physical information cues implicit in the built environment surrounding the user.

1. INTRODUCTION

Mobile guides are increasingly becoming a part of the way we operate in the physical world, and studies suggest that one of the major applications of future mobile information technology will be to digitally enhance user activities by giving them access to contextually adopted information through such guides. Hence the design of mobile guides has received considerable attention over the last decade within the field of HCI (see e.g. Abowd et al. 1996, Cheverst et al. 2000, Cheverst et al. 2002, Pospischil et al. 2002). The HCI research literature indicates a broad spectrum of different application areas for mobile guides. Mobile city guides provide the user with maps and other relevant information such as the location of restaurants, public offices, tourist sites etc. adapted to the user's location (e.g. Cheverst et al. 2000, Cheverst et al. 2002, Pospischil et al.

Originally published as Paay, J. and Kjeldskov, J. (2005) Understanding and Modelling the Built Environment for Mobile Guide Interface Design. *Behaviour and Information Technology,* 24(1), 21-35.

2002, Iacucci et al. 2003, Schmidt-Belz and Poslad 2003, Laakso et al. 2003). Mobile guides for specific tourist-sites provide users with the opportunity to follow guided tours or access additional information about the site being visited (e.g. Hermann and Heidmann 2002, Bornträger and Cheverst 2003, Bornträger et al. 2003) and mobile museum guides provide additional information about the specific items on display (e.g. Aoki and Woodruff 2000, Opperman and Specht 2000, Rocchi et al. 2003). Other mobile guide systems support the user's social or personal life, such as mobile personal guides and event planners keeping track of friends and upcoming social events (e.g. Fithian et al. 2003, Kolari and Virtanen 2003, Paulos and Goodman 2004) or mobile shopping assistants keeping track of shopping lists and informing the users about special offers within their vicinity (e.g. Randell and Muller 2000, Bohnenberger et al. 2002). Assisting people commuting and travelling, mobile navigation guides provide route planning information and directions for way-finding (e.g. Holland and Morse 2001, Chincholle et al. 2002, Kulju and Kaasinen 2002) and mobile travel planners keep track of users' itineraries, upcoming meetings etc. (e.g. Ricci et al. 2002, Kjeldskov et al. 2003).

Many of these mobile guide systems involve the user being situated in the built environment of a public space. Yet only a few references have investigated the challenges imposed and opportunities offered by the use of mobile guide systems in the context of buildings and other architectural structures in urban spaces. Exceptions include Vainio et al. (2002) who study navigation and way finding supported by a 3D model-based mobile city guide with clear architectural features such as landmarks, Kulju and Kaasinen (2002) who compare the use of photographs and semi-realistic 3D models augmented with textual information on mobile guides for supporting the user's experience of a physical space, and Laakso et al. (2003) who study the use of realistic 3D maps with topological as well as architectural features as a planning and navigational aid for leisure boat tourists at sea and in the cities they visit. These studies show that people often have a strong ability to make sense of the physical space in which they are situated and typically make extensive use of this when using mobile guides in a built environment. Consequently, it is argued that the use of mobile guides for navigation, way finding, etc. in environments with distinct architectural or other physical features (such as landmarks, noticeable physical structures or special topology) can benefit from taking these features into account by, for example, simply including visual representations of them using street perspective or bird's-eye views (Vainio et al. 2002, Laakso et al. 2003, Nakanishi et al. 2004). However, in order to exploit the user's ability to make sense of architectural features in their physical surroundings in HCI design for mobile guides, we need to better understand the role of the user's physical environment in defining their context and the contribution of existing information embedded into that environment to people's experience of it (Agre 2001, Bradley and Dunlop 2002, Tamminen et al. 2003). Also, we need to learn how to make a clear connection between the user's physical surroundings and the information presented on their mobile guide (Dix et al. 2000, Persson et al. 2002). So far, systematic methods for gathering, analyzing and understanding the properties of a built environment that help define the user's physical context, and appropriate analytical abstractions suitable for informing interaction design on the basis of this, have not yet been developed.

1.1. Indexicality: relating interfaces to their context

An interesting approach to making a clearer relation between mobile device interfaces and the user's context is to apply the idea of indexicality. Indexicality is a concept drawn from semiotics describing the relation between information representations and the context in which an interpreter (user) perceives them. Semiotics operates with three types of representations: symbolic (conventional), iconic (similarity) and indexical (material/causal). Symbols and icons are ways of representing information independent of context like e.g. text and graphical illustrations. Indexes, on the other hand, are ways of representing information with a strong relation to, for example, their spatial and/or temporal context exploiting information present in the interpreter's surroundings. Thus, indexical representations are highly context-specific and only make sense in particular situations; at a specific time, in a specific location, in relation to a specific activity, to specific people, etc. As an example, signposts and information boards are typically highly indexical in the sense that their meaning is tied to a specific location. By locating information in time and space, symbolic and iconic representations can be converted into temporal and spatial indexical representations (Andersen 2001). As shown by Kjeldskov (2002) increasing the level indexicality typically results in a significant reduction of required symbolic and iconic representations.

The concept of indexicality has previously been applied to the design of mobile guide interfaces (Kjeldskov 2002, Graham and Kjeldskov 2003, Kjeldskov et al. 2003) in order to streamline the information and functionality delivered to the user. The idea of applying indexicality to interface design for mobile guides is that if information and functionality on a mobile guide can be indexed to the user's context, then information already provided by the context becomes implicit and does not need to be displayed by the system. Hence, the user's environment becomes a vital part of the interface. This way of making use of 'knowledge-in-the-world' (Norman 1990) allows the limited screen real estate of mobile guides to be optimized to contain only the most vital content and the required user interaction with the mobile guide to be reduced. As a simplified example of this, an indexical mobile guide for patrons entering a cinema complex could be made temporally and spatially indexical by taking into account the time and location of the user, providing only information about the upcoming movies playing within a limited frame of time (temporal indexicality) in that specific cinema (spatial indexicality) (Kjeldskov 2002).

In the studies mentioned above, the mobile guide interface designs only explored temporal and spatial indexicality. However, the idea of indexical interfaces could also be broadened to exploit other aspects of the user's context. In addition to time and location, for example, factors such as existing information in our surroundings, the features of the built environment itself and the use of that environment by other people also play an important role in the context surrounding a mobile user (Paay 2003), and may also be indexed to in a mobile guide.

Relying heavily on the user's knowledge about, for instance, when, where and why they are situated, successful design of indexical interfaces for mobile guides, require that HCI designers have a strong understanding of the aspects of the user's context that they are indexing to. Hence, in order to include indexes to the built environment in the design of a mobile guide, the features of the built environment contributing to the user-experience of the physical space in which it will be used, needs to be analyzed and modelled in a way that extracts the essence of the place and provides an overview. Information embedded into the built environment and carrying parts of its meaning through its architectural features needs to be identified and proposals for how this can be applied to mobile guide design needs to be developed and evaluated.

This paper contributes to this discussion by presenting 1) a possible approach to enquiring into the features of a built environment informed by theories and methods derived from architecture, 2) a method for creating a descriptive abstraction on the basis of the field data, and 3) an example of how this abstraction can be used to inform the design of a mobile guide prototype. In section 2, we present and discuss the architectural theories and methods used in the study. In section 3, we report from a field study into the architectural and informational properties of a specific built environment, Federation Square in Melbourne, Australia. The idea and early results from this study are described by Paay (2002). The outcome of the field study is presented in the form of a descriptive framework for modelling and understanding the built environment called MIRANDA, which captures positive and negative architectural and informational properties of a physical space. In section 4, the use of this descriptive framework in HCI design for mobile guides is exemplified through the design of a mobile guide system for Federation Square, in which the interface is indexed to the built environment by exploiting information already implicitly present in the user's physical surroundings. Finally, section 5 concludes on our work and outlines our present and future directions of research.

2. ANALYZING THE BUILT ENVIRONMENT

Architectural design has a history of incorporation of social theories and user needs into design methods. Of special interest, Urban Planner, Kevin Lynch (1960) and Architect, Christopher Alexander (Alexander et al. 1977) both modelled built environments, specifically cities, with regard to the people that inhabit those places, hence implicitly including the users in their analysis of physical space.

Lynch (1960) developed a method for visual analysis of city precincts through descriptions of key aspects of the space held by people as they navigate and orient themselves within city precincts. This was done by diagramming the interplay of visible elements in the environment that contribute to a person's environmental image of a place. His purpose was to develop ideas and methods, rather than to prove facts in a final and determinate way. The method proved successful at assisting in the analysis of types of elements of a city, how they are put together, and what makes for strong identity. It also proved to be a useful technique for predicting the probable public image of that city. Lynch documented this technique for describing and understanding key aspects of the environmental image of space held by people as they navigate and orient themselves within city precincts and noted that it would be interesting to apply this method to an environment of a different scale. To understand the role of environmental images of cities in the lives of those who inhabit them, Lynch carried out two basic analyses. Firstly, interviews were conducted with people who either lived or worked in that area. The city dwellers image of the city was elicited from them using interview techniques which included asking participants to describe features of the city from memory, asking them to draw sketches of the city, and having them make imaginary trips to destinations within the city precinct. Secondly, a field reconnaissance was done by an architecturally trained observer, who mapped the presence of various elements of the physical environment, using categories that had proven significant in analysis of earlier pilot interviews: districts, landmarks, nodes, edges and paths. The observer made subjective judgments based on the immediate appearance of these elements in the field and about their visible contribution to the image of the city, which resulted in maps representing the visual form of a city in respect to these five elements. These maps were then compared to those maps derived from consensus information collected in the verbal interviews and the sketch maps made by the inhabitants to draw conclusions about 'imageability' of built environments.

Alexander et al. (1977) empirically investigated the interplay between architectural space and its inhabitants and identified architectural design problems in context and their impact on inhabitants of that environment. Drawn from observations of historical solutions to common design problems he created a method of analyzing aspects of the built environment and generated a 'checklist' of plausible solutions for design. This checklist was constituted by a collection of 253 hierarchically ordered patterns making up a 'Pattern Language' (Alexander et al. 1977) that begins with patterns defining towns and communities, through the design of individual buildings, down to the detailed design of building elements. The pattern language provides a framework for solutions to these common design problems. Each pattern describes the field of physical and social relationships required to solve a design problem in its stated context. These patterns in themselves are a form of composite pictures including: photographs, sketches, descriptive explanations detailing the context for the pattern, its relationship to parent patterns, a description of the problem, the empirical background of the pattern, evidence for its validity, and the design solution.

The methods of Lynch (1960) and Alexander et al. (1977) have not only proven their value within architecture, but have also demonstrated usefulness in HCI research in analyzing the interplay between architectural, informational and social space with the purpose of designing computerized information systems.

The work of Lynch (1960) is often referenced in relation to the design of computer systems supporting orientation and navigation in complex virtual information spaces, virtual environments and in the real world, and is claimed to be one of the most influential and useful pieces of research on spatial orientation and navigation for interaction designers (see e.g. Sparacino et al. 2000). Information City (Dieberger and Frank, 1998) is an ontology of spaces and connections based on Lynch's five major elements developed to support navigation in large information spaces where a spatial metaphor supports a user's sense of orientation by the use of recognizable landmarks. Similarly, Sparacino et al. (2000) present a 3D web browser using an urban-like information landscape following the guidelines proposed by Lynch (1960) to fetch and represent information from the web. Supporting navigation in virtual reality, Ingram and Benford (1996) apply urban planning principles to the design of virtual environments using Lynch's work to improve the navigation of a virtual environment including districts, landmarks, paths, nodes and edges. Similarly, Vinson (1999) proposes a series of guidelines for how designers can create and place landmarks in virtual environments for supporting navigation and orientation. Informing the design of a mobile guide system, Kulju and Kaasinen (2002) draw on the work of Lynch (1960) to include landmarks in a 3D city model on a mobile device and augment them with additional textual information to optimize spatial orientation.

The work of Alexander (Alexander et al. 1977) has been successfully applied within various areas of computing; from object-oriented software design (e.g. Gamma 1996) to human-computer interaction (e.g. Tidwell 1999, Borchers 2001). Inspired by, among others, Alexander et al. (1977), Harrison and Dourish (1996) define a distinction between space and place, and illustrate the influence of peoples sense of place on their interactive behavior. Designing information technologies for the domestic environment, Crabtree and Hemmings (2001) use an adapted pattern framework to inform requirements analysis and describe day to day use of technologies within that environment. Informing novel design of pervasive and mobile device services, Paulos and Goodman (2004) enquire into the application of new technologies by city inhabitants, and draw an analogy between Alexander's patterns and the movement and activities of people in urban settings. In line with this research, McCullough (2001) points out, that as interest in information technologies for physical contexts increases, what we already know from the work of people such as Alexander about the built environment and how people operate in it, can inform our understanding of peoples situated interactions, and guide the design of technologies that consider the user's physical context.

3. FIELD STUDY: FEDERATION SQUARE

Inspired by Lynch (1960) and Alexander et al. (1977), we conducted a field study of the architecturally designed built environment at Federation Square, Melbourne, Australia.

Figure 1. Federation Square.

Federation Square (figure 1) is a new civic structure, opened to the public in October 2002, to provide the people of Melbourne with a 'unifying square, a landmark, a civic focus' (official brochure), by bringing together a creative mix of attractions and public spaces. The design intention for the space was to incorporate digital technologies into the building fabric creating a meeting of virtual information space and physical building space for people to experience. We chose Federation Square for this study because it is a multi-modal public space with a mixture of distinct architectural features and embedded digital elements that provide a variety of activities to visitors.

The aim of the field study was to inquire into how informational and architectural elements of the built environment contribute to the visitor's experience of a public place and how this could subsequently be modelled. Thus, we wanted to identify important properties of the built environment as an inhabited public space and create an analytical abstraction, which could inform the design of a mobile guide supporting visitors to this place. The result of the field study is a descriptive framework called MIRANDA (Multilayer Info Related to Arch aNalysis Data Abstraction), representing the human experience of the informational (analogue and digital signage) and architectural properties of a physical space. The result of the subsequent prototype design is a proof of concept for a mobile guide, which takes these properties into consideration.

- The study consisted of the following sequential activities:
- Field inspection of Federation Square
- Data coding (based on Lynch and Alexander)
- Data analysis
- Data synthesizing: developing MIRANDA
- Development of mobile guide prototype design

The field inspection took approximately three hours, data coding took eight hours, and data analysis took approximately three days. Synthesizing the data took eight hours and developing the preliminary prototype design took approximately two days. Although these estimates can serve as guidelines for similar analysis by other mobile guide designers, the exact time spent on these activities would depend on the size of the physical site being analyzed. The details of these activities are described below.

3.1. Inspecting Federation Square

An initial visit was made to Federation Square by the first author, and resulted in the development of a scenario of the experience of a first-time visitor to the square. Observational expert audits were then made of Federation Square for both architectural and information space, in two separate field visits. These expert audits were inspired by the method outlined by Lynch (1960) in his visual analysis of city precincts for the purpose of identifying types of elements of a city. In his method, an architecturally trained observer maps the presence of various elements of the physical environment. In this adapted method the trained observer records, through photographs and field notes, the elements of the physical environment for later classification using significant categories. These visits, taking about three hours each to complete, resulted in a collection of 250 digital photographs of physical elements of the built environment. The location of each photograph was recorded on a map of Federation Square, and corresponding observations of the relationship between the elements being photographed and the environment were recorded in the form of field notes.

3.2. Coding the data

The photographs were stored electronically in a format where they could be directly associated with their corresponding observational field notes and sketches, and could be annotated further during the process of coding and analysis. For the information elements, each sign and media screen was sketched showing its relationship to building fabric, viewing direction and distance, and then annotated with general field notes explaining the sketches and detailing auditor observations of human activity in relationship to the signage. For the architectural elements, field notes recorded observations made by the trained auditor including general descriptions of elements and human responses to physical spaces within the architecturally designed environment. The coding of this data was necessarily a two-phase process, given that the original source data was a mixture of images and descriptive text. The first pass through the data used the encoding schemas to convert classified elements of the images into corresponding text. The second pass combined the classification-generated image descriptions with existing observational field data and highlighted those aspects of the resulting prose that contributed to the encoded abstraction of that data.

In coding the information elements, the graphic communication theories and concepts of Bowman (1968), Tufte (1990) and Bertin (2003) were used to derive the following set of classifications for the information elements: type, direction, distance, visibility, readability, and location. These categories were used as an encoding schema, to group and order descriptions of sketches and field notes in the electronic file. The ordered descriptions were then read through in their entirety several times, and repeating phrases and key concepts associated with the encoding schema were highlighted in the descriptions associated with the images. Bowman (1968), Bertin (1983) and Tufte (1990) were used to develop the categories for the analysis of information space because their graphic communication theories are based on human interpretation and human understanding of graphic elements in signage, and therefore necessarily incorporate human experience of information into the analysis.

The architectural elements were classified initially using Lynch's (1960) categories of: district, landmark, node, path, and edge, as an encoding schema. Using Lynch's detailed descriptions of what constitutes each of these categories, the key element of each image was classified in respect to the schema, by the architecturally trained auditor using a visual inspection technique. The photographs were annotated with a legend indicating their association with one or many of the five categories. An additional classification of this data was then conducted on the architectural elements using the 253 patterns of Alexander's pattern language as an encoding schema. Sketches and notes showing the applicability of each pattern were appended to the existing field notes. Each image was associated with one or many Alexandrian patterns, and the photograph annotated with the pattern number and pattern title of each associated pattern. The text descriptions associated with each architectural photograph were then read through several times to identify key concepts and repeating phrases in the loosely structured data that related to the encoding schema. Lynch and Alexander were used to derive the encoding schemas for this analysis because the categories that they provide were the outcomes of empirical studies of the interplay between architectural space and people, and this study aims to understand and model the human experience of a place.

3.3. Analyzing the data

Using a form of content analysis, adopted from rapid ethnographic method (Millen 2000), the highlighted concepts and themes were extracted from the descriptive prose, and analysed to see if some kind of 'vocabulary' of the space would emerge.

The architectural data was analysed first. The Lynchian categories derived from the image data were directly mapped and overlaid onto an existing two-dimensional map of Federation Square to produce a colour-coded abstraction indicating the location of the environmental categories. From this diagram it was clear that four key districts and four key landmarks could be identified in the complex. District 1 is a transit zone, a connector to the city, and has a focus on Landmark 1, the information centre. District 2 is the main plaza, an earthy uneven textured sloping open space with outdoors activities, which focuses on Landmark 2, the stage and large media screen. District 3 is the atrium precinct, sheltered, noisy, and constructed of machine-made materials, its focus is Landmark 3, the entrance, a large opening to this area. District 4 is the river precinct, which has the feeling of being at the back, flat, damp and lower than the rest of the spaces, Landmark 4, the river, is the focus of this district (see figure 2).

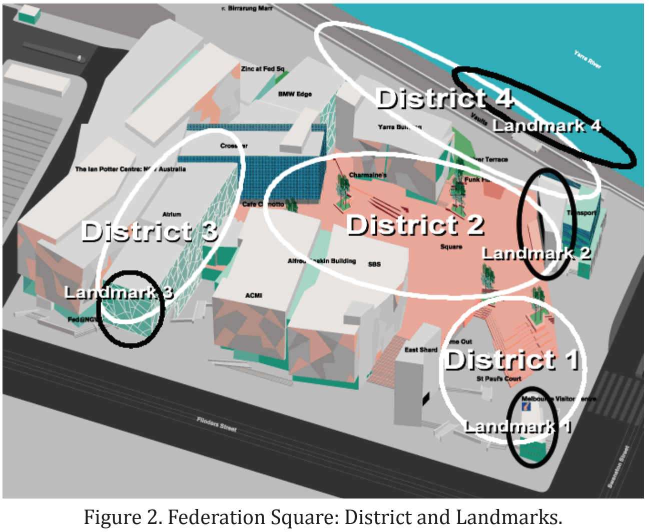


Figure 2. Federation Square: District and Landmarks.

The next stage was to use affinity diagramming adapted from the contextual design methodology (Beyer and Holtzblatt 1998) to group and refine the repeating phrases and key concepts that were identified in the data through the encoding process. Each phrase and word was written on a post-it note, and these notes were stuck on a large white board. After several iterations of grouping, regrouping, forming sets of words and refining words to a concise set of representative terms, the following themes emerged: words that described the space in terms of how it felt to people (affect), words that gave rich descriptions of architectural features, words that described the location of elements in relation to each other, and words that described human activity in the space. The emerging words were, by virtue of the encoding schema, influenced by the categories of Lynch (1960) and Alexander et al. (1977).

The same process of affinity diagramming was used with the informational data, and the following themes emerged: words that related to the readability of signs, words that related to the visibility of signs, words that related to the physical location of the sign, words that described how people were using the signs, and words that described the quality and usefulness of information given (including the understandability and effectiveness of the information). The emerging words were, by virtue of the encoding schema, influenced by the classifications and descriptions of graphic communication used by Bowman (1968), Bertin (1983) and Tufte (1990).

3.4. Synthesizing the data: MIRANDA

The process of grouping, sorting, dissecting and refining word sets using the affinity diagramming method had resulted in an emergent 'vocabulary', for both the architectural and the informational data, which described the physical environment using a concise set of words and clarified and identified the essence of the characteristics of the space. The process had been iterated until the syntax of a 'language' and a stable and concise set of words, representing the semantics of that language, evolved. This language was necessarily refined to a point where it could be used to completely replace the text descriptions associated with each of the images, and hence be used to represent a concise and abstracted description of that space, providing a minimalistic representation of human understanding of that space.

The syntax of the language, MIRANDA, became a signed word-pair:

[+ , / , -]<descriptor>.<place>

Explaining this syntax, the first set indicates that only one of these signs, '+', '/', or '-', applies to the phrase, where '+' indicates the positive form of the phrase, '/' indicates the in-part form of the phrase and '-' indicates the negative form of the phrase. The sign is then followed by a word-pair, the 'descriptor' and 'place' words, each chosen from a finite set of describing words, and a finite set of place words, which were defined during the context analysis phase of the space. Together, they give the adjective and noun for a 'language' statement that can represent the human understanding of the environmental element being described. For example, '+ inviting.path' indicates that an inviting path is pictured, whereas '- inviting.path' indicates that the path is uninviting to people. The two sets of words that represent the architectural characteristics of Federation Square, in respect to the expert audit can be seen in figure 3 and 4, where descriptor words are on the left-hand side, and place words on the right-hand side.

The same synthesis process was completed for the information elements. The syntax of the information language that was emerged was:

[+ , / , -]<descriptor>.<attribute>

To confirm the richness of MIRANDA, one or many language statements were then used to replace the detailed prose description associated with each photograph to ensure that the key element of every image could be comprehensively described using the language. For example, the description associated with figure 1 was replaced by the language statements: '- clear.path', '+ activity.floor', and '- activity.middle'. These collected statements can be interpreted to indicate that the paths in this area are undefined, the paving on the ground and the steps in this area are primarily used as places to sit and watch, and the middle of the space is clear of activity. In the interest of a concise descriptor set, the word 'activity' was a composite descriptor that emerged from the grouping process, representing the following different types of activities that were extracted from the image descriptions: walking, sitting, strolling, meeting, watching, being seen, shopping, and dining. Future iterations of the language may have to include these more detailed activity descriptions, depending on the granularity of its use. To confirm the succinctness of MIRANDA, word pairs were associated diagrammatically using the abstraction shown in figure 3 and 4 to visualize variations and frequencies of word pairs.

This was done using layered drawing software, with each different coloured layer representing the links from a signed place word, to its associated set of descriptors. This made it possible to view different combinations of word pairs for comparisons with each other, and to make it possible to view all pairs simultaneously. Each occurrence of a word pair within the dataset was plotted on the diagram as an additional one-point line width to the connecting link between the words. When all the language statements had been plotted on the diagram, the width of the word pair link therefore represented the number of occurrences of this word pair in the total dataset. This gave a visual form to the 'language' and helped to validate the choice of words in the finite word sets by indicating whether word-pairs were actually used in describing the space, the existence of a link, and frequency of use of those pairs, the width of the link.

This abstraction of architectural characteristics provided an additional benefit as a tool for visualising a summary of human experience of architectural elements of that environment, that is, an overview of the physical context of that space. Surveying the diagram, it is possible to draw summary conclusions about the space which would not be evident from viewing the original data, or from merely visiting the space, because it represents a composite view of that space, judiciously extracted from historical understanding of human experience of architectural space, as a lens for expert observation and analysis of this built environment.

MIRANDA was also applied to the information layer of Federation Square (analogue and digital signage) but for the purpose of clarity and exemplification of indexing to physical context in mobile guide interface design, the remainder of the paper will focus on the outcomes of the architectural analysis as illustrated in figure 3 and 4.

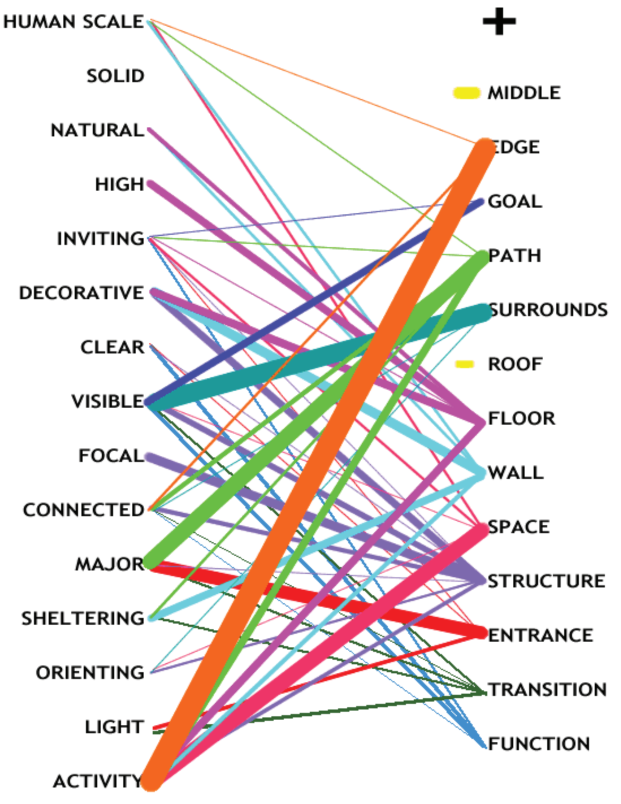


Figure 3. Abstraction of MIRANDA: positive description of architectural characteristics.

In analyzing figure 3 it is evident that activity in Federation Square occurs primarily along the edges and spaces generally have little activity in the middle. This can be surmised because the link between 'activity' and 'edge' is thick (+ activity.edge). We can also infer that the square looks out at the views of its surroundings, evidenced by the thick link between 'visible' and 'surrounds' (+ visible.surrounds). Many of the spaces have activities associated with them, such as places that are designed specifically for an activity such as sitting or dining, and so the link between 'activity' and 'space' is moderately thick (+ activity.space). Similarly, landmarks can be used to orient oneself, from most places around the square, indicated by the thick link between 'focal' and 'structure' (+ focal. structure). Other moderately thick links can be seen between 'major' and 'entrance' (+ major.entrance) and 'major' and 'path' (+ major.path), indicating that Federation Square has a few main entrances into buildings and main pathways around these structures, as opposed to a network of small pathways and multiple openings you might find in the centre of an old city.

Additionally, figure 4 enhances our understanding of the space by representing additional summary information about it in the form of the negation of these word-pairs. For instance, those major paths, that is, pedestrian links between two key destinations

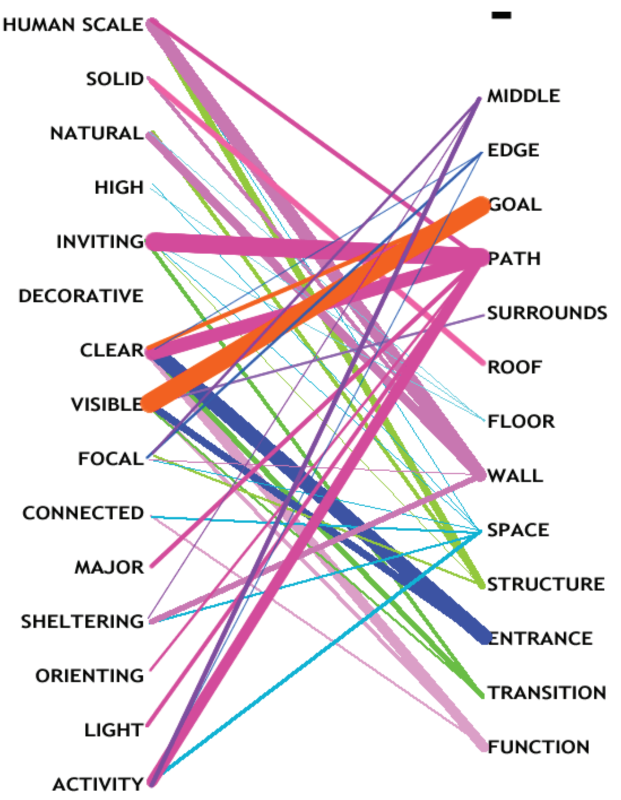


Figure 4. Abstraction of MIRANDA: negative description of architectural characteristics.

in Federation Square, are not inviting (- inviting.path). They are often either a narrow gap between two tall walls of adjacent buildings (- humanscale.wall) or they are down badly lit dark underground corridors (- light.path). When you are trying to find a specific place within the complex it is often difficult to see your destination, in several places you begin to follow a sign to a specific place, but the place that you are aiming for is not visible through the entrance indicated, leaving visitors unsure that it is the way to go (- visible.goal). In many places the pathway that you are supposed to take is unclear (-clear.path) because paths are not specifically indicated in open spaces.

This visual abstraction of MIRANDA represents a descriptive framework for understanding and modelling the physical context of a public place. MIRANDA makes available a vocabulary of a specific space that can then be used to understand the key physical characteristics of that built environment, and to describe the user's surroundings in a way that is grounded in human observation of that place, and formed with reference to collected knowledge about human understanding of architectural form and graphical communication. MIRANDA is unique, in so far as it is a systematic tool for creating an abstract representation of physical context, which does not exist in the literature surveyed, and yet that same literature often laments a lack of in-depth understanding of the physical context for which pervasive information systems are being designed. The aim of creating MIRANDA is to provide mobile system designers with knowledge about elements in the user's physical context, so that information that already exists in the world can be indexed to in the interface. The use of MIRANDA for this purpose is exemplified below.

4. MOBILE GUIDE DESIGN

To explore the contribution of MIRANDA to the design of physically indexed interfaces, we have designed a mobile guide system for use at Federation Square as a proof of concept. The design presented below is currently being implemented as a functional prototype running on handheld computers using Bluetooth for positioning within districts and GPRS for wireless access to the Internet. The prototype will be lab and field-tested early 2005.

The design of the prototype guide system is based on a combination of the findings from the Lynchian analysis of Federation Square and MIRANDA. The Lynchian analysis showed that clear districts can be identified within the space of Federation Square. MIRANDA showed for each of these districts, key architectural characteristics can be identified on a level of abstraction that allows a conceptual image of the space rather than a catalogue of specific physical elements.

Combining the findings from the Lynchian analysis and MIRANDA, we have developed three overall design ideas for a mobile guide for Federation Square, which exploits unique characteristics of the physical space and indexes to features of the built environment:

- The mobile guide responds to the user's location in terms of one of the defined districts rather than Cartesian coordinate
- Each district is represented in the mobile guide by an interactive photorealistic depiction of the physical surroundings augmented with textual or symbolic information needed to better understand the plac
- Locations and instructions for navigation are expressed through rich descriptions derived from the distinctive characteristics of the place rather than through Euclidian coordinates

The implementation of three design ideas in our mobile guide system is detailed below.

4.1. Location by district

In our mobile guide for Federation Square, the user's location is defined by four districts. These four districts were identified by the Lynchian analysis, each with their own distinct characteristics and a corresponding landmark. The location districts we are using make use of people's ability to make sense of the physical environment in which they are situated and are not defined by coordinates but by the human experience of the physical layout of the space. Basing positioning on location within a limited number of districts rather than on exact x, y coordinates also makes it possible to implement positioning using, for example, Bluetooth beacons rather than GPS. The location districts of the mobile guide and corresponding screens are illustrated in figure 5.

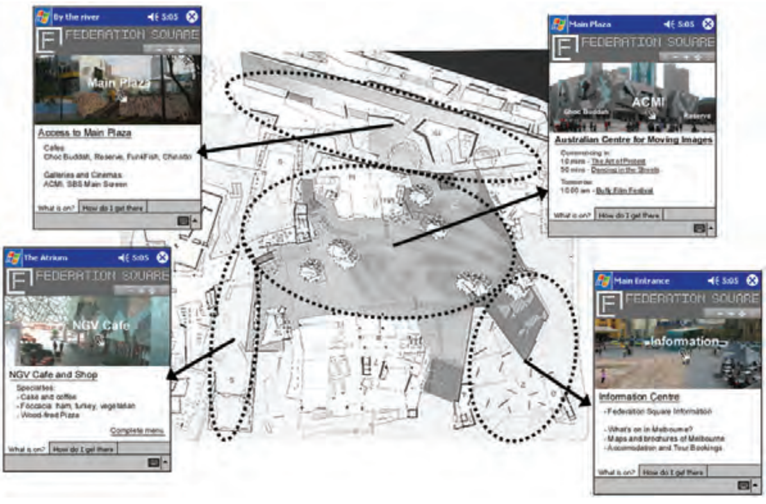


Figure 5. Location districts of the mobile guide for Federation Square and corresponding screens.

The information pushed to the mobile guide is tailored to information needs within a specific district. When a user moves into another district, their context changes, and so does the information that appears on the screen of their mobile device. To allow the user to align the information in the guide to their physical surroundings, the initial screen displays the corresponding landmark for that district.

In the following subsections, we describe the specific design of the screens associated with only one district, the main plaza.

4.2. Augmented interactive photorealistic depictions

When the user enters a district, an interactive photorealistic depiction augmented with textual and symbolic information pertinent to that district is pushed to their device. For the main plaza, MIRANDA identified six central architectural characteristics:

- Activity edges
- Focal structures
- No visible goals
- Visible surrounds
- Major entrances and paths
- No clear paths

These six features informed the interface design of the screens for the main plaza.

Wanting to represent the main plaza district in the system, we chose to use a 360° panoramic depiction of the plaza because MIRANDA tells us that the main plaza has 'activity edges' and is open in the middle. Focusing on the edges of the plaza will match the guide to the user's experience of the physical space and allow them to easily align the information presented in the system with their physical surroundings. We use photorealistic depictions in the design because we also know from MIRANDA that the main plaza has 'focal structures'. This enables users to align the image in the interface to the architectural features of the built environment in their current district (for example, as seen in Pospischil et al. 2002). Subsequently, they can easily scan through the virtual representation of the space and index the information presented in the mobile guide to the corresponding physical objects in their surroundings. Like when operating a QuickTime VR movie, panning is done by pointing on the panoramic depiction and dragging in the desired direction.

In a situation where the physical analysis of a district indicated that features of interest were, for example, clustered or spread out over a very large physical area, it might be more appropriate to use a different depiction such as a bird's eye view approaching the user's perception of that space (see e.g. Kulju and Kaasinen 2002, Laakso et al 2003).

The main screen for the main plaza is shown in figure 6.

Figure 6. Mobile Guide for Federation Square: Location and Event finding screen.

MIRANDA also identified a characteristic of the space as having 'no visible goals'. This feature of the space led to the augmentation of the 360° panoramic depiction of the plaza with white text labelling visitor destinations not clearly indicated by the architectural design. In many cases the actual location of activities lies behind the visible facades, but the facades function as clear 'focal structures' that can be indexed to in the system, linking information provided by the system to those visitor destinations. A location is automatically selected when it is positioned in the centre of the display. This is indicated by the enlargement of the text label of that location and by displaying the associated detailed information, as shown in figure 6. From this screen, the user can access additional details by clicking on the underlined links. Access to navigation information about that location is also now available by clicking on the "How do I get there" tab.

4.3. Rich descriptions for navigation

When the user clicks on the 'How do I get there?' tab from the main screen, they are presented with location and navigation information about the selected destination. The design of the navigation screen, as shown in figure 7, was informed by MIRANDA identifying the architectural characteristics of 'activity edges', 'visible surrounds', 'major entrances and paths' and 'no clear paths'. On the navigation screen, the user is presented with rich descriptions derived from the distinctive characteristics of the place. The descriptions used in the guide have two parts. The first describes the location in relation to focal structures. The second part describes the steps taken to get there.

Based on the knowledge from MIRANDA that activities are located on the edges, it follows that we are able to describe selected locations as being 'next to' and/or 'opposite' other locations on the edge of the main plaza. We use these terms in our rich descriptions, thereby indexing locations relative to one another. This is an alternative to absolute location descriptions typically used in mobile guides, for example 'located at the corner of Flinders and Swanston Streets', but holds more meaning to the users of the system, because it makes use of their understanding of the built environment.

Figure 7. Mobile Guide for Federation Square: Navigation Screen with rich descriptions.

The second part of the navigation information describes the path which the user should take to get there. We know from MIRANDA that paths in the main plaza are not clearly marked but that the urban surroundings are visible. This allows us to use descriptions, such as 'away from the train station' or 'towards the river', indexing to visible elements of the surrounding city. MIRANDA also tells us that access to the different locations surrounding the main plaza is afforded by major entrances. Based on this, the navigation screen contains an image of the entrance that the user must pass through to get there. Again, this indexes the information given by the mobile guide to the user's physical environment. Since some of the characteristics of a physical space sometimes may change in accordance to the time of day, season, special events etc. the rich descriptions in the system should be designed so that they can adapt accordingly.

5. CONCLUSIONS

As digital information becomes a part of the way we operate in the physical world, human computer interaction with mobile guides can benefit from making use of 'knowledge-in-the-world' by indexing to implicit information cues in the built environment. People have a strong ability to make sense of the physical space in which they are situated and often make use of this when using mobile guides in urban spaces. In order to exploit this ability in mobile guide design, we need to better understand the role of the user's physical environment in defining their context and the contribution of existing information embedded in that environment to people's experience of it. Also, we need to make clear connections between the user's physical surroundings and the information presented on their mobile guide.

This paper has presented the outcome of a field study into the properties of a built environment that contribute to people's experience of their physical context. Synthesizing the outcome of the field study, we have proposed a descriptive framework, MIRANDA, which provides a summarized abstraction of the fundamental architectural and informational features of a built environment. The usefulness of MIRANDA in interface design has been exemplified through the design of an indexical mobile guide system for Federation Square, Melbourne, Australia. Unlike other location-aware guide systems, the proposed design explicitly uses insight into user perceptions of architectural characteristics of the built environment to tailor the information presented to the physical context, in which it is being used.

While the outcomes of our analysis are specific to Federation Square, the empirical analysis of an environment used to create the MIRANDA framework could be applied to other cases of mobile guide design for the built environment. In physical spaces where a MIRANDA-based analysis reveals similar architectural characteristics to Federation Square, the specific design ideas of exploring indexicality presented above could also be applied.

The proposed design sketches are currently being implemented as a functional prototype. Through laboratory- and field-based evaluations at Federation Square, the prototype will allow us to validate the usefulness of MIRANDA for informing the design of mobile guides for use in built environments.

In future research, the proposed descriptive framework of MIRANDA should be refined and validated through further studies of other built environments both similar to and different from Federation Square. Additionally, it would be interesting to investigate further the dynamic aspects of physical spaces; characteristics changing according to time of day, season, special events etc. The proposed method does not take this issue explicitly into account during the data collection phase and would have to be extended for this purpose.

ACKNOWLEDGEMENTS

The authors thank Steve Howard and Bharat Dave for valuable supervision on the project and comments on the paper. The field study was designed and conducted by the first author, who was also responsible for the data analysis and development of MIRANDA. The design of the mobile guide for Federation Square was done in collaboration with the second author as a part of the "Design and Use of Context-Aware Mobile Multimedia Systems" and "Indexical Interaction Design for Context-Aware Mobile Computer Systems" research projects supported by the Danish Technical Research Council (project references 26-03-0341 and 26-04-006).

REFERENCES

Abowd, D., Atkeson, C., Hong, J., Long, S. and Pinkerton, M., 1996, Cyberguide: a mobile context-aware tour guide. Wireless Networks, 3(5), 421-433.

Agre, P., 2001, Changing places: contexts of awareness in computing. Human-Computer Interaction, 16, 177-192.

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I. Angel, S., 1977, A Pattern Language: Towns, Buildings, Construction (New York: Oxford University Press).

Andersen, P., 2002, Pervasive computing and space. In Organizational Semiotics: evolving a science of information systems, edited by L. Kecheng, R. J. Clarke, P.B. Andersen, R. Stamper, and El-Sayed Abou-Zeid, IFIP TC8/WG8.1, July 23-25 (Montreal, Quebec, Canada: Kluwer) pp. 133-152.

Aoki, P., and Woodruff, A., 2000, Improving electronic guidebook interfaces using a task-oriented design approach. Proceedings of DIS 2000 (New York, USA, ACM), pp. 319-325.

Bertin, J., 1983, Semiology of Graphics (Wisconsin: The University of Wisconsin Press).

Bohnenberger, T., Jameson, A., Krüger, A. and Butz, A., 2002, Location-aware shopping assistance: evaluation of a decision-theoretic approach. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 155-169.

Bornträger, C., Cheverst, K., Davies, N., Dix, A., Friday, A. and Seitz, J., 2003, Experiments with multi-modal interfaces in a context-aware city guide. Proceedings of Mobile HCI 2003 (Udine, Italy: LNCS, Springer-Verlag), pp. 116-130.

Bornträger, C., and Cheverst, K., 2003, Social and technical pitfalls designing a tourist guide system. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Bowman, W. J., 1968, Graphic Communication. (New York: John Wiley & Sons, Inc.).

Bradley, N., and Dunlop, M., 2002, Understanding contextual interactions. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 349-353.

Beyer, H., and Holtzblatt, K., 1998, Contextual Design: Defining Consumer-Centered Systems. (San Francisco: Morgan Kaufmann Publishers).

Borchers, J., 2001, A Pattern Approach to Interaction Design. (Chichester, England: John Wiley & Sons, Ltd).

Cheverst, K., Davies, N., Mitchell, K., 2002, Exploring context-aware information push. Personal and Ubiquitous Computing, 6, 276-281.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C., 2000 Developing a context-aware electronic tourist guide: some issues and experiences. Proceedings of CHI'00 (The Hague, Netherlands: ACM), pp. 17-24.

Chincholle, D., Goldstein, M., Nyberg, M. and Erikson, M., 2002, Lost or found? A usability evaluation of a mobile navigation and location-based service. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 211-224.

Dieberger, A., and Frank, A., 1998, A city metaphor to support navigation in complex information spaces. Journal of Visual Languages and Computing, 9, 597 – 622.

Dix, A., Rodden, T., Davies., Trevor, J., Friday, A., and Palfreyman K., 2000, Exploiting space and location as a design framework for interactive mobile systems. ACM Transactions on Computer-Human Interaction, 7(3), 285-321.

Fithian, R., Iachello, G., Moghazy, J., Pousman, Z. and Stasko, J., 2003, The design and evaluation of a mobile location-aware handheld event planner. Proceedings of Mobile HCI 2003 (Udine, Italy: LNCS, Springer-Verlag), pp. 145-160.

Gamma, E., 1997, Going beyond objects with design patterns. Proceedings ECOOP '97 - Object-Oriented Programming: 11th European Conference, June 1997 (Jyväskylä, Finland), p. 530.

Graham, C., and Kjeldskov, J., 2003, Indexical representations for context-aware mobile devices. Proceeding of The IADIS e-Society Conference (Lisbon, Portugal), pp. 373-380.

Hakkila, J., and Hexel, R., 2003, Interaction in location-aware messaging in a city environment. Proceedings of OZCHI2003: New Directions in Interaction: information environments, media & technology, December 2003 (Brisbane, Australia: CHISIG), pp. 84-93.

Harrison, S. and Dourish, P., 1996, Re-placing space: the roles of place and space in collaborative systems. Proceedings of the ACM CSCW 96 Conference on Computer-Supported Cooperative Work (New York: ACM), pp. 67-76.

Hermann, F., and Heidmann, F., 2002, User requirement analysis and interface conception for a mobile, location-based fair guide. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 388-392.

Holland, S., and Morse, D. R., 2001, AudioGPS: spatial audio in a minimal attention interface. Proceedings of Mobile HCI 2001 (Lille, France).

Iacucci, G., Kela, J. and Pehonen, P., 2004, Computational support to record and re-experience visits. Personal and Ubiquitous Computing, 8(2), 100-109.

Ingram, R., Benford, S. and Bowers, J., 1996, Building virtual cities: applying urban planning principles to the design of virtual environments. Proceedings of ACM symposium on Virtual Reality Software and Technology (VRST96), July 1996 (ACM), pp. 83-91.

Kjeldskov, J., Howard, S., Murphy, J., Carroll, J., Vetere, F. and Graham, C., 2003, Designing TramMate - a context aware mobile system supporting use of public transportation. Proceedings of DUX 2003 (San Francisco, CA, USA: ACM).

Kjeldskov, J., 2002, Just-in-Place: Information for mobile device interfaces. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 271-275.

Kolari, J., and Virtanen, T., 2003, In the zone: views through a context-aware mobile portal. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Kulju, E. and Kaasinen, E., 2002, Route guidance using a 3D city model on a mobile device. Proceedings of Mobile Tourism Support (Pisa, Italy: in conjunction with Mobile HCI 2002).

Laakso, K., Gjesdal, O., and Sulebak, J. R., 2003, Tourist information and navigation support by using 3D maps displayed on mobile devices. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Lynch, K., 1960, The Image of the City. (Cambridge, Mass.: MIT Press).

McCullough, M., 2001, On typologies of situated interaction. Human-Computer Interaction, 16, 337-349.

Millen, D. R., 2000, Rapid ethnography: time deepening strategies for HCI field research. Proceedings of DIS 2000 (Brooklyn, New York, USA: ACM), pp. 280-286.

Nakanishi, H., Koizumi, S., Ishida, T., and Ito, H., 2004, Transcendent communication: location-based guidance for large-scale public spaces. Proceedings of CHI 2004 (Vienna, Austria: ACM), pp. 655-662.

Norman, D., 1990, The Psychology of Everyday Things. (New York: Basic Books).

Opperman, R., and Specht, M., 2000, A context-sensitive nomadic information system as an exhibition guide. Proceedings of Handheld and Ubiquitous Computing Second International Symposium (Bristol, UK: Springer-Verlag), pp. 127-142.

Paay, J., 2003, Understanding and modeling physical environments for mobile location aware information services. Proceedings of Mobile HCI 2003 (Udine, Italy: LNCS, Springer-Verlag), pp. 405-410.

Paulos, E., and Goodman, E., 2004, The familiar stranger: anxiety, comfort and play in public places. Proceedings of CHI 2004 (Vienna, Austria: ACM), pp. 223-230.

Persson, P., Espinoza, F., Sandin A., and Coester, R., 2002, GeoNotes: a location-based information system for public spaces. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 151-173.

Pospischil, G., Umlauft, M. and Michlmayr, E., 2002, Designing LoL@, a mobile tourist guide for UMTS. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 140-154.

Randell, C., and Muller, H., 2000, The shopping jacket: wearable computing for the consumer. Personal and Ubiquitous Computing, 4, 241-244.

Ricci, F., Cavada, D., and Nguyen, Q. N., 2002, Integrating travel planning and on-tour support in a case-based recommender system. Proceedings of Mobile Tourism Support (Pisa, Italy: in conjunction with Mobile HCI 2002).

Rocchi, C., Stock, O., and Zancanaro, M., 2003, Semantic-based multimedia representations for the museum experience. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Sparacino, F., Davenport, G., and Pentland,, 2000, A city of news: cataloguing the world wide web through virtual architecture. KOS, Aug-Sept 2000, 179-180.

Schmidt-Belz, B. and Poslad, S., 2003, User validation of a mobile tourism service. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Tamminen, S., Oulasvirta, A., Toiskallio, K., and Kankainen, A., 2003, Understanding mobile contexts. Proceedings of Mobile HCI 2003 (Udine, Italy: LNCS, Springer-Verlag), pp. 17-31.

Tidwell, J., 1999, Common Ground: A Pattern Language for Human-Computer Interface Design. http://www.mit.edu/~jtidwell/interaction_patterns.html

Tufte, E., 1990, Envisioning Information. (Cheshire, Conn.: Graphics Press).

Vainio, T., Kotala, O., Rakkolainen, I., and Kupila, H., 2002, Towards Scalable User Interfaces in 3D City Information Systems. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 354-358.

Vinson N. G., 1999, Design Guidelines for Landmarks to Support Navigation in Virtual Environments. Proceedings of CHI 1999 (Pittsburgh, PA, USA: ACM), pp. 278-285.

Chapter 4

Social context

Jeni Paay and Jesper Kjeldskov

Abstract. Ubiquitous and mobile computer technologies are increasingly being appropriated to facilitate people's social life outside the work domain. Designing such social and collaborative technologies requires an understanding of peoples' physical and social context, and the interplay between these and their situated interactions. In response, this paper addresses the challenge of informing design of mobile services for fostering social connections by using the concept of place for studying and understanding peoples' social activities in a public built environment. We present a case study of social experience of a physical place providing an understanding of peoples' situated social interactions in public places of the city derived through a grounded analysis of small groups of friends socialising out on the town. Informed by this, we describe the design and evaluation of a mobile prototype system facilitating sociality in the city by 1) allowing people to share places, 2) indexing to places, and 3) augmenting places

1. INTRODUCTION

Mobile and ubiquitous computer technologies are increasingly being appropriated to facilitate people's social life outside the work domain linking people to people to places (Jones et al. 2004). Mobile phones, and especially SMS texting, have changed the way people communicate, interact in the physical world, and coordinate their social activities (Grinter and Eldridge 2001; Rheingold 2003). By embedding networked sensors into the built environment, adding advanced positioning technology and short range network capabilities (such as Bluetooth, RFID tags, etc.), context-aware mobile services are emerging that adapt their content to both the user's physical and social context.

When designing mobile services for fostering social connections and augmenting our physical built environment, system developers and interaction designers are faced with a series of new challenges. We need to understand better the physical and social context of the user's situated social interactions (McCullough 2004), the role of human

Originally published as Paay, J. and Kjeldskov, J. (2008) Situated Social Interactions: a Case Study of Public Places in the City. *Computer-Supported Cooperative Work,* 17(2-3), 275-290.

activity within the built environment (Ciolfi 2004) and the interplay between context and user actions (Dourish 2004). We also need to understand how physical and social affordances of a place influence the situated interactions that occur there, including the relationship between people, technology and interactions. Finally, we need to define useful and understandable ways of incorporating peoples' physical and social context in interaction design for context-aware mobile services.

Recent work in human-computer interaction (HCI), computer-supported cooperative work (CSCW) and interaction design has examined how the concept of *place* can contribute to our understanding of peoples' interactions within their physical environments and with ubiquitous computing technologies augmenting this environment, and how the notion of place can inform system and interaction design.

This paper addresses the challenge of informing ubiquitous and mobile technology design by using the concept of place as a central notion for studying and understanding peoples' social activities in a public built environment. We present a case study of social experience of a physical place providing an understanding of peoples' situated social interactions in public places of the city derived from a grounded analysis of small groups of friends socialising out on the town. Informed by this, we describe the design and evaluation of a context-aware prototype system facilitating sociality in the city by 1) allowing people to *share places,* 2) *indexing to places,* and 3) *augmenting places.*

The paper is structured in the following way. Section 2 discusses related work focusing on people, technology and interactions in place. It presents and discusses our understanding of place, ubiquitous technology use in city contexts, and introduces the concept and typology of situated interactions. In section 3 we present our field study of people socialising in public places, describing the details of our empirical method and data analysis. In section 4 we present the findings from our study of situated social interactions in public places. To illustrate the value of understanding social interactions in place for informing interaction design of mobile services, section 5 describes the design and evaluation of an implemented prototype system, which adapts to the user's physical and social context to foster social connections in that place. Section 6 concludes on our study.

2. BACKGROUND

2.1. People in place

The design of the city affects how people make sense of the social complexities of urban places. The architectural design of form in the built environment has traditionally occurred within the context of an explicit set of social and physical issues in respect of anticipated activities and historical expectations tied to particular institutions and building types (Agre 2001; Mitchell 1995). Physical and social affordances of a place have helped to define the social interactions that occur there (Gaver 1996). Physical space plays a constructive as well as a receptive role in shaping social interaction in urban places (Hillier and Netto 2002). Space is given significance and becomes place through its link to human activity. We are located in space, but we act in place. Our shared understanding of the physical world helps people in presenting and interpreting activity and behaviour (Harrison and Dourish 1996). The physical and social layers of a space form the context of interaction for its inhabitants, intimately connected to their activities (Donath 1996). Accumulated experience helps people to identify with a place and in turn gain an understanding of what is going on in that place. Understanding the context of social interactions is an important part of designing ubiquitous computing that delivers information to people in the places and activities of their daily life (Agre 2001).

2.2. Technology in place

Architectural ideas about the nature of place are being challenged as communication and computation devices begin to saturate the built environment (Rheingold 2003). Ubiquitous computing is breaking down the traditional mapping between activities and place, allowing people to participate in social interactions that are no longer tied to their current location by supporting continual presence in every place (Agre 2001). For example, cafés become corporate meeting rooms as users deal with business calls over lunch, without any changes to the physical fabric of a place. Technology is uncoupling the close relationship between activities and place previously imposed by architectural design allowing social interactions to extend beyond a person's current physical location. Places no longer define appropriate activities by their physical design alone: now every place can be for everything, all of the time (Agre 2001; Mitchell 2003).

Understanding how to design ubiquitous computing that meshes with human behaviour and the properties of place that structure human interaction is immensely important (Ciolfi 2004; Erickson 1993). People who are digitally connected to each other and to the elements of the city use that technology to deliver information that is "just in time" and "just in place", to guide them to where they want to go and inform them about possible activities. This digital layer not only helps to structure our social interactions, but also provides a social medium for facilitating and enriching every day interactions between individuals (Erickson 1993).

Mobile services are increasingly becoming a part of the way we operate in urban places. Context-aware mobile information systems provide access to contextually adapted information and can foster social connections by sensing and responding to groups of co-located people in a place. In essence, they are connected to and respond to the place in which they are operating. The design of context-aware mobile information systems covers a broad spectrum of application areas, many of these mobile information systems involve the user being situated in urban public places, and yet only a few have investigated the challenges imposed and the opportunities offered through a grounded understanding of the relationship between activity and place.

2.3. Interactions in place

Studying people's "everyday action" can provide designers with a sense of the meaning associated with user activities, knowledge about what they actually do in a particular situation, and an understanding of people's experience of place. As Ciolfi (2004, p. 39) says, "understanding the dynamics of interaction in a space can help us design more effective systems in responding to behaviour and to changes in the environment."

McCullough (2004) approaches this problem with the idea of using typology (the study of recurrent forms) as a design philosophy to provide types of everyday situations as a way of abstracting an understanding of the influence of place on interaction. Using typology as a design philosophy provides a framework for creativity, allowing design to be based on themes rather than arbitrary innovation. It acknowledges existing living patterns of an inhabited place and helps designers of digital technology to recognise situated interactions and make technology a simpler, more adaptive and more social part of those interactions. McCullough asserts that place becomes reconfigured by ubiquitous computing not replaced by it, and that technology then extends the living patterns of that place. This approach to information technology design focuses on the need to understand how people interact in place. Gaining that understanding can be used to facilitate human-centred design of mobile services for fostering social connections.

A rudimentary typology of 30 everyday situations that may be transformed by technologies is proposed by McCullough (2004). This typology classifies situational types, grouped to reflect the following categories of place: workplace, dwelling place, the "third place" for conviviality, and the "fourth place" of commuting and travel. By using this typology as an analytical lens in this study, the concept of place becomes an organising theme for the data collected. This also limits the focus of the fieldwork to a manageable range of recognisable situations, allowing for design variations to benefit from being based on a few appropriate themes (McCullough 2004). As derived from McCullough (2004), the situated interactions associated with places for conviviality, that is, being out "on the town" are: places for socializing, places to meet, places for seeing and being seen, places for insiders, places for recreational retailing, places for embodied play, places for cultural productions, and places for ritual.

3. FIELD STUDY: PEOPLE SOCIALISING IN A PUBLIC PLACE

Exploring the interplay between people, activity and place, we conducted an empirical field study of situated social interactions in the city. This study investigated the use of McCullough's (2004) typology of "on the town" everyday situations to guide fieldwork for informing interaction design of a mobile information system for a public place. The field study took place at Federation Square, Melbourne, Australia (figure 1). Federation Square is a new civic structure covering an entire city block, providing the people of Melbourne with places for a variety of activities including restaurants, cafés, bars, a museum, galleries, cinemas, retail shops and several public forums.

Figure 1. Federation Square, Melbourne, Australia, with surrounding skyline and river

3.1. Participants, procedure and data collection

The field study was conducted on location at Federation Square using the rapid ethnography method (Millen 2000). McCullough's (2004) typology focused the research scope at the beginning of the fieldwork by suggesting places for observations, and contextual interviews (Beyer and Holtzblatt 1998) facilitated interactive observation. Three different established social groups participated in the study as key informants. Each group consisted of three young urban people, mixed gender, between the ages of 20 and 35, who had a shared history of socialising at Federation Square. Each group met at Federation Square where they were not given any specific tasks but were asked to simply undertake the same activities that they would usually do as a group when socialising in the city. Each contextual interview and observation lasted approximately three hours (figure 2). Digital video was used to document all questions, responses, activities and movement of the group around the square.

Figure 2. Contextual Interview

Figure 3. Affinity Diagramming

3.2. Transcriptions and data analysis

Shortly after the field visits all recordings were reviewed and situated interactions transcribed. The analysis of the transcript involved open and axial coding adapted from the grounded theory method (Strauss and Corbin 1990) chosen for its structured bottom up approach to analysing data to generate themes, and affinity diagramming (Beyer and Holtzblatt 1998) as a method for hierarchical grouping of themes (figure 3). Grounded theory analysis produced 107 novel themes describing interactions and their relationship to place and activity. The affinity diagram refined these to a small set of high-level concepts, representing the essence of the data and encompassing all lower level themes, structured in a conceptual framework around the three key concepts of *knowledge, situation,* and *motivation* as described in section 4. Orthogonal to these concepts, three "place-related" design ideas of *sharing place, indexing to place,* and *augmenting place* were drawn, implemented and evaluated, as described in section 5.

4. SITUATED SOCIAL INTERACTIONS IN PUBLIC PLACES

The conceptual framework encapsulates a structured understanding of every day social interaction in the situation of a public place. It provides an understanding of the role of physical and social context in how people experience a physical place and how they interact with each other while socialising, in the form of a qualitative story woven around three key concepts: knowledge, situation and motivation.

4.1. Knowledge

Knowledge is an important part of how we operate while socialising in an urban environment. When interacting in urban places people use their understanding of the world around them to make sense of things.

In the study, participants operated using the physical affordances (Norman 1990) of elements, for example, assuming steps with high risers as being for sitting. They saw large open spaces as places for people to gather. If a space had a visual focal point then it was regarded as a good place for locating a special event or performance. Visible openings in facades indicated entrances, and architectural features such as low walls defined boundaries for sitting or walking and in this way confined activities. Participants drew on their history with that specific urban environment. Physical familiarity with a space meant that they approached familiar places using familiar paths, that is, the way that they "usually come". A familiar path was not perceived as the long way round, even if it was in terms of physical distance and they often assumed that others had the same familiar paths.

Participants also operated in public places using a set of social affordances. They looked to what other people were doing as cues for what to do in a place. Following crowds or people queuing was a way for them to decide where they might go. They looked at others to confirm what activities were acceptable in a place. Places where others were sitting made them feel they might sit there too. They read the presence of many people in an establishment as a recommendation that it was a good place to go. Participants expressed a desire to socialise where others were relaxing and enjoying themselves and were drawn into a place where they could see this happening from the outside. They also used social experience as a basis for selecting places to socialise with friends and their own past experience or shared group experience to index to past social events, for example "let's meet where we met last time". The impression of liking a place was based on successful past visits. Trying new places was based primarily on recommendations from friends or trusted media reviews. If they were socialising with a group of friends they met in the place where they usually met with those particular friends.

4.2. Situation

Situation is an important aspect of sociality in urban space. When socialising the presence of both friends and strangers influences the way that people behave and move through urban place.

In the study, friends maintained their sense of "group" by the way that they physically located themselves in a public place. As they moved through space they often walked abreast, or single file in crowded situations, but always very much together. When they stopped they gathered in a circle to discuss options and excluded the outside from the interaction.

Participants liked to be near others but not necessarily interacting directly with them. One participant called this "socialising by proximity" meaning that they wanted to be amongst others, often enjoying sharing a long table with several other groups in a place, but not feeling as if they had to talk directly to them. They liked to watch others, especially if they felt unobserved themselves. This generally meant being in an elevated position compared to the people they were watching or behind a low wall or plant box, to keep others at a distance. They mostly engaged in this activity when on their own.

Participants liked to wait for others in a place where they could see their friends arriving, specifically in a location that overlooked the entrance to a place, for example, at a table facing the door of a bar. The length of time that they had to wait affected the choice of meeting place. If their friend was going to be a long time (defined by participants as 30 minutes or more) they wanted an activity to do while waiting. If it was a short time (a few minutes) convenience to the meeting place was more important. Sitting outside at bars and cafes was perceived as more comfortable than waiting alone inside.

Setting influenced sociality. The presence of others and the types of people in a place influenced its acceptability. Participants expressed that they liked to socialise in places with similar types of people, i.e., age, dress, intentions. Environmental comfort was also important. Whether a place was sunny, sheltered, etc., influenced the choice of location to socialise or wait. Participants preferred sitting outside socialising in nice weather. They also preferred to sit in an elevated position with an interesting view out. The convenience of a place was also important. Participants preferred starting a "night out" in a location that had other activities they might like to do nearby.

Surroundings were an important part of situation and were often used as reference points. Participants indexed to things around them and to experiences shared with the friends they were with. They gave directions to a friend by referring to shared places and activities such as "next to the place we went last time where we sat in the sun". Participants also referred to visible elements, pointing to them or referring to generally known events or physical objects, including landmarks. For example, they would often use statements such as "through that opening", or index to landmarks in their surroundings, such as "it's near the big screen". Connecting stairs or pathways between physically separated spaces formed major transition points used in descriptions on how to get from one place to another.

4.3. Motivation

Reflection on current experience is part of socialising in a place. People try to size up the situation and like to get an overview of what is happening in a place. Before entering a place they stand back and familiarize with it and often pause before committing to a situation.

In the field, participants strived to make sense of things and places around them. Even if they had already decided to go to a familiar place, they would stand outside and review the menu before going in. Making sense of how things were organised was based on people's past experience with similar situations and by assessing the activities of others. Participants made very little use of signage, information kiosks or media screens in trying to do this sense making. Media screens while ostensibly informative were often regarded as decoration, something to make an environment more exciting. If they had a query, they usually asked a friend.

Participants gathered information about a place while socialising in it. Individuals required different levels of information for different activities. Those who required the most cursory level of information often set the pace of the group, others requiring more detail said they would only seek this depth of information when on their own. All participants wanted to know what was new in a place and if something special was happening.

In way finding, participants navigated by familiar paths and looked ahead for structures, objects and landmarks that they recognised and knew were near their destination. Participants discovered that urban spaces were dynamic, and paths were sometimes altered by the presence of crowds and temporary or new structures. In this situation, they avoided unfamiliar paths if they were not sure where they led, searching for the nearest familiar place and preferring to walk toward light rather than dark paths.

Extension of knowledge about a place often motivated social activity. Participants took part in exploration for the sake of it by wandering and browsing in a space. Sometimes they just wanted to know what was going on without any intention of joining activities. They enjoyed browsing as a group activity, allowing displays in shop windows to draw them in, and spent time negotiating what to do and where to go next.

5. DESIGNING FOR SITUATED SOCIAL INTERACTIONS

Inquiring into the usefulness of the understanding represented by the conceptual framework for informing interaction design, we designed, implemented and evaluated a prototype system for fostering social connections "in place". Firstly, we conducted a two-day design workshop to derive design ideas – or "design sensitivities" (Ciolfi and Bannon 2003) – for a context-aware mobile information system supporting sociality in the city. Following this, several iterations of paper-prototyping (Snyder 2003) turned the most promising ideas into a high-fidelity paper prototype. Subsequently, we implemented the paper prototype as a functional web application running in Microsoft Pocket Internet Explorer on HP iPAQ h5550 using mySQL, PHP, pushlets and server-side applications for handling context-awareness and dynamic generation of maps and graphics. The final system keeps track of the user's location, their current activity and friends within close proximity. It also keeps a history of the user's visits to places around the city. The technical details of the prototype are described in Kjeldskov and Paay (2005). The prototype system was evaluated by studying peoples' use of it for approximately 1 hour in either a laboratory or while socialising at Federation Square. The evaluation participants were 20 established social pairs familiar with Federation Square (10 in the lab and 10 in the field), and the prototype was pre-loaded with details about the participants' history of social interactions at Federation Square, together as well as on their own or with other people, derived from a pre-evaluation questionnaire.

In this section we focus on describing three of the seven design ideas emerging from the fieldwork to illustrate the resulting prototype design, and highlight feedback from the evaluations. Each design idea was drawn directly from themes and categories in the conceptual framework:

- Sharing place: recommendations based on history and context
- Indexing to place: way-finding referring to the familiar
- Augmenting place: representing people and activities in proximity

5.1. Sharing Place: Recommendations Based on History and Context

Evidenced in the data by the way people make decisions about where to go, was the importance of people's past experiences in terms of their existing knowledge, history of visits, social experience with places, and their current social group. This was explored using a sketch to examine the relationship of experience between two people, A and B (figure 4 left). Looking at the sketch from A's point of view, A has a past history which includes a number of familiar places. A subset of A's history is shared with B and represents shared experience which can be referred to through indexical relational descriptions such as "where we met last time". B also has a past history of familiar places that A has not been to. When A and B are socialising these places become recommendations from B for new places for A to go.

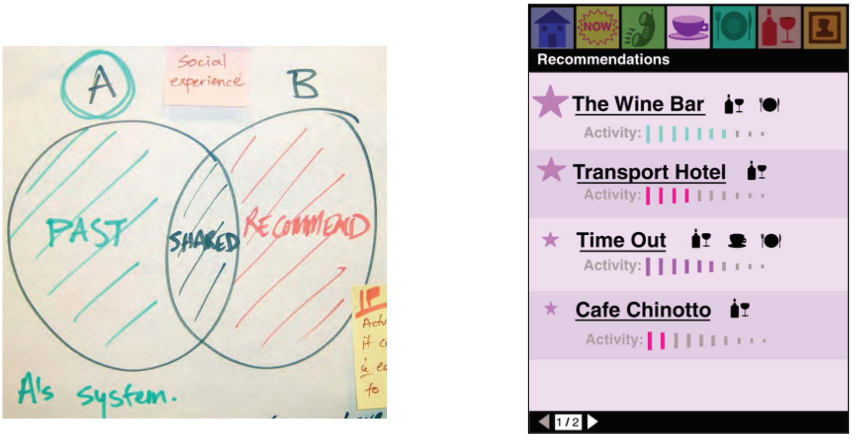


Figure 4. Design sketches: indexing to peoples' individual and shared histories

On the basis of the overall design idea of indexing content to the users' individual and shared histories, the prototype was designed to facilitate "sharing place" by ranking recommendations about places to go. When a member of a social group (the user) selects a specific activity on the device, for example, "having coffee", it presents a list of recommendations of places to go (figure 4 right), ranked on the basis of the systems knowledge about the user's familiar places (where the user has been to before together with these friends), current social setting (places that people in the current social group have been to before but not together), the current environmental setting (how well the weather situation of past visits to a place fits the current conditions), and convenience (places within the vicinity of the social group). Each place has an associated "activity-meter" displaying the current patronage and primary activity to accommodate the finding that setting matters in relation to the presence and similar intentions of others in a place. This gives the social group a chance to pause before committing to an activity or a place.

Studying the use of this feature in the evaluation of the system, we found that people generally thought it was interesting to be able to share information about places they liked to go to and also be able to explore new places in a space through implicit recommendations from the friends they were with. However, they also expressed that they would like to have more control over the system's methods for ranking of places. On the interaction design level, we found that while people generally understood that the system adapted information to their location in space and to the places around them, they were surprised that the system also adapted to their social context (who they were with) and had to have this explained to them indicating that this design lacked the necessary interface cues for them to fully understand it.

5.2. Indexing to Place: Way Finding Referring to the Familiar

The data collected shows that people seldom navigate by means of detailed maps and route descriptions when making their way around a space such as Federation Square as a part of a social group out on the town. Instead they use their history and especially physical familiarity with a space or place as well as physical affordances, such as visible places to enter and landmarks, to find their way around a space. They rely on simple indexing to their familiar places and prefer to follow their familiar paths from one place to another even if this may not be the most direct route. This finding was used to develop a sketch of the idea of basing wayfinding instructions on simple, indexical references to landmarks and familiar places with consideration to the user's history of familiar paths rather than the most direct route (figure 5 left).

Figure 5. Design sketches: indexing way finding to familiar places and paths

In the prototype, the "Getting There" option displays information to the user about how to get to a destination from their current location based on references to places where they have been before, for example, "Chocolate Buddha is located next to ACMI Cinemas opposite Arintji" (figure 5 right). If the destination is not in the vicinity of anything known by the user, the way-finding descriptions direct the user to the familiar place or landmark closest to the destination and give detailed directions from there. The way finding directions are combined with photographs of places, landmarks and transition points providing information that takes into consideration what people already know about places around them, and acknowledges their ability to make sense of an unfamiliar place on the basis of a few simple cues to familiar elements.

Studying the use of this feature, we found that people were highly capable of making sense of sometimes very reduced and fragmented information when it related to places they already knew. People were good at matching up objects, structures and outlines in their physical surroundings to images on the screen. Using pictures as reference points for both familiar places and for significant structures and elements of the surrounding space helped “fill in the gaps” in the way finding instructions.

5.3. Augmenting Place: indicating People and Activities in Proximity

Another important observation made from our field study was the importance of knowing about the existence of other people in a space and what they are doing. The interaction between a social group and the co-inhabitants of a space is complex. It involves a certain level of interaction between the group and others, either by proximity or by watching. Observing where other people are gathering and what they are doing there helps in getting an overview of a place, making sense of what is happening and sizing up the situation, which are an important part of pausing before committing to enter a place. This finding was used to sketch and develop the idea of representing current activities of others within close proximity (figure 6 left).

Figure 6. Design sketches: representing activities and people in proximity

In the prototype, when the user selects “NOW” in the main menu it displays a small map of the user’s immediate surroundings (figure 6 right). On this map superimposed, dynamically updated coloured circles indicate the clustering and activities of people within proximity. The radius of the circles indicates the number of people at a place while the colour represents their primary current activity (e.g. purple shows people “having coffee”). The map also shows the location of the user. By clicking on the coloured circles the user can access more information about each place.

Studying the use of this feature, we found that people were fascinated with the idea of knowing about people, places, and activities in the space immediately surrounding them. This was perceived as being of great interest and value for getting an overview of a public place and for informing discussions among the group about what to do and where to go next. People happily made detailed assumptions about the presence and activities of other people in the places around them based on this relatively simple graphical representation.

6. CONCLUSIONS

We have presented a case study of human experience of a physical place providing an understanding of peoples' situated social interactions in public places of the city derived from a field study of small groups of friends socialising out on the town. Based on a grounded theory analysis of our findings we have presented a qualitative conceptual framework of situated social interactions in a public place, and illustrated how this conceptual framework informed the design of a mobile context-aware prototype for supporting sociality in the city. This was achieved by providing a place-based understanding of people's situated social interactions in an abstract form inspiring design rather than specifying system requirements. Finally, we have presented preliminary empirical findings about the interplay between technology, people and place.

The literature calls for extended understanding of the contexts of everyday activities (Agre 2001; Ciolfi 2004; Dourish 2001; Erickson 1993; McCullough 2004). This is especially important when designing ubiquitous and mobile computer systems pervading the places and social activities of daily life. We need to understand better the user's physical and social context, their situated social interactions (McCullough 2004), the role of human activity within the built environment (Ciolfi 2004) and the interplay between context and user actions (Dourish 2004).

Understanding how people behave in public places can be interpreted by considering their social and physical context, that is, the roles of others and their surrounding environment. The presence and activities of people in the built environment gives locations in space cultural and social meaning, transforming spaces into places. The history of interactions in a place, and the experience of similar situations in other places, all influence peoples' perception and understanding of a place. To be able to design mobile services for fostering social connections in place, their situated social interactions need to be understood in respect to the physical and social context in which they occur.

Applying the notion of place to the study of peoples' situated social interactions in the city provides a useful lens and conceptual foundation for generating such understanding about the interplay between people, activities and place, and for informing the design of new ubiquitous and mobile technologies "augmenting the city".

ACKNOWLEDGEMENTS

This research is supported by the Danish Technical Research Council (26-04-0026), the Smart Internet CRC, Australia, and The University of Melbourne's David Hay Award program. The authors thank everyone participating in the field study and prototype evaluations. We also thank Steve Howard and Bharat Dave for valuable inpu.

REFERENCES

Agre, P. (2001): Changing Places - Contexts of Awareness in Computing. *Human-Computer Interaction*, vol. 16, pp. 177-192.

Beyer, H. and Holtzblatt, K. (1998): *Contextual Design - Defining Customer Centred Systems*. San Francisco: Morgan Kaufmann.

Ciolfi, L. and Bannon, L. (2003): Learning from Museum Visits: Shaping Design Sensitivities. *Proceedings of HCI International 2003, Crete, Greece, June 22 to 27, 2003*. London: Lawrence Erlbaum, pp. 63-67.

Ciolfi, L. (2004): Understanding Spaces as Places: Extending Interaction Design Paradigms. *Cognition Technology and Work,* vol. 6, no. 1, pp. 37-40.

Donath, J. (1996): *Inhabiting the Virtual City: The design of social environments for electronic communities*. Unpublished Thesis, School of Architecture and Planning, Massachusetts Institute of Technology.

Dourish, P. (2001): Seeking a Foundation for Context-Aware Computing. *Human-Computer Interaction*, vol. 16, pp. 229-241.

Dourish, P. (2004): What We Talk About When We Talk About Context. *Personal and Ubiquitous Computing,* vol. 8, no. 1, pp. 19-30.

Erickson, T. (1993): From Interface to Interplace: The Spatial Environment as a Medium for Interaction. *Proceedings of Conference on Spatial Information Theory, COSIT'93, Elba Island, Italy, September 19 to 22, 1993*. Berlin: Springer Verlag, pp. 391-405.

Gaver, B. (1996): Affordances for Interaction: The Social is Material for Design. *Ecological Psychology*, vol. 8, no. 2, pp. 111-129.

Grinter, R. E. and Eldridge, M. (2001): y do tngrs luv 2 txt msg? *Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work ECSCW'01, Bonn, Germany, September 16 to 20, 2001*. Dordrecht: Kluwer Academic Publishers, pp. 219-238.

Harrison, S. and Dourish, P. (1996): Re-Placing Space: The Roles of Place and Space in Collaborative Systems. *Proceedings of Computer Supported Cooperative Work '96, Boston, Massachusetts, USA, November 16 to 20, 1996.* Cambridge, MA: ACM Press, pp. 67-76.

Hillier, B. and Netto, V. (2002): Society seen through the prism of space: outline of a theory of society and space. *Urban Design International*, vol. 7, pp. 181-203.

Jones, Q., Sukeshini, G. A., Terveen, L. and Whittaker, S. (2004): People-to-People-to-Geographical-Places: The P3 Framework for Location-Based Community Systems. *Computer Supported Cooperative Work,* vol. 13, pp. 249-282.

Kjeldskov, J. and Paay, J. (2005): Just-for-Us: A Context-Aware Information System Facilitating Sociality. *Proceedings of Mobile HCI 2005, Salzburg, Austria, September 19 to 22, 2005*. New York: ACM Press, pp. 23-30.

McCullough, M. (2004): *Digital Ground - Architecture, Pervasive Computing, and Environmental Knowing.* Cambridge, Massachusetts: The MIT Press.

Millen, D. R. (2000): Rapid ethnography: time deepening strategies for HCI field research. *Proceedings of DIS 2000, New York, USA, August 17 to 19, 2000.* New York: ACM Press, pp. 280-286.

Mitchell, W. (2003): *ME++ The Cyborg Self and the Networked City*. Cambridge, MA: MIT Press.

Mitchell, W. (1995): *City of Bits: Space, Place and the Infobahn*. Cambridge, MA: MIT Press.

Norman, D. (1990): *The Design of Everyday Things.* New York: Basic Books.

Rheingold H. (2003): *Smart Mobs. The Next Social Revolution*. Cambridge, MA: Perseus Publishing.

Snyder C. (2003): *Paper Prototyping. The Fast and Easy Way to Design and Refine User Interfaces*. Amsterdam: Morgan Kaufmann.

Strauss, A. and Corbin, J. (1990): Basics of Qualitative Research. Newbury Park, California: Sage Publications.

Chapter 5

Personal context

Jesper Kjeldskov, Martin R. Gibbs, Frank Vetere, Steve Howard, Sonja Pedell, Karen Mecoles and Marcus Bunyan

Abstract. Intimacy is a crucial element of domestic life that has received insufficient attention from Human-Computer Interaction (HCI) researchers despite their rapidly growing interest in the design of interactive technologies for domestic use. Intimate acts differ from other activities, and there are unexplored opportunities to develop interactive technologies to support these acts. This paper presents the first phase of a two-part study exploring the potential of interactive technologies to support intimate relationships. We contribute to this uncharted domain of HCI research a literature review of concepts useful in understanding intimacy and methods for its investigation. We conclude with preliminary results and suggestive design ideas for interactive technologies intended to support intimacy.

1. INTRODUCTION

Interest in the design of interactive technologies for domestic use has been growing within the Human-Computer Interaction field (Hindus 1999). Recent research has fruitfully examined the instrumental activities involved in coordinating and scheduling family behaviour (Harper 2003), produced rich studies of the multiple meaning attached to domestic routines (Crabtree 2003), performed empirical and technological explorations of fun and leisure (Blythe et al. 2003) and developed proposals for aids to help family members stay in touch (Hofmeer 1999). Inspired by this research, we have been investigating and exploring a crucial element of domestic life that has received relatively scant attention to date from HCI researchers – intimacy. Our research is motivated by a desire to understand how intimate relationships between close family members might be supported by interactive technologies with a view to designing domestic and personal technologies for this very purpose.

Originally published as Kjeldskov J., Gibbs M., Vetere F., Howard S., Pedell S., Mecoles K. and Bunyan M. (2004) Using Cultural Probes to Explore Mediated Intimacy. *Australasian Journal of Information Systems,* 11(2), 102-115.

Intimate relationships are different from the kinds of relationships that have been typically studied by HCI researchers such as those found in the workplace or amongst social networks of friends. Intimate acts also differ from the domestic behaviours usually addressed in the literature (see Harper 2003 for a review) and attempting to study acts of intimacy presents the researcher with a number of unique and interesting challenges. Studying intimacy is challenging because intimate acts are ephemeral and transient yet ubiquitous and crucial to the ongoing life of an intimate relationship. They form the material and background of close personal relationships, yet occur in the doing and then often vanish unremarked. While the informational content of intimate acts may be low and seemingly trivial to outsiders, the act itself can be laden with emotional significance for those involved. Intimate acts often entail self-disclosure, and thus privacy is a concern. Much of what passes between intimates is unsaid and premised on deep knowledge and understanding of one another and occurs in the context of a rich, shared and sometimes idiosyncratic view of the world that may be difficult for others to fathom and comprehend. Intimacy also involves assumptions about commitment and mutuality, and carries nuanced expectations for reciprocity and exchange that are negotiated and arrived at over many years, yet remain fragile and are occasionally misjudged leading to misunderstandings and conflict. Finally, unlike instrumental tasks (e.g. coordination of family activities), or leisure activities (e.g. games) there is no generally accepted language for describing and discussing intimacy, especially in relation to designing technologies for its support.

In this paper we present our response to the interesting issues and challenges arising from our efforts to understand how intimate relationships can be supported with interactive technologies. Our study of 'mediated intimacy' has been divided into two major phases. In the first phase, we have focused on understanding current practice. To this end, we have adopted the 'cultural probes' techniques developed by Gaver et al. (1999) and the Equator team (Cheverst et al. 2003) and extended them with contextual interviews (Beyer and Holtzblatt 1998). We have used these methods to investigate how people appropriate and use the gamut of artefacts, devices and interactive technologies at their disposal to perform the various communicative acts that enable and sustain their intimate relationships. While we have adopted methods that are gaining prominence in the HCI community for investigative work in the domestic environment, how best to explore and understand intimacy remains an open and unexplored question. As a result, the first phase of our study has also had a strong methodological flavour with the aim of developing and refining techniques suitable for investigating these, and similar, phenomenon. The second phase of our study will take the insights and understandings generated in phase one and use them to design interactive technologies to support intimacy. These insights will feed a range of design activities such as the development of use scenarios, participative design workshops, prototype development and evaluation. At the time of writing, the first phase of study has been largely completed and the second phase is about to begin.

In the following section of this paper we present the results of a wide-ranging cross-disciplinary review of the intimacy literature. This review has two parts, the first being a theoretical discussion of the prominent and common themes and concepts found throughout the intimacy research literature. We also consider the use of artefacts to mediate intimacy, with a particular focus on interactive technologies. Mundane artefacts and existing interactive technologies are considered. In addition, a selection of recent prototypes and exploratory interactive technologies that attempt to mediate family relationships are reviewed. We describe in detail our research approach because of its potential for inspiring and informing (Crabtree et al. 2003) the design of interactive domestic technologies. We then present preliminary results from the first phase of our study of current practice, highlighting the key analytic themes that have emerged from our investigation of how interactive technologies are used within intimate relationships. Based on these finding, we offer three, broad-brush and indicative suggestions for the design of interactive technologies to support intimacy.

2. WHAT IS INTIMACY?

Although there are many interesting discussions and perspectives, there are no agreed-upon definitions of intimacy (Register and Henley 1992, Moss and Schwebel 1993, Robson and Robson 1998). In this section, we discuss what constitutes intimacy. We believe a clearer understanding of intimacy is important because it provides a basis for the design of artefacts that better support intimate relationships. With the purpose of understanding intimacy and its contributing factors, we conducted a literature review of conference papers, journal papers and book chapters. This section extracts some of the major themes from this review.

According to Cheal (1987), an intimate relationship consists of a private world of significant others, which needs to be continuously maintained. In intimate relationships the significant other is often reminded that "they are indeed significant". People remind each other through gestures, actions and gifts, some of which may be routine and unremarkable. Our study provides an opportunity to investigate these exchanges of intimacy. The literature provides a framework or 'lens' for conducting the observation.

While the literature provides no unanimous agreement about what constitutes intimacy, some common themes do appear. Our review offers eight prominent aspects of intimacy. These are: *physical intimacy, non-verbal communication, self-disclosure, presence, cognitive intimacy, affective intimacy, commitment* and *mutuality*.

Physical intimacy plays a central role in people's description of their own intimate behaviour (Robson and Robson 1998) and is acknowledged as an essential aspect of intimate relationships (Register and Henley 1992; Battarbee et al. 2002). Physical intimacy is the sharing of physical encounters ranging from close physical proximity to sexual contact (Moss and Schwebel 1993). Physical intimacy is not only concerned with bodily contact. Physical intimacy also includes the visceral experience of heightened awareness of ones own body or feelings of new bodily experiences (e.g. butterflies in the stomach, weak at the knees), arising from physical or mediated contact with another (Register and Henley 1992).

Non-verbal communication is identified by several researchers as a significant carrier of personal expression in intimate relationships. Non-verbal communication is communication by means of actions, gestures, facial mannerisms, close physical proximity or touch (Register and Henley 1992, Battarbee et al. 2002). Non-verbal communication is the aspect of intimacy that is better expressed through sensory evocations rather than linguistic forms, and helps to avoid the confusion that is sometimes caused by words (Register and Henley 1992).

Self-disclosure is a key characteristic that often differentiates intimate from non-intimate relationships (Register and Henley 1992, Moss and Schwebel 1993, Robson and Robson 1998). Self-disclosure is the act of revealing private information, such as the personal feelings of one person toward another. Self-disclosure includes the removal of boundary between oneself and an intimate other (physically and psychologically), getting inside the life of another, and/or allowing another to cross one's personal boundary (Register and Henley 1992). Self disclosure demands a certain degree of trust in the other and making oneself vulnerable. Furthermore, disclosing personal details often leads to an increased level of self-disclosure from an intimate other (Duck 1988). Thus, self-disclosure is an effective mechanism for maintaining and changing the level of intimacy in a relationship. Disclosing too little or too much can either escalate or de-escalate a relationship (Robson and Robson 1998).

Presence is the subjective feeling of another person being present in either a physical and/or a non-physical manner (Register and Henley 1992). The feeling of presence can be triggered by symbolic actions of the absent one(s) or the feelings can emerge spontaneously without any (objective) external cause. Thus, while being very powerful and contributing strongly to the feeling of intimacy with another, the feeling of another person being *present in absence* is very complex, subjective, and sometimes highly irrational. A feeling of presence is not so much due to being physically co-located, but due to re-living the pleasure related to being in the company of another person (Register and Henley 1992). Complementing this view, other researchers note that presence goes both ways. They stress the importance of feeling *oneself being present to another* (in either a physical or a non-physical manner) for creating and maintaining intimacy (IJsselsteijn et al. 2003, Biocca and Harms 2002).

Cognitive intimacy reflects the depth of awareness and knowledge intimates have of one another (Moss and Schwebel 1993). Cognitive intimacy is characterised by feelings of 'knowing' the other. Intimate friends, dating partners and spouses typically develop deep cognitive understandings of each other, and often share a range of personal information and preferences. This includes knowing one another's principles, values, strengths, weaknesses, hopes, fears and idiosyncrasies (Altman and Taylor 1973). Also, it has been shown that increasing the amount of cognitively exchanged information between spouses increases the level of intimacy they experienced. Thus being able to establish and maintain a shared cognitive life is a major requirement for building and sustaining an intimate relationship.

Affective intimacy is the reception and expression of emotion (Moss and Schwebel 1993). Affective intimacy involves a feeling characterized by a deep sense of love, caring, compassion and positive attraction for one another. Affective intimacy reflects the depth of awareness intimates have of one another's emotional world and the emotional exchanges they share. The level of affective closeness in friendships, serious dating relationships or marriage, is commonly reported as closely related to the level of intimacy of that relationship (Levinger and Senn 1967). Also, affective intimacy is often highlighted as a key differentiator between close friendships and relationships involving romantic love (Moss and Schwebel 1993).

Commitment is the extent to which partners in a relationship perceive their relationship as ongoing for an indefinite period (Chelune et al. 1984). Commitment includes acts intended to grow or maintain intimacy. Being in a committed relationship generates strong feelings of cohesion and connection (Moss and Schwebel 1993). Commitment is an important foundation for intimate relationships and often a precondition for other aspects of intimacy (e.g. self-disclosure) to flourish (Chelune et al. 1984). Misunderstandings in the expression of commitment and changes in one partner's belief in the commitment of the other may impede the growth of an intimate relationship or initiate its decline (Chelune et al. 1984, Duck 1981). Thus, being able to convey and experience commitment constitutes a crucial requirement for building and sustaining an intimate relationship.

Mutuality is considered the centre of any intimate relationship (Cheal 1987, Chelune et al. 1984). Mutuality is the assumption that intimate partners are co-engaged in a common cause. Mutuality originates from a process of exchange, interdependence and reciprocal expectations (Moss and Schwebel 1993). Mutuality exists when gifts or symbolic signs of value are exchanged (Cheal 1987). Mutuality is characterized by a sense of fairness shared by both partners in relation to the rewards and costs of their interactions (Chelune et al. 1984). While mutuality implies joint and shared interactions, it does not necessarily require similar or identical patterns of interaction. Rather, intimate relationships may involve both reciprocal interactions (partners showing similar behaviour) and complementary interactions (partners showing different behaviour that complements each other).

We believe these eight aspects of intimacy are useful for understanding how intimacy is constituted. However, it is important to note that these components do not exist or work independently, nor do they individually satisfy the intimate experience. Indeed, the themes overlap greatly and are highly interrelated (Moss and Schwebel 1993). The themes discussed above suggest that intimacy includes both a behavioural and an emotional level involving, on one side, *actions* caused by or causing a feeling of intimacy (such as mutuality, self-disclosure, non-verbal communication or physical intimacy) and, on the other side, *feelings* of intimacy (such as affective intimacy, cognitive intimacy, presence or commitment) caused by or causing these actions. Thus if technology is to support the mediation of an intimate relationship, it needs to facilitate these behaviours and emotions.

3. PROBING INTIMACY

This paper addresses the challenge of investigating intimacy and designing interactive technologies to support intimate acts. Our approach is a combination of ethnographic techniques and participatory design approaches that allow us to collaborative explore the phenomenon with a small group of 'participant-research' subjects. To this end we have adopted the 'cultural probes' techniques developed by Gaver et al. (1999) and the

Equator team (Cheverst et al. 2003) and extended them with contextual interviews (Beyer and Holtzblatt 1998). Cultural probes are a novel collection of techniques gaining prominence in interactive systems design. They are particularly suited to investigating people's everyday life in situations difficult to reach with traditional social science methods such as questionnaires, interviews, focus groups or participant-observation. Probes are designed to prompt and elicit information from people about their lives and 'local culture' (Gaver et al. 1999). In particular, probes are designed to garner an understanding of the playful character of human life and the multifaceted ways people 'explore, wonder, love, worship, and waste time' (Gaver 2001). Cultural probes gather insight from within the site in question, with the full cooperation and involvement of the participants concerned. The insights are gathered as activities are performed and while technology is in use, thus maintaining 'fidelity to the phenomenon' under investigation (Crabtree et al. 2003).

Like deep sea or planetary probes, cultural probes are 'sent-out' by researchers and return fragmentary data over time. Rather than relying on the presence and intervention of the researcher, cultural probes are designed to encourage and empower subjects to collect data themselves (Arnold 2004). This allows the collection of data from situations where researcher presence is problematic by giving participants the means to record everyday activities as they occur or shortly afterwards. It also allows research materials to be collected over longer periods in multiple locations compared with resource intensive methods such as traditional ethnographic approaches.

3.1. Cultural Probes

We assembled a collection of cultural probes into a 'probe pack' (figure 1) containing:

- Two scrapbooks providing the participants with an open format for creative and rich descriptions of both current practice and imagined future technologies
- Two diaries (one for each partner) allowing the participants to individually describe the temporal flow and routine of their lives throughout the period of study
- A digital camera and a photo printer allowing the participants to capture, print, edit and annotate up to 180 photographs for the scrapbooks, diaries and postcards
- Catchphrases (e.g. "I feel lonely when…" or "I really love it when…") printed onto sticky labels allowing the participants to stick them into the diaries or scrapbooks provoking reflection by the participants
- Various consumables including stamped pre-addressed postcards, coloured Post-It notes, pens, crayons, paper clips, glue and scissors etc. for use with the scrapbooks.
- Information for contacting the researchers via landline, voice-mail, e-mail and SMS throughout the study.

Probe packs were given to six couples. Participants were asked to use the probes to articulate the role technology played in their relationship. Probes were used to elicit: where, when, how and why they interacted; how they felt during interaction; and reflections on the usefulness of current technologies in these situations.

Figure 1. Contents of the probe pack

Figure 2. Scrapbook and diary from participants

The diaries were used individually to record daily communication and interaction activities. This included the form of communication (e.g. SMS) and other details such as time, date, location, the content and the feelings associated with it (e.g. urgency or dissatisfaction). The scrapbooks were used collaboratively to tell rich and evocative stories about communicative events and to express the technology wants, desires, likes and dislikes surrounding these events. Couples were encouraged to use coloured notes, pens, crayons, glue, photos, magazines clippings, drawings and so forth to form a montage of their intimate lives. The digital camera and printer were used to photograph and print significant events. The docking printer provided the immediacy of traditional Polaroid photographs with the convenience of lasting digital images. Participants were encouraged to take photographs of everyday artefacts or events that express some important dimension of their interactions: the answering machine at the time of receiving some unexpected news or of a child in a football final to share with an absent parent. Participants were asked to print, to annotate and to cut and paste the photographs into their scrapbooks or diaries or postcard, as they saw fit (figure 2). Finally, the participants were asked to read and reflect on the catchphrase labels, complete them, and stick them into the diary, scrapbook or on a post card. The stamped postcards could be used for short stories or images to be sent to the researchers or intimate other.

It was stressed that nothing in the probe pack was compulsory. The instructions were only suggestions. All the materials could be used in whatever way they wished. No time requirements were made, but it was suggested that the participants should spend about 20-30 minutes each day using the probe pack materials.

An additional probe element was introduced midway through the study. This new element consisted of small printed facsimiles of a variety of mobile device screens (e.g. mobile phones and personal digital assistants). This new element served to both refresh participants engagement in the study and encourage them to envision possibilities for future technologies to support their relationship. Participants were invited to use these mobile device screen facsimiles to note design ideas and to insert them into the scrapbooks or diaries (see figure 4).

3.2. Contextual Interviews

Our study combined cultural probes with a series of contextual interviews (figure 3). These interviews gave participants the opportunity to explain, clarify and expand upon the materials they had collected and collated through their work with the culture probes. In these interviews, the probes became a starting point for a conversation between researchers and participants that revolved around the participant's relationship, and the roles played by technologies in mediating (or not) their intimacy. The probes recorded fragments that were used during interviews to prompt memory, seek explanation, and encourage reflection. Through these interviews we were able to uncover previously unarticulated aspects of their relationship and intimate behaviours, routines and habits. Participants and researchers worked together to develop a shared understanding of how the relationship 'worked' and the roles technologies played and didn't play in mediating their engagement with each other. Interviews also allowed regular contact between researchers and participants. These meetings were used to 'tune' participant's engagement with the probe materials. Meetings were used to check on participants understanding of the study, to gently steer them toward the intended focus of the study, to ask them to focus more on certain probe activities and to introduce additional tasks. Returning from these interviews, researchers brought back a wealth of data that fuelled ongoing discussion, reflection and analysis by the research team.

Figure 3. Contextual interview with participants

Figure 4. Participant's children playing with facsimiles

3.3. Participants

The study involved six mixed-gender couples in long-term, stable relationships. All couples cohabitated, although work related travel occasionally required periods of separation. The age of participants ranged from late 20's to late 40's. Three couples had children, ranging in age from 18 months to 12 years of age. All couples used a variety of electronic media such as landline and mobile telephony, email, chat, SMS and fax to communicate with each other, although the exact mix of technologies used by each couple varied markedly.

3.4. Procedure

The couples were recruited through a screening process involving an initial, informal interview. Couples were selected to give diversity in both family situation (no children, very young children, older children); the degree to which they used ICTs to communicate with one another (once or twice a day up to 20 or more times a day); and types of technologies they used. Important in the selection of participants was the degree to which participants were willing and able to articulate, discuss and reflect on their relationship as well as their ability to engage in imaginative speculation about future ICTs. Two researchers were assigned to each couple and were responsible for introducing the cultural probe pack and conducting all contextual interviews with the couple. Interviews were predominantly conducted in the participants' homes, but some interviews were conducted at the University and several were conducted in a café. Choice of interview location was at the discretion of participants.

The cultural probes were deployed for a period of seven weeks. At the beginning of this period, an initial interview was carried out at the participant's homes. This included questions about the participants' backgrounds, their relationship, their communication habits and their use of technology. Following the interview, the researchers presented the cultural probe materials and instructed the participants on how to use them. After the first week of the study, the researchers visited the participants' homes again for a second interview. The purpose of this visit was to ensure the participants were 'on track' with the use of the probes and to investigate the activities of the first week through a conversation about the data collected in the diaries and scrapbooks.

Following the second interview, participants were left to work with the probe pack for three weeks, at which time, a third interview was conducted. The purpose of the third interview was to examine and review the materials collected through the probes since the last interview. Researchers and participants discussed the materials accumulated in the scrapbooks and diaries, and participants were invited and encouraged to clarify, elaborate and reflect on the materials they had recorded and composed over the previous three weeks. At the end of the third interview session, researchers introduced the small printed facsimiles of mobile device screens. Participants annotated these facsimiles to explore novel designs in situations where current technology inadequately supports personal communication and interaction (figure 4). After the third interview, the participants were again left to work with their probe materials for three weeks. In the fourth and final interview, researchers discussed the diaries, scrapbook and other materials composed over the previous three weeks. In addition, researchers and participants discussed design ideas produced on the mobile device screen facsimiles. This interview was also used to bring closure to the seven-week process. The researchers retrieved all materials gathered through the cultural probes at this time. It is our intention to return the probe pack diaries and scrapbooks to participants as a 'gift' after our analysis of the material is completed (copies will be retained for our records).

4. RESULTS

The contents of the scrapbooks, diaries and interviews were reviewed and discussed on a weekly basis by the research team. On the basis of our analysis of material collected in the first four weeks of the study, we present some preliminary findings.

A theme occurring across all couples was a strong need to support *presence in absence.* This is a feeling that the other is present, even though they may be far away. This is closely related to the issue of *presence* discussed earlier. Participants identified a desire to be in contact with each other while physically separated. This does not only happen when one partner is away for a long period of time, it also occurs during ordinary workdays, sometimes with only minutes separation. For example, couples who work in office buildings a few blocks apart or couples who work on different floors of the same building are in regular contact, up to 10 times a day, via phone, SMS, or email. Often the messages play the dual role of organising family affairs (e.g. "Who is picking up the children?" or "What is planned for dinner tonight?") and for declaring a caring presence ("I am here, thinking of you"). Other examples of creating presence-in-absence are through indirect, non-verbal communication "by proxy", for example, one partner preparing breakfast for the other before leaving for work. Presence is not only created by exchanges of message. Artefacts themselves carry presence. Merely carrying a mobile phone, because it affords the opportunity for immediate contact if desired, creates a sense of presence and feelings of security and comfort. Furthermore, absence need not be physical. In some instances, couples who were physically nearby but engaged in separate activities supported presence-in-absence by communicating with each other using SMS from the lounge room to the bed room or email from one side of the room to the other. In these examples, physical presence did not compensate for emotional absence. The couples felt a desire to be more present than afforded by physical closeness. This desire was mediated by technology.

Our study suggests that much of the communication that passes between intimates is *emotional rather than factual.* While intimate communication is not 'fact free' it often plays a role more akin to 'stroking and patting' than verbal conversation. The messages are often information poor, but laden with significance. Examples include spontaneous gift giving, messages of affection, flirtation, telephone calls to chat about nothing in particular. Many of these interactions serve to confirm that a person is thinking about, caring for and aware of their partner. It is often the act itself, rather than the explicit message it carries, that is significant.

Intimate acts are often *ambiguous and incomplete*, suggesting and hinting rather than explicating in detail. They occur in the context of a rich, shared and sometimes idiosyncratic view of the world. This *shared world view* is vital to feelings of intimacy and is also a resource artfully drawn on by intimates during interaction.

It comes as no surprise that our study indicates that intimate acts are often *private*. Our results suggest privacy has many dimensions. For example, intimacy entails *self-disclosure*. Intimates reveal something of themselves to each other and they may feel vulnerable if their interactions are exposed to the gaze of outsiders. As a result, intimate acts are often constructed to be hidden from the view of others. Participants in our study were selective in the communication media they choose for interaction, and partly based these decisions on their perception of the privacy and security of the various communication channels at their disposal. For example, some did not use work-based email for certain intimate exchanges because they were aware that system administrators can view their emails. Similarly, intimate information was not disclosed while on a crowded but quiet train car or over the telephone in an open plan office. Some participants, particularly those who regularly used workplace technologies to communicate with each other, reported taking efforts to obscure that they were engaged in communication with their partner.

Even when intimacy is displayed in public, these acts can be nuanced and imbued with *private meaning* difficult to see and interpret by outsiders. Some participants developed private 'codes' and short-hands to communicate with one another, such as an SMS of '146' for 'I love you' or calling the home telephone and allowing it to ring thrice at certain times of the day to signify 'I'm awake, I'm OK, I'm thinking of you'. Others engaged in flirting in situations were flirting was inappropriate – such as knowingly sending provocative SMS messages when their partner was in a work meeting – relying on the personal form of these messages to obscure their lack of proprietary. Some participants reported drawing on their *shared past* and detailed knowledge of one another to privately communicate in the full view of others, using oblique references that were presumably ambiguous or meaningless to others, such as when one end of a telephone conversation could be overheard. More generally, participants often drew on a repository of anecdotes, past conversations, knowledge of events and running jokes in their interactions. Indeed, much of what passes between intimates is unsaid and premised on deep knowledge and understanding of one another.

Actively constructing and maintaining a *shared history* was also important to participants. All participants collected mementos, photographs and other materials that evoked their past together. Some participants invested significant effort in constructing and maintaining these records in scrapbooks and photo albums, and often collaborated in reviewing and updating these repositories. These mementos were at once a public declaration of ongoing commitment and artefacts through which past intimate feeling could be rekindled. With the advent of digital photography, some participants had begun keeping digital archives of photographs burnt to CD-ROM for viewing and reviewing on the home DVD player, often as a joint activity.

Participants were also engaged in a *common journey* and shared the costs and rewards of lives together. Where a shared history brought comfort from the past, a common journey generates hope for the future. Both are related to the theme of *commitment*. Features of the participants' lives, such as raising children, caring for older family members, maintaining the household, or travelling to work, were all enterprises that were shared and, in the sharing, became vehicles for enacting, affirming and maintaining their relationship. Joint responsibility was taken for domestic life, including activities such as paying the bills, transporting children and preparing meals. These activities were often coordinated on the fly using interactive technology such as mobile phones. The division of labour within these relationships had regular patterns but was also fluid and renegotiated over the course of a day. In addition, responsibilities also provided occasions for affirming the relationship. For example, the success of a child at school was affirming "our" achievement as "good parents". Even activities that were the sole responsibility of one member of the relationship became joint enterprises. Participants reported drawing on their partner's help and skills for work related tasks such as database development, setting up a web-based email account or writing a job application. Being able to help and share common tasks affirmed the relationship and sense of moving through the world together as a team, rather than as atomised individuals.

Above all, our participants reported the need to find time to be *alone together* away from the hectic schedules of family and work related responsibilities that dominated their lives. Whether this time alone was found while driving a car with the children asleep, sharing a meal at the end of the day, or even working on separate projects but sharing the same physical space to do so, all felt that *physical closeness* was crucial to their relationship, and something that could not be adequately replaced or mediated by interactive technology.

All our intimate relationships involve expectations and assumptions about *reciprocity* and exchange between partners. In our study, we found interaction was founded on a *commitment* from both parties to reciprocate in both content and form. These expectations were nuanced, negotiated and arrived at over many years, yet remain fragile and are occasionally misjudged and misinterpreted leading to misunderstandings and conflict.

Our study has also taught us that it is important not to romanticise intimacy. Intimacy and the tight, emotionally charged bonds it entails, is fragile. While relationships may be robust, misunderstanding, and misinterpretations do occur. When they occur, these breakdowns can have serious repercussions; creating ill will and emotional hurt that can obstruct and undermine intimacy within the relationship. The fallout from a simple breakdown in understanding between partners can reverberate through the relationship for days, if not weeks. For these reasons, any interactive technology designed to support or mediate intimacy needs to mitigate against these forms of breakdown and allow for easy and rapid repair of them when they do occur.

Even though our results are still preliminary, they have given us a deeper understanding of how some of the themes identified in the literature contribute to intimacy, and how people use technologies to support and sustain them. Our challenge is to translate our observations and extended knowledge of the central themes to designs.

5. DESIGN IDEAS

We now present three preliminary design ideas that have been distilled from the data. The designs are intended to exploit opportunities for technologically mediating intimacy. The designs are not necessarily futuristic. Similar designs may already exist. It is not our intention to implement these design ideas in their current forms. Rather we intend to use the design ideas to seed ongoing design activities such as participant design workshops and scenario-based acting out (Howard et al. 2002).

Memorabilia Manager: Couples spend a great deal of time and energy organising their personal mementos. These are typically photographs and videos, but often include other items such as tickets (e.g. airline ticket from a honeymoon) and souvenirs (a sea shell from that special holiday). These have important value for intimate relationships. They signify common history and a shared journey and contribute to the broad theme of commitment. A 'Memorabilia Manager' should be very portable and allow the experience of creating and reviewing the memorabilia. It would allow couples to compose digital and non-digital forms into a meaningful mosaic. It would be a type of 'family blog' that facilitates simple recording of family events.

Constant Touch: Face-to-face communication was regarded by all participants as an authentic experience, while mediated interaction as somewhat impoverished, thin and 'abstracted'. However, when face-to-face interaction was not possible (or practical) intimates desired connectedness and presence. 'Constant Touch' is akin to walkie-talkie or 'push-to-talk' devices (Telstra 2004). It provides an open channel for constant updating throughout the day. It may contain a single point transducer (such as a light) that is activated when one wants to say "I'm thinking of you". The form is configurable to suggest a physical presence as if "She is with me all the time".

Family Digital Assistant: If a family had children, then the children tended to dominate all activities – including intimate ones. Children were an anchor to the family's life and a key mediator for the expression of intimacy. Rather than finding intimacy beyond family duties, it was through the routine of child-raising that intimacy was expressed. The 'Family Digital Assistant' (FDA) acknowledges that shared mundane experiences are part of the intimate experience. Where a PDA is for personal purposes, a FDA coordinates family activities. We observed that poor coordination can lead to emotional hurt and bad tempers. The FDA is a response to this need.

6. CONCLUSIONS

Intimate relationships involve an intricate and nuanced dance between partners involving a complex array of varied activities. Many of these take place in close physical proximity and involve touch and face-to-face communication. Others are mediated across space and time and involve activities such as talking on the phone or writing letters and postcards. While mediating intimacy is not a new phenomenon, new technologies influence how, when and why we interact with each other. However, little is known about the adoption and use of these technologies within intimate relationships and even less is known about how, when and why they provide good support for mediating intimacy and how, when and why they fail.

Perhaps more than the majority of domestic acts, intimacies display fragility when under examination. In this paper we have described an approach to the study of intimacy that stressed the need to empower participants; providing them with means to record intimacies during or soon after the acts themselves and means to describe intimacy in their own words. We have stressed the importance of hearing the 'voice of the intimates' given our current lack of understanding of this new domain of HCI. Although we have largely limited our report of our preliminary findings to the analysis and description of current practice, our approach has also aimed to be playful; we have worked to encourage and legitimise participants' exploration of possible futures, rather than strictly limiting them to reporting on their current practice. The insights generated through this playful imagining will be the topic of subsequent work.

ACKNOWLEDGMENTS

The authors would like to thank all couples participating in the study and acknowledge the support of the Smart Internet CRC and the Danish Technical Research Council (project reference 26-03-0341).

REFERENCES

Altman, I. and Taylor, D.A. (1973) Social Penetration: The development of interpersonal relationships. Holt, New York.

Arnold, M. (forthcoming 2004) The Connected Home: Probing the Effects and Affects of Domesticated ICTs, Proceedings of PDC2004, Toronto, July 27–31 2004, ACM.

ASTRA Project (2004) http://www.presence-research.org/Astra/ (Accessed May 7th 2004)

Battarbee, K., Baerten, N. and Loeber, S. (2002) Pools and Satellites - Intimacy in the City. Proceedings of DIS 2002, ACM, 237-245.

Beyer, H. and Holtzblatt, K. (1998) Contextual Design: Defining Customer-Centered Systems, Morgan Kaufmann, San Francisco.

Brave, S. and Dahley, A. (1997) inTouch: a medium for haptic interpersonal communication. Extended Abstracts from CHI'97, Atlanta, GA, USA, ACM.

Byrne, R. and Findlay, B. (2004) Preference for SMS versus Telephone Calls in Initiating Romantic Relationships. Australian Journal of Emerging Technologies and Society (Issue 2: Autumn 2004)

Biocca, F. and Harms, C. (2002) Defining and Measuring Social Presence: Contribution to the Networked Minds Theory and Measure. Proceedings of Presence 2002, ACM.

Chang, A., O'Modhrain, S., Jacob, R., Gunther, E. and Ishii, H. (2002) ComTouch: Design of a Vibrotactile Communication Device. Proceedings of DIS 2002, London, England, ACM.

Cheal, D. (1987) 'Showing them you love them': Gift giving and the dialectic of intimacy. The sociological Review, 35(1):150-169.

Chelune, G.J., Robison, J.T. and Kommor, M.J. (1984) A Cognitive Interactional Model of Intimate Relationships. In Derlega, V.J. (ed.) (1984) Communication, Intimacy, and Close Relationships. Academic Press, Inc., New York.

Cheverst, K., Clarke, K., Dewsbury, G., Hemmings, T., Kember, S., Rodden, T. and Rouncefield, M. (2003) Designing Assistive Technologies for Medication Regimes in Care Settings. International Journal of Universal Access in the Information Society (UAIS), 2003.

Crabtree, A., Hemmings, T., Rodden, T., Cheverst, K., Clarke, K., Dewsbury, G., Hughes, J. and Rouncefield, M. (2003) Designing with Care: Adapting Cultural Probes to Inform Design in Sensitive Settings. Proceedings of OzCHI 2003, Brisbane, Australia.

Crabtree, A. (2003) Remarks on the Social Organisation of Space and Place. Homo Oeconomicus, XIX(4), 591-605.

Donn, J.E. and Sherman, R.C. (2002) Attitudes and Practices Regarding the Formation of Romantic Relationships on the Internet. CyberPsychology & Behavior, 5(2), 107-123.

Duck, S. (1988) Relating to Others. Open University Press, London.

Gaver, B. (2002) Provocative Awareness. Computer Supported Cooperative Work, 11: 475-493.

Gaver, B. (2001) Designing for Ludic aspects of Everyday Life, ERCIM News, 47. http://www.ercim.org/publication/Ercim_News/enw47/gaver.html, Accessed 16 June 2004.

Gaver, B., Dunne, T. and Pacenti, E. (1999) Design: Cultural Probes. Interactions, 6(1), 21-29.

Goodman, E. and Misilim, M. (2003) The Sensing Bed. Proceedings of UbiComp 2003, Seattle Washington.

Grinter, R.E. and Eldrige, M.A. (2001) Y Do Tngrs Luv 2 Txt Msg? Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work, ECSCW '01, Bonn, Germany.

Harper, R. (ed.) (2003) Inside the Smart Home. Springer-Verlag, London.

Hindus, D., Mainwaring, S., Leduc, N., Hagstrom, A., and Bayley, O. (2001) Casablanca: Designing Social Communication Devices for the Home. Proceedings of CHI 2001, ACM.

Hindus, D. (1999) The Importance of Homes in Technology Research. Proceedings of the Second International Workshop on Cooperative Buildings (CoBuild'99), Pittsburgh USA, Springer-Verlag.

Hofmeer, K. (1999) The Digital Hug: Keeping Families Together. Interactions, 6(6).

Howard, S., Carroll Jennie, Vetere, F., Murphy, J. and Peck, J. (2002) Provoking Innovation: Acting out with contextual scenarios. Proceedings of BCS-HCI 2002, Springer-Verlag, London.

IJsselsteijn, W., van Baren, J. and van Lanen, F. (2003) Staying in Touch – Social Presence and Connectedness through Synchronous and Asynchronous Communication Media. Proceedings of the 10th International Conference on Human-Computer Interaction 2003, Crete, Greece, Lawrence Erlbaum Associates.

Levinger, G. and Senn, D.J. (1967) Disclosure of feelings in marriage. Merill-Palmer Quarterly, 13, 237-249.

Blythe, M.A., Overbeeke, C.J., Monk, A.F., and Wright, P.C. (Eds) (2003) Funology: From Usability to Enjoyment. Kluwer, Dordrecht.

Moss, B.F. and Schwebel, A.I (1993) Marriage and Romantic Relationships: Defining Intimacy in Romantic Relationships. Family Relations, 42(1): 31-37.

Register, L. and Henley, T. (192) The Phenomenology of Intimacy. Journal of Social and Personal Relationships, 9, 467-481.

Robson, D. and Robson, M. (1998) Intimacy and Computer Communication. British Journal of Guidance and Counselling, 26(1):33-42.

Taylor, A., and Harper, R. (2003) The Gift of the Gab?: A Design Oriented Sociology of Young People's Use of Mobiles. Computer Supported Cooperative Work, 12(3): 267-296.

Telstra (2004) What's Doing: Telstra completes Push To Talk trial with services to be available in June http://telstra.com/countrywide/Peninsula/localtelstra.asp?page= whatsdoing, Accessed 15 June 2004.

Tollmar, K., and Joakim, P. (2002) Understanding Remote Presence. Proceedings of NordiCHI 2002, Aarhus, Denmark, ACM.

van der Hoog, W., Keller, I., and Stappers, P.J. (2004) Gustbowl: Technology Supporting Affective Communication through Routine Ritual Interactions. Proceedings of CHI 2004, Vienna, Austria, ACM.

Chapter 6

Work context

Jesper Kjeldskov and Jan Stage

Abstract. Communication between physically distributed people in industrial and safety-critical domains is often spoken and mediated through walkie-talkies, or closed-circuit intercoms. Because this kind of communication is hampered by noise, radio interference, lack of persistency etc., vital information is sometimes lost. In response to this challenge, this paper discusses the use of "canned" text-based messaging as a supplement for improving such communication. Based on data from ethnographic studies of work activities in an industrial domain, and grounded in a theoretical model of communication, we have designed and evaluated a mobile canned communication prototype system facilitating exchange of predefined text messages, a persistent graphical representation of the operation in progress, and a filtered list of completed tasks. Results from two evaluations show that in the domain considered, canned text-based communication has a potential to supplement voice and assist in overcoming some of the inherent problems of spoken communication. Yet using a textual and persistent mode of communication also raises new challenges such as choice of modality, speed, flexibility and handling situations deviating from standard procedures.

1. INTRODUCTION

Industrial domains are potentially interesting areas of applications for mobile information and communication technologies. Work activities in industrial domains often involve a number of distributed collaborating actors who are mutually dependent on access to computer systems remote from their current location and on knowledge about the activities and strategies of their co-workers. Typically this is supported only by spoken communication mediated through mobile phones, walkie-talkies or closed-circuit intercoms. As spoken communication is highly sensitive to noise, radio interference, interruptions, lack of persistency etc., vital information may be lost in the transmission or missed by the receiver(s).

Originally published as Kjeldskov, J. and Stage, J. (2006) Exploring "Canned Communication" for Coordinating Distributed Mobile Work Activities. *Interacting with Computers,* 2006(18), 1310-1335.

Previous research has to some degree dealt with the extent to which distributed mobile users in industrial and safety-critical domains can benefit from handheld computer systems in situations where actors are concerned with computerized information and processes of critical importance remote from their current location. Examples include the use of mobile multimedia communication for telemedicine and early diagnosing in emergency ambulance services (van den Anker and Lichtveld 2000), distributed process control and error diagnosing in wastewater treatment plants (Nielsen and Søndergaard 2000), and the use of remotely controlled service robots for aiding disabled or elderly people (Hüttenrauch and Norman 2001). Also related to this, the limitations of voice-based communication by capturing spoken utterances and integrating it with other data for creating persistent graphical representations have been suggested within areas such as air traffic control (Fields and Paterno 1999), and fire-fighting (Champ et al. 2000).

In response to this growing area of interest within HCI, this paper explores the supplementary use of text-based "canned communication" in a prototype system for coordinating work activities on large container ships: the Handheld Maritime Communicator (see figure 1). The idea for the Handheld Maritime Communicator emerged from a multidisciplinary research project involving an ethnographic field study on work activities in the maritime domain involving computerized process control and information systems (Andersen 2000, Nielsen 2000). On the basis of this field study, we explored the usefulness of handheld computers for supporting communication by complementing existing spoken communication with the use of predefined (or "canned") text-based messages similarly to the way SMS and e-mail applications on mobile devices complement people's voice-based communication.

Figure 1. Canned communication on a mobile device for use on maritime bridges

In the study reported in this paper, we describe and discuss the lessons learned from a first step in the direction of "canned" communication aids designed for overcoming some of the limitations of spoken communication in industrial domains by supplementing it with a textual and persistent channel. Hence the aim of the study presented in this paper has not been to develop a final solution to a well-defined problem and deploy this solution in the domain studied, but to gain experience with and a deeper understanding of the use of canned textual communication as a supplement to spoken commands. This is done through experimental design and evaluation of a prototype system – using the prototype system as a sort of "technology probe" (Hutchinson et al. 2003) to prompt and study new communicational behaviour. This aim has influenced our research in several ways. Firstly, we did not try to change the structure and content of what was being communicated but intentionally replicated this directly in the prototype system focusing solely on the changed modality of these utterances (from audio to text) and the use of pre-defined commands (communication canning). Secondly, we did not pursue research into issues such as the physical form factor of the mobile communication device facilitating use in a potentially harsh outdoor environment, and the technical implementation of a network infrastructure working robustly within a physical environment dominated by large amounts of metal. We acknowledge that these (and other) issues are highly relevant for the development of new communication device solutions for the industrial domain, and welcome further research into these specific areas complementing our own endeavours within the use of canned textual communication.

In section 2, we introduce the industrial domain studied and the specific work activities supported by our prototype system. This includes highlighting findings from the field studies related to limitations in current means for communication and coordination. Section 3 presents our analysis of the field data. In section 4, we present the details of the design of the Handheld Maritime Communicator prototype, and in section 5 we present two evaluations involving usability experts and prospective users. The findings from these evaluations are discussed in section 6. Finally, section 7 concludes on our study and point out avenues for further research.

2. FIELD STUDY OF WORK ACTIVITIES

Maersk Line operates some of the world's largest container vessels of sizes equivalent to the length of five Boeing 747-400 Jumbo-Jets (figure 2). The operation of such vessels requires workers to be highly mobile and physically distributed. At the same time, however, work activities are often related to the use of computer systems located centrally on the ship. Thus a strong motivation exists for exploring the use of mobile computers for supporting distributed work activities in this domain. Designing useable mobile computer systems for the maritime domain is not trivial. Work activities on large container vessels are typically safety-critical and involve high risks in the case of errors. Especially when manoeuvring inside a harbour, erroneous actions may result in the vessel running aground, into the quay, or colliding with other ships. In either case, this would cause serious material damage, potentially severe injuries on personnel, and possible loss of human life.

Figure 2. Sine Maersk (347 m. long and 43 m. wide) in Gothenburg Container Terminal, Sweden

Qualitative investigations into work activities on a Maersk Line container vessel were carried out (Andersen 2000, Nielsen 2000). This included ethnographic observations of the application domain and interviews in situ from several voyages along the coastline of Europe. The field studies were documented through written notes and video recordings capturing overall views of the captain, harbour pilot and officers on the bridge as well as close-up views of the interaction with key instruments. The audio channel captured interpersonal communication on the bridge and VHF radio communication with the distributed crewmembers. In order to facilitate systematic analysis, a person with detailed insight into the application domain transcribed a selection of the video recordings. A partial transcription of the recordings from one short voyage within Europe amounts to approximately 200 pages.

Apart from informing new interface design for existing maritime instruments (Andersen and May 2001) a number of work activities were identified in which the use of mobile computer terminals could be desirable. These included diagnostic and maintenance work in the engine room, surveying the condition of reefers during voyages, locating personnel in case of accidents, and supporting various distributed collaborative work activities. Of particular interest to the interviewed crewmembers, our attention was brought to the processes of departing from and arriving at harbour including the operation of *letting go the mooring lines* before leaving the quay because this operation requires a high level of communication within a predefined pattern between actors that are physically distributed on the vessel. Currently, this communication is based on spoken commands being transmitted through handheld VHF radios. Through analysis of several video recordings and interviews with officers and captains, a series of limitations in the present means for communication and coordination were brought to our attention, some of which could potentially be overcome by the use of mobile computer technology. Project stakeholders from Maersk Line and the participating university researchers therefore agreed that supporting this particular operation would be a suitable starting point for experimenting with the use of canned text-based communication.

In the following sub-sections, the operation of letting go the lines as experienced through the field studies is described in detail. This description served as an overall context for the use of the envisioned Handheld Maritime Communicator and outlined a number of challenges, which had to be addressed in the design of the prototype.

2.1. The operation of letting go the lines

When a container vessel is ready for departure, the first step in leaving the quay is to let go of the mooring lines that are holding the ship in position fore and aft (figure 3). However, as physical space is restricted and means for precisely manoeuvring large vessels are limited, all lines cannot simply be released simultaneously.

When a line is let go, it will remain in the water for a period of time during which no means of propulsion is available due to the risk of lines getting sucked in and wrapped around a propeller or thruster. During this time, the vessel can only be manoeuvred by means of the remaining lines. Consequently, lines are released sequentially in accordance to specific need for manoeuvring in a given situation.

Figure 3. The aft mooring lines of Sally Maersk (sister ship of Sine Maersk)

Due to the huge size of the vessel, the work tasks involved when letting go the lines are distributed among a number of mobile actors located at strategic positions, as annotated on figure 2. On the bridge (1), the captain and other personnel control the rudder, propeller and thrusters. Fore (2) and aft (3), the first and second officers control the winches for heaving in the lines. Ashore, two teams of assistants lift the lines off the bollards. To insure the safety of the operation, individual work tasks are carefully coordinated and carried out under strict command of the captain in charge.

At present, communication between co-workers is spoken. While people on the bridge can see and hear each other directly, personnel on deck are, however, out of direct visual and audio contact and have to communicate with the captain via walkie-talkies. In order to carry out the operation of departure in a safe manner, the captain needs an overview and total control over the propulsion, direction and mooring of the ship. While information about the rudder, propeller and thrusters are available on dedicated instruments on the bridge no information about mooring is available. At present this information only exists as a mental model in the head of the captain based on his perception of the ongoing communication between bridge and deck. As this mental model is highly sensitive to errors or misunderstandings in the ongoing communication between bridge and deck, and since disparity between the captain's mental model and the real world may cause wrong decisions to be made, considerable cognitive resources are spent on establishing and maintaining *common ground* (Clark and Schaefer 1989) among the cooperating actors. By common ground, we refer to the principle of entering and maintaining a stage of mutual knowledge, beliefs, and assumptions among communicating participants through a collaborative process of *grounding,* during which common ground is updated in an orderly way, by each participant trying to establish that the others have understood their utterances well enough for the current purposes (McCarthy et al. 1991). While it has been pointed out that it is not necessary to fully ground *all* aspects of a conversation, it is essential that *"The contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose"* (Clark and Schaefer 1989:262). What constitutes this criterion, of course, depends on the context of the situation and will necessarily vary with the collaborators' goals. As a part of the process of grounding, communicating partners have different means of providing evidence of mutual understanding including displaying what has been understood, acknowledging utterances, continuing with the next expected step of a given process, as well as continued attention to the conversation.

Supporting reaching and maintaining common ground, established rules and formalized procedures exist for oral communication such as, for example, confirming status reports and commands by repeating them back to their sender. However, as the size of vessels and the use of technology increases so does the complexity of systems controlling the ship and the cognitive overhead and amount of parallel tracks of communication required for its operation.

2.2. Findings from field studies of "letting go the lines"

Through analysis of video recordings, transcriptions, and interviews with officers and captains, a number of key limitations experienced in the use of spoken communication for coordinating collaborative work activities were brought to our attention:

- Sound quality is often poor
- Communication is not persistent and
    - cannot be automated,
    - is time consuming,
    - suffers from bottlenecks on the bridge (multiple parallel tracks),
    - suffers from language barriers,
    - lacks integration with other systems.

As walkie-talkies and VHF-radios often lack sound quality, workers reported that misperceptions and misunderstandings between the actors often occur due to incomprehensible messages. This leads to a need for repeating statements and meta-communicating. Due to the ephemeral nature of spoken communication, the workers also reported that messages were easily and frequently missed because they were only available during the limited period of time when they were "in the air" and were not persistent. After an utterance had been communicated the information only existed in the memory of the actors taking part in the interaction and was not publicly available for others who had not received the utterance when first stated. In addition to this, workers reported that spoken coordination could not be automated but involved actors remembering sometimes highly complex workflow and continuously deciding for whom specific information may be relevant at which time. Workflows coordinated through spoken communication are hard to support and reducing the coordination workload was reportedly difficult. Workers also stated that spoken communication was very time consuming and that they tried to minimize time spent on communicating in order to maximize the time available for work tasks. As a part of this, workers reported that they would sometimes cut messages short and only communicate fragments of information and rely on implicit meaning in the given situation. While known for limiting the "air time" of communication this approach was also known for sometimes leaving people confused about the meaning of utterances requiring explanations and meta-communication. The use of spoken coordination was also reported to suffer from different types of bottlenecks. One type of bottleneck reported by the workers consisted of multiple people talking on top of each other on the same channel resulting in communication being cut up, and information being missed. Another, and more complex, bottleneck reported by the workers consisted of multiple parallel tracks of communication across different communication channels (e.g. radio messages between bridge and deck disturbing communication between people on the bridge and visa versa) complicating the regulation of turn taking. Due to the international nature of the domain, communication on board container vessels is usually conducted in a language different from the language being used by the local harbour pilot to communicate with other pilots, the pilot boat, tugboats, vessel traffic service etc. This results in the captain having limited immediate insight into the domain of the harbour pilot and vice versa and introduces a need for ongoing translations between the captain and the harbour pilot. Finally, workers raised the issue that information delivered through spoken communication cannot be integrated with the vast amount of other information sources in the ship's computerized systems. While the captain can, of course, take spoken information about, for example, distances, angles etc. to objects in the vessels immediate surroundings into consideration when looking at other instruments, this kind of information cannot automatically be made part of the computations regarding the ship's movements performed by the systems on the bridge. As a result, it was reported that the spoken information is usually not utilized to its full extent because it demands too many cognitive resources.

While some of these observed limitations may be unique for the studied context (e.g. language barriers), others apply generally to spoken communication within industrial domains. Overcoming or reducing these limitations served as an overall motivation for the experimental design of the Handheld Maritime Communicator.

Inspired by, amongst others, chat applications, newsgroups and Short Messaging Service (SMS) we speculated that a possible supplement to the present use of spoken communication could be the use of predefined, canned text messages on mobile devices. Hence, text offers some advantages over voice, it is a flexible communication channel requiring low cognitive overhead (Churchill and Bly 1999; Popolov et al. 2000), and it is not subject to the ephemeral nature of spoken utterances but is persistent. Furthermore, text-based communication can be done asynchronously as it fits in with other tasks or threads of communication and is not influenced by, for example, noise. On the basis of this, it was our expectation that some of the identified limitations could be eliminated or reduced by means of exchanging canned text-based messages and as a result more cognitive resources would be available for the other operations.

3. ANALYSIS OF COMMUNICATION

Motivated by the initial findings from the field study described above, we revisited the video recordings to investigate more thoroughly the communication and coordination of work activities related to the operation of letting go the lines and identifying structures and properties, which could help us overcome the identified limitations. Guiding this analysis, we focused particularly on the overall challenge of achieving persistency in communication.

Achieving persistency in communication means capturing the utterances of a conversation for later access (Erickson and Herring 2006). While audio or video recordings of spoken conversations can preserve a very rich picture, textual transcriptions capture the essence, are highly concentrated, and facilitate fast browsing. On the downside, raw transcriptions offer little support for linking related utterances and maintaining overview of present state or outcome of parallel tracks of conversations. This is similar to textual communication in chat-like applications where achieving common ground can be problematic as discussed in McCarthy et al. (1991). On mobile devices the usefulness of raw textual transcriptions is further limited by small screen sizes. By relatively simple means of formalization, however, some of these problems may be solved. Based on the analysis of our video recordings and guided by literature on the topic, we found that at least three properties of conversations exist, which may be used for improving the representation of textual communication on a mobile device: 1) the aspect and tense, 2) the object and 3) the structure of conversations.

3.1. Aspect and tense of conversations

On an overall level, a conversation can be categorized by aspect and tense (Andersen, 2000), hence, a conversation is either imminent (future tense) executing (present tense) or ended (past tense). While executing (present) conversations are still open for negotiation, ended conversations imply some kind of mutual agreement having been made among the communicating parties. Though the process by which this agreement was reached may be of interest, the essential properties of ended conversations are typically their outcome. Imminent (future) conversations are characterized by potentially being initiated when and if appropriate in relation to preceding conversations (ended and executing). In relation to interface design for persistent communication, this categorization enables us to separate different conversations and differentiate priority. In some situations, ended conversations may be important, while in others only executing tracks are of interest.

3.2. Objects of conversations

Communication consisting of a number of interweaved tracks of conversations can be difficult to overview when sorted from the sequence of utterances. This can be illustrated with the following transcription extract of three conversational tracks taking place in parallel:

```
1     <Captain>              you can let go the bow line
2     <1st officer>          let go bow line
3     <Captain>              and you can take the stern spring
4     <2nd officer>          letting go stern spring
5     <1st officer>          bow line let go
6     <Captain>              bow line let go
7     <2nd officer>          and stern spring let go
8     <Captain>              stern spring let go
9     <Captain>              you just let go the stern line also
10    <2nd officer>          let go line aft
11    <1st officer>          and we have the bow line home
12    <Captain>              ok
13    <2nd officer>          and all let go aft
14    <Captain>              all let go aft
```

Sorting these utterances by the objects of communication rather than their sequence, the following structure appears:

```
1     <Captain>              you can let go the bow line

2     <1st officer>          let go bow line
5     <1st officer>          bow line let go
6     <Captain>              bow line let go
```

```
11    <1st officer>       and we have the bow line home
12    <Captain>           ok

3     <Captain>           and you can take the stern spring
4     <2nd officer>       letting go stern spring
7     <2nd officer>       and stern spring let go
8     <Captain>           stern spring let go

9     <Captain>           you just let go the line aft also
10    <2nd officer>       let go line aft
13    <2nd officer>       and all let go aft
14    <Captain>           all let go aft
```

Grouping text in accordance to object rather than sequence thus enables the creation of a more comprehensible representation of communication threads as seen in e.g. email and newsgroups (Popolov et al., 2000). Designing for the limited space of a mobile device interface this principle is valuable, as it requires little or no extra space compared to the raw transcription. For a richer representation of sequence, absolute timestamps or timers may be needed.

3.3. Structure of conversations

A number of computer systems for communication have been designed on the basis of speech-act theory (see e.g. Winograd and Flores 1986, Frisse 1988, Alm et al. 1992; De Michelis and Grasso 1994, Jayaweera et al. 2001, Akhus 2001). The basic idea of these systems is that conversations follow an overall structure of recurrence. Formalizing and modelling this structure in a computer system, the state of a conversation and possible speech-acts at a given time can be identified. According to Winograd and Flores (1986:65), the basic course of a conversation for action can be described in a diagram with nine different states (figure 4).

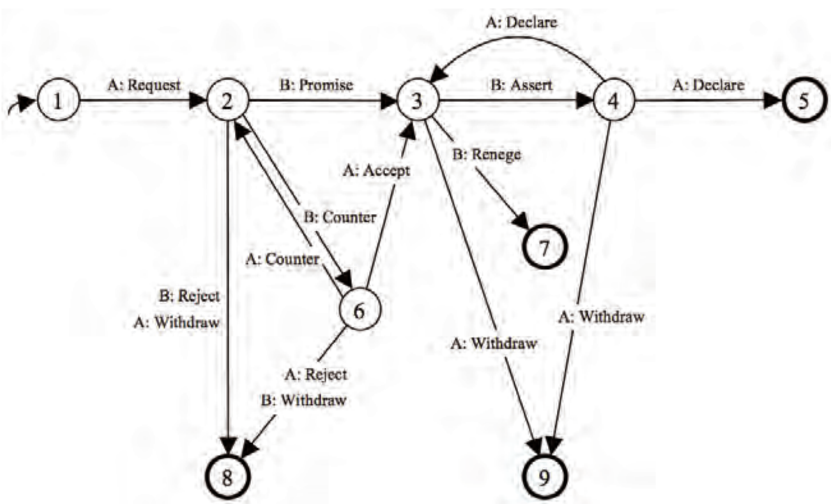


Figure 4. Winograd and Flores' conversation for action (1986)

The conversation for action model describes a generic pattern of communication, where one actor (A) makes a request to another actor (B). The model then describes how the conversation between A and B can develop over time through the performance of speech acts, resulting in a number of different intermediate states and end situations. As emphasized by Winograd and Flores (1986), the relevant regularities proposed by this model are not in the individual speech acts (exactly what is being said and how it is being said) but rather on the overall level of the conversation, in which successive speech acts are related to each other. For more detailed discussion of the speech-act theory approach to conversation modelling, see, for example, De Michelis and Grasso (1994), Winograd (1994), Suchman (1994), Denning (2003), or Goldkuhl (2003).

While we did not originally analyze our empirical data with the conversation for action model in mind, an object-oriented analysis (Mathiassen et al. 2000) of the problem and application domain produced, among others, a series of state-chart diagrams depicting structures in the tasks and communication patterns, which we quickly recognized from the work of Winograd and Flores (1986). Re-examining the video recordings in the light of the conversation for action model it became evident that the conversations taking place on the container vessel during the operation of, for example, letting go the lines, could indeed be mapped on to this structure. Hence, the conversation for action model was not enforced on the empirical data but emerged out of it. However, we also found that the recorded conversations between the distributed actors on the container ship did not involve rejection, withdrawal or counter orders. Thus states 6-9 in figure 4 were not encountered in our field data. This was highly unexpected, but was confirmed by reviewing all transcripts of real-life casting-off operations (as well as other operations) and through interviews with domain experts. When asked about this issue, one of the captains stated that "when I give an order, I mean it, and it is not up for negotiation". Discarding the options of rejection, withdrawal or counter orders, we thus reduced Winograd and Flores' conversation for action model to a five state model corresponding to the formalized procedure for communication about "execution of a direct command" observed in our field studies (figure 5).

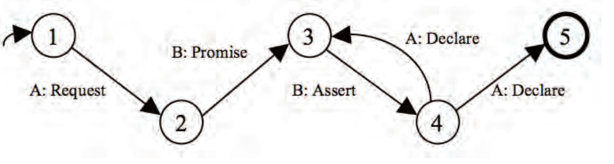


Figure 5. Execution of a direct command

The model of executing a direct command depicted in figure 5 applies to each track of conversation in the transcription from the field study above. Firstly, A requests the execution of a task (e.g. "you can let go the bow line"), which B promises to fulfil ("let go bow line"), taking the conversation to state 3. Having fulfilled the request, B asserts to A that the execution of the task has been completed ("bow line let go"). Finally, A declares contentment ("bow line let go"), and terminates the conversation (state 5) or goes back to state 3, waiting for additional asserts from B ("and we have the bow line home"). Note that states 1 and 5 in this diagram correspond to the categories of imminent and ended conversations. States 2 to 4 correspond to executing conversations.

Formalizing a conversation in relation to a structure such as the general conversation for action model (figure 4) or the execution of a direct command model (figure 5) has a number of advantages in relation persistency and "canning" of communication. First of all, knowing the state of executing conversations, this can be represented visually for persistent and fast access: has a request been met? has an agreement been made? etc. This information can then be integrated with other data sources on the bridge. In relation to canning communication, possible future utterances may be deduced and prioritized over others in the form of predefined standard phrases as seen on some SMS-enabled mobile phones. Thus demands for user-interaction may be simplified and reduced.

4. CANNED COMMUNICATION PROTOTYPE

Informed by the three principles described above, we designed and implemented a functional prototype, the Handheld Maritime Communicator, exploring the use of canned communication for supporting the operation of letting go the lines. This section describes the design of the prototype system.

The prototype was implemented in Microsoft eMbedded Visual Basic allowing it to work on any PDA running the Microsoft PocketPC operating system such as the Symbol PPT8800 industrial PDA series (figure 6). Apart from a touch screen, such industrial PDA's typically support interaction by means of large rubber buttons located below the display and suitable for one-handed interaction. Due to the potentially harsh conditions of the use domain, we decided that all interaction should be facilitated by the use of these buttons. On the Symbol 8800 this would mean that the two-way button would browse the list of possible commands and clicking on the large button below it would select the highlighted item. Though aware of the fact that the final system would probably need to run on a custom-built, solid and weather resistant device and also have a built-in radio for voice-based communication in emergency situations, we decided that for a proof-of-concept prototype, experimenting with off-the-shelf hardware would be sufficient.

4.1. System architecture

The application running on the captain's device works as a server containing a formalized representation of the operation and patterns of communication. All other devices (for example those on deck) log on to this server and identify their physical location following which an appropriate interface is displayed on them. During the operation, function calls and unique command identifiers are exchanged in real time over the network. Thus the problem of commands being missed in the air due to poor sound quality is eliminated. All network communication is broadcast to all devices on the network but processed and represented differently on each device in accordance with their physical location (bridge, fore or aft). While on the first prototype the devices all displayed the same representations (but different possible commands), this architecture would make it possible for us to change the representations and modality used on each individual device, in accordance with, for example, the location of the user, in future iterations without having to change the underlying code. Also, another feature of the exchange of unique command identifiers is that the desired language can be defined individually on each device, thus reducing potential language barriers between co-workers by commands being automatically translated. The desired language is specified in a simple text-file on each device and is thus easily extendable and modifiable.

4.2. Interface design

The Handheld Maritime Communicator prototype (figure 6) gives distributed workers on the container vessel access to a mobile text-based communication channel and provides a graphical representation of the ship and its mooring lines. Supporting bystanders “listening in” on the communication, all communication is broadcast on the network as it unfolds. The overall design was based on two key ideas: 1) to supplement verbal communication with exchange of predefined, canned, text messages, and 2) to provide a simple representation of the work activities in progress in order to improve the distributed co-workers’ reasoning about the ongoing operation as suggested by Rasmussen (1983). This supports human interaction rather than total system automation as discussed by Norman (1990). Meeting these suggestions, we designed an interface providing the user with access to a graphical representation of: the operation in progress; multi-threaded textual communication; a filtered list of completed tasks; and a selection of canned communication utterances.

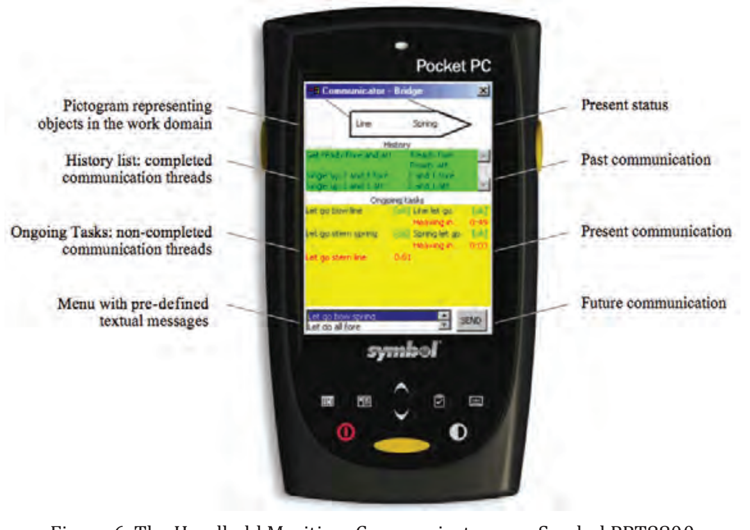


Figure 6. The Handheld Maritime Communicator on a Symbol PPT8800

The interface is divided into four sections resembling the tense of conversations:

1. Pictogram of ship and mooring (present);
2. List of completed communication threads (past);
3. List of ongoing communication threads (present);
4. List of unexecuted commands (future).

The user interface for the bridge is illustrated in figure 6. At the bottom of the screen (immediately above the navigation button) unexecuted pre-defined standard commands and pending confirmations are displayed on a list. The order of the list corresponds to the standard sequence of the overall operation, and possible utterances only appear when appropriate in relation to the state of the task and the location of the specific device (bridge, fore or aft). By default, the most likely next step of the operation is highlighted. The list can be browsed with the navigation button and the highlighted utterance is executed (sent) when pressing the select button. When a command is executed, it is removed from the list and the most likely next step is highlighted (as illustrated in figure 7). Thus the interaction required during standard procedures is limited to a minimum.

Figure 7. Executed commands being removed and new commands appearing

The list of ongoing tasks is perhaps the most important element of the interface. Here, inspired by newsgroups and multi-threaded chat applications, ongoing threads of communication are represented textually. As suggested in the discussion above, executing (present) conversations are grouped in accordance to the object to which they refer rather than by sequence (figure 6). Displaying parallel communication threads textually this way reduces the bottlenecks observed when multiple people speak simultaneously.

Figure 8. Commands being executed (a), confirmed (b), completed (c) and confirmed (d)

The representation of each thread of communication furthermore reflects the five stages of conversations for actions identified through the analysis. When a new command is executed (a request), it appears on the list of ongoing threads of communication representing uncompleted tasks. Next to it, a counter displays the time passed while waiting for confirmation (figure 8a). When a command is confirmed by repeating it back to the captain (a promise) the timer is substituted by the text "[ok]" followed by a description of the current activity (e.g. "Singling up..."). A counter next to this displays the time passed since confirmation (figure 8b). When a task is reported completed (an assert), a short statement (e.g. "1 and 1 fore") substitutes the description of activity and the captain is prompted for confirmation (figure 8c). When the completion of a task is confirmed by repeating it back to the deck (a declare) this is indicated by the text "[ok]" (figure 8d). As a feature of this design the list of ongoing tasks displays not only raw communication but also the present status of each command being executed.

When the captain confirms the completion of a task, the corresponding thread of communication is removed from the list of ongoing tasks and added at the bottom of the history list (figure 9). Hence it changes from present to past tense. The history list automatically filters itself to contain only the initiating commands and subsequent outcomes by removing information such as timers and confirmations (promises and declarations), thus reducing the complexity of the user interface. When the history list is full, it automatically scrolls the oldest commands and statements out of immediate sight. The list can be scrolled using the navigation button together with a function button.

Figure 9. Completed threads being added to history

At the top of the screen a simple pictogram graphically represents the lines attached to the quay for quick reference. Additionally, the overall status of mooring is shown textually (figure 10). This representation is generated from the formalized outcome of past and present threads of communication and supports different levels of abstraction in the interaction with the system. In future design, this information could be integrated with other systems on the ship or made available to others (e.g. harbour traffic control). While only containing redundant information that can also be deduced from the textual descriptions on the display, the graphical representation facilitates an overview of the present situation, which is not currently available on the vessel.

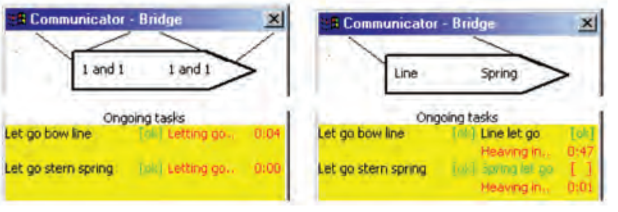


Figure. 10. Pictograms of current status of mooring

The interfaces on deck are very similar to that on the bridge thus providing the first and second officers with a view of the present status of the mooring and a list of all past and ongoing communication among the distributed actors. In the list of ongoing tasks, however, officers on deck are requested to confirm commands executed by the captain such as "let go bow spring". Correspondingly, the list of pre-defined commands only contains those appropriate at the specific location at given states of the operation – e.g. "confirm: let go bow spring" for confirmation of the latter command or "spring let go" for reporting the completion of this task.

5. EVALUATION STUDIES

Evaluating the use and usability of a mobile device for an industrial and potentially safety-critical domain is a major challenge. Firstly, field evaluations of an early stage prototype such as the one described here could cause a hazardous situation. Secondly, evaluating mobile systems in the field limits means of control and significantly complicates data collection (Kjeldskov and Stage 2003, Kjeldskov et al. 2004). Because of these issues, Maersk-Line did not want to evaluate the prototype on board their container vessels at this stage of the project but preferred studying the use of the prototype in the safe and controlled settings of a high-fidelity ship simulator used in their training programs. In addition to this, it was decided to complement the study of use by real users in a highly realistic simulation using a heuristic expert evaluation focusing on potential usability issues of the prototype design.

5.1. Heuristic inspection

For the expert evaluation, we applied an established method for heuristic inspection developed by Nielsen and Molich (1990). The aim of this approach was to test the basic design of an interface using few resources and without involving users.

The heuristic inspection was conducted by a team of three usability experts holding master degrees in computer science with specialization in HCI. The team was given a 15-minute introduction to the use domain covering basic maritime concepts, the operations to be supported, distribution of work, and present procedures of communication (figure 11, left and centre). They received no instructions on how to use the prototype. Aided by a standard heuristic for usability design (Dix et al. 1998), each person spent one hour checking for usability problems while using the prototype. Following the inspections, the team spent one hour producing a final list of problems.

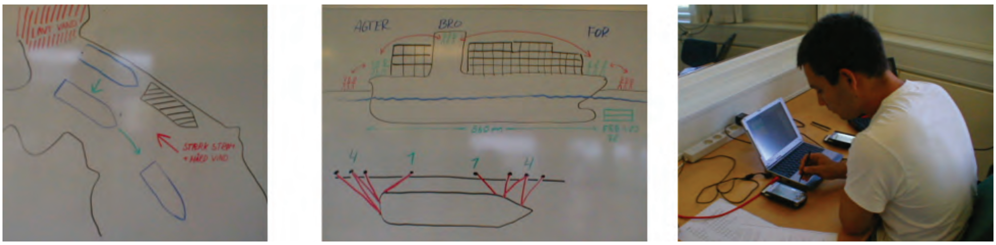


Figure 11. Heuristic inspection: instructions on whiteboards and inspection setup

Results

All three inspectors were able to use the prototype on their own and expressed that the interface design was intuitive and provided overview of ongoing activities and the status of the ship. Twenty-seven usability problems were identified. These were primarily about the graphical design and dialogue between users. Firstly, the history list was criticized for fragmented information, unclear direction of sorting and absence of timestamps. Furthermore, the expert team found the depiction of the ship and mooring lines lacking detail regarding activities (e.g. a line being heaved in). Secondly, errors occurred because the system did not prohibit some commands from being issued before the necessary preceding operations had been completed. On the other hand, further flexibility for deviating from standard commands and sequence was also found desirable. Especially options for withdrawing or correcting commands were found missing (which was, of course, interesting in the light of the discussion of that particular issue with the domain experts during the ethnographic field study). Thirdly, the inspectors were uncertain whether the users would understand compressed statements, different technical terms describing the same objects, and requirements for implicit knowledge. Finally, the difference between confirming receipt of a command and reporting the completion of the related task was found to be unclear because of linguistic similarity.

5.2. Evaluation with captains and officers in ship simulator

While the heuristic inspection provided us with input about the usability of the screen design of the prototype valuable for further refinements, the most interesting and important evaluation study was, of course, the one carried out with real users in the ship simulator. For this study we used a state-of-the-art ship simulator facilitating a fully equipped bridge and a high-fidelity interactive scenario of the operation of a large vessel to create a highly realistic yet controllable and safe environment thus combining strengths and benefits from both in situ and in vitro studies (figure 12). This approach is similar to other prototype evaluation studies of human-computer interaction design for industrial and potential safety-critical domains carried out in full-scope training simulators (for a recent example, see Norros and Nuutinen (2005) on the use of a nuclear power plant simulator for evaluating an experimental safety information and alarm system).

Figure 12. Use in a ship simulator: the simulator and the video recording

While we could not study the use of the prototype in the real world, we went to great lengths to ensure that the evaluation in the simulated use context was as realistic as possible. This was done by firstly, by developing a highly realistic and challenging scenario for the evaluation, simulating real world phenomena from the intended use context at a high level of fidelity. The scenario was developed in collaboration with Svendborg International Maritime Academy, which produces some of the most highly educated maritime officers in Denmark, and runs the simulator facility used for the evaluation. Secondly, the test subjects recruited were high-ranking real prospective users with several years of real life experience with the operation of very large commercial vessels as either captains or leading officers.

The ship simulator consisted of two separate rooms: a simulated bridge and a nearby control room. The bridge was fully equipped with controls for thrusters, propellers, rudder, etc., as well as instruments such as dobler log, echo sounder, electronic maps, radars, VHF radio, etc. From the control room, simulator operators could see the bridge on a closed circuit video surveillance system. The computer application driving the simulation made it possible to simulate the operation of any computer-modelled vessel at any modelled physical location. Also weather and dynamic traffic conditions could be included in the scenario.

The evaluation in the ship simulator involved six test subjects, divided into three teams of two experienced captains and maritime officers, given the task of departing from harbour using the prototype system for communication between bridge and deck. The captains were on the simulator bridge and the first officer was in a neighbouring room set up to simulate the fore deck. As a part of the realistic scenario developed for the evaluation, the captain had to consider all aspects of manoeuvring the ship. This included controlling the rudder, propellers and thrusters as well as communicating with personnel on the ship, harbour traffic control, etc., and keeping clear of and taking into consideration the movements of other vessels. The first officer on deck had to orally forward commands executed by the captain via the mobile device prototype to the operator of the simulation (acting as the team of assistants carrying out the actual tasks) and report back to the captain. The operator would then enter the commands into the simulation (making the vessel respond differently to controls on the bridge as it would in the real world), and report to the first officer when the requested operations (such as letting go a line) had been carried out. For simplicity, commands targeted at the second officer were fed directly into the simulation with feedback given by the operator. The duration of the evaluation sessions corresponded to the length of the operation if it had taken place in the real world.

The simulator was set up to imitate the operation of a large vessel in challenging weather and traffic conditions corresponding to a real world situation observed during the field studies (Nielsen 2000) merged with scenarios used for training at the Maritime Academy. During the evaluations, the captains and officers were asked to think-aloud, explaining their actions and their use of the prototype. Two evaluators located on the bridge and deck respectively observed the captains and officers and asked questions for clarification. Total views of the bridge, deck, simulator control room as well as close-up views of the mobile devices were captured by four video cameras and merged into one video signal providing a synchronized view of the whole setup (figure 12, right).

Following the evaluation sessions, a semi-structured group interview of 10-15 minutes was carried out reviewing the whole operation and discussing the use of canned communication as a supplement to spoken communication.

Data from the user-based evaluations was analyzed qualitatively and quantitatively by two researchers in collaboration producing a ranked list of usability problems as experienced by the captains and officers. Data from the interviews was analyzed qualitatively by 1) relating responses to associated usability problems and thus providing a more detailed account for them, and 2) extending the scope of the evaluation session findings (which traditionally tend to focus on problems rather than on potentials) with a set of themes related to user acceptance and ideas for extending and refining design.

Results

Observing the use of the prototype by prospective users performing a realistic operation in the simulator provided rich data on the usability of canned communication for coordinating distributed work tasks. First of all, the user-based evaluation showed that it was possible to communicate primarily by means of canned text messages while doing a real-world operation. Secondly, the captains and officers expressed that the text-based channel of communication and the graphical representation of the ship gave them advantages compared to the walkie-talkies. Generally, the captains and officers learned to operate the prototype within the completion of one to two threads of communication. The differentiation between future and present commands appeared intuitive as well as the use of parallel threads of communication and technical notions. The pictogram was highly appreciated for providing a quick overview. The desire for more detail about, for example, what people were *doing* fore and aft was expressed both during the evaluations and in the post-evaluation interviews. At the same time, however, the threaded strings of conversations were found to be indispensable supplements to the pictogram for more details while the history list was rarely used and thus took up valuable screen space. In response to the automatic translation of commands between two languages (Danish and English), the officers and captains reported in both the evaluation and in the post-evaluation interviews that this appeared completely transparent to them.

The simulator evaluation revealed 22 usability problems experienced by more than one user. Firstly, we identified a need for correcting or withdrawing commands (again, this was interesting in the light of the findings from the ethnographic field study). Also, we identified a need for requesting or reporting something out of the ordinary. Moreover, a number of standard commands were missing, for example, dismissing a team or requesting a status report. Thus, some captains and officers used the radio to retract a command or notify that a command could not be executed. Secondly, we observed that while all captains and officers executed commands or reports straightforwardly, the procedure for confirming having received a command or report was unclear in the current design of the prototype where those possible utterances appeared in the same list as possible commands. Most of the captains and officers did not immediately notice these new confirm options in the list of commands. One officer misinterpreted it until it was explained to him. One of the captains did not confirm reports from deck until four or five had piled up. On deck, this lack of feedback caused doubt as to whether or not reports had been successfully received. Finally, some officers on deck expressed that while textual communication supported overview and persistency, having to look at the device for reading an incoming command was not always ideal. In the post interview they expressed that they would like to be prompted by, for example, a synthetic voice in combination with the option of looking at the device to get a complete overview when it suited them.

6. DISCUSSION

The two usability evaluations and discussions with prospective users and domain experts revealed several interesting results about the use of canned communication and provided substantial input for refining the prototype system. Canned communicating reduced many of the problems listed in section 2.1. For example, it was obvious that the problems of poor sound quality and lack of persistence were eliminated and that partial automation by automatically suggesting commands proved possible. Furthermore, the graphical representation of the operation successfully supported maintenance of common ground. At the same time, however, interesting new challenges of canned communication also emerged. Below, we discuss some of these challenges and present some of our ideas for improving the use of canned communication.

6.1. Limitations of canned communication

Designing a text-based mobile messaging system for canned communication turned out to be an interesting challenge. As described above, the prototype is based on a reduced version of the conversation for action model. While this model provided a valuable foundation for structure and design, our evaluations also indicated a number of shortcomings, some of which have previously been discussed in the CSCW literature

(Winograd 1994, Suchman 1994). The current design of the communicator did not support the handling of three types of non-standard situations. The first was retraction of a command. Even though our field study and interviews found this to be unnecessary, the usability evaluations showed that this was not the case and that the captain may indeed want to modify or withdraw a command that had already been issued. A full implementation of the conversation for action (which would be technically trivial) would facilitate this. The second was error prevention. The change from continuous and open radio communication to discrete and closed text-based communication seemed to increase the risk of stating a wrong command (it is easier to select a wrong item on a list than to accidentally state a wrong command verbally). The third type was unanticipated communication. In an emergency situation, communication changes from asynchronous to immediate because the situation develops quickly. In this situation, the benefits of canned communication will be overshadowed by its limitations and communicating unconstrained from pre-defined messages will be necessary.

The evaluations also indicated some risk of task interference. When auditory communication is replaced with screen-based interaction some of the user's visual attention is diverted from the task being conducted to interacting with the mobile device. The captain and officers are already watching crewmembers, instruments, mooring lines, other ships, etc. In addition, they are watching their own physical actions. While relieving the highly busy auditory channel by introducing text messaging, having to look at a screen to handle communication puts additional burden on the users' vision, which in some situations may not be appropriate. In response, officers and captains suggested combining the two in a flexible manner.

6.2. Improving canned communication

The two evaluations and the post-evaluation interviews provided massive input for improving canned communication. Some of these are simple and closely related to the specific interface design of the prototype. Firstly, the history list may be hidden in a sub-menu thus freeing up valuable screen real estate. Secondly, the graphical representation could be extended to include more detail and, for example, reflect ongoing work activities. Thirdly, in order to bring attention to pending confirmations, these should be separated from commands, for example, by displaying them in two separate lists. Other problems are more general and complex, and require the development of new ideas and further evaluations.

Modifying and withdrawing commands

The lack of facilities for handling non-standard situations was found to be a key problem. Some of this problem would be solved trivially within the present overall design by implementing the full conversation for action model (figure 4) rather than the reduced execution of a direct command model (figure 5). A facility for modifying or withdrawing an issued command could then be included by introducing a special type of command that aborts an ongoing command and sends out a counter-order. If an error occurs frequently and is handled in a standardized manner, it can be integrated in a way that is similar to retraction. Otherwise, it must be handled as a case of unanticipated communication. The whole issue of modifying an issued command naturally raises the question of what happens if someone sends a wrong command? This can take time to correct and result in wrong actions to be taken. Essentially the need to modify or withdraw commands is not a new challenge emerging from the use of textual communication but also applies to spoken communication. Hence the use of canned communication does not introduce this challenge as such but merely inherits it. However, as canned communication based on selecting an utterance from a list rather than speaking it out may introduce a problem of commands being sent *by mistake*, a central challenge of building in the full flexibility of the communication for action model will be to produce a design that minimizes the risk of this happening, for example through a prompt for confirmation. What also needs to be considered in the design of communicator systems such as the one presented here is who has the overriding power to modify or withdraw a command? Does this privilege apply to everyone or just to the one issuing the order in the first place? In implementing the full conversation for action model, the ability to modify or withdraw a command would be distributed on all communicating actors through (potentially infinite) loops of negotiation.

Flexibility

Related to the need for modifying and withdrawing commands on a structural level, the use of canned communication also raises new challenges in relation to the flexibility of communication. Unanticipated communication, for example in emergency situations, seems to require flexibility and thus conflicts fundamentally with the motivation for using canned communication which takes advantage of the structures of standardized conversations to make it persistent, freeing up resources for other tasks. While new utterances may, of course, be "canned" in the system over time, it is very likely that new exceptions will also continue to emerge, thus eventually making the complexity handling the number of possible canned utterances outweigh the reduction of complexity gained from using them. Thus, other means of supporting flexibility should be considered.

As suggested throughout the paper, canned communication in industrial domains such as the one explored here should not be seen as a replacement of spoken communication but as a supplement. Hence one way of dealing with the flexibility issue is integrate facilities for radio communication directly in the mobile device and simply revert to this for communication out of the ordinary. In fact this was exactly what happened in the ship simulator. Combined with facilities for withdrawing a command, as discussed above, this would give the communicating partners full flexibility on top of a baseline persistent channel of communication. However, as human perception and action are to a large extent driven by expectation (Oatley, 1979), strong expectations about what channel to use in which situations is needed among the co-workers for such a multi-modal communication channel to work.

Another way of increasing the flexibility of canned communication is to look at the syntax of the individual speech-acts themselves. In the operation of letting go the lines, for example, commands typically consists of 1) *actions* (get ready, let go, etc.), 2) *objects* (spring, line, etc.) and 3) *locations* (fore, aft, etc.). Taking this differentiation into consideration, it would be possible to let the users create their own speech acts by combining a selection of actions, objects and locations rather than canning *complete* similar speech acts (such as "let go spring aft" and "let go line aft") as separate instances.

On the interface level, this could be done in a very simple manner by replacing the single menu of possible commands with three menus of possible actions, objects and locations.

Minimizing task interference
In order to reduce task interference, the use of alternative input and output devices could be considered. The requirement for visual attention towards a handheld device could be reduced by means of a wearable head-up displays or speech synthesis, and extending the device with a headset and voice recognition could support hands-free interaction while still maintaining the benefits of computerized persistent communication.

The discussion of alternative input and output media and modalities also raises an interesting question of whether all collaborating parties actually need the same representation of information, modalities and means of interaction or if differentiating between these in accordance to the context of each individual user would be preferable. Based on our findings from the evaluations, there are good reasons to explore solutions where, for example, officers on deck get commands delivered primarily through synthetic speech (allowing for the benefits of canned communication) while also having access to a secondary visual (and hence persistent) representation of the ongoing tracks of conversations in the form of text or a graphical representation of the vessel and mooring lines, or both. On the bridge, our field studies showed that during operations involving multiple physically distributed mobile collaborating actors the audio channel is the busiest one compared to the visual one. At the same time, it is the captain who has to keep track of most parallel tracks of communication. Hence the captains, not surprisingly, attached high value to the graphical representation of ongoing threads of communication and especially to the representation of the current status of the vessel. In relation to this, it would be interesting to investigate whether improving the graphical representation of the vessel could make the textual threads of utterances redundant. Also, as the bridge typically accommodates a physically co-located group of people (captain, harbour pilot, helmsmen, and sometimes machine engineers) clustered within a few meters, complementing their mobile devices with larger situated displays could provide a useful means for maintaining common ground and facilitating asynchronously "listening in" to secondary tracks of communication and thus staying "in the loop". Because the utterances are already persistent and formalized within the computer system, translating between input and output media and modalities would not be difficult to implement (see 4.1).

6.3. Canned communication in industrial domains revisited
The presented mobile device was designed and evaluated for a specific industrial domain: the operation of large container vessels. This domain is characterized by activities proceeding at regular pace leaving time for the involved actors to read and comprehend written communication. Also, established procedures exist and communication is formalized. In the case of emergencies and other non-standard events, this pattern is broken and workers should shift to spoken communication. Other industrial domains may have different characteristics. There may be more people involved, a much higher level of stress, and need for rapid concerted action. In such domains, the ephemeral nature of audio may sometimes be an advantage and actors may benefit more from sharing information by overhearing other people's communication than we observed

in the maritime domain where multiple clearly separated channels of communication were used to contain communication to specific and delimited groups of co-workers. The auditory communication also provides the listener with an impression of emotional state, identity of the speaker, and ambient noise, which may or may not be important issues to consider. These aspects illustrate that the design of a mobile communication device for one industrial domain cannot simply be transferred to another without further consideration. Comprehensive studies and analysis of the domain in question must be carried out, designs must be tailored to the unique features of this domain, and substantial evaluation studies must be carried out to validate the quality of the design and inform further refinements.

7. CONCLUSIONS

Based on a thorough ethnographic study we have explored the use of "canned communication" in an industrial and potentially safety-critical domain through the design and evaluation of a mobile prototype system. The prototype system supplements spoken communication with predefined, canned, text messages and provides a persistent graphical representation of an operation in progress and a filtered list of completed tasks. Two qualitative evaluations were conducted, a heuristic inspection and a user-based evaluation in a high-fidelity simulation of the use domain, providing rich and varied input on the potentials and limitations of canned communication for coordinating distributed and mobile work activities. Together, the two evaluations clearly indicate a series of advantages of canned text-based communication over spoken communication, but also bring attention to a number of challenges for canned communication regarding, for example, flexibility and task interference.

The mobile device was designed for a very specific and specialized domain but the concept of canned communication, as well as the central design ideas of the prototype system have value for the design of persistent mobile communication systems in general. In particular, the grouping of communication threads and the generation of a graphical representation, which integrated physical location, language, role and task proved to be highly useful. In order to increase the generality of our findings, additional studies of canned text-based communication on mobile devices should be conducted in both similar and different domains. The simulator-based usability evaluations should also be complemented with real-world evaluations over longer periods of time investigating the long-term use of canned communication. While requiring a very refined and stable prototype system, such evaluations might provide a basis for assessing other relevant factors such cognitive workload and possible reductions in time spent on communication, as well as identifying further benefits and challenges of canned communication.

ACKNOWLEDGEMENTS

The authors thank Peter Bøgh Andersen and Morten Nielsen who did the initial ethnographic studies, Svendborg International Maritime Academy and all test subjects. The evaluations were planned and carried out in collaboration with Mikael B. Skov and the heuristic evaluation was done by Rune Thaarup Høegh, Karsten Kryger Hansen and Søren Lauritsen. We also thank Steve Howard for comments on an earlier draft.

REFERENCES

Alm, N., Arnott, J.L., Newell, A.F., 1992. Prediction and Conversational Momentum in an Augmentative Communication System. Communications of the ACM 35 (6), 46-57.

Akhus, M., 2001. Designing reflective dialogue to support learning from experience. SIGGROUP Bulletin 22 (2), 29-37.

Andersen, P.B., 2000. Communication and work on maritime bridges. CHMI Research Report CHMI-1-2000, http://www.cs.auc.dk/~pba/ElasticSystems.

Andersen, P.B., May, M., 2001. Tearing up Interfaces. In: Liu, K., Clarke, R., Andersen, P.B., Stamper, R.K. (Eds.), Information, organization and technology – Studies in organizational semiotics. Klüwer, Boston, MA, pp. 299-338.

Champ, P.J., Hudson, J.M., Keldorph, R.B., Lewis, S., Mynatt, E.D., 2000. Supporting Communication and Collaboration Practices in Safety-Critical Situations. Proceedings of CHI'2000, The Hague, Netherlands, pp. 248-250.

Churchill, E.F., Bly, S., 1999. It's all in the words: Supporting work activities with lightweight tools. Proceedings of ACM SIGGROUP '99, Phoenix, Arizona, USA, pp. 40-49.

Clark, H.H., Schaefer, E.F., 1989. Contributing to discourse. Cognitive Science 13, 259-294.

De Michelis, G., Grasso, M.A., 1994. Situating Conversations within the Language/Action Perspective: The Milan Conversation Model. Proceedings of CSCW 1994, Chapel Hill, NC, USA, pp. 89-100.

Denning, P.J., 2003. The Profession of IT. Communications of the ACM 46 (7), 19-23.

Erickson, T., Herring, S., 2006. Persistent Conversation: Design and Analysis of CMC Systems. Proceedings HICSS-39 (minitrack on persistent conversation), Kauai, Hawaii.

Dix, A., Finlay, J., Abowd, G., Beale, R., 1998. Human-Computer Interaction. Prentice Hall Europe, London.

Fields, R., Paterno, F., Santoro, C., Tahmassebi, S., 1999. Comparing Design and Options for Allocating Communication Media in Cooperative Safety-Critical Contexts: A Method and a Case Study. ACM Transactions On CHI 6 (4), 370-398.

Frisse, M.E., 1988. Searching for Information in a Hypertext Medical Handbook. Communications of the ACM 31 (7), 880-886.

Goldkuhl, G., 2003. Conversational Analysis as a Theoretical Foundation for Language Action Approaches? Proceedings of the 8th International Conference on the Language Action Perspective on communication Modeling, LAP-2003, Tilburg, The Netherlands, pp. 51-69.

Hutchinson, H., Hansen, H., Roussel, N., Eiderbäck, B., Mackay, W., Westerlund, B., Bederson, B.B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., 2003. Technology probes: inspiring design for and with families. Proceedings of CHI'03, Fort Lauderdale, Florida, USA, pp. 17-24.

Hüttenrauch, H., Norman, M., 2001. PocketCERO – mobile interfaces for service robots. Proceedings of MobileHCI'01, IHM-HCI, Lille, France.

Jayaweera, P., Johannesson, P., Wohed, P., 2001. Collaborative Process Patterns for e-Business. SIGGROUP Bulletin 22 (2), 21-28.

Kjeldskov, J., Skov, M.B., Als, B., Høegh, R.T., 2004. Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. Proceedings of MobileHCI'04, Glasgow, Scotland, pp. 61-73.

Kjeldskov, J., Stage, J., 2003. New Techniques for Usability Evaluation of Mobile Systems. International Journal of Human-Computer Studies 60, 599-620.

Mathiassen, L., Munk-Madsen, A., Nielsen, P.A., Stage, J., 2000. Object-Oriented Analysis and Design. Marko, Aalborg, Denmark.

McCarthy, J.C., Miles, V.C., Monk, A.F., 1991. An experimental study of common ground in text-based communication. Proceedings of CHI'91, New Orleans, USA, pp. 209-215.

Nielsen, C., Søndergaard, A., 2000. Designing for mobility - an integration approach supporting multiple technologies. Proceedings of NordiCHI 2000, Royal Institute of Technology, Stockholm, Sweden.

Nielsen, J., Molich, R., 1990. Heuristic evaluation of user interfaces. Proceedings of CHI'90, Seattle, USA, pp. 249 - 256.

Nielsen, M., 2000. Letting go the lines: Departure from Felixstowe harbor. CHMI Research Report CHMI-4-2000, http://www.cs.auc.dk/~pba/ElasticSystems.

Norman, D., 1990. The 'Problem' With Automation: Inappropriate Feedback and Interaction Not Over-automation. In: Broadbent, D.E., Reason, J.T., Baddeley, A. (Eds.), Human Factors in Hazardous Situations. Clarendon Press, Oxford, pp. 137-145.

Norros, L., Nuutinen, M., 2005. Performance-based usability evaluation of a safety information and alarm system. International Journal of Human-Computer Studies 63, 328-361.

Oatley, K., 1979. Perceptions and Representations. Free Press, New York.

Popolov, D., Callaghan, M., Luker, P., 2000. Conversation Space: Visualising Multi-threaded Conversation. Proceedings of AVI2000, Palermo, Italy, pp. 246-249.

Rasmussen, J., 1983. Skills, Rules and Knowledge: Signals, Signs and Symbols and Other Distinctions in Human Performance Models. IEEE Transactions on Systems, Man and Cybernetics 13 (3), 335-350.

Suchman, L., 1994. Do categories have politics? the Language/Action Perspective Reconsidered. Computer-Supported Cooperative Work 2 (3), 177-190.

van den Anker, F.W.G., Lichtveld, R.A., 2000. Early Evaluation of New Technologies: The Case of Mobile Multimedia Communications for Emergency Medicine. In: Vincent, C., de Mal, B. (Eds.), Safety in Medicine. Elsevier Science, Oxford, pp. 193-215.

Winograd, T., 1994. Categories, disciplines and social coordination. Computer-Supported Cooperative Work 2 (3), 191-197.

Winograd, T., Flores, F., 1986. Understanding Computers and Cognition. Abblex Pub. Corp., Norwood, NJ, USA.

Part II

Designing and building

DESIGNING AND BUILDING

Part II addresses the question *how can we design and build interactive mobile systems grounded in context?* In order to make use of our empirical and theoretical understanding of context in the exploration of a design space, we need to develop ways of transferring this knowledge into the design activity, develop ways of grounding design ideation and exploration in such input, and support synthesising multitudes of ideas into concrete interactive systems. Four of my own contributions to this are included in chapters 7-10. These chapters adress and discuss possible ways of supporting the creative process of moving from abstract understanding towards concrete artefacts by using *form and context ensembles* as the central point of gravity for design exploration.

User- and technology-centeredness
Chapter 7 discusses user- and technology-centeredness. In this chapter we ground the design of mobile interactions in our experiences with processes that are either user- or technology-centred. The chapter provides a meta-commentary on the relative values and limitations of inspirationalistic and structuralistic creativity for addressing the same mobile interaction design brief as a part of the "TramMate" project in Melbourne, Australia. The chapter concludes that neither of the two approaches was superior, but that in concert they appeared to have potentials for complementing each other very well within a broader approach: the user-centred approach grounding design in the contextual complexity of current practice, and the technology-centred approach providing a counterpoint grounded in the context of future technical possibilities.

Socio-physical design
Chapter 8 explores a socio-physical approach to design. In this chapter we propose and illustrate an inter-disciplinary approach to designing mobile interactions combining rapid ethnography, architectural analysis, design sketching and paper prototyping. The socio-physical approach combines inspirationalistic and structuralistic creativity by, on one hand, informing unrestricted idea generation through inspirational and ambiguous representations of context and, on the other hand, offering a stepwise and structured process for translating the understanding of socio-physical context into specific design ideas. The approach was explored through a mobile interaction design case study following on from the empirical work presented in chapters 3 and 4 and leading to the design and implementation of an interactive mobile prototype system.

Sketches and mock-ups
Chapter 9 illustrates the use of sketches and mock-ups. In this chapter we revisit the study of intimacy from chapter 5, and describe how our understanding of personal context was used to inform mobile interaction design. The chapter discusses the use of interactive technologies to support intimacy and describes our activities of ideation and interaction design. These took place through a semi-structured process of sketching and mock-up development in collaboration with study participants, adding a level of structure and method to the design team's mainly inspirationalistic type of creativity. The diversity in scope and form of designs produced illustrate how sketches and mock-ups are a highly flexible means for exploring and communicating ideas in a heterogenous design team.

Ethnography and object-orientation

Chapter 10 explores the combination of ethnography and object-orientation. In this chapter we propose and describe an inter-disciplinary combination of methods for supporting the design and development of interactive mobile computer systems. This combines inspirationalistic and structuralistic creativity by integrating an open-ended method with a structured one. The chapter presents two case studies of mobile system development, following on from the work presented in chapter 6. The two studies show that ethnographic data is highly valuable for developing object-oriented models by providing contextual richness. In return, object-oriented analysis is a valuable way of working with ethnographic field data by providing structure. Combining the two we were able to strongly inform system design with our ethnographic field studies.

Chapter 7

User- and technology-centredness

Jesper Kjeldskov and Steve Howard

Abstract. We provide a meta-commentary on two approaches used for designing context-dependent mobile devices. On the basis of a 'user-centered' approach, consisting of interviews, observation of current practice and enactment of future scenarios in context, a number of non-functional design sketches were developed. While these sketches reflected a rich understanding of current work practices, they were little more than abstract speculations about future practice; *lacking in detail on usability and feasibility, and being largely reactive to current problem situations*. Conducted in parallel, the technology-centered study informed the design and implementation of a mature functional prototype. This facilitated a comprehensive usability evaluation revealing a series of technical challenges and problems related to mobile use. Though the technology-centered approach provided detailed input for refining the prototype, and an initial provocative break with current practice, it was less useful in supplying further original alternatives; *post-evaluation, the design discussion was largely reactive to the current prototype*. In concert, the two approaches complement each other well; the user-centered approach grounding design in current practice, in all its contextual complexity, and the technology-centered approach providing a counterpoint in technically detailed expressions of future possibilities.

1. INTRODUCTION

Mobile, pervasive and ubiquitous computing constitutes a challenge for human-computer interaction design. Like other emerging technologies, these technologies are characterized by being different from traditional desktop computers in their physical appearance and the contexts in which they are used. Consequently, they often imply interaction dissimilar from how computers are usually operated and are used for purposes beyond office and home computing. Many traditional approaches to

Originally published as Kjeldskov, J. and Howard, S. (2004) Envisioning Mobile Information Services: Combining User- and Technology-Centered Design. In *Proceedings of APCHI 2004,* Rotorua, New Zealand, LNCS (180-190). Berlin: Springer-Verlag. The epilogue was added in October 2010.

HCI design are evolutionary and based on incremental iterations of analysis, design, implementation and testing. These methods have proven valuable in relation to the development of computer systems with high usability based on well-known and widely adopted technologies such as desktop workstations and the Internet. In the light of emerging technologies, however, interaction design research has been stimulated into the development and use of design approaches that are less analytic and more creative. These approaches supplement traditional HCI methods by introducing an explicit focus on envisioning future technology and its use – so called 'blue-sky' research. The fundamental assumption is that just like the introduction of personal computing and computer networks have led to huge and unexpected changes in the work and leisure activities of many people, so will mobile, pervasive and ubiquitous computing in the decades to come. As the potential use of these technologies is, however, still unknown, researchers must develop new concepts and ideas for future technology: what it may look like, what it may be used for, and how people may interact with it.

HCI research promotes two overall approaches to concept design for emerging technologies: 1) experiments driven by the opportunities offered by new technology and 2) user-centered design driven by field studies. While a recent literature survey (Kjeldskov and Graham 2003) shows a tendency towards mobile HCI being driven by technology- rather than user-centered approaches, growing attention has lately been brought to concept development methods focusing on present and envisioned future user activities. Much of this research has its roots in the Participatory Design tradition (Greenbaum and Kyng 1991), and is motivated by the observation that "we can make amazing things, technically, but are often at a loss to understand what to make", as stated by Vygotsky (1978). In response to this, new methods are introducing techniques from theatre, such as actors, props and role-playing to HCI research. Sato and Salvador (1999) suggests a focus group approach modified to involve actors and an active audience acting-out use of existing and non-existing product concepts. The benefits from this approach are that the theatre approach creates a strong context with focus on interaction as well as an improvisational space for the participating users. However, it is also noted that it can be difficult for users to be creative and articulate on non-existing products in fictive use scenarios. Inspired by this work, Howard et al. (2002) have experimented with the use of 'endowed props' providing the design team with a means of directing the discourse between science- and plausible-fiction. Extending this line of thought, Iacucci et al. (2000) presents techniques where enacting the use of future technology is a part of a board game or takes place in the real world (the SPES technique), thus restricting discourse by means of a game master or the constraints of real life activity.

These and related techniques indicate a plausible approach to interaction design for emerging technologies. However, they are to a large extent disconnected from real technology. While on one side this frees designers from the technological constraints of the present, it also inhibits the exploration of new technological potentials and may blind insight into limitations of present and emerging technology. Hence, Rogers et al. (2002) promotes that instead of choosing between a user- and a technology-centered approach, blue-sky research should explicitly involve both playful visions of technology in their social context and innovation through 'technology inspiration'.

On the basis of this discussion, this paper presents a research project in which a user- and a technology-centered approach were combined for envisioning design of mobile information services for the same problem domain. The first section describes the motivation and methodological design of the project. Following this, the details of the user- and the technology-centered approaches are described and discussed. This includes a description of the empirical methods used, the designs produced and the evaluations conducted. Finally, the interplay between the two approaches is discussed and directions for further research are outlined.

2. SUPPORTING THE USE OF PUBLIC TRANSPORT

Between October 2002 and April 2003 we conducted a research project focusing on the potential use of context-aware mobile computing for supporting the use of public transportation in Melbourne, Australia. The project was motivated by discussions among consultants and sales staff of a large IT company about alternatives to the use of cars for traveling in city meetings with clients. In large cities where traffic is often very dense, traveling by car can be highly time-consuming, necessitating much planning. Using Melbourne's tram-based public transport would not only be more environmental, but might also be more effective if supported by a mobile information service providing travelers with relevant information at the right time and place.

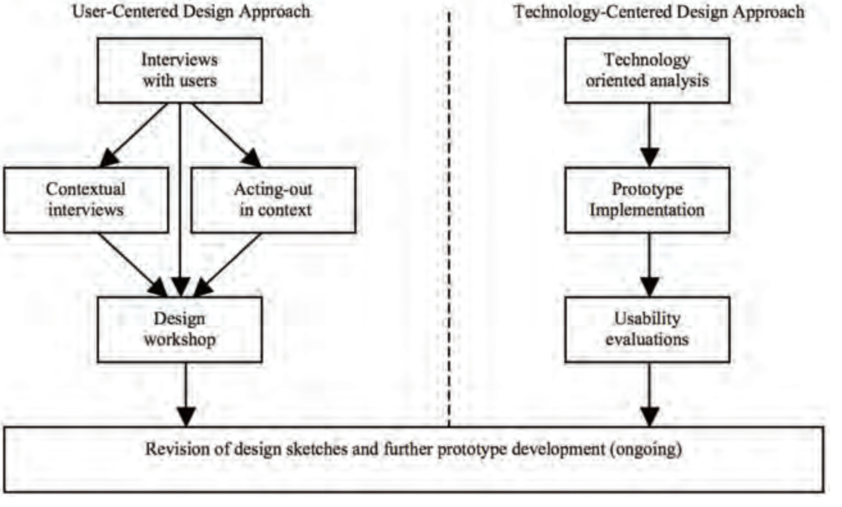


Figure. 1. Combining user- (left) and technology- (right) centered approaches to interaction design

Over the course of six months, a team consisting of three senior researchers, five Ph.D. students, two research assistants and an industrial interaction designer worked on different parts of the project. The first author of this paper managed the project and participated in the majority of activities. The researchers were all involved in HCI design for mobile computing but had different specific individual research interests ranging from ethnographic user-studies to prototype development and usability evaluation. On the basis of the different points of interest, the project consisted of two parallel tracks of research: a user- and a technology-centered track (figure 1).

The user-centered track grew out of the initiating discussions with the IT professionals. The technology-centered track grew out of an already ongoing activity to develop a mobile route-planning service for the tram system of Melbourne. The researchers in each track remained mutually unaware of each other's activities for the first 3 months of the project. Thus the user studies were conducted without any knowledge of the prototype development and the design of the first experimental prototype was not informed by the empirical user studies. Following the design workshop, researchers joined across the tracks and conducted a usability evaluation of the prototype. Finally, the outcomes from the two tracks of research were merged. The following sections describe the activities in the two parallel tracks in detail.

3. THE USER-CENTERED APPROACH

Based on the discussion above, we designed a user study to investigate into the current and possible future work practices of our nomadic users. In order to construct a rich understanding of travel and work practices, and building on project members' experience with techniques for acting-out future scenarios from previous research (Howard et al. 2002) we combined three empirical methods: 1) Semi-structured interviews with users (at the office), 2) Contextual interviews and observation of current practice (driving cars in city), and 3) Acting-out future scenarios in context (on board trams in city)

These methods all involved users talking about their perceptions, activities and needs. While the semi-structured interviews were separated from context, the contextual interviews and acting-out sessions took place while the users were carrying out real work activities. During the contextual interviews and acting-out sessions, we observed current practice, including the use of different artifacts in relation to the work activities being carried out. In the acting-out sessions we furthermore provided a series of non-functional props to assist the users' imagination. The three methods and a summary of their primary outcomes are described below.

3.1. Interviews

To establish the context, four semi-structured interviews were conducted with the IT professionals, who frequently traveled for work activities in the city. The interviews focused on the users' perceptions of pros and cons of traveling by car and tram. Each interview was conducted by a pair of researchers and lasted 30-60 minutes.

It was clear from the interviews that the users were willing to use public transport when attending meetings in the city, but that the use of the local tram system had a number of limitations. Firstly, uncertainties were related to route planning: which routes to take, when to change trams, how many changes would be needed, would there be a wait at each change, how close was the tram stop to the desired destination…? Often users would tend towards their car, or a taxi, if tram changes were needed, or if the journey was to an unfamiliar destination. Being able to predict the time of arrival of the tram rather precisely was seen as critical. The interviewees attempted to avoid being either late or too early (thus wasting time for their clients, or themselves respectively). To assist in timely arrival at appointments the interviewees needed to know exactly when to leave their office in order to arrive at their meeting just at the right time. When using the car this estimation was reportedly done based on knowledge about normal traffic conditions in the city and was not very precise. Uncertainty about combining routes, trams not running on time, finding the nearest stop and having to walk all constituted barriers to tram use.

3.2. Contextual Interviews and Observation of Current Practice

Following the interviews, the next phase consisted of observing and enquiring into current practice (figure 2). Here we were trying to understand both hurdles and enablers to car and tram travel. The observations were conducted by shadowing the users during travel from their offices to meetings in the city. One researcher asked questions about pros and cons of using trams while another video taped the session.

From the observations we learned that, although driving to meetings in the city could be rather time consuming, the car was seen as flexible and provided a useful 'office on the move', a semi-private space for limited work activities. Thus time spent in the car was not necessarily 'wasted' time but often used for preparing the upcoming meeting or coordinating other activities over the phone.

Figure 2. Contextual interview and observation of current practice

3.3. Acting-Out in Context

The final phase of our field study consisted of a number of sessions in which our nomadic users acted-out future scenarios of using mobile information technology to support travel by tram (figure 3). The acting-out approach was adapted from previous research by project participants (Howard et al. 2002)

When acting-out a future scenario, people are asked to envision and enact situations involving future use of technology, based on an overall frame of context and supported by simple props, which are, attributed the desired functionality. For the present project, the original approach was modified in a number of ways, resembling aspects of the SPES technique (Iacucci et al. 2000). First, real users instead of actors did the acting-out. Secondly, the acting-out was done in context: using the trams for attending a real meeting in the city instead of performing fictive tasks in a studio. Two researchers facilitated the sessions and took notes; a third recorded the sessions on video.

From the acting-out sessions we learned a lot about information needs and desires and how these varied across different situations and locations. Also criteria for assessing the appropriateness of different form factors were revealed, for example the need for privacy when working in a public space. Specifically, the acting-out in context sessions revealed that before catching a tram into the city, estimation of travel time was essential to determine when to go. On the tram, the primary information needed was when and where to get off and what to do next. In addition, users envisioned that they could prepare a meeting, browse the Internet or listen to music. One user also envisioned that the device would automatically buy an electronic ticket from the tram.

Figure 3. Acting-out in context

3.4. Design Workshop

Following the field studies, we conducted a one-day design workshop. The purpose of this was to recap and discuss the outcome of the empirical user studies and to produce a number of ideas for mobile systems supporting current and/or envisioned future practice. The workshop involved three senior researchers, three Ph.D. students, two research assistants and one industrial interaction designer. The latter managed the workshop in collaboration with the first author. First the three teams of researchers who had conducted the interviews, contextual interviews and acting-out sessions presented their findings. This was followed by a one-hour joint discussion. The participants were then divided into three groups, mixed so that they did not resemble the teams who did the user studies. The groups then spent 2½ hour on producing a design sketch of their own choice. No restriction was put on their focus, but the design had to address issues identified in the user studies. Following this, each group presented their design sketches using whiteboards and paper drawings. The workshop resulted in four design concepts. The ideas were highly creative and diverse: from a foldable mobile office to an MP3 player with voice-based route-planning, capabilities for presenting PowerPoint slides and wireless connectivity to stationary I/O devices at the customer's site. The envisioned concepts all addressed different aspects and challenges of the mobile work activities studied and were informed by both input from the interviews, the contextual interviews and the sessions of acting-out. Due to limited resources and time, only one of the concepts, described in detail below, was developed further.

3.4. TramMate

TramMate supports the use of public transportation by means of a context-aware mobile calendar application. On the basis of the field study, the basic idea of TramMate is to 1) relate traveling information directly to appointments, 2) provide route planning for the tram system based on current location 3) alert when it is time to depart, and 4) provide easy access to travel time, walking distance and number of route changes. Elaborating on ideas of context awareness (Cheverst et al. 2000) and indexicality (Kjeldskov 2002), the interface of TramMate is indexed to contextual factors such as time, location, and activity (Kjeldskov et al. 2003).

Figure 4. Refined design sketches for TramMate

The rationale behind TramMate is not to impose too much additional complexity on the user. Accomplishing this, an extended electronic calendar provides dynamic route planning information related to the user's schedule for the day. TramMate thus requires very little additional interaction. When a new appointment is made, the user is asked to specify the physical location of it, following which TramMate schedules a time slot for getting there. When an appointment is coming up, this timeslot adjusts itself in accordance with the location of the user and the estimated time needed to get there based on real time information about the public transport system (figure 4 left). Apart from specifying the first step of the route plan, the calendar view also provides additional details on the suggested route: estimated travel time, required walking distance and the number of route changes. During traveling, the TramMate timeslot continuously updates itself with information about the next leg of the route. From the calendar view, the user also has access to a screen providing his current location on a map with directions to the nearest tram stop. This screen also outlines the full route (figure 4 center). Based on the time required to walk from the user's current location to the first tram stop, TramMate notifies when it is time to leave in order to make the appointment. The reminder contains simple information about the appointment, what tram to catch, how soon it leaves, where it leaves from and how to get there (figure 4 right). On the tram, TramMate notifies when to get off and what next step to take. Arriving at the destination, a map provides the location of the appointment.

4. THE TECHNOLOGY-CENTERED APPROACH

In parallel with the user-centered approach, researchers at the University of Melbourne's Department of Geomatics conducted a technology-centered track of research with the purpose of developing a functional prototype for Melbourne's tram system. The prototype addressed the same overall use-situation involving nomadic workers in the city that motivated the user study, but was not influenced by the findings in this track of research. Instead, it was driven by a desire to explore the potentials for providing location-based route planning on GPS and GPRS enabled handheld computers.

The prototype had a number of differences from and similarities to the TramMate concept. The prototype was similar in providing route-planning facilities for the tram system based on the user's current location as a mix of textual instructions and annotated maps. However, it provided a fundamentally different user interface, requiring the user to actively look up travel information rather than relating it to planned activities, and did not alert the user before or while in transit. Also, the prototype did not automatically reflect contextual changes such as the user changing location. This was only taken into account when the user actively looked up timetable information.

4.1. Usability Evaluations

Whereas the design sketches produced in the workshop were non-functional, the prototype enabled us to study the use of real technology. We conducted two usability evaluations of the prototype: in the field and in a usability lab, involving ten subjects. The evaluations were identical in terms of tasks and the profiles of test subjects. The subjects had to complete three realistic tasks involving route planning while traveling to appointments in the city derived from the user studies with IT professionals.

Figure 5. Field evaluation

The field evaluation focused on use in realistic settings. The test subjects had to lookup necessary information on the device according to the tasks and then perform the tasks "for real" (e.g. catching a tram to a specific destination). The prototype accessed live timetable information on Internet but GPS positioning had to be simulated. One researcher encouraged thinking-aloud, one took notes and one recorded the evaluation on video (figure 5). In the laboratory evaluation, the subjects were seated at a desk, with the mobile device in their hand (figure 6). To ensure a good video recording, the subject held the device within a limited area indicated on the table. Two researchers observed through a one-way mirror and took notes.

Figure 6. Laboratory evaluation

Studying the use of real technology revealed a number of problems concerning the design of context-aware mobile information services for supporting the use of the tram system. Overall, two critical and nine serious themes of usability problems were revealed (Pedell et al 2003). One of the major problems identified concerned *the relation between information in the system and in the world.* During use, the context changed constantly; the users moved, the trams moved, time went by etc. While this was to some extent reflected in the system, the granularity of this relation was not sufficient. Consequently, the information on the screen would often be often 'out of sync' with the real world. Another major problem concerned the *graphical design of maps.* All users wanted to use the maps a lot. However, the maps turned out to have three significant limitations. Firstly, the level of detail was generally not appropriate. Either the screen would be cluttered or it would provide second to no information. Secondly, the maps lacked annotation of key information such as landmarks and precise indication of the user's location, route and destination. Thirdly, the users had serious problems relating the orientation of the map to the real world and would frequently rotate the device. As a consequence, screen text was often viewed upside down.

Other problems identified included difficulties entering data while being mobile, missing vital information due to lack of visual attention, and lack of functional transparency when information on the screen changed to reflect contextual changes.

5. DISCUSSION AND CONCLUSIONS

It is interesting to compare the outcomes of the two parallel tracks of R&D. The user-centered design approach facilitated the development of design concepts representing considerable insight into present work practices, and some speculations about the future not envisioned by the prototype designers. However, the usability and feasibility of these concepts were unclear. The technology-centered approach revealed specific technical challenges and problems related to the design and use of context-aware mobile information services not considered in the user-studies. However, aside from the initial visionary prototype, the technology-centered approach largely provided input for refining the implemented design rather than informing original alternatives, thus endorsing a trial-and-error approach (Kjeldskov and Graham 2003).

With careful interleaving however, the two approaches complemented each other in the revision of design sketches and development of further prototypes by providing both playful visions and technology inspiration based on a mix of contextual richness and technical detail. Fig. 7 illustrates this interleaving, describing design as the exploration of the relation between an understanding of current practice (acquired in this case through the interviews and contextual interviews) and speculations about future practice (expressed as the technology prototype, and explored in the evaluations).

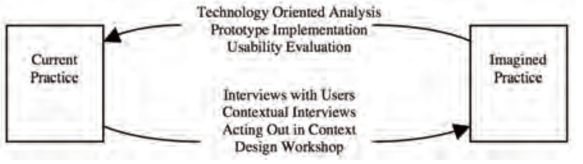


Figure 7. Interleaving user- (bottom) and technology- (top) centered interaction design

In the technology-centered approach, usability evaluation is used as a way of examining the impact, on current practice, of technology prototypes. Design proceeds by 'imposing' the design on the current situation of use during the evaluations, thereby exploring the design's impact on that use. In the user-centered approach, design moves from an understanding of current practice to imagined future use, thus driving design with user needs. Employing both a user- and technology-centred approach, results in a design process that attempts the co-evolution of an understanding of both current and future use, delivering innovative technology that is grounded in use (Howard et al. 2002b).

The technology-centered approach informed the revision of the TramMate concept by identifying those envisioned design solutions that were challenging to implement or simply not feasible. Some examples may illustrate the interplay between the approaches. The critical user need (clear from the user-centered study) of maintaining a close relation between the system and the real world was not to be satisfied by the granularity of current GPS technology (clear from the evaluations of the technical prototype); especially on trams, other means of positioning would be required. Further, studying use of real technology showed that requirements for the user to devote their full attention to the device were excessive, as envisaged in the acting-out sessions. Also, an unexpected amount of effort would have to be put into the design of the maps; including pointers to landmarks, reflecting the orientation of the user etc. Finally, additional functionality was suggested such as manual timetable lookup. The user-centered approach assisted in the revision of the prototype by indicating and prioritizing information needs, outlining realistic use scenarios and proposing a novel interaction design concept integrating the system into another application. In short, the user-studies taught us a lot about what the system should do, while the evaluations of the prototype provided valuable insight into what could be done and feedback on the usability of what a specific design solution did do.

The presented project points towards many opportunities for future research. On a system design level, the revised design concepts should be implemented and evaluated. On the methodological level, it would, among others, be interesting to enquire into the potentials of managing discourse in blue-sky research envisioning future design concepts by introducing functional prototypes into acting-out sessions.

EPILOGUE (OCTOBER 2010)

Looking back at the TramMate project now eight years later, it is interesting to reflect further on the potential interplay between technology- and user-centered approaches in the design of mobile interactions. Taking a step back and comparing the outcomes of the two approaches we were able to provide a meta-commentary on the relative weaknesses and strengths of technology- and user-centred design of interactive mobile systems. This highlighted that each of the two approaches were missing something that the other one had. In short, the user-centred approach was very useful in producing design ideas that accommodated current practice and suggested some radical new concepts. However, the real world usability and feasibility of these could not be evaluated. The technology-centred approach identified important technical constraints and challenges and produced an application that could be evaluated in realistic settings. However, the original design was far less radical, and the usability evaluation of was largely reactive to the current system. Hence, in concert, we found that the two approaches complemented each other well. The user-centred approach grounded the design in current practice, in all its contextual complexity, and resulted in novel design ideas. The technology-centred approach provided a counterpoint in technically detailed expressions of future possibilities.

Both approaches were creative processes. They were, however, different in their type of creativity. Using the terms of Shneiderman (2000), the user-centred design approach was dominated by *inspirationalist* creativity, while the technology-centred approach was dominated by *structuralist* creativity. Inspirationalist creativity is largely intuition-based in concert with preparation and incubation leading to moments of illumination. It promotes techniques such as brainstorming, free association, lateral thinking and divergence to support idea generation (Luther and Diakopoulos 2007) but would also welcome the use of scientific visualization and other ways to help understand a problem space and explore potential solutions. Although the outcome of inspirationalistic creativity is not necessarily art, much art is driven by inspirationalistic creativity. Structuralist creativity is driven by more systematic approaches and highlights the importance of methodical exploration of possible solutions by systematically breaking down the problem and structuring the solution accordingly. It promotes techniques such as flow charts, decision trees, and other structured diagrams (Shneiderman 2000), and would also include software engineering modelling such as object-oriented analysis. Although the outcome of structuralistic creativity is not necessarily science, much science is driven by structuralistic creativity.

This difference in types of creativity, and the fact that both were present and useful in the TramMate example mobile system development process emphasises an important aspect of the mobile interaction and user-experience design discipline. It is a combination of applied art and applied science and therefore also involves different types of thinking, exploring and creating. What is highlighted from the TramMate project (Jones and Marsden 2005) is that Mobile Interaction Design should combine technology- and user-centred approaches in attempts to envision future mobile systems and services. In essence what this means is that mobile interaction and user experience design requires both inspirationalistic and structuralistic creativity and is a challenge of applying art *and* science. But how can this be done in practice? How can the strengths of the two types of creativity be combined? What does it mean to combine user- and technology-centred design? And what techniques can be used to support this?

First of all, the combination of user- and technology-centred design should probably not be done the way we did it in the TramMate project. This parallel and mutually isolated approach was only useful for identifying individual strengths and weaknesses of two approaches to a similar design problem without methodological contamination, but resulted in two designs that were difficult to merge afterwards. If done in parallel, at the very least, the two approaches should be carried out as intertwined rather than separate processes. We would like to make a slightly different proposal though. Looking at the two approaches through Buxton (2007) and Pugh's (1990) lens of design as a "funnel" of concept generation and selection that takes place over several iterations of concept generation and controlled convergence, what really happened in the TramMate project was that each of the two approaches did half of the process from ideation to refined design, but lacked the other half. The user-centred approach had a strong ideation phase but lacked technological realism and refinement. The technology-centred approach was strong in its refinement, but lacked originality in its design. Consequently the two approaches would have made a better joint result if combined *in sequence*.

This sequence of concept generation and controlled convergence from a wide range of design ideas towards detailed refinement is exactly what we believe Buxton's sketching approach offers. The central point here is that user- and technology-centred design can *both* benefit from sketching as a key technique. Sketches can capture outcomes from user-centred activities, as well as outcomes from technology-centred activities. Sketches can obviously facilitate production of a broad range of ideas informed by, for example, user studies, ethnography, etc. But sketches can *also* facilitate and capture outcomes from experimentation with technological possibilities through, for example, interactive mock-ups or smoke-and-mirrors approaches. Following this line of thought, what we labelled as a "prototype" in the TramMate project was in fact not, even though it was created in program code. It was, instead, a design representation produced to facilitate a certain type of reflection and to get a certain type of feedback. It was never meant to evolve into a real system in that exact form (or even in a slightly modified one) but was meant to be a research vehicle externalizing a design idea in order to generate feedback and new knowledge. It was a proof-of-concept implementation – or simply, an interactive sketch.

ACKNOWLEDGEMENTS

Thanks to colleagues on the TramMate project: Jennie Carroll, John Murphy, Jeni Paay and Frank Vetere (user-centered component), Connor Graham, Sonja Pedell and Jessica Davies (technology-centered approach). Thanks also go to professionals who participated in the user studies and to the usability evaluation test subjects.

REFERENCES

Buxton, B. (2007) Sketching User Experiences: getting the design right and the right design. Morgan Kaufmann, San Francisco.

Cheverst K., Davies N., Mitchell K., Friday A. and Efstratio C. (2000) Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences. Proceedings of CHI'00, The Netherlands, ACM

Greenbaum J. and Kyng M. (eds.) (1991) Design at Work: Cooperative Design of Computer Systems. London, Lawrence Erlbaum Associates

Howard S., Carroll Jennie., Murphy J. and Peck J. (2002a) Endowed props in scenario based design. Proceedings of NordiCHI'02, Denmark, ACM

Howard, S., Carroll, Jennie., Murphy, J., and Peck, J. (2002b), Managing Innovation in Scenario Based Design. In Proceedings of Human Factors 2002, Melbourne.

Iacucci, G., Kuutti, K. and Ranta, M. (2000) On the Move with a Magic Thing: Role Playing in Concept Design of Mobile Services and Devices. Proceedings of DIS'00, ACM

Jones, M. and Marsden, G. (2006) Mobile Interaction Design. Glasgow: John Wiley & Sons Ltd.

Kjeldskov J., Howard S., Murphy J., Carroll Jennie, Vetere F. and Graham C. (2003) Designing TramMate - a context aware mobile system supporting use of public transportation. Proceedings of DUX'03, San Francisco, CA, ACM

Kjeldskov J. and Graham C. (2003) A Review of MobileHCI Research Methods. Proceedings of Mobile HCI'03, Italy. Lecture Notes in Computer Science, Springer-Verlag

Kjeldskov J. (2002) Just-In-Place Information for Mobile Device Interfaces. Proceedings of Mobile HCI'02, Italy, Lecture Notes in Computer Science, Springer-Verlag

Luther, K. and Diakopoulos, N. (2007) Distributed Creativity. Creativity and Cognition Workshop on Supporting Creative Acts Beyond Dissemination. Washington, DC. June 2007.

Pedell S., Graham C., Kjeldskov J. and Davies J. (2003) Mobile Evaluation: What the Metadata and the data told us. Proceedings of OzCHI 2003, Brisbane, Australia

Rogers Y., Scaife M., Harris E., et al. (2002) Things aren't what they seem to be: innovation through technology inspiration. Proceedings of DIS'02, London, ACM

Pugh, S. (1990) Total design: Integrated methods for successful product engineering. Addison-Wesley, Reading MA.

Sato S. and Salvador T. (1999) Methods & tools: Playacting and focus troupes: theater techniques for creating quick, intense, immersive, and engaging focus group sessions. Interactions of the ACM, 6 (5), pp 35-41

Shneiderman, B. (2000) Creating creativity: user interfaces for supporting innovation. ACM Transactions on Computer-Human Interaction, 7(1): 114-138.

Vygotsky L. (1978) Mind in Society. MIT Press, Cambridge, MA

Chapter 8

Socio-physical design

Jeni Paay, Jesper Kjeldskov, Steve Howard and Bharat Dave

Abstract. As urban environments become increasingly hybridized, mixing the social, built and digital in interesting ways, designing for computing in the city presents new challenges – how do we understand such hybridization, and then respond to it as designers? Here we synthesize earlier work in human-computer interaction, sociology and architecture in order to deliberately influence the design of digital systems with an understanding of their built and social context of use. We propose, illustrate and evaluate a multi-disciplinary approach combining rapid ethnography, architectural analysis, design sketching and paper prototyping. Following the approach we are able to provide empirically grounded representations of the socio-physical context of use, in this case people socializing in urban spaces. We then use this understanding to influence the design of a context aware system to be used whilst 'out on the town'. We believe that the approach is of value more generally, particularly when achieving powerfully situated interactions is the design ambition.

1. INTRODUCTION

1.1. The Emergence of Hybrid Environments

Pervasive computing is increasingly becoming a part of our everyday lives: at work, at home, and out on the town. It is blurring the boundary between physical, social and digital layers of our inhabited spaces, providing users with highly localized contextual information. Our physical, virtual and social worlds are colliding, merging and coordinating (Rheingold 2003). We operate in the built environment using a combination of fixed devices, such as digital information screens embedded in the fabric of the environment, and hand held personal mobile devices, such as mobile phones.

Originally published as Paay, J., Kjeldskov, J., Howard S. and Dave, B. (2009) Out on the town: a socio-physical approach to the design of a context aware urban guide. *Transactions on Computer-Human Interaction,* 16(2), 7-34.

These digital devices provide both communication and computation capabilities, and therefore offer an exciting opportunity for computing to augment and enhance the way that we socialize.

Pervasive computing exploits our familiarity with the everyday environment (Dourish 2001) and breaks down the traditional mapping between activities and places, allowing people to be continually present in every place (Agre 2001). For example, cafes become corporate meeting rooms as users deal with business calls over lunch, corporate meeting rooms become social arenas while participants text loved ones unobtrusively, streets become guided walks, and plazas become information kiosks, all without any changes to the built fabric.

As our lived world becomes increasingly hybrid physical, social and digital spaces (Graham and Marvin 1996, Mitchell 1995), the intersecting issues of spatial context, sociality and pervasive digital technologies need to be understood in order to design for their inhabitant's interactions (Agre 2001). This introduces a new set of issues for analysis, including studies of the connections between physical and social space, reconsideration of existing design practices, and extension and enhancement of current HCI and CSCW methodologies (Ciolfi 2004). By understanding the influences of both the physical environment and the human activities that unfold in that context, designers will be better equipped to provide specialized computation to support likely situated interaction (Ciolfi 2004, Dourish 2001, Erickson and Kellogg 2000).

1.2. The notion of "layered" space

There are three distinct layers of space in this approach: physical, social and digital.

In the physical, architectural design has traditionally taken place within the context of an explicit set of physical and social issues in respect of anticipated activities and historical expectations of certain building types (Agre 2001, Mitchell 1995). In architecturally designed environments people make assumptions about the kinds of activities and social interactions that are supported there. This is partly determined by the physical affordances of the environment and partly determined by people's prior experiences. Physical spaces are formed to support the way that people do activities, and similarly, people's situated interactions in an architecturally designed space are shaped by cues in the physical environment about what is possible there.

In the social, our shared understanding of the physical world and the presence and activity of others helps people interpret activity and behavior (Harrison and Dourish 1996). Traditionally, the opportunities of the social layer of space have been determined by qualities of the architectural design of the physical layer. Today this configuration also includes digital elements, which by their ephemeral nature facilitate fluidity, serendipity and presence that takes us beyond the limitations of a physically built environment. In the design of hybrid spaces, architecture has acquired a new, digital, layer of expression and design extending its capabilities to facilitate and organize social interactions (McCullough 2004).

In the digital, pervasive computing is a relatively new phenomenon strongly influenced by the uptake of mobile computing technologies facilitating people's social life outside the work domain (Rheingold 2003). By extending what the built environment offers its inhabitants, pervasive computing provides new opportunities for sociality (Jensen and Lenskjold 2004). People who are digitally connected to each other, and to the elements of the city, are now less reliant on fixed signage. They use technology to deliver relevant information 'just in time' and 'just in place', guiding them to where they want to go and what they want to do. By allowing people to make inferences about the activities of others, digital systems create environments in which new social forms can evolve (Erickson and Kellogg 2000). This digital layer not only helps structure our social interactions, but also provides a medium for facilitating and enriching everyday interactions between individuals (Erickson 1993).

In this study, the built environment is viewed as comprising these three interrelated layers: 1) the physical (material) layer comprising the architecturally designed buildings, structures, paths, signage and spaces; 2) the social layer comprising social interaction between people moving around that space, queuing, gathering, meeting etc; and 3) a digital, context aware, layer. We advocate that the digital layer should be designed based on a rich understanding of the physical and social layers of a space. Only through this will it form a direct relationship with the existing social and physical context of use.

1.3. Context Awareness

Design is about creating a 'fit' between form and context (Alexander 1964). One way of creating such a fit is to make systems context-aware, automatically adapting to the setting in which it is being used. This could be a mobile phone that automatically switches itself to silent when the user is in a meeting, or a laptop that automatically adjusts to local time. Within pervasive computing context awareness is an area of research that has received a great deal of attention. Many prototype systems have been developed and evaluated, and it has been demonstrated that value can be added to the user experience by adapting information to contextual factors such as people's location. However, a context-aware device situated in an environment should be aware, not only of its location, but also of factors like the physical and social context of that location (Agre 2001, Bell and Dourish 2004, Bradley and Dunlop 2002, Dourish 2004, Cheverst et al. 2001, Goodman and Gray 2003, McCullough 2004, Schmidt et al. 1999, Tamminen et al. 2003).

The details of what constitutes physical and social contexts of use are not well understood, at least not in a way that provides design traction. Technology rather than user studies are currently driving the development of computation and communication systems pervading our physical and social worlds (Mitchell 1999). Current software development methods and design techniques could be augmented with methods that provide detail and thoroughness in terms of understanding human experience of physical space and of the situated social interactions taking place there.

When dealing with context-awareness for mobile and pervasive computing systems, design should be based on field studies of existing situations of use, as done by, for example, Ciolfi (2004), Cheverst et al. (2000b), Paulos and Goodman (2004), and Tamminen et al. (2003). If we want a system to fit well to its physical and social context, we need to understand these contexts and their interrelationships better, and explore how such understanding can be represented in ways that are useful in informing design.

In response to this, the research presented in this paper is grounded in a human-centered empirical study of physical and social context. We demonstrate how understandings of the user's physical and social context can be achieved and represented through a structured socio-physical approach, and how such understandings can then inform interaction design. This is relevant if you are building, for example, situated display or mobile device systems with content that is indexing strongly to static and dynamic elements in users' physical surroundings including buildings and people. This could be, for example, an information system in a train station or airport, a mobile tourist guide for an historic town, or a social networking system for a new housing development.

If the factors we are going to present in this paper are not taken into consideration in the design of such systems, there is a risk of ending up with designs that, at best, do not fit well with their physical and social context of use. Worse, they may simply not get used because they get in the way of people going about their business. They could even impact people negatively by, for example, not giving them necessary information about the socio-physical character of a particular place of interest.

1.4. A Multidisciplinary Approach

Despite many projects looking at issues associated with designing context aware computing (e.g. Borntrager et al. 2003, Bradley and Dunlop 2002, Cheverst et al. 2002, Iacucci et al. 2004, Paulos and Goodman 2004, Randell and Muller 2000, Tamminen et al. 2003), only a few projects have explored the orchestrated use of information presented across multiple sources in the user's surroundings including non-digital ones, i.e. digital, physical and social information (e.g. Cheverst et al. 2000a, Kulju and Kaasinen 2002, Laakso et al. 2003, Vainio et al. 2002). To do this, we need to learn how to create relationships between the user's physical and social surroundings and the information presented digitally (Dix et al. 2000, Dourish 2004, Persson et al. 2003).

The study presented in this paper adapted and combined qualitative research methods to analyze and represent people's understanding of existing physical and social contexts in urban environments. This was then used to derive '*design ideas*' for the incorporation of physical and social context into interaction design. This approach adapted existing research methods from Architecture and Sociology to provide an understanding of physical and social contexts and representations for use in interaction design.

Architectural and urban planning methods can be used to explore the interrelationships between physical spaces and social interaction (Erickson and Kellogg 2000). Architectural research is concerned with the user's experience of the built form in the context of the activities that they are involved in. Representations of physical context can be used as theoretical apparatus to answer questions about the interdependence of technology, space and society (Hillier 1996). Hence, designers of pervasive and ubiquitous computing environments have turned to architecture and urban planning to provide a basis for devising methods for understanding physical environments. The work of architecture and urban planning involves observing how people socialize in everyday spatial environments. This has been used to draw out models and metaphors for incorporating similar combinations of physicality and sociality into digital information systems (Dieberger et al. 2000). The design of urban environments affects the degree to which those spaces encourage social encounters between inhabitants (Ingram et al. 1996). Buildings carry social ideas within their spatial forms; in this way, spaces link to human behavior. Their configuration creates expectations about people that guide our behavior (Hillier, 1996).

Understanding the human experience of built form provides a basis for understanding the context of activities within that space and helps interaction designers to provide digital links between people's activities and their current environment.

Ethnographic methods from Sociology can be used to help developers conceptualize and reveal opportunities for pervasive computing design, and suggest system requirements (Crabtree and Rodden 2004). The study presented in this paper shows that the creation of analytical and conceptual frameworks resulting from ethnographic field studies can be used to sensitize designers to the social aspects of technology use, and support the design of relevant and appropriate technologies. In terms of conceptualizing opportunities for such design, sociological research has shown that people are constantly communicating social cues. This is so that others can perceive our social networks by the patterns of activities and the affiliations that we have (Donath 1996). As part of our social identity we have a way of aligning ourselves with particular groups (Goffman 1963). Being aware of others and the activities that they are involved in influences the choices we make about our own activities. We are also aware that the activities that we are involved in provide information to others (Erickson 2002).

By understanding how people operate socially in public places we can identify opportunities for useful digital augmentation of these spaces.

2. CASE STUDY: UNDERSTANDING PHYSICAL AND SOCIAL CONTEXT

For the purpose of understanding physical and social context, we chose a compelling urban environment, a newly opened and geographically delimited civic space in the city centre of Melbourne, Australia called Federation Square (Figure 1) (Paay 2005).

Figure 1. Images from Federation Square, Melbourne, Australia

Federation Square was chosen because it was a relatively new civic structure, opened to the public in October 2002. It covers an entire city block and provides the people of Melbourne with a creative mix of attractions and public spaces for socializing including restaurants, cafes, bars, a museum, galleries, cinemas, retail shops and several public forums. In just a few years, Federation Square has become a highly popular place to socialize for all Melbournians. It is open from early until late, every day of the week, and it hosts a rich range of planned and ad hoc activities. Located in the centre of the city, on major tram routes, and adjacent to the main train station, Federation Square is easily accessible, is considered a landmark in itself, and is a convenient place for people to arrange to meet up in the beginning of a night out on the town. One of the design intentions for the public space of Federation Square was to incorporate digital technologies into the building fabric, creating a combination of virtual information space and physical building space for people to experience.

2.1. Investigating Physical Context

Physical context, as characterized by both Agre (2001) and McCullough (2004), consists of architectural structures and elements of the built environment that people use in every day life to orient themselves and to operate in that environment. This includes the use of landmarks as reference points, identifying legible pathways in the landscape as indication of the way to go and reading the design of doorways as places to enter. This physical context is created in response to the situated activities that occur there (Erickson 1993), and with regard for human perception of that place. For example, a landmark only becomes one in response to use as a reference point by people inhabiting that space, or a place description such as “the sitting steps” only has meaning through an understood activity that occurs there. The investigation of physicality of an environment, that is, our physical interactions with the world (Dix 2004), provides a practical understanding of physical context.

An investigation of physical context was conducted to understand the physicality of urban space as defined by the material elements of an urban environment that contribute to visitor experience of an urban space. It involved the identification of important characteristics of the physical context of an inhabited urban environment, and the creation of an analytical abstraction useful for informing interaction design. This is a somewhat novel approach in HCI to the problem of understanding context in urban space for interaction design of pervasive computing. Sociological observational studies have been made of people inhabiting urban spaces (e.g., Whyte 1980) and conceptual models have been developed to capture the nature of digital cities in urban planning (e.g., Graham and Marvin 1996). An HCI methodology developed for creating an analytical representation of people’s understanding of urban environments for the purpose of interaction design could be built on these studies.

Our investigation of physical context, PIA (Physical Interaction Abstraction), resulted in visual representations representing the physicality of an urban environment including a layered map diagram. PIA combines two existing methods from the disciplines of urban planning and architecture for analysis of space: 1) an environmental image map identifying landmarks, districts, nodes, edges and paths, a representation devised by Lynch (1960); and 2) an analysis of space using the sketches and descriptions from Alexander et al.’s (1977) Pattern Language.

Lynch (1960) developed a method for visual analysis of city precincts through descriptions of key aspects of the space held by people as they navigate and orient themselves within city precincts. This was done by diagramming the interplay of visible elements in the environment that contribute to a person’s environmental image of a place. From these studies grew the categories of landmarks, districts, nodes, edges and paths as key descriptors of the image of the city held by its inhabitants. The method has proved successful at assisting in the analysis of types of elements of a city, how they are put together, and what makes for strong identity. It has also proved to be a useful technique for predicting the probable public image of that city.

Alexander et al. (1977) empirically investigated the interplay between architectural space and its inhabitants and identified architectural design problems in context and their impact on inhabitants of that environment. Drawn from observations of historical solutions to common design problems, he created a method of analyzing aspects of the built environment. This led to a collection of 253 hierarchically ordered patterns of plausible solutions making up a Pattern Language for design. Each pattern consists of photographs, sketches, descriptive explanations detailing the context for the pattern, its relationship to parent patterns, a description of the problem, the empirical background of the pattern, evidence for its validity, and the design solution.

The investigation of physical context began with an exploratory study of the physical elements of urban space at Federation Square. The analytical methods of Lynch (1960) and Alexander et al. (1977) were combined and adapted to provide a novel method for analyzing and representing qualities of physical space to provide a story about physical context of an urban environment. Lynch's method provided guidance on techniques for conducting an audit of physical elements of a space. It defines the following classification categories for those elements: district, landmark, node, path, and edge. Alexander's patterns provided a window on recurrent and complementary 'fit' between functional and spatial patterns. As Lynch and Alexander et al. both viewed their analyses of built environments from the perspective of the people who inhabit those spaces, our physical audit was undertaken from the perspective of people's perception of and interaction with the physical elements.

An observational expert audit, based on the expert audit from Lynch's (1960) method, was undertaken in single field visit to Federation Square. In Lynch's method, an architecturally trained observer maps in-situ the presence of various elements of the physical environment to create an environmental image map. In the adapted method of this investigation, the architecturally trained observer recorded through photographs and field notes the elements of the physical environment for later mapping and classification using content analysis (Neuman 2003). A total of 124 photographs recorded the material elements (building fabric, cladding, structures, surfaces, building elements, entire buildings, public spaces, paths, entrances, media screens, etc.) of Federation Square for the purpose of documenting the physical elements in the "public" areas of the space. The location, from which each shot was taken, including direction faced, was recorded on a map of Federation Square. Human activities associated with physical spaces, and the way people were using and responding to elements, were recorded in field notes. This documented every architectural element in Federation Square and its relationship to its surrounding context, including the people who inhabited the space (see Figure 2).

Figure 2. One out of the 124 photographs and notes on physical elements of Federation Square

2.2. Investigating Social Context

Social context, as characterized by Dourish (2001), includes interaction with and the influence and behavior of people in an environment. Dourish (2004) regards context as a central concept in social analyses of interaction and that social and cultural factors affect how the user makes decisions about actions and interprets a system. In understanding the social context of a place there needs to be a way to understand the social processes and human activities through studying everyday interactions. The sociality of a place reflects the social interactions that occur there and investigation of this sociality situated in an environment should provide a practical understanding of social context.

An investigation of social context of an urban space using the SOPHIA (SOcial PHysical Interaction Analysis) method was initiated to inquire into social interaction in that environment. It identified those aspects of a person's social environment that represent their understanding of social context (Paay and Kjeldskov 2008). The outcome from this investigation is a conceptual framework representing situated social context in the urban environment of Federation Square in Melbourne, complementing the understanding generated from the PIA analysis of physicality.

The PIA and SOPHIA approaches are different in the sense that the PIA map emphasizes the physical context for understanding socializing in an urban environment while the SOPHIA table provides an hierarchical representation of activities emphasizing the social context of being physically situated in an urban environment.

Our investigation developed an approach, including the SOPHIA and PIA methods, for gaining an understanding of the social context of an urban environment using rapid ethnographic methods (Millen 2000). This related social interactions to the physical environment in which they were taking place. Understanding the physical aspects of human experience of spaces can best be achieved through studying people situated in place (Ciolfi 2004). To support our approach, the work of McCullough (2001) provided insight into the situated nature of social interactions through his typology of everyday situations. This provided a framework from which to view social interactions as related to the situation in which they occur, a view in turn influenced by the physicality of the space. McCullough (2004) draws together the concerns of architects and interaction designers by acknowledging that interaction design for pervasive computing has a direct relationship with and impact on the environment and the inhabitants of a place. As an example, digital devices can give you social information you can't see physically, such as where a crowd has gathered outside or beyond your immediate field of view.

The method of investigating social context involved observation in the built environment to identify the *social affordances* of a space. This included, for example, where people tend to go, where they tend to gather, and what they tend to do at different places etc., which contribute to the people's understanding of that space.

This method takes a grounded approach to understanding the existing social situation. It involved accompanying three groups each of three people (9 participants in all, mixed gender, young urban professionals aged between 20 and 35, and all familiar with the location), on a typical social visit to the Square. The aim was to observe and record the group interactions using McCullough's (2001) typology of 'on the town' everyday situations as a theoretical lens through which to view and guide social interactions in the field, and as a sensitizing concept in analyzing social interactions in urban environments. The categories of interest to this investigation were: eating, drinking, talking (places for socializing); gathering (places to meet); cruising (places for seeing and being seen); belonging (places for insiders); shopping (places for recreational retailing); sporting (places for embodied play); attending (places for cultural productions); and commemorating (places for ritual).

Participants were required to be familiar with Federation Square, thereby acting as key informants (Millen 2000), capturing representative interactions of an established social group. Each field visit was used to observe the social interactions of the group, the activities they participated in and how they were affected by physical space and the presence of others. Participants were instructed to go about their usual socializing practices in Federation Square. The group determined the activities undertaken and the social interactions that they engaged in. Contextual interviews from the contextual design method (Beyer and Holtzblatt 1998) were used in the field, combined with observational ethnographic methods.

Prior to each field visit the group received a 10-minute introduction to the investigation followed by a 20-minute interview about general socializing experiences as a group. This introduction occurred at a place familiar to the group, where they might meet before socializing in the city. This encouraged them to reflect on past social interactions, to relax about the visit, and gave the interviewer insight into the situated interactions that the group typically participated in. At the start of each field visit, one member of the group was taken to Federation Square and asked to contact the others to meet them there. This was designed into the method so that their meeting up processes could be observed. All groups used mobile phones for contacting their friends, as they usually would.

The contextual interviews and observations lasted approximately three hours for each visit, allowing the group to participate in many varied activities (or situated interactions) during the field visit. Reflection on past visits by participants gave additional data about their responses to alternative activities in that space, without taking the time to do them during the visit. This reflection also gave access to the group's history of interactions in the space. The total amount of time spent in the field with the three groups was 11 hours, which proved to be a sufficient amount of time to observe that very few novel observations were occurring with the last group. The outcome from the field visits amounted to 1) approximately eight hours of digital video (Figure 3) recording situated interactions, all questions and responses, the initiation of activities and movement of the group around the square, 2) notes of ethnographic field observations, and 3) diary of reflections on visits recorded immediately after each visit.

Figure 3. Video recording the social study in Federation Square

3. REPRESENTING PHYSICAL AND SOCIAL CONTEXT

3.1. Representing Physical Context

Before content analysis of the photographs collected in the field commenced, coding of the elements of each photograph was done by assigning one or more of Lynch's (1960) five categories (landmarks, districts, nodes, edges, paths) to the focal element in each image. Using the recorded locations of the photographs, these coded elements were then used to create an environmental image map for Federation Square (see the grayed under layer of Figure 4) showing landmarks, districts, nodes, edges and paths as perceived by the architecturally trained observer.

In addition to this, one or more of Alexander's (Alexander et al. 1977) 253 patterns of the Pattern Language were associated with each photograph, which was then annotated with the pattern number and pattern title of each associated pattern. Sketches and notes showing the applicability of each pattern were appended to the existing descriptions of each image. Content analysis was then conducted on this data, involving coding it by classifying, sorting and grouping concepts in the written descriptions, and refining the themes emerging from that process. After several iterations of grouping, regrouping, forming sets of words and refining words, a concise set of representative terms emerged. These were influenced by the categories of Alexander et al. (1977) by virtue of the encoding schemas. Each of these representative terms was then related back to the original photographs, giving them the additional quality of locatability. The coding and content analysis was done by one architecturally trained researcher alone, and was subsequently validated by a second researcher experienced with the works and methods of Lynch and Alexander.

PIA Layered Map

The visual representation that came from combining these two analyses was a layered map diagram. This showed an overview of the different Lynchian-based elements evident in the space (drawn as a grey sketched 'under layer') and the related, and specifically located, Alexandrian-based descriptions of the space (placed as a text over layer). This map, depicted in Figure 4, is specific to Federation Square but the method of creating it would be applicable to the analysis of other spaces in a similar way to derive a visual representation of an urban environment. The PIA layered map represents an abstraction of the physical characteristics of Federation Square. It provides a visual overview of the inhabited physical context of that space.

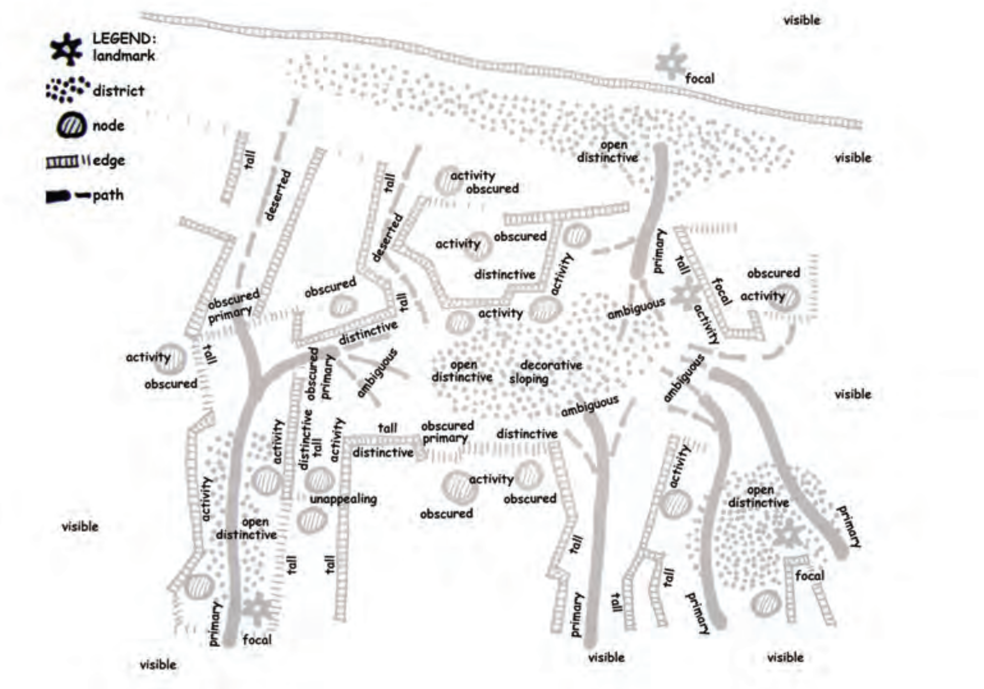


Figure 4. Layered map of Federation Square generated from the PIA approach

Surveying the layered map, it is possible to draw summary conclusions about the space, which would not be evident from viewing the original data, or from merely visiting the space. This is because it represents a composite view of that space, judiciously extracted from historical understanding of human experience of physical space through observational expert audit and analysis of urban space.

The PIA layered map provides a specific representation of those key contextual factors that characterize the environment at Federation Square (see Table 1), including the following key understandings of its physicality:

- Federation Square has: four key districts with distinctly different characteristics, each with an associated landmark.
- Federation Square has: open spaces with activity edges; distinctive and tall structures and walls; and obscured places where activities are also happening - so it is not clear from the middle of the main spaces what is around, and how to find where you want to go.
- Federation Square has: visible surrounds; some primary paths that become ambiguous when they lead into open space or are deserted or unappealing when wedged between tall walls; primary entrances that are obscured; and focal structures - so people need to use the structures and surrounds in finding their way around the space.

Table 1. Understandings from PIA related to Design Ideas

	DESIGN IDEA:	1. Location by District	2. Augmented Photos	3. Rich Descriptions	4. Use of History	5. Wayfinding	6. People and Activities	7. Meeting and Waiting
UNDERSTANDING: PIA								
Landmarks:	structures-focal	●						
Districts:	spaces-distinctive	●						
	spaces-open		●					
	surrounds-visible			●				
	ground-sloping							
	ground-decorative							
Nodes:	structures-distinctive			●				
	structures-tall			●				
	places-activity						●	
	places-obscured		●					
Edges:	edges-activity		●	●				
	entrances-obscured		●					
	entrances-primary		●					
	walls-distinctive			●				
	walls-tall			●				
Paths:	paths-primary					●		
	paths-ambiguous			●		●		
	paths-unappealing							
	paths-deserted							

PIA's significance lies in its ability to capture, in a readily accessible form, inhabited physical context. It makes available a visual representation describing inhabited space that can be used for identifying the key physical characteristics of any built environment.

This is done grounded in human observation of that place and formed with reference to collected knowledge about human understanding of architectural form.

The PIA approach is novel, in so far as it is a combination of two well respected architectural schools of thought that have both been applied in HCI before, but not in combination, and not with such adherence to their architectural origins.

3.2. Representing Social Context

The eight hours of the digital video collected during the investigation of social context at Federation Square resulted in a 60-page transcription of situated interactions 'out on the town'. General conversations about participant's families and work issues were not transcribed and were regarded as being outside the bounds of this investigation. The analysis of the transcript involved open and axial coding from the grounded theory methodology (Neuman 2003, Strauss and Corbin 1990) and the affinity diagramming method from the contextual design methodology (Beyer and Holtzblatt 1998).

Open coding is that part of the analysis concerned with identifying, naming, categorizing and describing phenomena found in the text by assigning codes to them. This involved underlining key words and repeating phrases in the transcript, identifying the phenomenon that participants described, and coding that as a category-property-dimension triplet. For example, a conversation about how a group chose to meet up resulted in the code: "MEET PLACE – choice – familiar to all". These codes were then entered into a table, recording when in the transcript this phenomenon occurred, and a longer description of it, supporting the meaning behind the code. During this process, codes were modified and merged as related and similar situations were found in the remainder of the transcript. After coding of the transcripts, category names were consolidated and refined using McCullough's situated interactions as a theoretical lens. This resulted in 214 distinct codes, grouped under the following 17 descriptive categories: Contacting, Meeting, Approaching, Entering, Eating, Drinking, Sitting, Watching, Being Seen, Attending, Viewing, Shopping, Deciding, Directing, Walking, Exploring and Locating. In the previous example, the code became: "MEETING – place – familiar", and was grouped with many other phenomenon related to people meeting up when socializing.

Axial coding is the part of the process that relates codes to each other using a combination of inductive and deductive thinking. To achieve a higher level of abstraction with our data, axial coding was used to draw a set of overarching themes from the outcomes of the open coding process. For inter-coder reliability, two researchers reviewed the codes in respect to the underlined transcript to identify overarching themes emerging from the data. The two sets of themes produced were then merged and consolidated as a joint activity to produce a group of 21 agreed themes as the outcome of the fieldwork. These themes formed the starting point for the process of further abstraction, using affinity diagramming (Beyer and Holtzblatt 1998).

Affinity diagramming drew successively higher levels of abstraction from the data through a process of grouping and sorting the 21 themes until a set of three high-level key aspects emerged. These represented the essence of the data and encompassed all lower level themes as concepts supporting these key aspects. This resulted in a conceptual framework encapsulating a structured understanding of the context of every day social interaction providing a rich story of sociality in urban environments showing how people experience physical space and how they interact with each other while socializing in these spaces.

SOPHIA Conceptual Framework

The conceptual framework generated through the SOPHIA approach, provides a hierarchical summary of key social characteristics affecting social interaction in urban space. It consists of three key aspects of social interaction in urban environments: knowledge, situation and intention. These key aspects and their related concepts tell the following story about situated social interaction.

Prior Experience and Expectations: When interacting in urban space people draw on their knowledge in the world. They recognize entrances, and they see large open spaces as places for people to gather. They use landmarks as reference points. People operate in public places using a set of social affordances. They look to what others are doing as cues for what to do in a place. Following crowds or people queuing is a way to decide where they might go. Places where others are sitting make them feel they can sit there too. People draw on their history with a specific urban environment. Physical familiarity with a space means that they approach familiar places using familiar paths, that is, the way that they "*usually come*". They use past social experience of places as a basis for selecting places to socialize with friends this time, for example "*let's eat where we ate last time*", or require trusted recommendations to try a new place. Often they will have a personal preference for why they choose a particular place. People relate activities with establishments, that is, a place for drinking, based on their past experiences with it. If it is a place where they "usually sit outside" then it becomes the place to go in fine weather. If they are socializing with a particular group of friends they like to start the social outing in a place that they share a common experience of and often arrange to meet in the place where they "*usually meet*".

Situations, places and spaces: Situation is an important aspect of sociality in urban space. When socializing the presence of other people influences the way that people behave and move through urban space. Friends show they are a "group" by maintaining close physically proximity, e.g., walking abreast, as they move through a public place. People like to be near others but not necessarily interacting directly with them. For example, they like to share a table with others in a bar, yet not talk with them; they are "*socializing by proximity*". People like to watch others, especially if they feel unobserved themselves. The length of time that someone has to wait for a friend influences the choice of meeting place. The setting in which a particular activity takes place matters. The presence of others and the types of people in a place influences its acceptability. Generally, people like to socialize in places with similar types of people, i.e., age, dress, intentions. Whether a place is sunny, sheltered, etc., influences the choice of location to socialize or wait. The convenience and location of a place is also important. People prefer a place to eat that is near other places to eat. Surroundings are an important part of people's situation and are often used as reference points. They index to things around them, including buildings, e.g. "*the railway station*", or distinctive elements, e.g. "*that big white umbrella*". They describe a location of an unknown place to a friend by referring to the places and activities of shared experience that they hold with that person. People describe a location as "*next to the place we went where we sat under those heat lamps*". They might also refer to a place in terms of a past event there, i.e. "*where we saw the world cup*".

Sensing-making: Sense-making is an important part of socializing in a place. People try to size up the situation. They like to get an overview of their environment. They strive to make sense of things and places around them. People make sense of what is happening in a place by assessing the activities of others. Before entering a place they tend to stand on the outside and familiarize themselves with the situation before committing to enter or join in. People gather information in an urban environment while socializing, and require differing levels of information for different activities. They view information/media screens as decoration, and if they have a query, they are most likely to ask a friend. They don't like to interact in places where they are unsure of how things operate. They want to know what is new and if there is something special happening in a place, especially their familiar places. Movement through an environment is part of their social activity. People explore places just for the fun of it, often wandering and browsing without a specific goal exploring both physical space through movement and shared knowledge through conversation. At other times they are trying to find their way to a specific place, involving transition through spaces preferring paths that have people and activities of interest along the way. Places are dynamic, and familiar and preferred paths are sometimes blocked or altered by the presence of ad-hoc structures or large crowds. People can get lost when taking an unknown route and get frustrated when signage is not helpful. In this situation, they look ahead for familiar objects. Friends spend time negotiating on places to go, and will make decisions by discussing options until they reach consensus, or someone leading the group.

This outcome of the SOPHIA analysis of social context, represented as a conceptual framework, provides a hierarchical representation of the social layer of an urban environment, and can be seen in Table 2.

This conceptual framework represents social influences at work at Federation Square, including the following key understandings of situated social context:

- At Federation Square: people's past experience with places and people (familiar places and shared experiences) and the situation of these experiences are important in choosing places and activities to socialize;
- At Federation Square: people give directions by referring to shared experiences and visible elements, and use their history and physical familiarity with a place to find their way around using familiar paths;
- At Federation Square: people like getting an overview of what is happening and want to know about the presence of other people in places and what they are doing; and
- At Federation Square people typically coordinate meeting up with friends in an ad-hoc manner, depending on activity and shared history with those friends.

Table 2: Understandings from SOPHIA related to Design Ideas

		DESIGN IDEA:	1. Location by District	2. Augmented Photos	3. Rich Descriptions	4. Use of History	5. Wayfinding	6. People and Activities	7. Meeting and Waiting
UNDERSTANDING: SOPHIA									
Knowledge: in-the-world	physical affordances	places to enter	●						
		places for gathering							
		landmarks as focal points	●						
	social affordances	cues for what to do						●	
		cues for where to go						●	
Knowledge: history	physical familiarity	familiar paths			●		●		
		familiar places				●	●		●
	social experience	past experience				●	●		●
		shared experience				●			●
		recommendations from others				●			
		personal preferences							
Situation: people	us and them	interaction by maintaining group							●
		interaction by proximity						●	
		interaction by watching						●	
		discomfort of waiting						●	●
Situation: setting	setting matters	others (social)				●			
		environment (physical)				●			
		convenience to current location				●		●	
Situation: surroundings	indexing to surroundings	index to shared knowledge				●			
		index to visible elements		●					
		index to events							●
		index to physical objects		●					
Intention: sense-making	sizing up the situation	getting an overview						●	
		pausing before committing						●	
		making sense of a place						●	
		making sense of what's happening						●	
	gathering information	different levels of information		●					
		media screens as decoration							
		what's new		●					
		uncertainty (lack of info)		●					
Intention: movement	exploring	exploration for the sake of it						●	
		wandering and browsing							
	wayfinding	transition through spaces					●		
		dynamics of a place					●	●	
		getting lost (unclear signage)					●		

Some of the observations described above were surprising to us. For instance, we were surprised about the observed influence of familiar places and paths on peoples' socializing and navigational behavior. People would rather revisit well-known places than exploring new ones, and would knowingly prefer the long way between two places to a newly found shorter path. Related to this, we were also surprised about the importance of peoples' history of socializing in Federation Square and how past and present interactions were not perceived as a random set of disjointed events, but rather as interwoven parts of a continuous experience over time. In terms of the way people communicated about places outside their immediate view, we were surprised with the extent of references made to activities and earlier interactions there, rather than to its physical properties. As a final example, we were surprised about the huge importance of social affordances of places when venturing into unfamiliar areas.

Many of the findings presented above involve, and relate, social and physical aspects of context. It is our belief that these findings would not have been noticed as strongly without SOPHIA. The SOPHIA approach can be used to analyze any urban environment and provide an analytical representation of that urban space in respect to those elements of situated social contexts that are most strongly represented there. It makes available a set of concepts representing sociality in urban space that can be used to identify key social characteristics of any built environment. It describes the user's social situation in a way that is grounded in human observation of people socially interacting in place, collected through ethnographic study of situated social interaction.

SOPHIA is unique, in so far as it represents a grounded approach to providing a widely sought representation of social context in urban environments.

4. INFORMING INTERACTION DESIGN

Using the PIA and SOPHIA methods and resulting representations of the physical and social layers of an urban environment, a pervasive computing prototype was designed for the intangible goal of 'enriching people's experience of Federation Square'.

The design process used in this study involved identification of a method, where no generally accepted one exists (Ciolfi and Bannon 2003), for taking knowledge gained during a grounded analysis of context of a space through to the specification of design requirements for a prototype system. The process of transition from field data to prototype design is a difficult one (Cheverst et al. 2005, Ciolfi and Bannon 2003, Kuutti 1996). The study presented here used a method of drawing design ideas from the PIA layered map and SOPHIA conceptual framework using a technique of design sketching (Buxton 2007) to make this link between the analysis and design processes. Two researchers reflected on these field investigation outcomes and used design sketching to extract design ideas from them. This method is a combination of idea sketching as used in, for example, architectural design (Yee 1997), interaction design (Buxton 2007, Sharp et al. 2007), and software design (Checkland 1981), and empirically grounded identification of considerations relevant for design (Ciolfi and Bannon 2003).

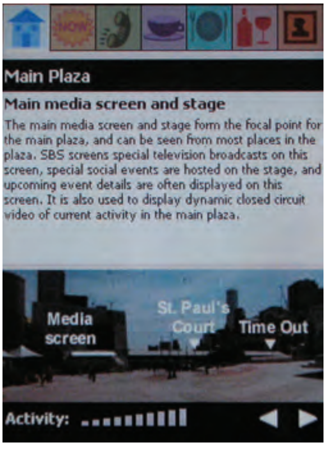
During the design process, the PIA layered map and SOPHIA conceptual framework were continually revisited and used to inspire seven design ideas. Using the design processes of storyboarding and paper prototyping, these design ideas evolved into the design ideas that guided the creative design of an operational prototype. This prototype was then used to verify the ability of conceptual frameworks to inform the design process and also to evaluate the usefulness and understandability of references to the user's current context in the human computer interface. The following seven design ideas identify key aspects of the understandings of inhabited physical context and situated social context of the urban environment studied:

1. *Location by District*: The system responds to the users' location in terms of one of the defined districts from the PIA layered map. Importantly the understanding of location is imbued with social meaning. The information provided by the system is tailored to information needs within that specific district. To help the user relate the information in the interface to their physical surroundings, the initial screen displays the corresponding landmark for that district, using physical indexicality to align the system with the real world.
2. *Augmented Photorealistic Depictions*: Each district is represented in the system by an interactive photorealistic depiction of the physical surroundings of the user augmented with textual or symbolic information needed to better understand the place. The outcomes from PIA tell us that the space has activity edges but primary entrances and destinations that are obscured. The interaction design matches the user's experience of this physical space and facilitates aligning the information presented in the system with their physical surroundings. The augmentation of these images helps people to know what is located behind the visible facades, and to identify primary entrances serving several different places.
3. *Rich Descriptions for Navigation*: Locations and instructions for navigation are expressed through rich descriptions derived from the distinctive characteristics of the place. Based on knowledge generated from PIA that activities are located on the edges, it follows that selected locations could be described as being "next to" or "opposite" other locations. These terms are used in the rich descriptions, thereby referring to locations relative to one another. The navigation information also indicates the path that the user should take to get to a place. The outcomes from PIA show that paths in many areas of Federation Square are ambiguous and not clearly indicated, but the space has visible surrounds. It therefore follows that descriptions, such as "away from the train station" or "towards the river", can be used to refer to visible elements of the surrounding city.
4. *Use of History*: The system keeps a record of the user's history of visits to Federation Square and visits of accompanying friends and uses this to deliver socially appropriate information about things to do and places to go. From SOPHIA we know that people use their history with a place and shared experiences with others when socializing. The past visits to places that they share with particular friends affect where they choose to socialize. Also, the current situation in a place, including environmental conditions and the presence and activity of others affects choices of where to go and what to do. The system makes use of the social experiences and history of the user to give recommendations to them based on this database of past visits.

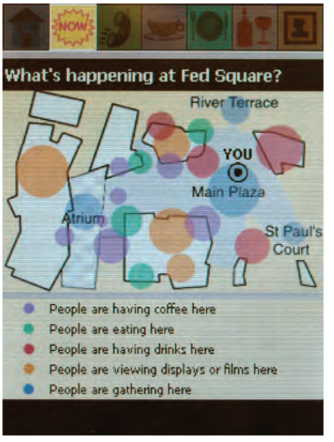
5. *Wayfinding*: The system supports wayfinding using people's familiar paths and indexing to their familiar places. The outcomes from SOPHIA shows that familiar places and familiar paths are important to people. Also that if signage in a place is inadequate people get lost going to new places. A system that knows a person's familiar places can present a series of them, along with key landmarks and distinctive building features as identified through PIA, to guide a person to an unknown destination. The person will use familiar paths to get to each point thereby reducing the need for detailed step-by-step movement directions. This uses people's social experience and history with a place.
6. *Representation of People and Activities*: The activity and location of others in Federation Square is represented to the user so that they can make activity choices based on assumptions that they make about this information. Representation of what is happening, including both people and activities, helps people in making sense of a space. The outcomes from SOPHIA show that interaction between the group and others by proximity and by watching is important when socializing. There is a desire to know where other people are gathered. People are drawn to new places by the presence of others, or explore where others are out of interest. This lets users "see" important aspects of the current social context of a place that they would be otherwise unable to access from their current location.
7. *Meeting and Waiting*: The process of ad-hoc meeting up with friends is streamlined through the use of familiar places, identified groups of friends and their proximity, and information about how long a person will need to wait. This requires knowledge of the history and shared social experiences of friends. As indicated by the outcomes from SOPHIA, people coordinate meetings based on places familiar to the group, how long they will take to meet up, and what activity the group want to do. This indexes to past social interactions of a group of friends.

These design ideas were directly derived from the understandings generated through the PIA and SOPHIA methods, and can be directly traced to the outcomes from PIA and SOPHIA that inspired them as shown in Tables 1 and 2.

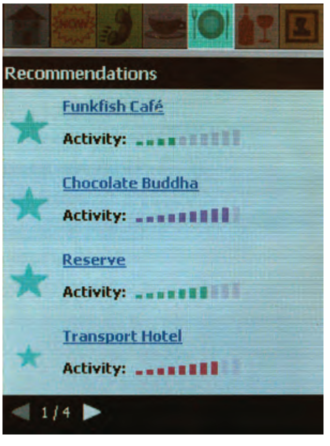
Informed by the seven design ideas, and with respect to the established functionality typically provided by, for example, mobile guides systems, a pervasive computer system prototype was developed for access through mobile devices (Figure 5).



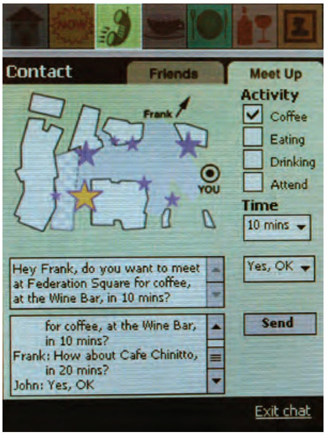
"Home" screen



"Now" screen



"Recommendations" screen



"Contact" screen

Figure 5. The "Just-for-Us" prototype system - four example screens.

The "Just-for-Us" prototype was created as a proof of concept, to evaluate the usefulness of the design ideas emerging from PIA and SOPHIA in situ. The four example screens shown in figure 5 have design elements that can be directly linked to the design ideas detailed above. For example, the *NOW* screen was inspired by design idea 6, *Representation of People and Activities,* where the location and activities of other people within proximity are visually represented on a dynamically updated map. This represents the current situation, allowing people to make sense of a place through the social affordances provided by the presence of other people.

When a user enters one of the square's four districts, the system pushes data about that district, including a panoramic photograph, to the mobile device (the Home screen). Clicking on an annotation brings up a brief description of the item. By clicking on the arrow icons at the bottom of the screen the user can rotate the view and learn about other locations in the area. When the user enters a new district the corresponding panoramic photograph is automatically pushed to the device.

Clicking on the Now icon at the top of the home screen brings up a small map showing the user's approximate location and dynamically updated colored circles. The radius of each circle indicates the number of people present, while the color represents their prevalent activity - for example, having coffee, eating, having drinks, or attending a cultural event. Clicking on a circle calls up detailed descriptions and images of the place, for example, a particular restaurant, along with information such as way-finding directions and menus.

Just-for-Us was implemented as a server-side web application that can be accessed through a mobile web browser. For the first prototype, we used a series of HP iPAQ h5550's connected to the Internet through WLAN or a GPRS connection. The content of Just-for-Us is generated from a mySQL database containing information about the physical layout of Federation Square (derived from the architectural field study). It also holds descriptions and photographs of landmarks and transition points, and information about the different establishments and businesses in the precinct. Additionally, the database is continuously updated with information about people's current context (location, activity, social group, etc.) and keeps a history of their interactions in Federation Square. PHP is used to generate web pages on the basis of the information in the database, and JavaScript is used for handling client-side interaction and information push. Supporting the web application, a number of server-side programs perform specific sub-tasks such as pushing information to the user when appropriate and dynamically generating maps and annotated photographs. The system scans for other Bluetooth-enabled mobile devices to identify nearby friends, and uses Bluetooth beacons in the built environment for positioning. Hence, it does not know peoples' exact geographical coordinates but only their approximated location, for example if they are in a specific café or in the main square. The location of people appears anonymous. For full details of the prototype design and implementation of Just-for-Us see Kjeldskov and Paay (2005).

5. EVALUATING JUST-FOR-US

To investigate the usefulness and understandability of the prototype system, and to extend our understanding of physical and social context in urban environments, an empirical user-based evaluation was conducted. Unlike many other mobile information systems, the proposed design is built on insights into user perceptions of the built environment gained from empirical fieldwork. This facilitate presentation of information in the interface that refers to elements of the user's physical and social contexts.

The evaluations of the Just-for-Us prototype involved 20 participant pairs of mixed gender, with a history of socializing in Federation Square. Ten evaluations took place in the field at Federation Square, and ten in a laboratory. Using a mixture of field and laboratory evaluations made it possible for us to powerfully introduce the situation of use into the assessment (in the field) whilst allowing for a degree of control (in the laboratory); in doing so we asked questions of use in the field, and questions of usability in the laboratory, integrating our insights as we went. As previous research has stressed the value of researcher control in field evaluations (Kjeldskov et al. 2004), users were given a number of overall tasks to prompt use of specific parts of the system that related to the users physical context and social context, in respect to the six design ideas that were implemented in the operational prototype. A set of tasks and field questions were devised to ensure that these parts of the system were evaluated. Supporting this approach, users were asked to validate the relevance and realism of these tasks in relation to the activity of socializing out on the town. Before taking part in a visit, each participant pair jointly completed a history survey of their previous visits to Federation Square to simulate history data that the real system would have collected automatically. For the purpose of the field evaluation, the user's position, people and friends in vicinity etc. were "Wizard of Oz'ed" (Dahlbäck et al. 1993, Buxton 2007) with data being entered manually behind the scenes without the knowledge of the test subjects.

The investigation was an evaluation of use but borrowing techniques from traditional usability studies such as specific tasks, think aloud protocol, and the data collection methods of video and audio recording. Inspired by the co-discovery testing approach to thinking-aloud studies with more than one user (Snyder 2003), pairs were asked to discuss their perception of and interaction with the system with each other. The researcher read the tasks and asked questions about participant's interactions for clarification. Each evaluation took approximately 1.5 hours. The evaluation was documented on digital video both in the laboratory and in the field.

Due to the fact that this was not a theory building exercise but an exploration of the use of the PIA and SOPHIA understandings in the design of Just-for-Us, a detailed grounded analysis of this data was deemed unnecessary. Instead the rapid ethnography method of collaborative data analysis (Millen 2000) with two researchers provided the level of feedback sought from the use evaluations. The collaborative data analysis approach was combined with the analytical technique of identifying critical incidents to produce a list of observations (Sharp et al. 2007) with each observation associated with one of the five major tasks.

The outcome of this analysis was a list of 74 issues related to user experience and comprehension of the system, for example "People want to use the map representation overview to make activity choices". These issues were then associated with specific outcome elements of the PIA and SOPHIA tables to extend the understanding represented by these frameworks.

In terms of inter-coder reliability, a total of 1390 instances of the 74 issues were coded across the 20 participant pairs. Out of these, the two researchers independently identified and coded 1318 matching instances of these issues, which shows a high level of reliability (94.9%).

6. FINDINGS

The user study provided rich data on the use of a public pervasive information system within an urban context. On a general level, the study showed that people could easily operate the system; find what they were looking for and understand the presented information and functionality. They found the design of the system attractive, streamlined and professional looking, and trusted its content to be true. On a more specific level, most users reported that providing a public digital layer of information augmenting the city on their mobile device was "very cool", "useful" and "fun". In particular, people were fascinated that the system knew their current physical location, who they were with, and where other people in the civic space were currently gathering. They were also fascinated by the ability to access information about the places around them from both businesses and from other people, and they perceived the service as a credible source of information augmenting their surroundings.

The user study also provided us with the following detailed findings that can be tied to specific themes captured in the understandings generated through PIA and SOPHIA. Informed by the empirical findings of the evaluation, we can confirm that *landmarks*, that is, features that are distinct from their environment, worked well as anchor points for matching information in the system with information in the real world. Landmarks are an important focal point for people operating in urban environments.

Even though the accuracy of the system's positioning was limited to only knowing what district or place a user was at, rather than knowing their precise geographical coordinates, this proved to be specific enough and matched with people's perception of their current location. Within each district it was quite natural for people to use the visible surroundings for aligning the system with the real world by matching the outline of buildings and distant skylines with images on the screen. In situating themselves people would use distinctive elements and structures to make quick confirmations that they were in the right place using images on their screen to index to their surroundings. We were surprised that while doing this many users stated that they did not need such a detailed image and would be capable of using, would even prefer, a line drawing or an outline with a few detailed features. In doing this matching however, it did surprise us that even though people seemed perfectly capable of identifying and matching on a more abstract level than we had designed for, they wanted the virtual world to automatically correspond and align with their exact orientation in the real world - they expected the system to "know which way I'm facing" and would even relocate themselves a few meters to one side so that this alignment was achieved.

We found that people did treat the space as a series of interconnected *nodes*, that is, places of activity and interest, which they were keen to know more about. Given the distinctive structures of places, we found that using a physical quality of a place as a descriptor in instructions worked very well, for example, telling users to "walk toward the black building" negated the need for detailed distance and vector-based wayfinding instructions. Surprisingly the same success was not achieved using activity as a descriptor, for example, "walk past the sitting steps". Although all participants said they were well aware which steps we meant, they were worried that activity was transient, and not necessarily happening at all times of the day.

Generally people navigated and oriented themselves using the perceived *edges* of the space. Although we understood people to use physical affordances to determine places to enter, we also knew that the space had primary entrances that were often obscured. We found these transition points were vital to people finding their way in the space, and that augmenting an image of the activity edges of the space with text indicating the location of entrances and places supported improved wayfinding in the space.

As noted in the analysis, *paths* in the space were both very large and generally lead into the square and were not clearly indicated inside the space. This did not prove to be something that needed to be augmented by the system because as expected, people used their familiar paths, and therefore only needed fragmented detail to get to the vicinity of a new place, with more specific detail when up close.

In terms of *prior experience* used when socializing with the system, it was confirmed that people do navigate very successfully using physical familiarity. They adapt to fragmented wayfinding descriptions finding their own path as far as possible by navigating to places that the system knows are familiar to the user. However in giving these instructions we found that user's often did not know the formal name for their familiar places, and in this case resorted to viewing and recognizing a picture of their destination. In our evaluation we confirmed that it made sense to people to know about the activity and number of people at places nearby, and influenced their decision to go there. The system provided the kinds of cues, or social affordances needed for deciding what to do and where to go, but for places beyond the users current visual range. Many parts of the system relied on knowledge of the user's past social experience with the space to give directions and recommendations on places to go. Although this was successful in most cases, and people do generally favor returning to familiar places, we found that frequency of visits to a place is not a universal indicator that people want to go there this time. The system would give recommendations based on familiarity and current social situation, and it would surprise users when the system adapted to their history – most users expressed uncertainty about how to control this ability and a desire to do so. In our design we had overlooked that fact that when places are being recommended by a system, people want to know the factors influencing that recommendation. We did however have it confirmed that people regard the favorite place of a friend as a form of recommendation.

The *situation* for socializing was also confirmed as important. Knowing about the activity of people in a place influenced the decision to go there. It was interesting to note that for some individuals a busy area made them want to go there and "check it out", showing an interest in interacting by proximity, with users stating that the presence of others is a sign that a place is good. For others large numbers of people clearly indicated a place to be avoided, although some clarified this choice by telling us that it depended on the mood they were in. It was our understanding from the social study that spatial convenience was also important, but in the evaluation users said that this did not matter – this may have simply been a factor of the small-scale testing space.

People's *sense-making* when socializing was the most difficult to confirm. It was evident in our use study that people do spend a significant amount of their socializing time on making sense of their surroundings and sizing up the situation. People really loved getting an overview of other people and their activities in their surrounding environment. They appreciated having information presented to them about events that were about to happen at places around them – suiting the serendipitous form of socializing that most users engaged in. People also really liked it when the system automatically gave them relevant information about the activity they were about to do at a place. For example, having menu information pushed to their device when in proximity of a café gave them the opportunity to pause before committing to the place, which was perceived less of a commitment than going inside to read their menu. However, what we did not foresee was that at this point users wanted to make a quick comparison between this place and others nearby without moving. Nor did we predict the level of detail that they required, including: food type, price, genre, ambience, outdoor spaces, and type of people there. Generally, we found that people exploring and making sense of an urban space wanted to be able to access differing levels of informational detail, in sequences that was difficult to predict, constantly changing, and not simply related to their current location. When delivering specific information adapted to their locational context we found that people still wanted to be able to access non-context specific information relatively easily. We also found that people required quite different information about their favorite places, such as "what is new since the last time I was there", but more general information about new places. In the wayfinding sections of the system we found that the dynamic nature of urban spaces presented a problem when using detailed photographs for matching between virtual and real world. The presence of new or temporary structures changed the look of the physical environment (and the social interactions that occurred there) to such an extent that people found it difficult to make the match. Perhaps the more abstract line-drawing-type representation, that users suggested would be acceptable, could help alleviate this problem.

7. DISCUSSION: THE VALUE OF A MULTI-DISCIPLINARY APPROACH

7.1. Understanding the socio-physical context of urban environments

One of the guiding questions of the study presented in this paper was how we could understand and represent the socio-physical context of urban environments. In response to this question, we have proposed a multi-disciplinary approach combining empirical and analytical methods and techniques from the fields of human-computer interaction, sociology and architecture. The result is a rich understanding of the social and physical properties of an inhabited urban environment, and the interplay between the two.

It is extremely difficult to separate the physical context of a space from the people inhabiting that environment. Physical context is more complex than a mere catalogue of physical elements in a specific environment, and in analyzing a built environment it is important to regard responses of the inhabitants of that space to the physical elements around them. Likewise, social context in urban spaces is more than just the existence of people and their immediate interactions with each other. Interactions occur in place and are influenced by the configuration of physical spaces. The situation of these interactions, the history of interactions in that place and experience of similar situations all influence people's understanding of social as well as physical affordances of a space. The dynamics of an interaction are very much influenced by the configuration and population of the urban environment in which they are taking place. This interrelationship between physical and social aspects of an environment is an important part of the understanding of socio-physical context of urban environments we have presented in this paper.

In informing our understanding of the socio-physical context of Federation Square, the investigation of both physical and social context produced holistic representations of the interrelationship between physical and social aspects of an urban environment. By making one part of the investigation focus primarily on the physical context of socializing in an urban environment, represented as a map, and the other part primarily on the social context of being physical situated there, represented as a table of activities, the overall investigation of the urban environment of Federation Square captures not only the unique properties of physical and social context but also the interplay between the two. The PIA layered map and hierarchy of architectural features provides an understanding of the physical context of an inhabited urban environment in the form of a graphical, people centered, representation. It highlights key properties of a particular space in an understandable and easily extractable way. The SOPHIA conceptual framework provides an understanding of the social context of a built urban environment in the form of a hierarchy of themes. It describes key properties of situated social interactions and through a qualitative, but structured, story about how people experience physical space and how they interact with each other while socializing in these spaces. Combined, PIA and SOPHIA provides a method for understanding of the socio-physical context of an urban environment.

7.2. Informing interaction design for a socio-physical context

The second guiding question for the research presented in this paper was how understanding the user's socio-physical context could inform design of a digital layer of pervasive computing for urban environments. In response to this question we have proposed an iterative, creative process as a way for interaction designers to incorporate this knowledge into their design. This process is based on design sketching in combination with systematic development of design ideas from PIA- and SOPHIA-type representations of socio-physical context. Through the development and evaluation of our prototype application, Just-for-Us, we have reflected on the value of this approach.

The design of the Just-for-Us prototype is very tightly coupled the understanding of the socio-physical context of the urban environment it was intended for. As illustrated in table 1 and 2, each of the seven design ideas making up the basics of the Just-for-Us system can be traced back to specific elements of the understanding generated through PIA and SOPHIA. Five of the seven design ideas were developed on the basis of the combined architecturally and sociologically derived understanding of the socio-physical context of the urban environment of Federation Square. *Location by district (1)* and the use of *augmented photorealistic depictions (2)* both respond to the combined socio-physical understanding of the space. They do so by utilizing, for example, the presence of distinctive and focal structures, districts, obscured places and entrances, open spaces with visible surroundings and activities around the edges *as well as* the way people use physical affordances, past experiences with a space. They index to visible elements and objects to collectively deal with uncertainty and gather information about the environment around them. Without the combined understanding represented in the outcomes from PIA and SOPHIA, these two design ideas, and their specific implementation in the Just-for-Us prototype system, would not have had the same depth and richness. Only from the combined qualitative understanding of peoples' use of the physical space, and the quantitative understanding of the actual properties of that physical space was it possible to develop these ideas.

The value of the combined socio-physical understanding of the context of Federation Square also comes to show in relation to the development of the idea of using *rich descriptions for navigation (3).* This idea can be quantitatively strongly traced back to understandings about nodes, edges and paths originating from the architectural part of the investigation as generated through the PIA method (see tables 1 and 2). The understanding of people's use of familiar paths, as a part of their physical familiarity with a place, was derived from the sociological part of the investigation. Both played equally important roles in the development of a design that takes into consideration people's actual use of a space over time. Not only could our rich descriptions for navigation refer to the prominent physical properties of a space, they could also refer to well-known interactions there in terms of places and paths that people are familiar with. Similarly, the idea of *representation of people and activities (6)* can be traced back most strongly to understandings originating from the sociological part of the investigation generated through the SOPHIA method. However, the development of this design idea was as strongly influenced by one of the observations captured by PIA highlighting a major presence of "activity places" at Federation Square. Again, in informing the design of a pervasive computing system, the strength of the proposed approach lies in the richness of the combined, socio-physical understanding of the urban environment context being designed for.

Only two of the seven design ideas, *use of history (4),* and *meeting and waiting (7),* were not derived from a combination of outcomes from PIA and SOPHIA. These two ideas respond to understanding captured in the SOPHIA framework dealing largely with experience *over time* and of *other people* in an urban environment rather than with the physical space itself. Naturally, these highly human-centered factors are hard to capture with pure architectural methods and techniques. This emphasizes the importance of a combined socio-physical approach sensitive to both aspects of context when designing pervasive computer systems for urban environments.

Based on the findings from our user-based evaluation of the Just-for-Us prototype system, we found that the design ideas developed on the basis of our socio-physical understanding of context at Federation Square were, indeed, successful at delivering the user experiences aimed for. The implemented design successfully created a pervasive digital layer of information for an urban environment that tied directly into the existing physical and social layers of that space, acknowledging and reinforcing the interrelationship between the three. Positive user experiences of the prototype system could be traced back to elements in the interaction design of Just-for-Us that were informed directly by our understanding of both physical and social context as represented in the outcomes from PIA and SOPHIA. When given access to the described digital layer of pervasive computing at Federation Square, people were able to overcome some of the limitations of the physical layer of the environment. At the same time they used elements of this layer as anchor points for the social and digital layers. They were able to access otherwise invisible information about places and people around them and obtain an overview of what was happening and where people were gathering. They were also able to string together their situated interactions at Federation Square over time, and share this with their friends. The socio-physical understanding of urban environment context provided by the PIA and SOPHIA methods was necessary for creating these user experiences.

7.3. A socio-physically informed development process

Based on our experiences, we propose a socio-physically informed approach to pervasive computing interaction design for urban environments (figure 6).

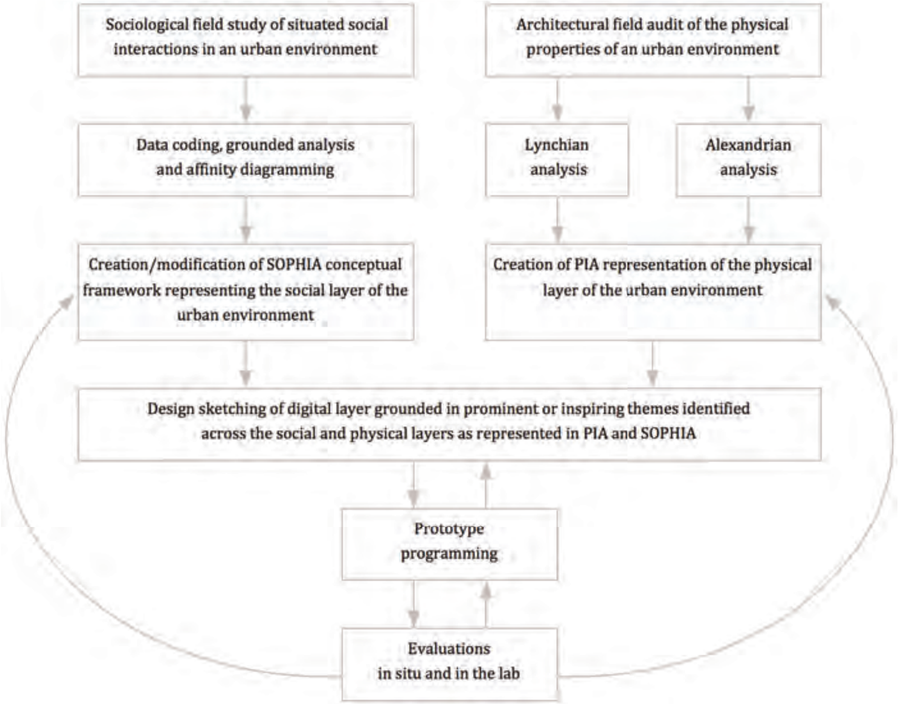


Figure 6. Socio-physical development process for pervasive computing in urban environments

The process depicted in figure 6 combine architectural and sociological streams of field studies and analysis towards the creation of PIA- and SOPHIA-type representations of the socio-physical context of an urban environment. The architectural stream involves field audits and analysis guided by the techniques of Lynch and Alexander. The sociological stream involves field observations and contextual interviews. This activity could be guided by, for example, the SOPHIA conceptual framework presented in this paper, McCullough's (2001) typology of "on the town", or it could be open-ended in its focus. The data collection activity leads to an activity of grounded data analysis and affinity diagramming. The two streams of research then feed understanding of the socio-physical context of the urban environment into a creative process of grounded design sketching of a digital layer of pervasive computing. Following on from this activity, the process takes the shape of a traditional prototyping process through which design ideas are iteratively implemented, evaluated, and refined until a satisfactory product outcome has been reached. However, as illustrated in figure 3, and described in the findings section above, studying the user experience of a pervasive computing prototype system in situ often leads to more knowledge about the socio-physical context for which it was designed. This means that additional understanding may have to be fed back into the PIA and SOPHIA conceptual frameworks, which again may lead to the emergence of new design ideas or refinement of existing ones.

7.4. The open-endedness of PIA, SOPHIA and the seven design ideas

It is important to notice that PIA and SOPHIA are not complete methods and have not generated complete frameworks and representations. Neither is the list of design ideas, which emerged from these frameworks, a complete collection of design outcomes possible to derive from the understanding encapsulated herein.

PIA and SOPHIA have both generated specific summaries of the context of Federation Square, and we do not claim that their outcomes are generally valid for all urban environments. The level of generalizability of the PIA and SOPHIA methods and their outcomes can only be determined through repeated studies in similar as well as different types of urban environments. In terms of the outcome, repeating the described socio-physical investigation in other urban environments would possibly confirm aspects of the presented conceptual frameworks, but would most likely also extend them with further concepts and categories. In terms of the methods, it is our belief that the two approaches would be able to capture the essence of the socio-physical context of sites other than Federation Square.

In terms of generalizability, however, it is also important to notice that PIA and SOPHIA differ fundamentally in what they are capturing. PIA provides understanding that is very specific to a certain physical environment being designed for. The combination of Lynch's and Alexander's approaches to architectural analysis supports this by providing a set of physical features to look for and PIA provides a combined way to represent the outcome of such analysis. In contrast, SOPHIA provides understanding that is potentially more generally applicable to situated interactions in urban environments. This is the case because the SOPHIA empirical study is based on a broader empirical foundation, and because the data analysis process was grounded towards the creation of general concepts rather than top-down from a set of pre-determined ones. The value of combining this top-down architectural and bottom-up sociological approach lies in the potential to provide focus as well as scope. The architectural part of the investigation adds focus on the physicality of situated interactions to the sociological analysis. In return, the sociological part of the investigation broadens the scope of the architectural analysis by highlighting the role of sociality.

The seven design ideas presented are, in a similar way, specific to the particular team of designers working on the Just-for-Us prototype and their creativity at the time. Other designers, or even repeated design sessions with the same designers, would most likely generate more ideas. Hence, the design-idea dimensions of tables 1 and 2 are not complete, and can never be. Adding to this open-endedness, extending the outcomes provided from PIA and SOPHIA with additional understanding through further investigations of socio-physical context of urban environments would undoubtedly expand the design-space with more ideas.

8. CONCLUSIONS

Hybridized spaces, in that they blend the physical digital and social, challenge our current conceptions of technology, and our approaches to understanding and designing. As we learn from our reference disciplines, sociology and architecture for example, we can return the favor with our own insights and collectively strengthen our response to a significant digital challenge; the 'turn to the social' and indeed physical that is implicit in the pervasive agenda. Later work should examine the value of our approach more generally, and contribute to our collective ability to compellingly situate interaction in the built, the social, and the increasingly occupied digital space.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of the David Hay Postgraduate Writing-Up Award, the Danish Technical Research Council through the Indexical Interaction Design project, the participants themselves and Federation Square itself. We would also like to thank Keith Cheverst and Malcom McCullough for their constructive reviews of the PhD thesis associated with this project.

REFERENCES

Agre, P. 2001. Changing Places: Contexts of Awareness in Computing. *Human-Computer Interaction 16*, 177-192.

Alexander, C. 1964. *Notes on the Synthesis of Form*. Harvard University Press, London.

Alexander, C., Ishikawa, S., Silverstein, M., with Jacobson, M., Fiksdahl-King, I., and Angel, S. 1977. *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press, New York.

Bell, G., and Dourish, P. 2004. Getting Out of the City: Meaning and Structure in Everyday Encounters with Space. In *Proceedings of Workshop on Ubiquitous Computing in the Urban Frontier, UbiComp, 2004*, retrieved March 2005, from http://www.urban-atmospheres.net/UbiComp2004/.

Beyer, H., and Holtzblatt, K. 1998. *Contextual design: Defining customer-centred systems*. Morgan Kaufmann, San Francisco.

Borntrager, C., Cheverst, K., Davies, N., Dix, A., Friday, A., and Seitz, J. 2003. Experiments in Multi-modal Interfaces. In *Proceedings of Mobile HCI 2003*, Udine, Italy, September 2003, Springer Verlag, Berlin, 116-130.

Bradley, N. A., and Dunlop, M. D. 2002. Understanding Contextual Interactions to Design Navigational Context-Aware Applications. In *Proceedings of Mobile HCI 2002*, Pisa, Italy, September 2002, Springer Verlag, Berlin, 349-353).

Buscher, M., and Hughes, J. 1999. Screen Scenery: Transposing Aesthetic Principles from Real to Electronic Environments. In *Social Navigation of Information Space*, A. Munro, K. Hook and D. Benyon, Eds. Springer Verlag, London, 90-111.

Buxton, B. 2007. *Sketching User Experiences: getting the design right and the right design*. Morgan Kaufmann, San Francisco.

Checkland, P. 1981. *Systems thinking, systems practice*. John Wiley & Sons, Chichester, Sussex.

Cheverst, K., Davies, N., and Mitchell, K. 2002. Exploring context-aware information push. *Personal and Ubiquitous Computing 6*, 276-281.

Cheverst, K., Davies, N., Mitchell, K., and Efstratiou, C. 2001. Using Context as a Crystal Ball: Rewards and Pitfalls. *Personal and Ubiquitous Computing 5*, 8-11.

Cheverst, K., Davies, N., Mitchell, K., Friday, A., and Efstratiou, C. 2000a. Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences. In *Proceedings of CHI 2000*, The Hague, Netherlands, ACM, 17-24.

Cheverst, K., Mitchell, K., and Davies, N. 2000b. Exploiting Context to Support Social Awareness and Social Navigation. *SIGGRAPH Bulletin 3*, 43-48.

Cheverst, K., Gibbs, M., Graham, C., Randall, D., and Rouncefield, M. 2005. Fieldwork and Interdisciplinary Design. Supplementary Notes for Tutorial at *OZCHI 2005*, retrieved November 2005, from http://www.comp.lancs.ac.uk/computing/users/rouncefi/Tutout.html.

Ciolfi, L. 2004. Understanding Spaces and Places: Extending Interaction Design Paradigms. *Cognition Technology and Work 6(1)*, 37-40.

Ciolfi, L., and Bannon, L. 2003. Learning from Museum Visits: Shaping Design Sensitivities. In *Proceedings of HCI International 2003,* Crete, Greece, Lawrence Erlbaum, 63-67.

Crabtree, A., and Rodden, T. 2004. Domestic Routines and Design for the Home. *Computer Supported Cooperative Work 13,* 191-220.

Dahlbäck, N., Jönsson, N. and Ahrenberg, L., 1993. Wizard of Oz studies: why and how. In: W.D. Gray, W.E. Hefley and D. Murray, eds. Proceedings of the 1st international conference on intelligent user interfaces, 4-7 January 1993 Orlando. New York: ACM, 193-200.

Dieberger, A., Dourish, P., Hook, K., Resnick, P., and Wexelblat, A. 2000. Social Navigation: Techniques for Building More Usable Systems. *Interactions November + December*, 37-45.

Dix, A. 2004. Physicality, Rationality and Imagination. Keynote Address at *Interaccion 2004*, retrieved August 2005, from http://www.hcibook.com/alan/papers/i2004-imagination.

Dix, A., Rodden, T., Davies, N., Trevor, J., Friday, A., and Palfreyman, K. 2000. Exploiting Space and Location as a Design Framework for Interactive Mobile Systems. *ACM Transactions on Computer-Human Interaction 7(3)*, 285-321.

Donath, J. 1996. *Inhabiting the Virtual City: The design of social environments for electronic communities*. Unpublished doctoral dissertation. Massachusetts Institute of Technology, Massachusetts, USA.

Dourish, P. 2001. Seeking a Foundation for Context-Aware Computing. *Human-Computer Interaction 16*, 229-241.

Dourish, P. 2004. What we talk about when we talk about context. *Personal and Ubiquitous Computing 8(1)*, 19-30.

Erickson, T. 1993. From Interface to Interplace: The Spatial Environment as a Medium for Interaction. In *Proceedings of Conference on Spatial Information Theory*, Elba, Italy, Springer Verlag, Berlin, 391-405.

Erickson, T. 2002. Some Problems with the Notion of Context-Aware Computing. *Communications of the ACM 45(2)*, 102-104.

Erickson, T., and Kellogg, W. A. 2000. Social Translucence: An Approach to Designing Systems that Support Social Processes. *ACM Transactions on Computer-Human Interaction 7(1)*, 59-83.

Goffman, E. 1963. *Behaviour in Public Places - Notes on The Social Organization of Gatherings*. Greenwood Press, Publishers, Westport, Connecticut.

Goodman, J., and Gray, P. 2003. A Design Space for Location-Sensitive Aids for Older Users. In *Proceedings of Workshop on HCI in Mobile Guides at Mobile HCI 2003*, retrieved March 2004, from http://www.mguides.info/.

Graham, S., and Marvin, S. 1996. *Telecommunications and the City: Electronic spaces, urban places*. Routledge, London.

Harrison, S., and Dourish, P. 1996. Re-Placing Space: The Roles of Place and Space in Collaborative Systems. In *Proceedings of Computer Supported Cooperative Work '96*, ACM, Cambridge, MA, 67-76.

Hillier, B. 1996. *Space is the Machine*. Cambridge University Press, Cambridge, UK.

Iacucci, G., Kela, J., and Pehonen, P. 2004. Computational support to record and re-experience visits. *Personal and Ubiquitous Computing 8(2)*, 100-109.

Ingram, R., Benford, S., and Bowers, J. 1996. Building Virtual Cities: applying urban planning principles to the design of virtual environments. In *Proceedings of VRST '96,* Hong Kong, ACM, 83-91.

Jensen, R., and Lenskjold, T. 2004. Designing for social friction: Exploring ubiquitous computing as a means of cultural interventions in urban space. *Proceedings of CADE2004*, http://aso.cbs.dk/cade2004/proceedings.htm.

Kjeldskov, J., and Paay, J. 2005. Just-for-Us: A Context-Aware Mobile Information System Facilitating Sociality. In *Proceedings of Mobile HCI 2005*, Salzburg, Austria, September 2005, ACM, 23-30.

Kjeldskov, J., Skov, M., Als, B, and Hoegh, R. 2004. Is it Worth the Hassle? Exploring the Added Value of Evaluating Usability of Context-Aware Mobile Systems in the Field. In *Proceedings of Mobile HCI 2004*, Glasgow, Scotland, Springer Verlag, 61-73.

Kulju, M., and Kaasinen, E. 2002. Route Guidance Using a 3D City Model on a Mobile Device. In *Proceedings of Mobile Tourism Support Workshop at Mobile HCI 2002*, retrieved March 2004, from http://www.mguides.info/.

Kuutti, K. 1996. Activity Theory as a Potential Framework for Human-Computer Interaction Research. In *Context and Consciousness: Activity Theory and Human-Computer Interaction* B. Nardi, Ed. MIT Press, Cambridge, MA, 17-44.

Laakso, K., Gjesdal, O., and Sulebak, J. 2003. Tourist information and navigation support by using 3D maps displayed on mobile devices. In *Proceedings of Workshop on HCI in Mobile Guides at Mobile HCI 2003*, retrieved March 2004, from http://www.mguides.info/.

Lynch, K. 1960. *The Image of the City*. MIT Press, Cambridge, MA.

McCullough, M. 2001. On Typologies of Situated Interaction. *Human-Computer Interaction 16*, 337-349.

McCullough, M. 2004. *Digital Ground - Architecture, Pervasive Computing, and Environmental Knowing*. The MIT Press, Cambridge, MA.

Millen, D. R. 2000. Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In *Proceedings of DIS '00*, Brooklyn, New York, ACM, 280-286.

Mitchell, W. 1995. *City of Bits: Space, Place and the Infobahn*. The MIT Press, Cambridge, MA.

Mitchell, W. 1999. *e-topia: "urban life, jim - but not as we know it"*. The MIT Press, Cambridge, MA.

Neuman, W. L. 2003. *Social Research Methods: Qualitative and Quantitative Approaches*. Allyn and Bacon, Boston.

Paay J. and Kjeldskov J. 2008. Situated Social Situations: A Case Study of Public Places in the City. *Computer Supported Cooperative Work* 17(2-3),275-290

Paay, J. 2005. "Where We Met Last Time": A Study of Sociality in the City. *In Proceedings of Australasian Computer Human Interaction Conference, OZCHI 2005,* Canberra, Australia, ACM, 8-17.

Paulos, E., and Goodman, E. 2004. The Familiar Stranger: Anxiety, Comfort, and Play in Public Places. In *Proceedings of CHI 2004*, Vienna, Austria, ACM, 223-230.

Persson, P., Espinoza, F., Fagerberg, P., Sandin, A., and Coster, R. 2003. GeoNotes: A Location-based Information System for Public Spaces. In *Designing Information Spaces: The Social Navigation Approach* K. Hook, D. Benyon and A. Munro, Eds. Springer Verlag, London, 151-173.

Randell, C., and Muller, H. 2000. The shopping jacket: wearable computing for the consumer. *Personal and Ubiquitous Computing 4*, 241-244.

Rheingold, H. 2003. *Smart Mobs - The Next Social Revolution*. Perseus Publishing, Cambridge, MA.

Schmidt, A., Beigl, M., and Gellersen, H. 1999. There is More to Context than Location. *Computers and Graphics Journal 23(6)*, 893-902.

Sharp, H., Rogers, Y., Preece, J. 2007. *Interaction Design: beyond human-computer interaction, 2nd Edition*. John Wiley & Sons, Chichester, Sussex.

Snyder, C. 2003. *Paper Prototyping*. Morgan Kaufmann Publishers, Amsterdam.

Strauss, A. L., and Corbin, J. 1990. *Basics of Qualitative Research*. Sage Publications, Newbury Park, CA.

Tamminen, S., Oulasvirta, A., Toiskallio, K., and Kankainen, A. 2003. Understanding Mobile Contexts. In *Proceedings of Mobile HCI 2003*, Udine, Italy, September 2003, Springer Verlag, Berlin, 17-31.

Vainio, T., Kotala, O., Rakkolainen, I., and Kupila, H. 2002. Towards Scalable User Interfaces in 3D City Information Systems. In *Proceedings of Mobile HCI 2002*, Pisa, Italy, September 2002, Springer Verlag, Berlin, 354-358.

Whyte, W. 1980. *The social life of small urban spaces*. Conservation Foundation, Washington DC.

Yee, R. 1997. *Architectural Drawing: A Visual Compendium of Types and Methods*. John Wiley & Sons, Canada.

Chapter 9

Sketches and mock-ups

Frank Vetere, Martin R. Gibbs, Jesper Kjeldskov, Steve Howard, Florian Mueller, Sonja Pedell, Karen Mecoles and Marcus Bunyan

Abstract. Intimacy is a crucial element of domestic life, and many interactive technologies designed for other purposes have been appropriated for use within intimate relationships. However, there is a deficit in current understandings of how technologies are used within intimate relationships, and how to design technologies to support intimate acts. In this paper we report on work that has addressed these deficits. We used cultural probes and contextual interviews and other ethnographically informed techniques to investigate how interactive technologies are used within intimate relationships. From this empirical work we generated a thematic understanding of intimacy and the use of interactional technologies to support intimate acts. We used this understanding to inform the design of intimate technologies. A selection of our design concepts is also presented.

1. INTRODUCTION

Artifacts have commonly been used to mediate intimate relationships. Over the ages, intimacy has been mediated through symbols of affection such as flowers, missives and love letters. New and emerging technologies have also been appropriated to mediate close personal relationships. In particular, we observe this appropriation with the Internet and mobile phones. It is now commonplace for family members separated by distance to maintain contact via the Internet. Matchmaking and online dating are popular Internet services (Donn and Sherman 2002). SMS (Short Message Service), i.e. sending a text message via mobile phone, is increasingly being used to forge new romantic relationships (Byrne and Findlay 2004) and to coordinate activities with intimate friends (Grinter and Eldrige 2001).

Even though an intimate relationship often requires no mediation, new technologies are regularly manipulated to help us feel connected with those for whom we care.

Originally published as Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoles, K. and Bunyan, M. (2005) Mediating Intimacy: Designing Technologies to Support Strong-Tie Relationships. In *Proceedings of CHI 2005,* Portland, Oregon, USA (pp. 471-480). New York: ACM.

So strong is this desire, that we will spend significant amounts of money on communication technologies, and will be inconvenienced by poor usability so that our personal relationships are nurtured and maintained. For example, young people will endure SMS even though it lacks expressiveness, it has confusing syntax and it is prone to errors (Grinter and Eldrige 2001). However, it satisfies an important social and personal need to feel connected. The strength of this desire together with the inadequacy of current technologies to support expressions of intimacy, offer a unique research opportunity.

In this paper we are interested in phenomena that are recognizably intimate – expressions of tenderness, acts of devotion and habits of demonstrable affection. We are interested in understanding how technologies are used to support these phenomena and in investigating methods for designing new ICT (Information and Communication Technologies) for mediating intimacy. This is not simply a question of creating representations for expressing emotions, as for example emoticons attempt to do. Rather it is a broader exploration of the role of ICT in people's intimate lives in order to create more fulfilling designs.

2. ATTEMPTS TO MEDIATE INTIMACY

There is a growing interest in technologies that support relationships with intimate others (Bell et al. 2003). For instance Gaver (2002) proposes provocative ideas for connecting people in close relationships. He describes technologies that provide a feeling of presence of remote lovers through peripheral awareness. Examples of awareness technologies are: a feather in a plastic cone that floats when the distant partner picks up a picture frame of the couple (Strong and Gaver 1996); a light 'orb' that glows in New York when a family member walks into their London apartment (Tollmar and Joakim 2002); and two sets of cylinders that roll and rotate in unison as they are manipulated by separated partners (Brave and Dahley 1997). Through a critical analysis and review of awareness technologies, Gaver identifies three typical characteristics: (i) the designs often make use of evocative materials (such as feathers and scents); (ii) mappings are more likely to make use of literary rather than didactic metaphors (e.g. rolling cylinders that evolve into tactile languages); and (iii) objects have a unique physicality (e.g. a real feature is more poetic than one simulated on a screen).

These characteristics are common in technologies for mediating intimacy. For instance, Kaye and Goulding (2004) presents three Intimate Objects that address the problem of close couples trying to maintain feelings of intimacy when separated by large distances. These objects are intended to be used by a specific couple (i.e. custom built) for communicating intimacy specifically (e.g. not used for work-related activities). The first object is How do I love thee, a pair of abacus, whose beads are synchronized over the internet. Next is Hand Holding, devices that simulate the warm touch of holding hands. Finally is Love Egg that rolls on a concave dish when an intimate message is left on its messaging system. These objects display each of Gaver's three characteristics. They use evocative materials (soft silicone in Hand Holding), exploit non-didactic metaphors (positioning abacus beads) and have a unique physicality (shape and texture of an egg).

Another intimate object with a dominant physicality is the Sensing Bed (Goodman and Misilim 2003). The Sensing Bed comes in pairs, and is intended for romantic couples who are not co-located. The bed senses the body position of one person and transmits warmth to congruent parts of the lovers' bed. The Sensing Bed is similar to the Bench Object (Gaver 2002), which has a seat that warms when somebody sits on a partner bench far away. Unlike the Sensing Bed, the Bench Object is public. The warmth on the Bench Object is generated by an unknown other (and so the effect is often disturbing). Even though the Sensing Bed and the Bench Object are technologically similar and both exploit peripheral awareness and physicality, their intent is different. The first is intended for couples in an established relationship, the latter is intended to provoke a visceral response between strangers in a public place. It is with the former that our work is more closely aligned.

A study concerning intimate relationships of a different kind comes from a group of students motivated by the routine of emptying their pockets of coins and keys when arriving home. They recall that the clanging of keys on the table is often followed by the cry "Mom, I'm home". The researchers exploit the association between the clang of keys and the arrival call to create the Gustbowl (van der Hoog et al. 2004). When items are thrown into the Gustbowl, the bowl wobbles and takes a picture of the items. The picture and the movement are transmitted over the Internet to an identical bowl that wobbles and receives the picture. The Gustbowl is intended to support the strong-tie relationship between mother and grown-up son. It is non-verbal, using pictures and sounds to trigger meaning from established routines and behaviors.

A comparable study, the ASTRA project (Markopoulos et al. 2004), is also concerned with family members living apart, but focused on the interrelationships within the family rather than one-to-one relationships. Their motivation is to understand the role of peripheral awareness systems in social interactions within families.

3. CHALLENGES FOR STUDYING MEDIATED INTIMACY

Intimate relationships are different from the kinds of relationships that have been typically studied by HCI research. Intimate relationships are different to those found in the workplace or amongst social networks of friends. Intimate acts also differ from the domestic behaviours usually addressed in the literature [see 15 for a review]. Attempting to study acts of intimacy presents the researcher with a number of unique and interesting challenges (Kjeldskov et al. 2004). Studying intimacy is challenging because intimate acts are ephemeral and transient yet ubiquitous and crucial to the ongoing life of an intimate relationship. They form the material and background of close personal relationships, yet occur in the doing and then often vanish unremarked. While the informational content of intimate acts may be low and seemingly trivial to outsiders, the act itself can be laden with emotional significance for those involved. Intimate acts often entail self-disclosure, and thus privacy is a concern. Much of what passes between intimates is unsaid and premised on deep knowledge and understanding of one another and occurs in the context of a rich, shared and sometimes idiosyncratic view of the world that may be difficult for others to fathom and comprehend. Intimacy also involves assumptions about commitment and mutuality. It carries nuanced expectations for reciprocity and exchange that are negotiated and arrived at over many years, yet remains fragile and is occasionally misjudged leading to misunderstandings and conflict. Finally, unlike instrumental tasks (e.g. coordination of family activities), or leisure activities (e.g. games) there is no generally accepted language for describing and discussing intimacy, especially in relation to designing technologies for its support.

In designing our research, we were mindful and sensitive to the challenges presented by studying intimate acts in the domestic setting. In responding to these challenges we wanted an approach that enabled us to maintain a longitudinal presence in the field so we could study the ongoing life of relationships. We were keen to capture the ephemeral and "unsaid" (Geertz 1973) aspects of intimate exchange and needed an approach that enabled intimate acts to be recorded as they occurred, or soon after. Given the private character of intimate relationships, we were keen to empower our subjects as participants in the research and to give them control over what was, and was not, revealed to us. Since intimate acts are highly nuanced and often subtly ambiguous to outsiders, we wanted to provide participants with materials enabling them to interpret and explain their practices to us. Finally, we wanted an approach that allowed us to carry out an ongoing conversation with participants and through this conversation arrive at a shared understanding of intimacy and the place of ICT in mediating intimate acts. We describe this approach in the next section.

4. RESEARCH DESIGN

Other researchers who have had to confront the difficulty of investigating personal relationships have adopted a range of research methods. These include: online questionnaires and surveys (Byrne and Findlay 2004, Donn and Sherman 2002); data logs (Grinter and Eldrige 2001); longitudinal focus group (Taylor and Harper 2003); interviews (Kaye and Goulding 2004); and written reflections (Pedell et al. 2003). Since our goal is not only to understand intimacy, but to design for it, we adopted a suite of methods and techniques. Our research plan is represented in Figure 1. In the first phase of our research we sought to understand current practice. We undertook ethnographically informed field studies using cultural probes (Gaver et al. 1999) and contextual interviews to understand how people use ICT in their intimate relationships. In the second phase of our research we sought to use the insights into current practice gleaned from phase one to provoke and inform the design of future ICT to support intimate acts. In this phase, we engaged in a variety of design activities including expert design workshops, participative design workshops, scenario development and the development of fidelity prototypes.

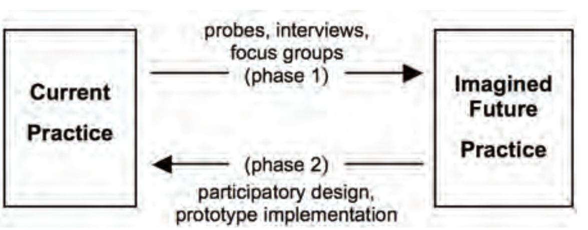


Figure 1. Cycle of use-centered innovation (Howard et al. 2002)

4.1. Method

Our approach extends the work of Gaver et al. (1999) and the Equator team (Cheverst et al. 2003) by combining cultural probes with contextual interviews and focus groups. We assembled a collection of cultural probes into a ‘probe pack’ (Figure 2). The probe pack contained: diaries and scrapbooks; digital camera with printer; postcards; pens, glue, and catch-phrase stickers (e.g. “I feel alone when …”, “I feel supported when …”).

Figure 2. Contents of Probe Pack

The diaries were used individually to record daily communication and interaction activities. This included the form of communication (e.g. SMS, telephone, email, letters, notes, tokens, gifts) and other details such as time, date, location, the content and the feelings associated with it (e.g. urgency or dissatisfaction).

The scrapbooks were used to tell rich and evocative stories about communicative events and to express the technology wants, desires, likes and dislikes surrounding these events. Couples were encouraged to work together with pens, crayons, glue, photos, magazine clippings, drawings to form a montage of their intimate lives.

The digital camera and printer were used to photograph and print significant events. The docking printer provided the immediacy of traditional Polaroid photographs with the convenience of lasting digital images. Participants were encouraged to take photographs of everyday artifacts and events that express some important dimension of their interactions: e.g. the answering machine at the time of receiving some unexpected news or of a child in a football final to share with an absent parent. Participants were asked to print, annotate and cut-and-paste the photographs into their scrapbooks or diaries or postcard, as they saw fit.

The participants were asked to read and reflect on the catchphrase labels, complete them, and stick them into the diary, scrapbook or on a post card. The stamped addressed postcards were used for sending short stories or images to the researchers or to the intimate other.

Finally, an additional probe element was introduced midway through the study. This new element consisted of small printed facsimiles of a variety of mobile device screens (e.g. mobile phones and PDAs). This new element served to both stimulate engagement in the study and encourage participants to envision possibilities for future technologies to support their relationship. Participants were invited to use these facsimiles to note design ideas and to insert them into the scrapbooks or diaries.

All the materials could be used in whatever way participants wished. The instructions were only suggestions. No time requirements were made, but it was suggested they should spend about 20-30 minutes each day using the probe pack materials.

4.2. Participants

The study involved six Caucasian couples in stable heterosexual relationships spanning between 4 and 16 years in duration. All couples cohabitated, although work related travel occasionally required periods of separation. The age of participants ranged from late 20's to late 40's. Three couples had children, ranging in age from 18 months to 12 years of age. In one case the children were from a previous marriage. The couples could be broadly described as middle-class, leading professional lives in urban environments. The couples were neither very rich nor very poor, though some were reliant on government subsidies. The participants all had tertiary qualifications with occupations as Social Worker, Charity Worker, Business Analyst, Industrial Relations consultant, IT consultant, IT trainer and Journalist. Two participants were undertaking full time higher education, and one worked as a casual administration officer while caring for young children. All couples had access to the Internet and mobile telephony. They used a variety of electronic media such as landline and mobile phones, email, Internet chat, SMS and fax to communicate with each other, although the exact mix of technologies used by each couple varied markedly. The number of participants in our study is comparable to that in other probe research (Gaver et al. 1999; Hutchinson et al. 2003).

The participants were recruited through calls via email, posters and personal contacts. The participants were allowed to keep the digital camera and docking printer as appreciation for their participation in the study.

4.3. Data Collection

Cultural probes were deployed for a period of seven weeks. This period was followed by focus groups and design workshops.

Week 0: Probe pack distribution and initial interview

An initial interview was carried out at the participants' homes. This included questions about the participants' backgrounds, their relationship, their communication habits and their use of technology. Following the interview, the researchers presented the cultural probe materials and informed participants of their use.

Week 1: Interview and process checking/steering

A week after the probes were distributed, researchers visited the participants' homes for a second time to answer questions about their use of the probes and to discuss activities of the first week.

Week 4: Interview and addition of new probe element

Researchers and participants discussed the materials accumulated through probe activities, and participants were encouraged to clarify, elaborate and reflect on the materials they composed. At the end of the third session, researchers introduced the printed mobile device screen facsimiles.

Week 7: Interview and probe collections
Diaries, scrapbook and other materials such as design ideas produced on the mobile device screen facsimiles were discussed. This session was used to bring closure to the seven-week process. All materials gathered through the cultural probes were retrieved at this time

Week 9: Focus groups
Several weeks after the probes were collected, participants in the study were invited to take part in one of two focus group discussions. In these sessions, we presented our preliminary themes to participants and invited them to comment and reflect on them. Their contribution was to refine and consolidate the intimacy themes.

Week 12-15: Design Activities
Intimacy themes identified during data analysis (see below) were used to motivate design ideas through three activities. We firstly conducted a series of brainstorming sessions, then a design workshop with HCI experts, and finally a participatory design with those involved in the study. The outputs of these activities were high-level descriptions of technologies to mediate intimacy, and development of a prototype to demonstrate a subset of design ideas. The design process is only partially reported in this paper.

4.4. Data Analysis

Six researchers were involved in data collection and analysis. Researchers worked in pairs with participants. Analysis and interpretation of data began immediately. Following each meeting with participants, the researchers who conducted the interview would meet to discuss the probe traces and interview contents using an open-ended process of rapid reflection (Pedell et al. 2003) to identify important themes, ideas and concepts that emerged from their encounter with the participants. They would then compose a summary of the interview and this summary, along with the themes identified was reported back to the larger research team at weekly meetings.

At these weekly meetings themes emerging from the data were explored, debated and discussed. The probe data was naturally incomplete, unclear and biased (Gaver et al. 2004). This inevitably led to subjective interpretations where the data was often discussed in terms of the researchers' own experiences of intimacy. These ongoing discussions around the data and its analysis continued throughout the project and the themes identified in the data were refined iteratively. Data integration meetings were held at critical stages of the project. At these meetings, data and interim analyses were presented and debated. These meetings helped to develop and refine dominant themes. Finally, focus group discussions were held with participants in the study. These focus groups allowed the participants to make a final contribution and comment on our observations and findings.

5. FINDINGS

Despite numerous social science studies of intimacy and the exchanges that occur within intimate relationships, a universally acknowledged definition of intimacy has yet to surface. Of those definitions that have emerged (for example Cheal 1987; Moss and Schwebel 1993; Register and Henley 1992; Robson, D. & Maggie 1998) few provide any significant design traction.

Figure 3 presents a schematic view of the primary themes induced from our qualitative data. The themes are placed according to affinity with their neighbors. Whilst each theme is evident in our data, they overlap to a significant degree, and the affinities we have tried to reveal in the structure of Figure 3 are complex and multidimensional; in truth the diagram could be rendered in many ways, with each alternative giving preference to different themes.

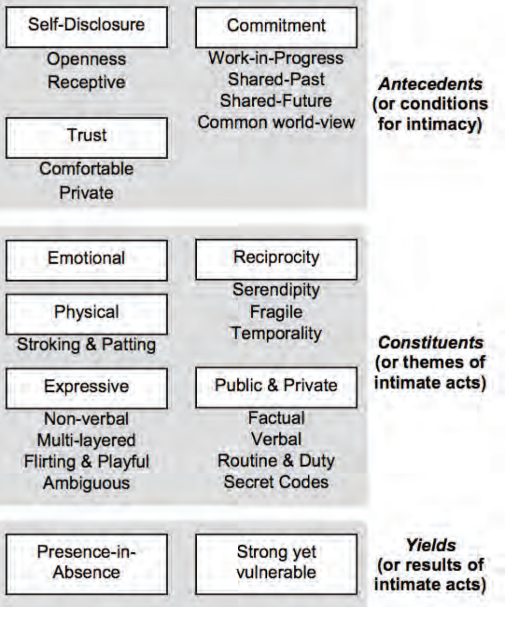


Figure 3. Antecedents, Constituents and Yields of Intimacy

We've found it useful to think of intimacy in terms of those themes that precede its experience, Antecedents; those themes that characterize the act(s) itself, Constituents; and those themes that reflect the consequences of an intimate exchange, Yield. Figure 3 is structured according to these three clusters, before, during and after. So, for example, 'Commitment' is an antecedent or a pre-condition of intimacy. Being 'Emotional' is better considered an expression of intimacy, and so characterizes something about the intimate act itself. 'Presence-in-absence' is a feeling that results from an intimate exchange or an intimate thought, and so is a yield of intimacy. Of course, Antecedents and Constituents exist in a recursive relationship, and our so-called Antecedents can, over time, also be a consequence of intimate acts. The sub-themes are grouped and clustered around a central theme (shown as outlined boxes in Figure 3).We now illustrate the central themes, drawing on our data where useful.

5.1. Before Intimacy: Antecedents

Self disclosure

The price of intimacy is revelation and mutual openness (Monsour 1992; Moss and Schwebel 1993). It is during self disclosure that we 'get to the heart of the matter'. An act of self disclosure carries with it an expectation that it be returned in-kind, and a failure to reciprocate is felt keenly as a breakdown in the maintenance of intimacy. One participant completed the catchphrases "I misunderstand my partner when…" with "… he goes quiet." The maintenance of 'partner awareness' rests on an almost constant background of chatter and stroking that both reveals and acknowledges each others' internal state; "I really need to tell my partner…" is completed with"…what happens during my day and for him to tell me about his; it feels right."

Trust

Intimacy requires trust and sincerity. The self disclosure that so keenly illustrates an intimate act risks both the self and the other, and that risk demands trust. Intimacy also requires a commitment to the relationship as something bigger than oneself (Cheal 1987). The trust that characterizes strong-ties is deep and resilient, allowing each partner to be playful and flirtatious without any real risk to the relationship. This is exemplified by a participant who sent a sexually provocative text message to her partner, knowing that he was in a pressured business meeting. She was being playful and feeling sufficiently secure to risk provocative behavior. Their shared commitment extended beyond the act itself. The levels of trust and mutual commitment we observed provide a robustness against the negative effects of communication breakdowns; with trust comes tolerance for the inevitable little fractures in the intimate dialogue (e.g. failure to return a call within an expected time window) but a raised expectation for a deeper reciprocity.

Commitment

Commitment is the extent to which partners in a relationship perceive their relationship as ongoing for an indefinite period, and is a precondition for other aspects of intimacy to flourish (Chelune et al. 1984). Some of our participants framed this as being on a "shared" or "common journey" together. They shared both the costs and rewards of lives together. Features of participants' lives, such as raising children, caring for older relatives, maintaining the household, or traveling to work, were all enterprises that were shared and, in the sharing, became vehicles for enacting, affirming and maintaining their relationship. Joint responsibility was taken for domestic life, including activities such as paying the bills, transporting children and preparing meals. These activities were often coordinated on the fly using interactive technology such as mobile phones. The division of labor within these relationships had regular patterns but was also fluid and renegotiated over the course of a day. Responsibilities also provided occasions for affirming the relationship. For example, the success of a child at school was affirming "our" achievement as "good parents". Even activities that were the sole responsibility of one member of the relationship became joint enterprises. Participants reported drawing on their partner's help and skills for work related tasks such as database development, setting up a web-based email account or writing a job application. Being able to help and share common tasks affirmed the relationship and sense of moving through the world together as a team, rather than as individuals.

5.2. During Intimacy: Constituent Themes

Emotional

In contrast to the routine and dutiful exchanges that characterize much work activity and family life, for example the coordinative activities that occupy much of parenting, intimate acts mediate emotion (Moss and Schwebel 1993). In intimate acts the medium can indeed be said to be the message; within most relationships, there is little new information exchanged in saying "I love you", but few would argue against the value contained in its saying!

Reciprocity

Though many intimate exchanges are asynchronous, there remains a strong sense of reciprocal binding. At the atomic level, the utterance "I love you" demands to be answered in like terms; it is unimaginable that such an exchange be met with ignorance. In more general terms, intimacy depends on intimates who are co-engaged in a common cause (Moss and Schwebel 1993).

An example of reciprocity is the 'goodnight message' (Grinter and Eldrige 2001). One couple would insist on sending each other text messages late at night when one person was working a night shift. The recipient lies in bed, turns the phone volume down and awaits the incoming message. The preparation and expectation is part of the reciprocity of their imitate relationship.

Expressive

These are non-verbal but highly expressive interactions often involving playful, even ambiguous exchanges. Though intimate acts are not fact-free acts, the centrality of emotion to their meaning gives them a special quality. Intimate exchanges can themselves become the subject of 'personal innovation'; our intimates strive to keep the conversation of intimacy alive and fresh by changing its form or medium, often in playful ways. For example, Figure 4 shows a section of a scrapbook, where a participant used pictures to express his love. In another case, a participant would hide messages in places their partners are likely to visit (e.g. underwear drawer). These gifts would be discovered later serendipitously; the element of chance involved in the discovery merely added to the experience of receiving the gift, rather than being something to be 'designed out' in an attempt to make the gift giving process more 'efficient' or 'effective'. "Having fun, being creative and using humor is something we like to do".

Figure 4. Being creative and expressive.

A clear distinction existed between those participants who were parents, and those who were childless. The presence of children in a household colors all interaction, including intimacy. In comparison to child free couples, working parents are time poor to a highly significant degree. Where child free couples wish to experiment with the language of intimacy in playful and creative ways, working parents strive to reclaim space to 'be just us again'; their need is less for playful new ways of saying 'I love you' than it is for reclaiming opportunities to interact as a couple. Couples with children routinized time together, often late in the evening after children are bathed and put to bed.

Physical, involving stroking and patting

Intimates communicate in often non-verbal but nevertheless highly expressive and nuanced ways; ways not well supported by current technologies that are biased towards verbal language and encoding information. The physical expression of intimacy includes but is certainly not limited to sexual relations. Our intimates stroked and touched each other in gentle ways when co-present, and exchanged gifts that had physicality and could therefore act as a proxy when the partners were distant in space and/or time.

Public & Private

While much that is associated with intimate acts is quintessentially private in character, intimate relationships also have public expression. Couples present a 'face' to the world, that both affirms to themselves, and to others, their status as a couple (Goffman 1959). This can occur through ephemeral public displays of affection, such as touching and kissing in public, or more concretely through the public and legal declaration of love and commitment entailed in marriage. While public, these expressions of self were also often highly nuanced and coded, allowing for the private communication of meaning in a public forum.

Intimate relations are rich sources of private covert language. Often obscure ("NYUM" meaning yummy or delicious and '143' meaning 'I love you') and sometimes containing hidden meanings ("It is very simple really" said during difficult times to refocus the couple onto what is really important in life, i.e. 'us') they define their partners as distinct from the world and mediate their relations, allowing private interactions in public places whilst limiting the risk of being 'overheard'.

5.3. Consequences of Intimacy: Yield

Presence-in-absence

We frequently observed intimate exchanges when couples were face-to-face, and indeed it was surprisingly common for those exchanges to be mediated by technology. For example, participants sending an intimate text message from the living room to the bedroom, or from one floor of an office building to another floor, or sending an email to a partner who is working on another computer a few meters away. Email and SMS were used to facilitate stroking and patting behavior. However, perhaps the most vivid expression of the centrality of intimacy to our participants' lives came with the elevation of 'presence-in-absence' to a core need. When separated by distance or time, our intimates described a strong sense of presence-in-absence, a sense of the other despite their physical absence (Register and Henley 1992). Our participants invested considerable time, effort and emotion in ensuring that their partners stay with them, at the forefront of their hearts and minds, throughout daily life. A good example of where presence-in-absence was missing is shown in Figure 5. The child is proud of having built a house of cards, and would very much like to show his father who was away on work duties. The ephemeral tower was captured by a photograph to be shown to him later.

Strong yet vulnerable
Our study has also taught us that it is important not to romanticize intimacy. Intimacy and the tight emotionally charged bonds it entails are strong, yet can be strangely and unexpectedly fragile.

While relationships may be robust, misunderstanding and misinterpretations do occur. When they occur, these breakdowns can have serious repercussions, creating ill will and emotional hurt that can obstruct and undermine intimacy within the relationship. When one participant was angry with his wife, he would put his mobile phone on the mantelpiece, declaring through his actions that he was unavailable. The phone represented a connection to his wife because it was used almost exclusively to talk to her. By placing it on the mantelpiece, he indicated that he did not wish to 'carry' her. Another couple had a routine of coordinating their departure from work each evening so that they could catch the train together. When as a result of an unexpected meeting one person caught the train without notifying the other, resentment and anger ensued. Such a fallout from a simple breakdown in understanding between partners can reverberate through the relationship for days, if not weeks. For these reasons, any interactive technology designed to support or mediate intimacy should mitigate against these forms of breakdown, allow repair of them when they do occur, and contribute to relationship stability.

Figure 5. Entry in one of the diaries. "Hi Dad. I've been making card houses and mums been taking photos. She wants to talk to you …"

6. IMPLICATIONS FOR AN 'INTIMATE TECHNOLOGY'

Based on the findings, we have developed a series of design ideas. This process had three phases. First, the research team conducted a brainstorm session to produce a series of design sketches. These designs were highly abstract, often consisting of a schematic drawing, a short description and keywords linking the ideas to the themes of intimacy emerging from our empirical data (Figure 3). Second, we conducted two one-day design workshops. One workshop was with experts in human-computer interaction design and the other workshop with couples participating in our earlier ethnographic study. Third, we developed and implemented a functional prototype to demonstrate some of our ideas and concepts. A selection of our design ideas are presented below.

6.1. Preliminary Design Sketches

The brainstorming of design ideas resulted in the development of 22 design sketches. The designs produced at this time were highly diverse, addressing a range of requirements, opportunities and challenges identified in our user study. Most of the ideas involved mobile as well as stationary devices, visual as well as auditory and tactile interfaces and explicit as well as ambient interaction. For example, we discussed several personal devices that would enable partners to interact unseen with each other throughout the day via low-fidelity communication channels. One of such devices was the "Secret Touch" (Figure 6), which enables partners to "virtually hold hands" while physically separated. The Secret Touch device allows partners to exchange tactile impulses over the Internet by padding or squeezing the device in their hand or pocket. At the receiving end, these pads and squeezes would be emulated as vibrations, heat or pressure.

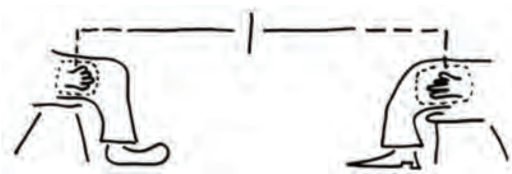


Figure 6. Secret Touch: handheld devices allowing partners to make secret tactile exchanges while physically separated.

6.2. Design Sketches from Workshops

The design workshops we conducted with interaction design experts and study participants produced significant results. Over two separate days, six groups of 4-5 people produced a refined design sketch that included: (i) a description of the functionality of their system, (ii) a central screen design, (iii) a basic interaction design, and (iv) a scenario of use illustrated with drawings, pictures and text.

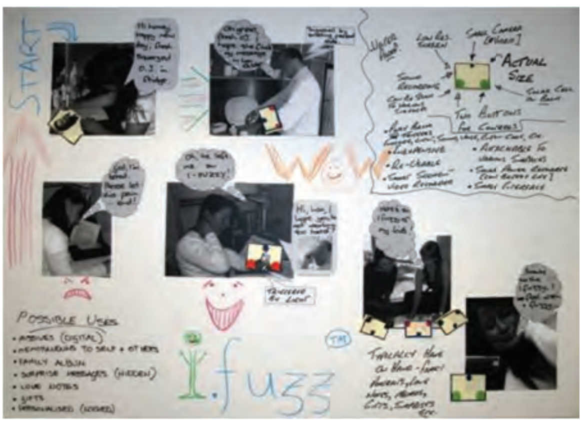


Figure 7. Poster for the mobile device 'i.Fuzz' produced in the participatory design workshop

One of the designs was a mobile device called i.Fuzz (Figure 7). The idea of the i.Fuzz concept is to provide a simple digital analogy of common analogue media, such as postcards and post-it notes. The i.Fuzz would be a cheap, light-weight, semi-disposable multimedia appliance that enabled pre-recorded messages to be left at different physical locations to be serendipitously discovered by one's intimate other. The i.Fuzz devices should thus include memory, processing power as well as facilities for recording and playing back audio and video. An example scenario, derived from our user study, illustrates i.Fuzz use. A missive for one's partner is left in their sock-drawer on an i.Fuzz pre-programmed to start playback when the drawer is opened in the morning and the i.Fuzz exposed to light. Other envisioned uses of the i.Fuzz included using it as a digital keep-sake or to carry video clips of one's children.

Importantly, this scenario is taken from the participants. A participant left a message of affection on a piece of paper hidden in a sock drawer before leaving for work. This story was an inspiration for the design workshops. It nicely encapsulated the playful, the emotional and the private themes of intimacy, and led to a feeling of presence-in-absence when the message was eventually found.

On the basis of the design sketches produced by the research team and in the two workshops, we produced a number of more refined designs and implemented a functional prototype. Some of these are described below.

6.3. SynchroMate

A more refined design responded to the user study observation that partners would sometimes send and receive messages simultaneously via SMS or email. The messages seem to "cross each other in the air". These coincidences, of receiving a message from someone while simultaneously sending that person a message, had a lasting impact on the couples. The coincidence was attributed with almost metaphysical significance, such as "a stroke of faith" or "indicating a special personal connection".

SynchroMate is a round device attached to the palm that aims to assist this "faith" a little and to make this metaphysical event more likely. It supports serendipitous synchronous communication by exchanging not only the message itself but also that a message is being composed. The SynchroMate allows intimate couples to 'stretch time' by allowing the serendipitous moment to last a little longer and to provoke anticipation. This device helps to express themes of reciprocity and serendipity.

6.4. Hug Over a Distance

Another of our early ideas took a step towards wearable computing and smart clothes; equipping partners with jackets providing couples with an open, physical and ambient channel of interaction, and enabling them to exchange a 'hug' while physically separated.

In the original idea for the Hug Over a Distance a person would be able to initiate a hug causing their remote partner's jacket to emulate, in some fashion, the feeling of a hug. This remote hug could then be rejected or reciprocated by the remote partner. When reciprocated, the jackets would remain synchronized in 'hug mode' until either of the parties chose to end the hug. This design idea was inspired by several of the themes of intimacy identified through our field study. It supports stroking and patting in a non-verbal but physical way. It is flirtatious, playful and it involves a high degree of reciprocity. Also, it permits couples to express private and discrete signs of affection in public places; a behavior identified by several of our participants.

We chose to implement a simplified, one-way version of Hug Over a Distance as a functional prototype (Figure 8) because it allows an immediate tactile interaction with an emotional content (similar to holding hands or giving a hug) and it can be used anywhere, anytime, yet discreetly and unobtrusively.

Figure 8. Hug Over a Distance prototype: showing pump, valve, battery, relay controller and wireless PDA. During use, these components are hidden inside the vest.

Hug Over a Distance uses two PDA devices connected wirelessly through TCP/IP via WiFi or through mobile phones via Bluetooth. One PDA is embedded in an inflatable, yet tastefully tailored sleeveless jacket. The other PDA is carried by the partner. If the PDA in the jacket receives a hug request, air channels embedded in the jacket are inflated to create a light, but palpable pressure on the body. After four seconds of pressure, an electronic valve opens and releases the air. The PDA sends an acknowledgement saying the hug was received, and thanks the sender with a kissing sound.

It is important to emphasize that we are not intending to recreate accurately the physical experience of a real hug using technology. Rather, we want to demonstrate, using a piece of wearable computing, that it is possible to send an emotional "ping" to a remote loved one with a tactile and unobtrusive interface. We are referring to a "hug", because the feeling of light pressure surrounding the body combined with the associated warmth is most closely related to a hug.

7. DISCUSSION

In this paper we have reported on ongoing work exploring designs ideas and concepts for technologies to support intimate acts. The significance of our contribution is threefold. First, we have presented an effective method for studying mediated intimacy in the 'wilds' of everyday life. Our approach used and interlaced cultural probes, contextual interviews, technology provocations, and participant observations to produce rich qualitative data of everyday situations of use. This approach is also applicable to the study of other forms of practice in the domestic environment.

Secondly, we have used these empirical materials to produce a nuanced and detailed understanding of mediated intimacy. This analysis and the thematic model produced have led to profound insights into the current role of technologies in intimate lives. It has also provided an empirically grounded springboard for inspiring and informing the design of future intimate technologies.

Finally, our various design activities, which included brainstorming sessions, and expert and participant design workshops generated a large number of design ideas. We have presented a small selection of these design ideas and concepts. In addition, one of these – Hug Over a Distance – was developed as a demonstrator prototype.

While these designs for future technologies have not, as yet, been evaluated, they do emerge from empirically grounded understandings of current practice. Their validity is thus grounded in the real and observable practices of everyday life, as people use a variety of appropriated technologies to mediate their intimate lives. In the future, as indicated in our user-centered design approach (Figure 1), we intend to place prototypes of varying fidelity back in the field with participants in the study, as well as with other intimate couples, for evaluation and refinement of our ideas and methods.

8. LIMITATIONS

The schematic presented in Figure 3 is not a theory of intimacy. It does not explain why or how expressions of intimacy occur. Nor is it comprehensive. We do not consider likely influences on expressions of intimacy due to culture, sexual orientation, social class, age or geography. Nor is the thematic model a list of requirements. The themes are grounded concepts intended to trigger and inspire design ideas. For example, what would it mean to design for “stroking and patting”, “routine and duty” and “work in progress”? These are not requirements in any traditional sense, but they may provoke further exploration of user needs.

We agree that probes are intended to “elicit inspirational responses from people– not comprehensive information about them, but fragmentary clues and their lives and thoughts” (Gaver et al. 2004). However, because we have chosen to visualize our probe data as themes, our results appear less fragmentary. We believe this helps to communicate our work without losing the power of provocation.

9. CONCLUSION

As Gaver points out communication of emotion is often not in the device output, but in the dynamics of use (Gaver 2002). Similarly the innovations presented in this paper are not specifically conceptual innovations (for example, Rosella and Sakai (2002) suggests the building of a Hug Over A Distance underwear device), but innovations of use. In other words, our design ideas and concepts are often a reconfiguration of existing technology, rather than speculation on future technology. Importantly, this reconfiguration is motivated by our understanding of current practices of intimate behavior. The innovation exists in the extensive analysis of the fieldwork and in the new uses that emerge from this reconfiguration.

Technology will never replace the physicality and immediacy of face-to-face contact. However, there is still much to understand about how interactive technologies can further augment, extend and support intimate experiences. Our work takes an important step along this path to understanding, and we hope it will encourage further work in the development of technologies expressly designed for mediating intimacy.

ACKNOWLEDGMENTS

The authors would like to thank all couples participating in the study and acknowledge the support of the Smart Internet CRC (project UE-06) and the Danish Technical Research Council (project 26-03-0341).

REFERENCES

Bell, G., Brooke, T., Churchill, E. & Paulos, E. Intimate Ubiquitous Computing. Proc. UbiComp Workshop (2003), 3-6

Brave, S. & Dahley, A. inTouch: A Medium for Haptic Interpersonal Communication. Proc. CHI 1997 (Ext. Abstracts - Short Talks), ACM Press (1997), 363-364

Byrne, R. & Findlay, B. Preference for SMS versus Telephone Calls in Initiating Romantic Relationships. Australian Journal of Emerging Technologies and Society, Issue 2, Autumn (2004).

Cheal, D. Showing them you love them: gift giving and the dialectic of intimacy. The Sociological Review, 35, 1 (1987), 150-169.

Chelune, G., Robison, J., Kommor & Martin. A Cognitive Interactional Model of intimate Relationships, in Communication, Intimacy, Intimacy and Close Relationships, V. Derlega, Editor. Academic Press: Orlando, FL. (1984), 11-40.

Cheverst, K., Clarke, K., Dewsbury, G., Hemmings, T., Kember, S., Rodden, T. & Rouncefield, M. Designing Assistive Technologies for Medication Regimes in Care Settings. International Journal of Universal Access in the Information Society, (2003)

Donn, J.E. & Sherman, R.C. Attitudes and Practices Regarding the Formation of Romantic Relationships on the Internet. CyberPsychology & Behavior, 5, 2 (2002), 107-123.

Gaver, B. Provocative Awareness. Computer Supported Cooperative Work, 11, (2002), 475-493.

Gaver, B., Boucher, A., Pennington, S. & Walker, B. Cultural Probes and the Value of Uncertainty. Interactions, XI, 5 (2004), 53-56.

Gaver, B., Dunne, T. & Pacenti, E. Design: Cultural Probes. Interactions, 6, 1 (1999).

Geertz, C. The Interpretation of Culture. 1973, New York: Basic Books.

Goffman, E. The Presentation of Self in Everyday Life. 1959, New York: Doubleday.

Goodman, E. & Misilim, M. The Sensing Bed. Proc. UbiComp 2003 Workshop (2003)

Grinter, R.E. & Eldrige, M.A. Y Do Tngrs Luv 2 Txt Msg? Proc. ECSCW '01, Kluwer Academic (2001), 219-238

Harper, R., ed. Inside the Smart Home. 2003, Springer: London.

Howard, S., Carroll, J., Murphy, J. & Peck, J. Managing Innovation in Scenario Based Design. Proc. Human Factors, CHISIG (2002)

Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., Eiderback, B., Lindquist, S. & Sundblad, Y. Technology Probes: Inspiring Design for and with Families. Proc. CHI 2003 (2003), 17-24

Kaye, J.J. & Goulding, L. Intimate Objects. Proc. DIS, ACM Press (2004)

Kjeldskov, J., Gibbs, M.R., Vetere, F., Howard, S., Pedell, S., Mecoles, K. & Bunyan, M. Using Cultural Probes to Explore Mediated Intimacy. Proc. OZCHI (2004)

Markopoulos, P., Romero, N., van Baren, J., de Ruyter, B. & Farshchian, B. Keeping in Touch with the Family: Home and Away with the ASTRA Awareness System. Proc. Late Breaking CHI 2004, ACM Press (2004), 1351-1354

Monsour, M. Meanings of intimacy in cross- and same-sex friendships. Journal of Social and Personal Relationships, 9, (1992), 277-295

Moss, B.F. & Schwebel, A.I. Defining Intimacy in Romantic Relationships. Family Relations, 42, 1 (1993), 31-37

Pedell, S., Graham, C., Kjeldskov, J. & Davies, J. Mobile Evaluation: What the MetaData and the Data Told Us. Proc. OZCHI 2003, CHISIG (2003), 96-105

Register, L. & Henley, T. The Phenomenology of Intimacy. Journal of Social and Personal Relationships, 9, (1992), 467-481

Robson, D. & Maggie, R. Intimacy and computer communication. Vol 26, Issue 1, Feb (1998), 33-42.

Rosella, F. & Sakai, R. F+R HUGS. http://www.interaction-ivrea.it/en/news/press/ releases/2002/bigtorino/hugs/index.asp

Strong, R. & Gaver, B. Feather, Scent and Shaker: Supporting Simple Intimacy. Proc. CSCW'96, ACM Press (1996)

Taylor, A. & Harper, R. The Gift of the Gab?: A Design Oriented Sociology of Young People's Use of Mobiles. Computer Supported Cooperative Work, 12, 3 (2003), 267-296.

Tollmar, K. & Joakim, P. Understanding Remote Presence. Proc. NordiCHI (2002)

van der Hoog, W., Keller, I. & Stappers, P.J. Gustbowl: Technology Supporting Affective Communication through Routine Ritual Interactions. Proc. CHI 2004 (Ext. Abstracts - Late Breaking), ACM Press (2004), 775-776

Chapter 10

Ethnography and object-orientation

Jesper Kjeldskov and Jan Stage

Abstract. There has been a lot of interest in ethnography within human-computer interaction over the last two decades, and its relevance within systems development is today beyond question. However, one of the challenges reported is that ethnography generates findings and knowledge with such contextual richness that it can be hard to transfer into system design. In the light of recent years' push for the use of ethnography within the area of *mobile* human-computer interaction, this challenge has resurfaced and is of renewed importance to the research field. In this article we describe an interdisciplinary combination of ethnography with a structured software engineering method supporting the transition from collected data to design and implementation. We explore this combination through two case studies of mobile system development for supporting distributed work activities within industrial process control. We show that when developing mobile systems ethnographic data is a highly valuable source of input for developing object-oriented models by providing contextual richness, and that in turn, objected-oriented analysis is a highly valuable method for working with ethnographic field data in systems development by supporting the creation of abstract models. Combining the two, we have a method where ethnographic field studies inform core system design.

1. INTRODUCTION

In a literature review from the early 2000s (Kjeldskov and Graham 2003), it was found that the majority of research within the area of mobile human-computer interaction was based on applied research and laboratory experiments. One of the consequences of this was that research was poorly grounded in real world activities and subscribed to the assumption that we already knew what problems to solve – just not exactly how to solve them. In response to this, the mobile HCI community was urged to explore the

Originally published as Kjeldskov, J. and Stage, J. (2012) Combining ethnography and object-orientation: contextual richness and abstract models for mobile interaction design. *International Journal of Human-Computer Studies,* 70(3), 197–217.

use of more contextually grounded methods, such as ethnography, as illustrated by, for example, the work on mobility in collaboration by Luff and Heath (1998). In a follow-up review from 2010 (Kjeldskov 2012) it was found that although fieldwork is now much more common, there is still not a strong link to design and development. One of the problems with ethnography for informing design and development of computer systems is that it sometimes generates findings and knowledge with such contextual richness that it can be hard to translate into system design.

There has been a lot of interest within human-computer interaction over the last two decades in how ethnography can inform system design (Crabtree et al. 2009), and the challenge of transition from collected field data to requirements analysis has been described many times in the literature for over a decade (Paay et al. 2009, Kjeldskov et al. 2006, Constantine et al. 2003, Viller and Sommerville 2000, Hughes et al. 1994). With the push towards using ethnography within the area of *mobile* HCI this challenge has resurfaced and is of renewed importance. Hence the aims of our work presented here are similar to that of the references above, but with a particular focus on mobile computing. As our scope for this research we wish to, firstly, address the overall interplay between ethnography and existing object-orientated methods, rather than the development of new specific modelling techniques as in, or example, Viller and Sommerville (1999) and Iqbal et al. (2005). The latter, however, is an apparent opportunity for further research in light of the case studies presented here.

When designing and developing a computer system, a level of formalization is always required at some point in the process. Very often, this is where findings from ethnographic studies are lost or summarised beyond recognition, simply because they do not translate well directly into system specifications. As a consequence, even though ethnographic field studies might teach us a lot about the use context of our future systems, this knowledge is often difficult to trace through to the resulting designs. This is also the case for mobile systems.

In response to this, we have explored an interdisciplinary combination of ethnography with a structured software engineering methodology that is known for effectively supporting actual system development, namely that of *object-oriented analysis and design* (OOA&D) as devised by Mathiassen et al. (2000). The object-oriented analysis and design methodology has proven its value in systems development by producing formalized and abstract models that can be directly implemented in an object-oriented programming language (for an insightful discussion of different methodologies and representations in human-computer systems development including object-orientation see Benyon 2002). In the specific OOA&D method described by Mathiassen et al. (2000) the development of these models is done as a part of an analysis of a system's *application domain* and its *problem domain* respectively. Based on concrete experience with this method, it is our position that ethnographic field study data is highly valuable as a source of input for such application and problem domain analyses by providing *contextual richness.* In turn, objected-oriented analysis is highly valuable as a method for working with ethnographic field study data by supporting the creation of *abstract models.* By explicitly combining the two, we have a method where ethnographic field studies can influence the core of a system design. This is what we intend to show in this article.

We present our experiences with the development of two mobile systems using a combination of ethnography and OOA&D. The first case study concerned the operation of a large container ship (Kjeldskov and Stage 2003). The second one concerned the operation of a large power plant (Kjeldskov et al. 2006). We present the two development process case studies in detail, showing how they each followed the same overall process from ethnographic field studies, through object-oriented analysis and design, to conception of prototype systems that could be evaluated with prospective users. The second case study naturally built on the lessons learned from the first one, and we show how this led to modifications to the method applied.

The paper is structured in the following way. First we introduce related work. We then describe the overall method applied for our two case studies. Sections 4-9 take you through the five phases of our development process for each of the two case studies and describe the two resulting functional prototypes. At the end of each of these sections, we outline what was learned in that phase, and how the method was modified for the second case study based on the experiences with the first one. Section 10 summarizes and discusses the lessons learned for the different phases of the method, and for the application of the method as a whole. Section 11 concludes and outlines further work.

2. RELATED WORK

Mobile and wearable computer devices and applications are being developed for a broad variety of use areas, and recent research has devoted much attention to investigating how in particular distributed mobile workers can benefit from mobile computer systems. This includes work settings where people are concerned with computerized information and processes of critical importance remote from their current location. Early examples include distributed process control and error diagnosing in wastewater treatment plants (Nielsen and Søndergaard 2000), and the use of mobile multimedia for telemedicine and early diagnosing in emergency ambulance services (van den Anker and Lichtveld, 2000). More recent examples include logistic vehicle management at automobile terminals (Rügge et al. 2009), vessel control within the fishing industry (Lumsden et al. 2008), and mobile technologies in hospital settings (Tang and Carpendale 2008, Skov and Høegh 2006, Bardram and Hansen 2004). In order to avoid mistakes and accidents in such safety critical work domains, it is important to understand the state of the systems being operated. This is a classic challenge within human-computer interaction (HCI) and Software Engineering (SE) in response to which Rasmussen (1983, 1986) suggests that computer interfaces should be designed to improve operators' reasoning about the *domain of operation*, and Norman (1990) suggest that systems should support *human interaction* rather than total system automation. These are good, general, recommendations – but in order to achieve this, it is important that system developers have a great deal of rich, detailed, and structured understanding of the use domain, and that the structures, relationships and processes of the use domain are clearly reflected in the core of resulting system designs.

Analysis, design, and prototyping methods are general means for supporting development of user interfaces and software systems, and a lot of attention has been given to their development and refinement for traditional, stationary, computer technologies.

For mobile systems, however, although it has been argued that these necessitate re-thinking of established systems development methods (Krogstie et al. 2004), generally very little has been published on the topic. Instead the field of mobile computing is characterized by research describing specific systems, their technical implementation, their context, and their use (Kjeldskov 2012, Hosbond and Nielsen 2005, Kjeldskov and Graham 2003). Exceptions include Sharples et al. (2002) and Mikkonen et al. (2002) who outline a number of differences from the development of traditional software within the area of HCI, and (Hosbond 2005) who presents a case study of development practice in seven companies within the mobile industry. Given the inherently "situated" characteristics of mobile systems use contexts, an obvious candidate approach for informing technology development is ethnography.

Ethnography in HCI is a methodology originating from the social sciences and used to provide descriptions of human activities in a particular context through studies in their fullest possible context, such as observational field studies and contextual interviews. Although there is currently a trend towards returning ethnography in HCI to its anthropological origins (Crabtree et al. 2009) and focus on the broader study of culture, it is the use of ethnography for informing systems design that we are interested in here. Like Crabtree et al. (2009) we believe that this is done best through structured in-depth analysis rather than through literary accounts based on selective observation and reporting.

In reviewing the literature from the last half a decade it is clear that ethnography, in various forms, has increasingly been applied as a research methodology within the area of mobile human-computer interaction since 2003 (Kjeldskov 2012). In our opinion this is a positive trend for the research field as it indicates a shift in perspective from a singular focus on technology to a more holistic view on technology-in-use. However, one of the problems with ethnography for informing design and development of computer systems, mobile or not, is that it generates primarily descriptive findings and knowledge with such contextual richness that it can be hard to transfer into specific system design. Looking at recent literature, which reports only very few accounts of any kinds of field work directly informing design and implementation of mobile systems, this still appears to be a challenge (Kjeldskov 2012).

There has been a lot of research into bridging the gap between ethnography and software engineering over the last decade (Paay et al. 2009). The essence of this challenge is that it can be difficult to transfer findings and knowledge from ethnographic studies to system design and implementation (Hughes et al. 1997). This is partly caused by a fundamental difference in thinking and working within the social sciences and engineering. As pointed out by Viller and Sommerville (1999), ethnography deals in "the particular" while software engineering deals in "the abstract". Extending this further ethnographers work with rich and concrete descriptions and understandings, while software engineers work with formal models and abstractions (Paay et al. 2009). In respect to their outputs, "Ethnographers avoid judgements; designers make them" (Paay 2008).

Hence, HCI researchers have struggled with the challenge of making these two communities of practice function well together in the creation of software systems that are both grounded in contextual richness and abstract models. From a team composition perspective, approaches have been explored that aim to integrate social analyses better into design processes through creation of multi-disciplinary teams comprising both ethnographers and software engineers (Diggins and Tolmie 2003, Hughes et al. 1995). From a data representation perspective, approaches have been explored that present outcomes from ethnographic field studies in software engineering terms, such as UML notation (Viller and Sommerville, 1999). However, bridging the gap between ethnography and software development remains a current research challenge (Schraefel et al., 2004; Walenstein, 2003; Wiltshire, 2003), as pointed out in the overview of related literature provided by Paay et al. (2009). Motivating further research within this topic, there are several examples where ethnographically based methods have demonstrated usefulness in the requirements gathering phase of software development projects, for example Crabtree (2004), Viller and Sommerville (2000), Simonsen and Kensing (1998), and Hughes et al. (1994).

3. AN EMPIRICAL STUDY OF BRIDGING THE GAP

Between 2001 and 2006 we conducted two related case studies with the aim of developing mobile information and communication systems for supporting distributed work activities within industrial process control. The second case study built on the experiences from the first one in terms of the system developed and the process followed. Both case studies explored the development and use of mobile computer systems devices for making communication persistent and providing formalized representations of a shared, safety-critical, work domain. Both case studies followed the same overall process of combining ethnography object-oriented analysis and design for obtaining contextual richness as well as abstract models as foundations for the design of functional prototype systems. The central activities of the process, and how they influenced each other, are illustrated in figure 1.

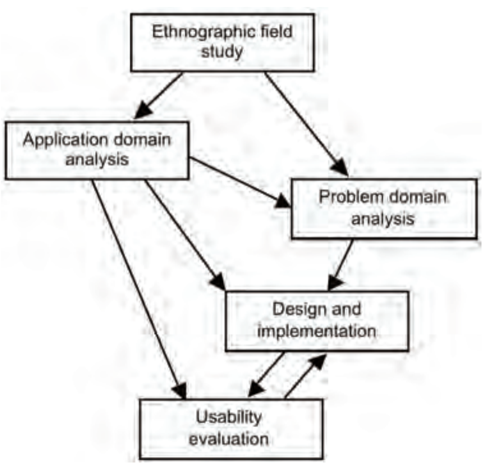


Figure 1. The five major activities in the two development processes

The development processes involved 1) ethnographic field studies, 2) application domain analysis, 3) problem domain analysis, 4) design and implementation of a functional prototype, and 5) usability evaluation of the system in use. The ethnographic field study activities involved observational studies and contextual interviews (Beyer and Holtzblatt 1998). The object-oriented analysis and design activities followed the method described by Mathiassen et al. (2000) with additional input from Nunes and Cunha (2001a, 2001b) and Dayton et al. (1998). The usablity evalutions followed the methods described by Kjeldskov and Stage (2004) and Kjeldskov and Skov (2007). In the following sections, we describe each of these five steps for each of the two case studies. We describe the two case studies in parallel, structured by the five steps of the development process in order to show their similarities, how the second study built on the first, and how it was modified (and improved) on the basis of lessons learned.

4. ETHNOGRAPHIC FIELD STUDIES

The starting point for the two development processes was an ethnographic field study of work activities in the specific domain of industrial process control in focus. These are described below.

4.1. Case study 1: the container ship

The first case study emerged from a larger study into work activities in the maritime domain involving computerized process control and information systems on board Maersk-Sealand container ships. Maersk-Sealand operates some of the world's largest container ships of sizes equivalent to 3½ soccer fields (figure 2). The operation of such a ship is safety-critical. Especially when manoeuvring inside a harbour, erroneous actions may result in the ship running aground, into the quay, or colliding with other ships. In either case, such collisions would cause serious material damage, potentially severe injuries for personnel, and possible loss of human life.

Figure 2. Sine Maersk at the terminal. Numbers indicate positions of cooperating teams

The field studies were conducted over a period of several months in 2000 and 2001 and involved researchers taking part in a number of voyages on board large Maersk-Sealand container ships and other ships. During these voyages, data consisting of high quality video and audio recordings and written notes regarding work activities, communication and technology use during operation of the ships was gathered through observations and interviews in situ.

Apart from informing new interface design for existing maritime instruments (Andersen and May 2001), a number of work activities were identified in which the use of mobile computer systems could be useful. These included diagnostic and maintenance work in the engine room, surveying the condition of reefers during voyages, locating personnel in case of accidents, and supporting various distributed collaborative work activities. Of particular interest to us, the field studies identified some general limitations in means of communication and coordination in relation to the operation of "letting go the mooring lines" before departing from harbour. This particular operation was therefore chosen for further investigation, leading to the design and development of the "Maritime Communicator".

4.2. Case study 2: the power plant

The second case study grew out of the first study and an interest in our work on mobile computing from a regional power plant (figure 3). The Northern Jutland power plant is a large coal-based power plant situated on the outskirts of Aalborg, Denmark, that produces central heating and electricity to the region. The operation of the power plant is safety-critical because it involves the operation of large industrial machinery and there are particular fire dangers. The collaboration with the Northern Jutland power plant allowed us to further explore some of the design ideas embodied in the Maritime Communicator, and the methodological approach used to develop it, but in a different domain. This led to the design and development of the "Power Plant Communicator".

Figure 3. The Northern Jutland power plant

For the field studies we took the same approach as with the container ships. We wanted to gather rich empirical data about the use domain before engaging in design, which required spending long periods of time in the field – quite literally getting our hands dirty. After initial enquiries, it was suggested by our collaborators at the power plant that we focus on work activities in the fuel department, as the operation of this part of the plant is essential for ensuring continuous production of energy. As the fuel department also involves workers distributed over a large physical area, relying on centralized computerized controls and coordination based on spoken communication in a noisy environment, this was comparable to the operation of "letting go the lines" on board the container ships.

The field studies consisted of a series of visits to the fuel department of the power plant interviewing the workers in situ and observing their work areas and tasks. This took place over a period of two months. The visits were documented with photographs of work places and artefacts, and through video recordings of the way key tasks were carried out. This provided a rich understanding of the application domain and detailed insight into the communication problems experienced during the work.

5. APPLICATION DOMAIN ANALYSIS

One of the key characteristics of the OOA&D method devised by Mathiassen et al. (2000) is the distinction between *problem* and *application* domain in the analysis phase. Mathiassen et al. (2000) defines the problem domain as "the part of a context that is administrated, monitored, or controlled by a system". Accordingly the application domain is defined as "an organization that administrates, monitors, or controls a problem domain". This distinction is introduced in order to guide the focus of the analysis activity towards the broader use context (application domain) and the specific part of this context that the system is going to be concerned with (problem domain) respectively. The sequence of the application and problem domain analyses is not prescribed but should be informed by the specific system development case and process at hand. When building on top of an ethnographic study, starting with the application domain analysis facilitates a gradual narrowing of focus from the broader context to the specific system, and was found most suitable.

The central principles of the application domain analysis are to describe the current situation in terms of processes, structures, and problems, and to identify actors and use cases. These are described for both case studies below.

5.1. Case study 1: the container ship

When a container ship is ready for departure, the first step in leaving the harbour is "letting go the lines" that holds it in a fixed position (figure 4) and heaving these lines onboard the ship.

Figure 4. The aft mooring of Sally Maersk

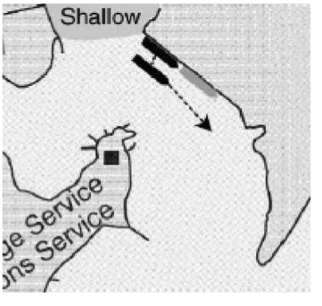


Figure 5. Sally Maersk leaving Felixstowe harbour

While this might seem a trivial procedure, it is in fact both complex and dangerous. As the physical space inside harbours is quite restricted and means for precisely manoeuvring large ships are limited, all lines cannot simply be released simultaneously. Furthermore, when a line is let go, it will remain in the water for a period of time during which no means of propulsion is available due to the risk of lines getting sucked under water and wrapped around the propeller or into the thrusters. Instead the ship can be pulled ahead or astern by means of operating the winches on the remaining lines. Due to these premises, mooring lines are released sequentially in accordance to the need for manoeuvring in a given situation, bringing the ship out side wards and then going ahead when clear of the quay and other ships nearby (illustrated in figure 5).

Due to the huge size of container ships, the work tasks involved when letting go the lines are distributed among a number of co-workers located at strategic positions (see figure 2). On the bridge (1) chief officers control the rudder, propeller and thrusters; and fore and aft (2 and 3) the first and second officers control the winches for heaving in the lines. Ashore, two teams of assistants lift the lines off the bollards. To insure the safety of the operation, individual work tasks are carefully coordinated and carried out under strict command of the captain. Communication between co-workers in the maritime domain is primarily spoken. While people on the bridge can see and hear each other directly, personnel on deck are, however, out of direct visual and audio contact, and communicate via walkie-talkies.

In order to carry out the operation of letting go the lines in a safe manner, the captain needs overview and control over the propulsion, direction and mooring of the ship. While information about the rudder, propeller and thrusters are available on dedicated instruments on the bridge no information is available about mooring. This only exists as a mental model in the head of the captain. As this mental model is highly sensitive to errors or misunderstandings in the ongoing communication between bridge and deck, and since disparity between the captain's mental model and the real world may cause wrong decisions to be made, considerable resources are spent on establishing and maintaining common ground (Clark and Schaefer, 1989) among the distributed co-workers.

Observed communication problems

Supporting operations on board the container ship, well-established rules and procedures exist for oral communication such as confirming status reports and commands by repeating them back to their sender. To a large extend these procedures work very well. However, from the observations and interviews a number of key limitations emerged related to the use of spoken communication for coordination of collaborative work activities: Sound quality is often poor, communication is not persistent but is time consuming and suffers from bottlenecks due to multiple parallel treads.

Overcoming or reducing (at least some of) these limitations served as an overall motivation for the design of the Maritime Communicator. Inspired by chat applications, newsgroups and Short Messaging Service (SMS) it was our belief that shifting to text-based communication on mobile devices would provide an asynchronous channel making communication persistent and requiring low cognitive overhead (Churchill and Bly 1999, Popolov et al. 2000).

Communication structures

In order to understand the communication happening while letting go the lines better, video recordings from the bridge of Sally Maersk during multiple instances of manoeuvring inside harbours were transcribed, coded and analyzed.

This was used to derive the use cases involved with coordinating and carrying out the operation in terms of the sequence of commands issued by the captain and the structures of subsequent communication between bridge and deck. Furthermore, a complete set of utterances: commands, confirmations and status reports, necessary for coordinating the whole operation was produced. Analyzing the typical sequences in the transcriptions of commands issued by the captain also revealed that some commands would always precede others. Also, the possible next commands at any point of the operation could always be deduced. Not surprisingly, this structure was found to correspond to the basics of the "conversation for action" model in speech-act theory (Winograd and Flores, 1986). Hence we realized that if parts of this structure were modelled in the system, the current stage of each step of the operation could be formalized and, for example, represented graphically or integrated with other computer-based data. Also, the potential next stages of any step of the operation could be identified, and possible utterances by any of the communicating actors be deduced, and possibly prioritized.

5.2. Case study 2: the power plant

The power plant is divided into two separate production plants (locations #7 in figure 6). The fuel for the two plants is supplied from a large central coal storage area (locations #2 and #3).

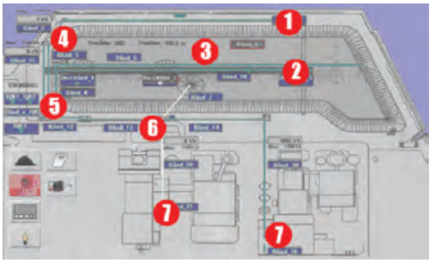


Figure 6. Key locations of the fuel department

Figure 7. Coal storage area

The fuel department is responsible for delivering the coal used in the two production plants, amounting daily to 5000 tons of coal for each. The employees in the fuel department continuously monitor and control the transportation of coal and must ensure that the correct amount of coal arrives at the correct location, and that the coal has certain properties and quality. In order to ensure this, the coal is filtered and grinded (locations #4 and #5). After the coal is processed, it is transported to the two production plants (locations #7) by means of underground conveyer belts. Another important task for the workers in the fuel department is to prevent the coal from self-combusting in the storage area.

Communication to support coordination

The workers perform a variety of different tasks to ensure that the amount of coal needed is delivered to the two production plants. In order to coordinate the many tasks described above, quick and easy communication is important, and in some cases even essential, in order to carry out the job in a safe and efficient manner. Communication happens via VHF radios, DECT wireless phones, and some times mobile phones. Currently, the control tower (location #6) is the only place where all necessary information is accessible. It is also the place where workers operate and control most of the machinery.

Figure 8. Coordination between workers in the control tower and on the ground using VHF radios and mobile phones

The overall operation of coal transport can be controlled by means of a computerized process control system. However, when a problem occurs that cannot be solved from the control room, for example in the grinder building (location #5), the personnel sent to the site to solve the problem do not have direct access to the control system. Conversely, the specific parts of the individual machinery distributed throughout the plant (for example physical handles on the grinder) can only be operated on site. Consequently, full control of the plant requires communication and coordination between personnel on site and in the control room (figure 8).

Observed communication problems

Although DECT phones, VHF radios, and mobile phones offer flexibility and portability, several problems were reported and observed. Firstly, since the conveyor belts run underground and many machines are located inside solid concrete buildings, radio communication is not always reliable due to lack of signal strength. Secondly, there is typically a deafening noise in the tunnels under the plant and inside the buildings, which makes talking to each other difficult and the use of any kind of mobile device for verbal communication virtually impossible. In summary, the workers experience communication problems related to three issues. One is loss of signal, another is noise, and a third is lack of information access on the ground.

In terms of communication structures, similar sequences and confirmation procedures were observed as on board the container ships. This supported our belief that a text-based communication device might also be of value here. It was, however, observed that due to a larger number of locations and objects (i.e. machinery) a wider vocabulary of nouns was used in the operation of the fuel department than in the operation of letting go the lines. The vocabulary of verbs was, however, similarly small, and the same verbs were used in relation to a lot of different machinery (e.g. start, stop, reverse, assist).

5.3. What did we learn from the two application domain analyses?

The two application domain analyses provided a rich, yet structured, understanding of the use context that we were dealing with. Grounding the analyses in ethnographic field studies meant that a lot of detail was captured, and that our initial design ideas were directly inspired by the use context.

In terms of the application domain analyses there was no notable difference between the two case studies. Both mapped out the overall use context, specific work tasks, means and procedures of coordination, revealed problems related to the use of spoken communication, and discovered structures in communication.

The application domain analyses provided input for the problem domain analysis phase in the form of focus and data, for the design and implementation phase in the form of system context, and for the usability evaluations in the form of information for recreating realistic use contexts and developing realistic and relevant tasks.

6. PROBLEM DOMAIN ANALYSIS

At this point in the two case studies, we had identified specific work areas of focus, overall division of work, and sequences and structures in communication. As is often the case in projects like these, we had also already begun thinking about design of new technology solutions for the work activities and problems observed. However, rather than moving straight into the design phase, we conducted a second round of analysis as prescribed by Mathiassen et al.'s (2000) OOA&D method. The second round of analysis focused on the problem domain, defined as "the part of the context that is administrated, monitored, or controlled by a system" (Mathiassen et al., 2000). The aim of the problem domain analysis is to support later design and implementation by identifying objects, classes and relations, and providing detailed descriptions of the states that these objects go through. The outcome of the analysis is a description of these objects, classes, structures, and behaviors. The outcomes of the two problem domain analyses are described below, including some of their implications for the design of the resulting prototype systems.

6.1. Case study 1: the container ship

The two major activities of the problem domain analysis for the container ship case study were the modelling of a class diagram for the problem domain (figure 9) and a state chart diagram for the *Let go* class.

Class diagram

Modelling the class diagram was a challenge primarily because we extended the problem domain to include additional work activities related to manoeuvring inside the harbour other than the operation of letting go the lines. This was done to support extending the scope of the Maritime Communicator later without having to change the model.

From the class diagram, we identified a number of issues, which directly informed the design of our prototype system. Firstly, the class diagram illustrated that letting go the lines is one of three similar operations (or tasks) related to manoeuvring inside the harbour. On the basis of this, we found that the Maritime Communicator should have facilities for tailoring an interface supporting each of these specific operations on the basis of a general design for supporting communication and coordination. Thus, general interface elements should be associated to the overall Task class while more specialized interface elements, such as specific graphical representations, should be associated with their appropriate specialization.

Secondly, we found that since the *Commanding officer* class and the *Officer* class aggregate from fundamentally different classes (*Ship* and *Team* respectively), different interfaces for these actors might also be appropriate. Thus the interface for the commanding officer should possibly include information about the ship while the interface for the officers on deck should include information about the team instead.

Finally the class diagram helped us realize the role of the *Location* class as a mediator between the *Ship* and *Task* classes. Thus knowing the location of a mobile device would enable us to deduce the role of its user and adapt the interface accordingly.

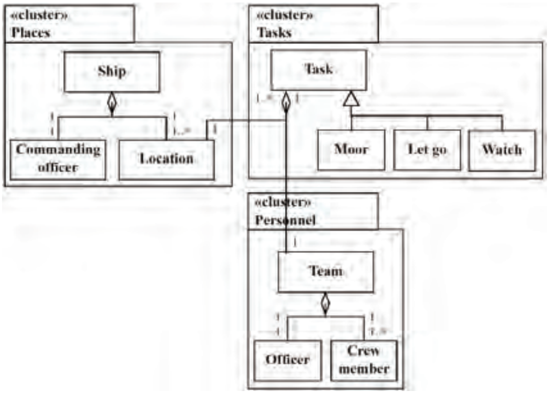


Figure 9. Class diagram for the problem domain

State chart diagram

As the *Let go* class would contain the structure of the operation and communication, a lot of effort was put into modelling its state chart diagram (figure 10). Even though valuable input for this already existed from the application domain analysis, making a precise model was a challenge. Firstly, a number of variations in the sequence of the operation had to be modelled. Secondly, we had to model some commands as implicitly containing others, depending on the situation. Finally, we had to model that several commands were being executed in parallel, and could have influence on each other. The state chart diagram was created on the basis of the video recordings and transcripts of communication gathered through the field studies.

The state chart diagram for the Let go class provided a detailed view of the operation of letting go the lines. First of all, it provided both the details and abstraction needed to actually implement the idea of partially automating sequence and communication that emerged from the application domain analysis. Secondly, it showed that each command goes through three overall temporal stages of being imminent, executing, and ended. While in some situations ended commands may be important, in other situations only executing commands are vital. From this finding, we were later able to divide the interface into three corresponding areas reflecting enabling a very simple differentiation of priority.

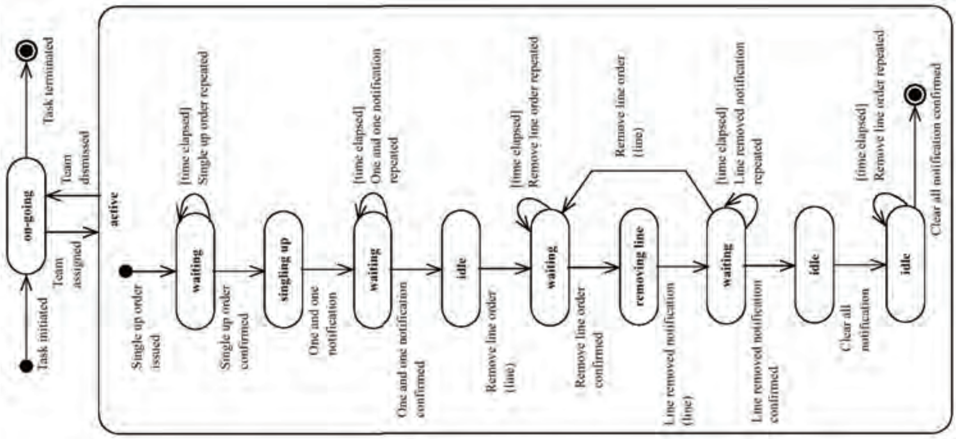


Figure 10. State chart diagram for the Let go class

6.2. Case study 2: the power plant

In the power plant case study, we modelled the problem domain in a class diagram that captured the most central objects in the use context of a mobile information and communication system for workers in the fuel department. Informing the development of the class diagram, we also did user profiling in order to describe in more detail the workers whose activities needed to be supported by the system. The outcome of this was the identification of the two primary roles of *controller* and *field worker.*

Providing additional detail for the subsequent design and implementation phase, the problem domain analysis was extended with additional modelling techniques from the related Wisdom method (Nunes and Cunha 2001a, 2001b). This prescribed the identification of, for example, essential tasks and associated state chart diagrams. Modelling these essential tasks in state chart diagrams, however, turned out to be more difficult because many tasks at the power plant were only carried out rarely, and not all of them had happened during our ethnographic field study. In order to overcome this problem, we decided to intervene and stage a series of situations in which the workers *acted-out* (Howard et al. 2002) work activities, communication and coordination in real world settings. This allowed us to observe these rare situations, adding to the ethnographic data. On the basis of the joint body of data, we identified nine essential tasks that the system needed to support. For each of these tasks, a state chart diagram capturing the flow of the task was produced (see figure 11a).

Motivated by the lessons learned from the container ship case study, where the usability evaluation of the functional prototype had later revealed a problem with one of the state chart models from the problem domain analysis, we wanted to validate these models with the users. As we were also interested in some early feedback on a number of design ideas spawned from our analysis, the model validation was integrated into a paper prototype evaluation (Snyder, 2003). The paper prototype consisted of a series of screens drawn on paper, which could be placed on top of a PDA (figure 11b). It was evaluated with real users on site at the fuel department through a series of sessions where a number of workers acted-out their use of the prototype. This provided feedback on both the envisioned functionality and the prototype's structural design, enabling us to evaluate the applicability of our overall design ideas, as well as the underlying models. The evaluation resulted in some modifications to the essential task models, and yielded a number of useful new design ideas. The final class diagram included 9 classes, and the functional list for the system comprised 22 functions.

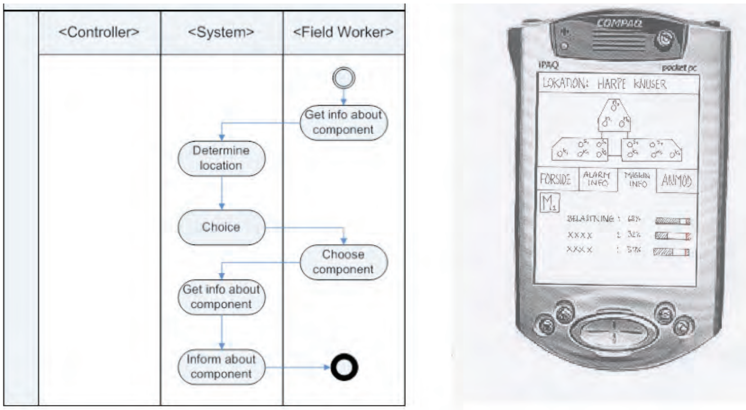


Figure 11. State chart diagram for getting information about a component (a) and the corresponding user interface of the first paper prototype (b)

6.3. What did we learn from the two problem domain analyses?

The two problem domain analyses reduced the complexity of information from the application domain analyses to a series of abstract models, while still preserving the essence of contextual grounding obtained through the ethnographic field study. As described above, creating object-oriented models based on our ethnographic data and application domain analysis not only formalized important aspects of the immediate system context but also generated additional insight into this context than had been gained from the previous phases alone. We learned more about the specific division of labour in the two work domains, the complexities of tasks, and the structures of operations and communication. This additional insight was essential for the subsequent design of our mobile prototype systems.

The two problem domain analyses differed in some specific aspects. While the overall focus and purpose was the same, the second case study deployed additional techniques from the Wisdom method (Nunes and Cunha, 2001a; 2001b) prompting the development of additional models, for example of essential tasks. This was done to explore the added value of these models when working with ethnographic data and mobile human-computer interaction design. As a result of this addition, we were able to see that some information was missing from our ethnographic field studies, and go back and gather additional data at a relatively early stage of the development process.

7. DESIGN AND IMPLEMENTATION

The models created in the problem domain analysis formed the basis for the design and implementation phases of the two case studies.

7.1. Case study 1: the container ship

Following the analysis phases, a team consisting of the authors of this article engaged in the design and implementation of the Maritime Communicator. The design and implementation process consisted of four iterations each producing design documents on different levels of abstraction (figure 12). Apart from being directly informed by the application and problem domain analyses, the designs produced were inspired by related literature on text-based communication such as Popolov et al. (2000), Smith et al. (2000), Smith and Fiore (2001) and related systems such as chat, newsgroups and the Short Messaging Service (SMS).

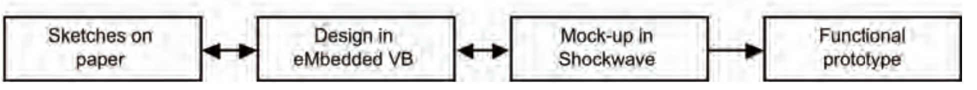


Figure 12. The four stages of our 1st design and implementation phase

Sketches on paper

First, a series of paper based sketches of possible interface designs were produced on the basis of the analysis and inspired by the use of multi-threaded communication and pre-defined messages in other systems. These paper mockups facilitated fundamental discussions of the basic interface design and led to an overall concept providing: 1) a graphical representation of the ship and its mooring lines; 2) a transcription of the communication sorted by objects of reference; and 3) facilitating communication through pre-defined textual commands selected from lists.

Design in eMbedded Visual Basic

On the basis of the paper sketches, detailed design was produced in Microsoft eMbedded Visual Basic. This forced the design team to work within the limitations of the target platform in terms of screen size and graphical interface elements supported in the development toolkit. Consequently, a number of shortcomings arose related to the specific division of screen real estate and the desired level of detail of graphical representations. While most of the subsequent refinements of the design were done directly in Visual Basic, larger design issues such as how to support the textual representation of multiple parallel threads of commands, in what turned out to be a *very* limited graphical space, temporarily forced the design team back to pen and paper.

Mock-up in Shockwave

Screen exports from eMbedded Visual Basic (modified in Adobe Photoshop) were used to produce a Shockwave-based mock-up in Macromedia Director showing a possible sequence of communication. Adding life to the Visual Basic-based design in this "quick-and-dirty" fashion facilitated further discussions within the development team and resulted in minor design modifications before doing any real programming. For example, it was decided that the user should not be prompted for confirmations through pop-up screens (as these would temporarily cover vital information). Instead, confirmations were included in the existing list of possible utterances.

Functional prototype
After a final design had been agreed upon, this was implemented in a functional prototype using Microsoft eMbedded Visual Basic. First, the underlying data structure of the system was implemented in accordance with the class diagram and state chart diagrams produced in the object-oriented analysis. Following this, the user interface was implemented in accordance to the final design. Once the user interface and interaction worked properly we implemented network interfaces and protocols for distributing the system on multiple devices. Implementation amounted to a total of 15 person-days.

7.2. Case study 2: the power plant

The design and implementation phase for the Power Plant Communicator followed a slightly different process (figure 13). System design and implementation was delegated to a group of four developers under the management of the authors, and the design of the Maritime Communicator, along with evaluation data, was provided as a starting point. Instead of design sketching we included an additional object-oriented modelling activity based on the Bridge method (Dayton et al. 1998), and used this activity to direct our specific user interface design.

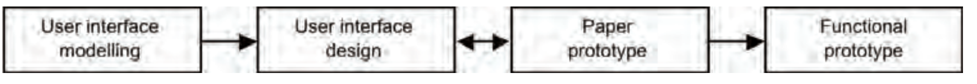


Figure 13. The four stages of our 2nd design and implementation phase

User interface modelling
As indicated by its name, the Bridge method *bridges* between problem domain analysis and user interface design. This is done through the development of three related models of the user interface on different levels of abstraction: 1) *interaction model,* 2) *presentation model,* and 3) *dialogue model.*

The interaction model describes the interaction spaces and tasks that the system should support. An interaction space is an abstract user interface element, which still has no specific graphical properties. A task is an activity that is carried out by a user employing one or more interaction spaces (figure 14). The left column of the interaction model contains the interaction spaces. The second column contains the tasks, and the three columns on the right represent internal elements of the system.

The presentation model describes each interaction space in more detail, specifying on an abstract level what each individual screen should contain in terms of output elements, input elements and possible user actions but without saying anything concrete about what these elements should look like (figure 15). The attributes of the presentation model classes are defined on the basis of the class diagram for the problem domain. The operations are defined by distributing the function list developed in the problem domain analysis between the individual presentation model classes.

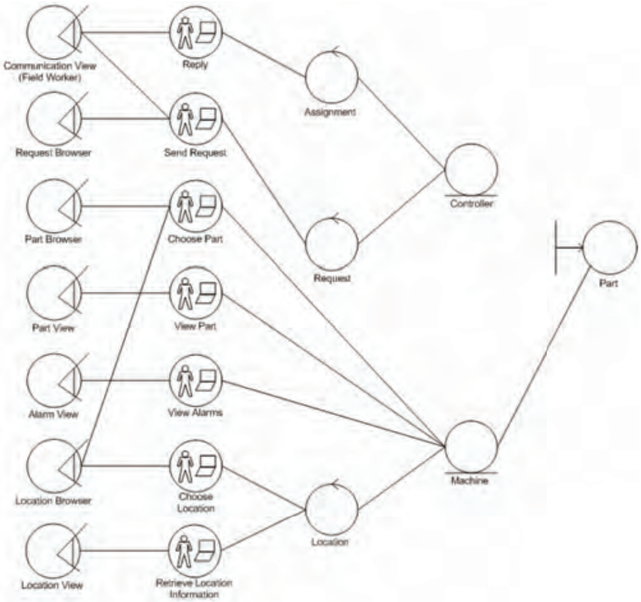


Figure 14. Interaction model for the Power Plant Communicator

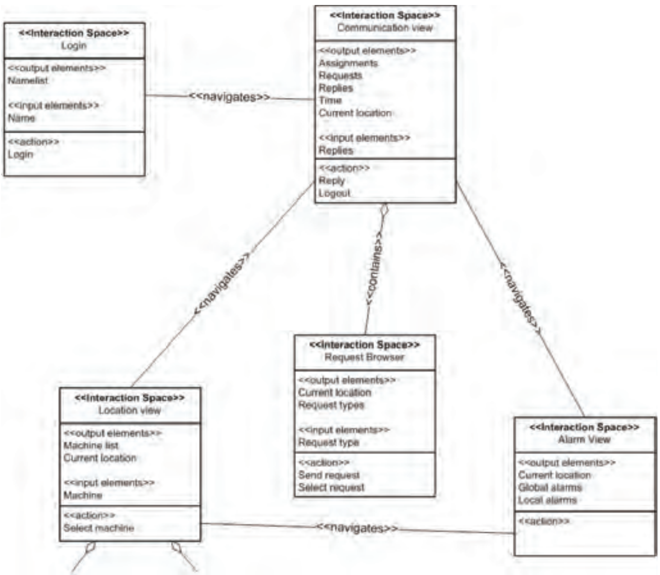


Figure 15. Part of the presentation model for the Power Plant Communicator

User interface design

The first step of the user interface design activity was to transform the interaction, presentation and dialogue models into concrete design. Informing this process, the three models from the Bridge method described above provided important input on different levels. The interaction model provided input on the overall structure and relationship between individual screens. The presentation model provided a detailed list of elements that we had to include on each specific screen. The dialogue model described the sequence by which the workers would typically interact with the screens.

As illustrated in figure 16, knowing what information and functionality to provide, and when, proved to be highly valuable information in order to maximize the use of the small screen real estate of the target platform. In order to reduce complexity, figure 16 only depicts some of the connections between interaction spaces and screen design.

Figure 16. Mapping of presentation model classes to user interface design

Paper Prototype

The user interface design developed on the basis of the interaction, presentation and dialogue models was transformed into a second paper prototype with a high level of fidelity (figure 16 right). The paper prototype was then evaluated in situ at the power plant with prospective users acting-out a series of realistic use scenarios. Observations from the paper prototype evaluations led to refinements of the user interface design.

Functional Prototype

Following the paper prototype evaluations, a functional prototype that matched the final design specifications was implemented in C# using Microsoft Visual Studio .Net 2003 Professional and the .Net Compact Framework. The implementation activity amounted to a total of 20 person-days.

7.3. What did we learn from the two design and implementation processes?

While the two implementation activities were very similar in process and the amount of time spent, the design activities differed notably from case study 1 to case study 2. The differences reflect that the design activity was more exploratory in nature in the first case study and more structured in the second one. Both processes have their strengths and weaknesses and are applicable depending on the nature of the specific project. The design process for the Maritime Communicator represents an iterative approach where design is carried out with an increasing level of detail using different means of expression such as paper, interface toolkits, and mock-ups. In this approach, the diagrams produced in the problem domain analysis were important for reference, but design was very much driven by creativity. The design process of the Power Plant Communicator, on the other hand, represents a more structured approach where the outcomes from initial analyses are directly elaborated on in the form of additional models defining abstract properties of the user interface. The "creative" part of the design phase is therefore postponed for longer but in effect informed more directly by ethnographic field study through the models developed. In our experience, the additional object-oriented models focussing explicitly on the user interface helped in keeping the design of the Power Plant Communicator grounded in our field data. In particular, we found that the Interaction Space models provided structured input into the design of each screen of the system without prescribing their look and feel in a way that impeded creativity.

The second major difference between the two design phases was that case study 2 involved user feedback on a paper prototype by subsequent field visits to the context of use prior to implementation. Again, this helped us maintain the grounding of our design process in the initial ethnographic field study to a stronger degree than in case study 1.

8. FUNCTIONAL PROTOTYPES

In this section we describe the two functional mobile device prototypes resulting from the field study, analysis, design, and implementation processes described above: the Maritime Communicator and the Power Plant Communicator. The prototype systems are not the contribution of this article but serve as illustrations of outcome. The two systems are described in detail in (Kjeldskov and Stage 2006) and (Kjeldskov et al. 2006).

8.1. The Maritime Communicator

The Handheld Communicator was targeted at PDAs running the Microsoft PocketPC operating system such as the industrial grade Symbol PPT8800 series (figure 17). Apart from a touch screen, such industrial devices support interaction by means of large rubber buttons located below the display and suitable for one-handed interaction. Due to the potentially harsh conditions of use, in which pen-based interaction might be problematic, all interaction took place through these buttons.

The system setup consisted of three PDAs connected through a wireless network. One device was intended for the captain on the bridge while the other two were for the 1st and 2nd officers on the fore and aft deck respectively.

Overall Design

The Maritime Communicator explores the ideas described in section 5 replacing verbal communication with an exchange of predefined text messages grouped by objects and the state of tasks, and providing an updated representation of the shared work domain.

Figure 17 shows the user interface of the Handheld Communicator for the captain. At the bottom of the screen there is a list of unexecuted commands and confirmations. The order of the list corresponds to the standard sequence of the overall operation, and possible utterances only appear when appropriate in relation to the state of the task. By default, the most likely next step is highlighted (based on the state chart diagrams).

Figure 17. The Maritime Communicator on a Symbol PPT8000

The most important element of the interface is the list of "ongoing tasks". When a command is executed, it appears on this list along with a counter displaying the time passed since it appeared. When a command is confirmed the timer is substituted by the text "[ok]" followed by a description of the current activity (e.g. "Singling up..."). When a task is reported completed, a short statement (e.g. "1 and 1 fore") substitutes the description of activity and the captain is prompted for confirmation. When the captain confirms the completion of a task, the communication thread is moved to the "history" list. When the history list is full, it scrolls the oldest commands and statements out of sight. At the top of the screen a simple pictogram displays the current state of the operation for quick reference: the lines still attached to the quay and the current status of fore and aft.

On deck, the interface for the officers is very similar to that on the bridge thus providing a view on the present status of the mooring and a list of all past and ongoing communication among the distributed co-workers. The main difference is the list of available commands and utterances, which is filtered on the basis of their location and task, and the indications under "ongoing tasks" of commands with responses pending.

Implementation

The application running on the captain's device worked as a server containing a formalized representation of the communication pattern of the task. The devices on deck logged on to this server and identified their physical location. During operation, function calls and unique command identifiers were exchanged over the wireless network. All network traffic was broadcast but processed and represented differently on each device in accordance to their physical location (bridge, fore or aft).

8.2. Case study 2: the power plant

Like the Maritime Communicator, the Power Plant Communicator was targeted at industrial grade PDAs running the Microsoft PocketPC operating system. Interaction was facilitated by the device's function keys and through finger-based input on the touch screen. In addition to the handheld terminals used by the workers moving around the plant, a desktop PC interface was designed for use in the control tower.

Overall Design

The Power Plant Communicator provides mobile distributed workers with access to information in the central computer system about the general status of the plant and about the specific machinery within their proximity. It also gives them a simple text-based communication channel for coordinating certain work activities. The system is divided into three overall screens with a number of associated sub-screens: 1) Communication screen, 2) Alarm screen, and 3) Status screen. At the top of the display, the system indicates who is logged in, where the user is located and what time it is. This information is important for the field workers because it gives them a frame of reference for interpreting the information and functionality provided by the system.

The communication screen

The communication screen (figure 18) provides workers with a text-based communication channel (inspired by the Maritime Communicator). Using the two lists next to each other above the panes the user can compose a message by combining a verb with a noun, for example "stop production". Above these two lists there are three buttons for sending the composed message, or for sending a standard reply that you Accept (Accepter) or Reject (Afvis) a particular request. To avoid pressing the wrong button by mistake, the buttons for accepting or rejecting are placed furthest apart. Above these three buttons the ongoing conversation is displayed on a list. The list is divided into a series of conversation threads, grouping communication about the same object or task together. To make a clear difference between requests and confirmations, the latter are indented and have their first word (for example ACCEPT or REJECT) in capital letters. To add a new message to a thread the user selects an utterance on the list, for example "Check Harper/Knuser 2, and sends a response.

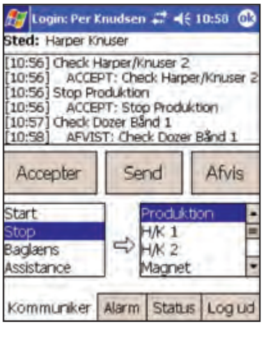


Figure 18. Communication screen

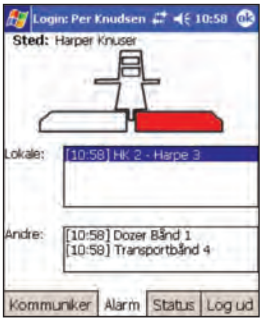


Figure 19. Alarm screen

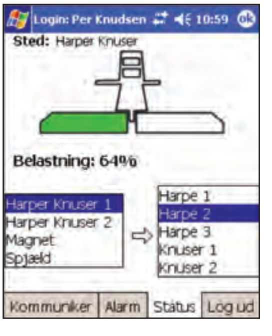


Figure 20. Status screen

A main difference in this design compared to the Maritime Communicator is that messages are composed from two lists rather than one. This design emerged from the object-oriented models revealing that the number of possible utterances was very large, but that they were all composed from a common set of verbs and a set of nouns related to the users' location. Because the system knows where the workers are physically located at the power plant, it can deduce which machinery they are most likely to be communicating about (generating the list of nouns). Knowing about the functionality of the machinery the system can deduce what actions the user can request (generating the list of verbs).

The alarm screen

The alarm screen (figure 19) alerts the workers if something is wrong at the power plant and provides them with the available details about the problem. If an alarm is associated with the users location, detailed information is displayed immediately. The alarm screen contains a graphical representation of the machinery in question, with the area of the problem highlighted in red. Below the graphical representation, there are two lists displaying alarms at the users current location and other alarms along the production line. These lists were included because our models showed that problems along the production line often cause other problems to happen, requiring co-located workers to start coordinating a shared strategy. Using these lists, the workers can get an overview of the cause and/or effects of a specific problem, thus assisting them in assessing its criticality. Each alarm has a timestamp that enables the workers to determine their sequence and current relevance.

The status screen

The status screen (figure 20) provides workers with access to information about machinery in the production line within their close physical proximity. Below the graphical representation, the users can choose which of the machines within proximity they want to work with (the list on the left), and what specific part of it they want to access information about (the list on the right). The screen can also function as a remote control to some machines basic operations such as start, stop and reverse. The status screen looks similar to the alarm screen in many ways. At the top of the screen, there is a graphical representation of the machine being accessed, with the specific part chosen being highlighted. If the chosen part of the machine is functioning or the load on the part is at a normal level, it is highlighted in green. If the load climbs towards a critical level, the colour changes to orange. If reaching a critical level, the colour changes to red and an alarm is activated. These states were revealed through our models.

Implementation

The technical implementation of the Power Plant Communicator was similar to the Maritime Communicator. The computer in the control tower (a desktop PC) worked as a server containing a formalized representation of the power plant and typical work activities. The handheld terminals logged on to this server and identified their physical location, following which an appropriate interface was displayed on them. During use, function calls and commands were exchanged over the network, using handshake to confirm delivery, thus eliminating commands being "lost in the air". All network traffic was broadcast but processed and represented differently on each device in respect to their location.

9. EVALUATIONS AND USER FEEDBACK

The final phase of our two case studies was to evaluate the usability of the produced prototype systems. In the context of this article the evaluations below serve as validations of the outcome of the development processes as well as illustrations of different approaches to "closing the circle" by going back to the domains where our ethnographic field studies were originally conducted. Due to the nature of the two use domains, both evaluations were challenging to carry out in a realistic, but safe, manner. This challenge was addressed in two different ways: by carrying out the evaluation in a high-fidelity ship simulator and by acting out use scenarios at the power plant.

9.1. Case study 1: the container ship

The Maritime Communicator was evaluated in a state-of-the-art ship simulator at Svendborg International Maritime Academy in Denmark (Kjeldskov and Skov 2007). The evaluation involved three teams of two experienced maritime officers as test subjects. The officers were given the task of letting go the lines and departing from harbour using the Maritime Communicator for communication and coordination. One subject user acted as captain on the bridge (figure 21) while the other acted as 1st officer on the fore deck in a neighbouring room.

The ship simulator was set up to imitate the operation of a large ship in challenging weather and traffic conditions. The scenario was developed in collaboration with the simulator division of Svendborg International Maritime Academy, and corresponded to a real world situation observed during our field studies described above (see figure 5).

During the evaluation the captain had to consider all aspects of manoeuvring the ship. This included controlling the rudder, propellers and thrusters as well as communicating with personnel on the ship, harbour traffic control and taking into consideration the movements of other ships. The 1st officer on deck verbally forwarded commands to his team of assistants on the quay and manning the winches and reported progress back to the captain. The team of assistants were a part of the digital simulation.

Figure 21. Evaluation of the Maritime Communicator in a ship simulator

During the evaluation, the users were asked to think-aloud, explaining their experience of and interaction with the prototype. Two evaluators located on the bridge and on the deck respectively observed the users and asked questions for clarification. Following each evaluation session, a group interview of 10-15 minutes was carried out. The evaluation sessions were recorded on digital video from four camera sources and audio capturing close-up views of the handheld devices and overall views of the use context.

Highlights from evaluation

The evaluation of the Maritime Communicator in the ship simulator provided rich data on the usability of the design. First of all, the user study showed that the prototype could actually replace the majority of spoken communication between the captain and officers. Also, the users expressed that the text-based interface gave them a simple channel of communication in which they could easily monitor the ongoing communication and progress of the operation. As a supplement to this, some officers on deck expressed a wish for having commands read out by a synthetic voice as an option when they were busy with their hands elsewhere. Generally, the users learned what to do and what feedback to expect within the completion of one or two threads of communication and reported that the design was very intuitive to use. The differentiation between commands appeared straightforward as did the grouping and progress of ongoing threads. The graphical representation of the ship was highly appreciated for overview, as this was not currently provided.

Apart from the positive feedback, the user study also revealed 22 usability problems experienced by the users. First of all, we identified a need for being able to request or report something out of the ordinary, and to correct or withdraw an utterance. This finding indicated that our ethnographic field study had missed situations where there was a need to communicate things out of the ordinary, and consequently this was also missing from our problem domain analysis. This lesson learned led us to modify the analysis procedure in the subsequent Power Plant Communicator case study to involve a validation of these models, facilitated by an initial paper prototype evaluation, after the problem domain analysis, as described earlier.

9.2. Case study 2: the power plant

The Power Plant Communicator was evaluated in situ at the fuel department. The evaluation involved five workers with several years of work experience with this particular plant. The prototype system was used during live operation of the power plant while carrying out a series of typical work tasks. In addition to this, a series of less common work activities were acted-out in situ as we had done with the evaluation of the paper prototypes described earlier.

Figure 22. Field evaluation of the Power Plant Communicator in situ

One researcher managed the evaluation and asked questions for clarification. Another researcher recorded the evaluation on a handheld video camera shifting focus between the workers, the settings, and the screen on the handheld device (figures 22 and 23).

The users were asked to think-aloud, but because of the noise level in some buildings, this did not always work well. In these situations, a post-use interview was conducted outside the building immediately afterwards.

Figure 23. Close-up of the communicator during field evaluation.

Highlights from evaluation

The results of the evaluation showed us that the system was indeed usable and that the workers were satisfied with using it. The users had no problems understanding and adopting the basic functionality of the system. Some compared it to sending SMS messages on their mobile phone. Others compared it to a remote control. In relation to the communication screen, the users reported that they liked to be able to combine text-based messages and spoken communication over the VHF radio. Text-based communication was primarily found useful when noise prevented spoken commands but was also reported very useful in complex situations where they otherwise needed to remember what had been said when and by whom.

Regarding the alarm screen, the users reported that the annotated graphical representation would help them greatly in locating and fixing problems quickly. Some reported that they would like even more detailed graphics and preferably a plant overview as well. On the status screen the users found the simple access to information about machines and control of it within physical proximity compelling and highly useful for their daily operation of the plant. Furthermore, not being able to operate machinery out of proximity was perceived as a significant safety advantage. Apart from the positive observations and feedback the evaluation also revealed 14 usability problems experienced by the users, to be addressed in the next iteration of design and development.

10. SUMMARY AND DISCUSSION

The challenge we set out to respond to by combining ethnography with Mathiassen et al.'s (2000) object-oriented analysis method was that mobile software development requires insight into rich and dynamic use contexts, but that ethnographically based approaches, which are good for studying such contexts, provide data and insights that can be hard to translate into system design. This led us to devise an approach that explicitly combines ethnography with a structured analysis and design approach, enabling us to obtain both contextual richness and abstract models for informing the development of mobile systems. In the following, we summarize and discuss some of the lessons learned from the two case studies of using this approach presented above.

10.1. Value of application domain analysis

The primary value of applying an object-oriented application domain analysis to the ethnographic field data was that this approach provided a specific focus for the data. Rather than doing an open-ended or grounded type of analysis, which outcomes can be difficult to translate into a system development process, the application domain analysis approach prompted us to specifically address the question of what the system is going to be about, and to generate a set of overall requirements for what we were going to build. Apart from the obvious value of these tangible outcomes, the value of the application domain analysis also lay in the process of doing these analyses as a shared activity within the project teams. This prompted focused discussions about what was observed in the field, and through these discussions we gradually understood more and more the domains we were working with, and were essentially developing systems for. As pointed out above, the outcome of this analysis was able to serve as a common frame of reference for the project teams throughout the rest of the project, representing one of the outcomes from the ethnographic field study.

10.2. Value of problem domain analysis

The two object-oriented problem domain analyses reduced the complexity of information from the application domain analyses to a series of abstract models, while still preserving the essence of contextual grounding obtained through the ethnographic field study. It was at the time of the problem domain analysis that the added value of the object-oriented approach particularly began showing its strengths for the process of system development by prompting the development teams to formalize structures and processes observed in the ethnographies into abstract models. During this process we learned more about the specific division of labour in the two work domains, the complexities of tasks, and the structures of operations and communication than had been obvious from the ethnographies. This additional insight was essential for the subsequent design processes, and the models created were also essential for the later implementation phases. In the power plant case a particular valuable experience with the problem domain analysis was that this activity revealed that some information was missing from our ethnographic field studies, and made us go back and gather this additional field data at a relatively early stage of the development process (as discussed more below). This exemplifies how structured modelling can act as a verification of the scope of the much more unstructured ethnographic study, and illustrates a strength of combining these two approaches in system design.

10.3. Informing the design process

The value of the object-oriented approach for the design process was slightly different for the two case studies. The design process of the Maritime Communicator represented an approach where the diagrams produced in the problem domain analysis were important as structural reference for an iterative and highly creative design activity. By comparison, the design process of the Power Plant Communicator represented an approach where outcomes from the object-oriented analyses were directly elaborated on through additional abstract modelling before any concrete design activity was carried out. Although this meant that actual interface design was done later in the process, this was at the same time informed more directly by the ethnographic field study. By developing object-oriented models focussing explicitly on the structural design of the user interface, such as Interaction Space models, the design of the Power Plant Communicator remained strongly grounded in data from the ethnography.

10.4. Supporting the implementation process

The value of the object-oriented analysis and design efforts for the implementation of our prototype systems were two-fold. Firstly the structured and iterative approach had informed the creation of a well-specified system and user interface. This minimized effort spent during the implementation phase on essentially doing design, and ensured that focus could be kept on the task of coding. Secondly, and equally importantly, the abstract models developed during the problem domain analysis, and for the power plant case also during the design phase, directly informed the structure of the software programs, and contained all the information required to build the system. This minimized time spent during the implementation phase on translating design specifications into structures that could be coded, and ensured that the code correlated well with the specified design.

10.5. Value of ethnography in mobile HCI

Taking a step back, the ethnographic field studies described in section 4 played an invaluable role in the two system development processes, confirming the potential value of ethnography in mobile HCI as suggested by Kjeldskov and Graham (2003), and illustrated by, for example, Luff and Heath (1998) in their work on mobility in collaboration back in the 1990s. Both of the studies presented in this paper generated rich insights into the use contexts and work activities that would unlikely have been reached otherwise. Firstly, the ethnography on board the container ships generated insights and inspiration that, amongst others, led to the conception of the text-based communicator systems in the first place. In comparison, the ethnography at the power plant was more focused, but confirmed observations from the maritime domain, and led to the identification of different, yet related work activities where similar technology might be applicable, as well as identification of new, unique type of work activities not observed on board the container ships.

Secondly, the two ethnographies generated valuable data for contextually rich application and problem domain analyses. As can be seen in section 5 and 6, these analyses of the general use context, and the specific part of that context that is directly related to the use of the proposed systems, were particular rich in details and strongly grounded in real world observations. This made them extremely useful for the further development process as a frame of reference representing the data from the ethnographies in a tangible and organized form.

Thirdly, the two ethnographies provided insight into the use contexts that was critical for the recreation of realistic test environments for the evaluation of prototype systems. For both cases this included the creation of realistic and relevant tasks. For the container ship case it also included informing the scenario programmed into the ship simulator.

The only limitation of the ethnographic approach experienced related to the issue of missing out of observing a rare event. This was partly due to the inherent naturalistic and non-intrusive nature of the ethnographic methodology, and partly due to the length of time spent in the field. One approach to overcome this issue is, of course, to spend longer time in the field until all possible events have happened and been observed. However, apart from the intrinsic uncertainty about exactly when this criterion has been reached, this is typically not a realistic approach in a software development project. As a response to this, various time-optimized field study techniques such as "rapid ethnography" (Millen 2000) and "rapid contextual design" (Holtzblatt et al. 2005) have emerged, which, amongst others, propose that special effort is made to identify and cover those unusual situations also. The next question here is, of course, *how* to do that? As one possible answer, we experienced that the systematic nature of the objected-oriented analysis provided valuable input for testing the completeness of the ethnographic data gathered, and subsequently guiding the focus of final fieldwork. Consequently, it can be argued that ethnography and object-orientation is best combined iteratively, with going back and forward between field studies and analysis being an explicit part of the approach.

11. CONCLUSIONS AND FURTHER WORK

We have presented two case studies of mobile system development for supporting distributed collaborative work activities within industrial process control. Based on the lessons learned from these two case studies, we have described and discussed how ethnography and object-orientation can be successfully combined for obtaining contextual richness and abstract models for mobile interaction design.

One of the challenges reported is that ethnography generates findings and knowledge with such contextual richness that it can be hard to transfer into system design. By combining ethnography with the software engineering method of object-oriented analysis and design, we have shown how ethnographic field data is a highly valuable source of input for developing object-oriented models by providing contextual richness and that, in turn, objected-oriented analysis is a highly valuable method for working with ethnographic field data in systems development by supporting the creation of abstract models. Combining the two, we have demonstrated a method where the use of ethnographic field studies in system design is supported by the use of a structured analytical approach, and where abstract modelling of systems is informed by rich contextual data. By applying this method, ethnographies are empowered in their capability for informing core system design, and system design is consequently strongly grounded in their use context.

We have applied the proposed methodological approach to two different mobile system development processes over a period of 5 years. In doing this we found that the experienced value of the method generalised from one case study to the other, and we were even able to make some improvements to the specific analytical activities involved. While on the basis of our experiences we believe that the described process has general value for the development of other mobile systems, this should, however, be evaluated through other cases of mobile interaction design applying the method of explicitly combining ethnography and object-oriented analysis and design.

ACKNOWLEDGEMENTS

The authors would like to thank our collaborators at Maersk and the Northern Jutland Power Plant including everyone who participated in the field studies and usability evaluations. We would also like to thank Jeni Paay for valuable discussions of the work and comments on the manuscript. The Power Plant Communicator was developed by Christian Monrad Nielsen, Michael Overgaard, Michael Bach Pedersen and Sigge Stenild, who also collected field data in the second case study. We also thank Svendborg International Maritime Academy for access to their simulator facilities. The usability evaluation of the Maritime Communicator was planed and conducted in collaboration with Mikael B. Skov. Finally, we would like to thank the anonymous reviewers who's comments and questions have contributed to improving the manuscript.

REFERENCES

Andersen, P. B. and May, M., 2001. Tearing up Interfaces. In Liu K. et al (eds.) Information, organization and technology. Studies in organizational semiotics. Boston, Klüwer.

Bardram, J. E. and Hansen, T. R., 2004. The AWARE architecture: supporting context-mediated social awareness in mobile cooperation. In Proceedings of CSCW'04. ACM, 192-201

Benyon, D. 2002. Representations in Human-Computer Systems Development. Cognition, Technology & Work, 4(2002): 180-196.

Holtzblatt, K., Wendell, J. B., and Wood S. 2005. Rapid Contextual Design – a how-to guide to key techniques for user centered design. San Francisco: Morgan Kaufmann.

Beyer, H., & Holtzblatt, K., 1998. Contextual design – Defining customer centred systems. San Francisco: Morgan Kaufmann.

Churchill, E. F. and Bly, S., 1999. It's all in the words: Supporting work activities with lightweight tools. In Proceedings of ACM Siggroup'99, 40-49.

Clark, H.H. and Schaefer, E. F., 1989. Contributing to discourse. Cognitive Science, 13, 259-294.

Constantine, L., Biddle, R., Noble, J., 2003. Usage-centered design and software engineering: models for integration. In: Proceedings of ICSE 2003, IFIP, Portland, USA, pp. 106–113.

Crabtree, A., Rodden, T., Tolmie, P., and Button, G., 2009. Ethnography considered harmful. In Proceedings of CHI '09. ACM, 879-888.

Crabtree, A., 2004. Design in the absence of practice: breaching experiments. In: Proceedings of DIS 2004. Cambridge, MA, USA, pp. 59–68.

Dayton, T., McFarland, A. and Kramer, J., 1998. Bridging User Needs To Object Oriented GUI Prototype Via Task Object Design. In Wood, L. E. (ed.) User Interface Design – Bridging the Gap from User Requirements and Design. CRC Press LLC.

Diggins, T., Tolmie, P., 2003. The 'adequate' design of ethnographic outputs for practice: some explorations of the characteristics of design resources. Personal and Ubiquitous Computing 7, 147–158.

Hosbond, J. H. 2005. Mobile Systems Development: Challenges, Implications and Issues. In Proceedings of MOBIS 2005, Leeds, UK, IFIP TC8.

Hosbond, J.H. and Nielsen, P.A., 2005. Mobile Systems Development - A literature review. In Proceedings of IFIP 8.2: Designing Ubiquitous Information Environments. Cleveland, Ohio, IEEE.

Howard, S., Carroll, J., Vetere, F., Murphy, J., and Peck, J., 2002. Provoking Innovation: Acting out with contextual scenarios. Proceedings of HCI 2002, BCS-HCI.

Hughes, J., King, V., Rodden, T., Andersen, H., 1995. The role of ethnography in interactive systems design. Interactions 2 (2), 56–65.

Hughes, J., King, V., Rodden, T., Andersen, H., 1994. Moving out from the control room: ethnography in system design. In: Proceedings of CSCW ‘94. ACM, Chapel Hill, USA, pp. 429–439.

Iqbal, R., James, A., and Gatward, R., 2005. Designing with ethnography: An integrative approach to CSCW design. Advanced Engineering Informatics, 19(2005): 81-92

Kjeldskov, J. (2012) A longitudinal review of mobile HCI research methods. (Forthcoming)

Kjeldskov, J. and Skov, M.B., 2007. Studying Usability in Sitro: Simulating Real World Phenomena in Controlled Environments. International Journal of Human-Computer Interaction (IJHCI), 22(1):7-37.

Kjeldskov, J., Nielsen, C.M., Overgaard, M., Pedersen, M.B., Stage, J. and Stenild, S., 2006. Designing a Mobile Communicator: Combining Ethnography and Object-Oriented Design. Proceedings of OzCHI 2006, Sydney, Australia, ACM and CHISIG, 95-103.

Kjeldskov, J. and Stage, J., 2006. Exploring “Canned Communication” for Coordinating Distributed Mobile Work Activities. Interacting with Computers, 2006(18): 1310-1335.

Kjeldskov, J. and Stage, J., 2004. New Techniques for Usability Evaluation of Mobile Systems. International Journal of Human-Computer Studies (IJHCS), 60(2004): 599-620

Kjeldskov, J. and Graham, C., 2003. A Review of Mobile HCI Research Methods. Proceedings of Mobile HCI 2003. Springer-Verlag, 317-335.

Kjeldskov, J. and Stage, J., 2003. Designing the Handheld Maritime Communicator. Proceedings of the 1st Conference on Designing User Experiences, DUX 2003. San Francisco, CA, USA. ACM, 1-15.

Krogstie, J., Lyytinen, K., Opdahl, A.L., Pernici, B., Siau, K., and Smoland K., (2004) Research areas and challenges for mobile information systems. International Journal of Mobile Communications, 2(3): 220-234.

Luff, P, and Heath, C. (1998) Mobility in Collaboration. Proceedings of CSCW’98, Seattle, Washington, ACM, 305-314.

Lumsden, J., Langton, N., and Kondratova, I., 2008. Evaluating the appropriateness of speech input in marine applications: a field evaluation. In Proceedings of MobileHCI ‘08. ACM, 343-346.

Mathiassen, L., Munk-Madsen, A., Nielsen, P. A. and Stage, J., 2000. Object-Oriented Analysis & Design. Aalborg: Marko Publishing.

Mikkonen, M., Vayrynen, S., Ikonen, V. and Heikkila, O., 2002. User and Concept Studies as Tools in Developing Mobile Communication Services, Personal and Ubiquitous Computing, 2002(6), 113-124, Springer-Verlag.

Millen, D.R., 2000. Rapid ethnography: time deepening strategies for HCI field research. In Proceedings of the 3rd conference on Designing interactive systems (DIS ‘00). ACM, New York, NY, USA, 280-286.

Nielsen, C. and Søndergaard, A., 2000. Designing for mobility - an integration approach supporting multiple technologies. In Proceedings of NordiCHI 2000, Stockholm, Sweden.

Norman, D., 1990. The ‘Problem’ With Automation: Inappropriate Feedback And Interaction Not Over automation, in Broadbent D.E. et al., (eds.) Human Factors In Hazardous Situations, Oxford, Clarendon Press, 137-145.

Nunes, N. J. and Cunha, J. F., 2001a. Wisdom: A Software Engineering Method for Small Software Development Companies. IEEE Software.

Nunes, N. J. and Cunha, J. F., 2001b. Wisdom - Whitewater Interactive System Development with Object Models. In van Harmelen, M. (ed.) Object Modeling and User Interface Design, New York: Addison-Wesley.

Paay, J., Sterling, L., Vetere, F., Howard, S. and Boettcher, A., 2009. Engineering the social: The role of shared artifacts. International Journal of Human-Computer Studies 67, 437-454.

Paay, J., 2008. From ethnography to interface design. In: Lumsden, J. (Ed.), Handbook of Research on User Interface Design and Evaluation for Mobile Technology. Idea Group Inc (IGI), PA, USA, pp. 1–15.

Popolov, D. Callaghan, M. and Luker, P., 2000. Conversation Space: Visualising Multi-threaded Conversation. In Proceedings of AVI2000, Palermo, Italy, ACM, 246-249.

Rasmussen, J., 1983. Skills, Rules and Knowledge: Signals, Signs and Symbols and Other Distinctions in Human Performance Models, IEEE Transactions on Systems, Man and Cybernetics 13(3).

Rasmussen, J., 1986. Information Processing and Human-Machine Interaction. New York, North-Holland.

Rügge, I., Ruthenbeck, C., Piotrowski, J., Meinecke, C., and Böse, F. 2009. Supporting mobile work processes in logistics with wearable computing. In Proceedings of MobileHCI '09. ACM, 1-2.

Schraefel, M., Hughes, G., Mills, H., Smith, G., Frey, J., 2004. Making tea: iterative design through analogy. In: Proceedings of DIS '04, Cambridge, MA, USA, pp. 49–58.

Sharples, M, Corlett, D. and Westmancott, O., 2002. The Design and Implementation of a Mobile Learning Resource, Personal and Ubiquitous Computing 2002(6), 220-234, Springer-Verlag.

Simonsen, J., Kensing, F., 1998. Make room for ethnography in design!. Journal of Computer Documentation 22 (1), 20–30.

Skov, M.B. and Høegh, R.T. 2006. Supporting information access in a hospital ward by a context-aware mobile electronic patient record. Personal Ubiquitous Computing 10, 4, 205-214.

Smith, M., Cadiz, J. J. and Burkhalter, B., 2000. Conversation Trees and Threaded Chats. In Proceeding of CSCW'00, Philadelphia, USA, ACM, 97-105.

Smith, M. A. and Fiore, A. T., 2001. Visualization Components for Persistent Conversations. In Proceedings of CHI 01, Seattle, USA, ACM, 136-143.

Snyder, C., 2003. Paper Prototyping. The Fast and Easy Way to Design and Refine User Interfaces. Amsterdam: Morgan Kaufmann.

Tang, C. and Carpendale, S., 2008. Evaluating the deployment of a mobile technology in a hospital ward. In Proceedings of CSCW '08. ACM, 205-214.

van den Anker, F. W. G. and Lichtveld, R. A., 2000. Early Evaluation of New Technologies: The Case of Mobile Multimedia Communications For Emergency Medicine. In Vincent C. and de Mal B (eds.) Safety in Medicine. Qxford: Elsewier Science.

Viller, S., Sommerville, I., 1999. Coherence: an approach to representing ethnographic analyses in systems design. Human–Computer Interaction 14, 9–41.

Viller, S., Sommerville, I., 2000. Ethnographically informed analysis for software engineers. International Journal of Human–Computer Studies 53, 169–196.

Walenstein, A., 2003. Finding boundary objects in SE and HCI: an approach through engineering-oriented design theories. In: Proceedings of ICSE 2003. Workshop on Bridging the Gap between SE and HCI. IFIP, Portland, USA, pp. 92–99.

Wiltshire, M., 2003. Where SE and HCI meet: a position paper. In: Proceedings of ICSE 2003, Workshop on Bridging the Gap between SE and HCI. IFIP, Portland, USA, 57–60.

Winograd, T. and Flores, F., 1986. Understanding Computers and Cognition. Addison-Wesley.

Volume II

Designing mobile interactions

the continual convergence of form and context

Jesper Kjeldskov

Designing mobile interactions

the continual convergence of form and context

Jesper Kjeldskov

Volume II

Designing mobile interactions - the continual
convergence of form and context, Volume II

Printed by UniPrint, Aalborg University

Denne afhandling er af Akademisk Råd ved Det Teknisk-Naturvidenskabelige Fakultet ved Aalborg Universitet antaget til forsvar for doktorgraden i naturvidenskab. Forsvaret finder sted på Aalborg Universitet i Auditorium 0.1.95, Selma Lagerlöfs Vej 300, den 29. maj 2013 kl 13.00.

Aalborg den 31 januar, 2013

Eskild Holm Nielsen
Dekan

This thesis has been accepted by the Faculty of Engineering and Science at Aalborg University for public defence in fulfillment of the requirements for the doctoral degree in science The defence will take place at Aalborg University in Auditorium 0.1.95, Selma Lagerlöfs Vej 300, on 29 May 2013 at 1.00 pm.

Aalborg, 31 January, 2013

Eskild Holm Nielsen
Dean

Table of contents

Part III

Improving evaluation

IMPROVING EVALUATION

Part III addresses the question *how can we improve our techniques for studying the user experience of mobile interaction design in context?* In order to realistically assess the quality of mobile interaction designs we need to systematically study prototypes under conditions that are appropriately representative of the future use context. Four of my own contributions to this are included in chapters 11-14. These chapters illustrate different ways of evaluating mobile interaction design with both ecological validity and control. They show how elements of context can be simulated in controlled artificial setting research environments, and how means for guiding focus and enabling high quality data collection can be applied to natural setting research environments.

Simulating mobility
Chapter 11 investigates how mobility can be simulated in a controlled setting. In this chapter we aim to increase the ecological validity of laboratory evaluations by simulating the user being physically mobile during use. The chapter presents and evaluates five lab techniques involving various aspects of physical motion combined with either needs for navigation in physical space or division of attention, using the case of walking down a pedestrian street as base line reference. The findings from the study show that each of the proposed techniques had similarities to the real world reference, but that none of them were completely identical to the field condition. The best simulation of mobility in the real world context was obtained when using a treadmill running at varying speed.

Simulating the domain
Chapter 12 investigates how use domains can be simulated in controlled settings. In this chapter we aim to increase the ecological validity of laboratory evaluations by simulating the use domains of interactive mobile systems. The chapter presents two case studies of mobile interaction design evaluations in controlled high-fidelity simulations of the real world. Findings show that simulating the domain provided better results than the traditional non-contextual lab approach, but that the field evaluation provided additional insight into real world use. The study concludes that although not as ecologically valid as a field study, it is possible to obtain a higher level of ecological validity by simulating significant elements of a use domain in controlled laboratory settings.

Bringing the system into the field
Chapter 13 investigates how a structured evaluation can be carried out in the field. In this chapter we aim to facilitate increased control in field evaluations without compromising its ecological validity. The chapter presents a multi-method evaluation of an interactive mobile system with the purpose of investigating how a field-based evaluation guided by tasks performs against other evaluation methods. Findings show that each of the methods had its own benefits. In relation to context, however, the field evaluation was able to uniquely highlight a number of issues of real world use. The study concludes that a field study, with a high level of ecological validity, can advantageously be combined with techniques from lab-based evaluations if more control and replicability is needed.

Taking the lab with you

Chapter 14 investigates how data collection can be improved in the field. In this chapter we aim at facilitating increased control and better data collection in field studies by exploring the use of small wireless cameras attached to users and their mobile devices. The chapter presents the development and use of a "field-laboratory" over four years of evaluating mobile interaction design in the field. It describes the current setup and explains the rationales for key decisions on technology and form factors. The study shows that it is possible to collect ecologically valid field data about mobile interaction design in use in a quality matching stationary usability laboratories by means of a field-lab. It also shows that field-labs can be made small, lightweight, and operational for hours.

Chapter 11

Simulating mobility

Jesper Kjeldskov and Jan Stage

Abstract. Usability evaluation of systems for mobile computers and devices is an emerging area of research. This paper presents and evaluates six techniques for evaluating the usability of mobile computer systems in laboratory settings. The purpose of these techniques is to facilitate systematic data collection in a controlled environment and support the identification of usability problems that are experienced in mobile use. The proposed techniques involve various aspects of physical motion combined with either needs for navigation in physical space or division of attention. The six techniques are evaluated through two usability experiments where walking in a pedestrian street was used as a reference. Each of the proposed techniques had some similarities to testing in the pedestrian street, but none of them turned out to be completely comparable to that form of field-evaluation. Seating the test subjects at a table supported identification of significantly more usability problems than any of the other proposed techniques. However a large number of the additional problems identified using this technique were categorized as cosmetic. When increasing the amount of physical activity, the test subjects also experienced a significantly increased subjective workload.

1. INTRODUCTION

Usability evaluation of systems for stationary computers has grown to be an established discipline within human-computer interaction. Debates are still taking place, but they are often based on a shared understanding of basic concepts. For example, there is a basic distinction between field and laboratory-based evaluations. The majority of literature accepts that both approaches are important and necessary, and for each of them many authors have contributed with methods and evaluation techniques as well as empirically documented experience with their use.

Originally published as Kjeldskov, J. and Stage, J. (2004) New Techniques for Usability Evaluation of Mobile Systems. *International Journal of Human-Computer Studies,* 60(5-6), 599-620.

Extensive guidelines exist that describe how usability evaluations in laboratory settings should be conducted (e.g. Dumas and Reddish 1999, Nielsen 1993, Rubin 1994). This is complemented with experimental evaluations of the relative strengths and weaknesses of different techniques that can be applied in a usability evaluation (e.g. Bailey et al. 1992, Henderson et al. 1995, Karat et al. 1992, Molich et al. 1998).

Established concepts, methodologies, and approaches in human-computer interaction are being challenged by the increasing focus on systems for wearable, handheld, and mobile computing devices. This move beyond office, home, and other stationary use settings has created a need for new approaches to design and evaluate useful and usable systems (see e.g. Luff and Heath 1998).

Mobile systems are typically used in highly dynamic contexts. Moreover, their use often involves several people distributed in the user's physical surroundings (Danesh et al. 2001, Kjeldskov and Skov 2003a, Kjeldskov and Skov 2003b). Therefore, field-based evaluations seem like an appealing, or even indispensable, approach for evaluating the usability of a mobile system. Yet evaluating usability in the field is not easy (Brewster 2002, Nielsen 1998). Three fundamental difficulties are reported in the literature. Firstly, it can be complicated to establish realistic studies that capture key situations in the use-context described above (Rantanen et al. 2002, Pascoe et al. 2000). Secondly, it is far from trivial to apply established evaluation techniques such as observation and think-aloud when an evaluation is conducted in a field setting (Sawhney and Schmandt 2000). Thirdly, field evaluations complicate data collection and limits control since users are moving physically in an environment with a number of unknown variables potentially affecting the set-up (Johnson 1998, Petrie et al. 1998).

In a laboratory setting, these difficulties are significantly reduced. When usability evaluations are conducted in a laboratory setting, experimental control and collection of high quality data is not a problem. Yet one of the drawbacks of this setting is the lack of realism. Existing approaches to laboratory-based usability evaluations of stationary computer systems try to solve this problem by recreating or imitating the real context of use in the laboratory by, for example, furnishing it as an office (Rubin 1994). However, when mobile systems are evaluated in a laboratory setting, mobility of the user and activities in the user's physical surroundings can be difficult to recreate realistically (Pirhonen et al. 2002, Thomas et al. 2002). Evaluation of mobile system usability is increasingly being reported (see e.g. Brewster and Murray 2000, Sharples et al. 2002). A recent survey of mobile human-computer interaction research has shown that laboratory experiments are presently the most prevalent method for evaluating mobile systems (Kjeldskov and Graham 2003). This study reveals that 41% of the surveyed research on mobile human-computer interaction between 2000 and 2002 involves evaluations of system designs. 71% of these evaluations are conducted by means of laboratory experiments and very few of these laboratory evaluations involve special techniques being applied to meet the challenges of evaluating a *mobile* system.

In the light of this, the purpose of this paper is to explore new techniques for evaluating the usability of mobile systems in laboratory settings that addresses the limitations discussed above while preserving the advantages of a laboratory experiment.

In section 2, we present the results from a comprehensive study of existing literature on usability testing of mobile systems. Section 3 presents ideas for new techniques that can recreate or imitate real-world use situations of a mobile system in a laboratory setting. Section 4 describes the design of two experiments, inquiring into the qualities of these techniques. The purpose of these experiments was to explore and compare several techniques, rather than evaluating a single technique. Section 5 and 6 presents and discusses the results from these experiments and section 7 provides the conclusion.

2. RELATED WORK

The literature on human-computer interaction contains a number of contributions on techniques for evaluating the usability of mobile systems. In order to identify proposals for new techniques, we have searched part of that literature.

2.1. Literature Survey

To identify literature that deal with usability evaluation of mobile systems, we conducted a systematic literature search in the following journals and conference proceedings:

- Proceedings of Mobile HCI: 1998, 1999, 2001
- Proceedings ACM CHI: 1996-2002
- ACM Transactions on Computer-Human Interaction (TOCHI): 1996-2002

While books and other journals and conference proceeding series exists, which also report interesting research in mobile human-computer interaction (see e.g. Kjeldskov and Graham 2003), we found that this selection provided a representative overview of state-of-the art in usability evaluation techniques for mobile systems. A total of 636 papers were examined, resulting in the identification of 114 papers dealing with human-computer interaction for mobile systems. These papers are categorized in table 1.

Table 1. Distribution of papers dealing with human-computer interaction for mobile systems

		CHI	TOCHI	Mobile HCI	Total
A	General aspects of usability evaluations	0	1	1	2
B	Usability evaluations on device simulator	5	2	4	11
C	Usability evaluations with traditional techniques	34	3	7	44
D	Usability evaluations with new techniques	3	0	3	6
E	Usability evaluations not described	6	0	9	15
F	No usability evaluations performed	3	3	30	36
	Total	51	9	54	114

The 2 papers in category A deal with general aspects of usability evaluations of mobile systems and provide practical advice. In the 11 papers in category B, usability evaluations were carried out based on simulations of mobile systems on desktop personal computers. The 44 papers in category C typically deals with design of mobile applications, and employs traditional usability evaluation methods such as heuristic inspection and think-aloud in laboratory settings. Many of them employ a technique where test subjects are being seated at a table in order to test a device that was intended for use in a mobile situation. Contrary to this, the 6 papers in category D present and apply new techniques for usability evaluations in order to reflect or recreate a mobile use situation. Below, we describe these proposals for new techniques in more detail.

The 15 papers in category E mention that usability evaluations have been carried out but do not describe them. In the 36 papers in category F, no usability evaluations were performed.

2.2. Proposed Techniques

The six papers in category D employ new and different techniques for increasing the realism of the evaluation situation. In these papers, there are two basic categories of techniques.

In the first category, the test subjects were required to walk while using the mobile system being evaluated. This would either take place on a treadmill or on a specifically defined track in a laboratory setup (Pirhonen et al. 2002) or on a real world route, which also involved way finding etc. (Petrie et al. 1998). Both of these settings facilitated the collection of a magnitude of qualitative and quantitative data such as task completion time, error rate, heart rate, perceived cognitive workload and deviation from preferred walking speed, etc.

In the second category, the test subjects were using a mobile system while driving a car simulator. The type of car simulator that was used ranges from low-fidelity personal computer-based simulations (Graham and Carter 1999, Koppinen 1999) to high-fidelity simulators with large projection screens involving real dashboards (Lai et al. 2001) or even real cars (Salvucci 2001). This technique does not involve the user being physically mobile to the same degree as when walking, but it facilitates the evaluation of mobile system use while simultaneously engaged in a demanding cognitive activity. The simulator-based technique facilitated both quantitative and qualitative data to be collected and a huge number of test sessions to be conducted within limited time frames.

Only two of the six papers in category D (Pirhonen et al. 2002, Graham and Carter 1999) employ multiple techniques or variations of proposed techniques. This is done to systematically measure their relative applicability and ability to support the identification of usability problems in the mobile systems being evaluated.

Overall, the literature study reveals that only a limited amount of human-computer interaction research involves usability evaluations of mobile systems. Out of 114 papers dealing with human-computer interaction for mobile systems, less than half of them report results from a usability evaluation of the presented design (category C and D combined). This is consistent with the tendency identified in (Kjeldskov and Graham 2003). Furthermore, the majority of usability evaluations of mobile systems employ only traditional techniques (category C) and little variety exists within the studies employing new techniques. This raises two key questions about usability evaluations of mobile systems. 1) Are traditional techniques optimal for evaluating the usability of mobile systems? 2) What new techniques might be suggested?

3. IDEAS FOR NEW TECHNIQUES

To complement the traditional techniques for usability evaluation, we have developed a number of alternatives. In order to make this effort more systematic, we used the literature on mobility as our point of departure. The problem is, however, that much of this literature deal with mobile systems on a level that is much more abstract than the physical activity of a person using a mobile system in a specific context. For example, mobility has been described in terms of application, mobility type, and context, where mobility type then is sub-divided into visiting, travelling, and wandering (Kristoffersen & Ljungberg, 1999). Yet when looking at the different ways in which a mobile phone is used in order to recreate these situations in a laboratory, such a framework is not particularly helpful, and it is very difficult to use it for generating alternatives to traditional usability evaluation techniques.

Based on these experiences, we changed our focus to theories on human information processing and their description of attention and conscious action (a summary can be found in Preece et al., 1994). Based on these theories, we developed two different frameworks for mobile use.

3.1. Framework A

Framework A focussed on the different ways in which a user could be moving physically while using a mobile system. Describing this, we used the following two dimensions:

- Type of body motion: none, constant, varying
- Attention needed to navigate: none, conscious

By juxtaposing these two dimensions we ended up with a two by three matrix of different overall configurations for incorporating mobility in laboratory usability evaluation setups (table 2). If there is no motion, no navigation in physical space is necessary, which leaves one cell empty.

Table 2. The five configurations based on motion and need for navigation

		Attention needed to navigate	
		None	Conscious
Body motion	None	1. Sitting at a table or standing	n/a
	Constant	2. Walking on a treadmill with constant speed or stepping on a stepping machine	4. Walking at constant speed on a track that is changing because obstructions are moved
	Varying	3. Walking on a treadmill with varying speed	5. Walking at varying speed on a track that is changing because obstructions are moved

In relation to the previous research presented in section 2, configuration 1 in table 2 is the traditional evaluation situation, where a user is sitting at a table or standing still while using a mobile system. This corresponds to the research in category C in table 1. The user is not moving through physical space, but the configuration is still being used in many usability evaluations of mobile systems. The research in category D in table 1 can also be related to the configurations in table 2. For example, it has been suggested to use a hallway with fixed obstructions as the experimental setting, which corresponds to configuration 4, or a stepping machine that allows the use to walk without moving (Pirhonen et al. 2000), which corresponds to configuration 2 in table 2.

3.2. Framework B

Framework B was based on the notion of divided attention. When people are using a mobile system while being mobile, their attention is divided between physical motion and the use of the system. Similar to the studies applying car simulators for usability evaluations of mobile systems, we thus aimed at creating a configuration that replicated a division of attention between a demanding cognitive task and the use of a mobile system. Unlike studies determining, for example, the deviation from the user's preferred walking speed during the evaluation (Pirhonen et al. 2002) we were not interested in measuring the user's performance on the secondary task. Our aim of the dual task approach was only that it should serve as a distracting factor in the laboratory experiment in order to simulate divided attention directing the user's focus away from the use of the mobile system in a real world setting.

4. EXPERIMENTAL DESIGN

We conducted two different experiments, each based on one of the two frameworks described above. The two experiments evaluated two different types of mobile systems for text-based communication. In this section, we describe the design of these experiments.

4.1. Experiment A

The first experiment was based on framework A and the five configurations described in table 2 (Beck et al. 2002). The purpose of this experiment was to inquire into the relative strengths and weaknesses of the different configurations when used as techniques for usability evaluations in a laboratory. In addition, we wanted to compare these to a typical use situation in the field. Thus the experiment involved the following six techniques, of which the first five match the five configurations described earlier:

1. Sitting on a chair at a table.
2. Walking on a treadmill at constant speed.
3. Walking on a treadmill at varying speed.
4. Walking at constant speed on an course that is constantly changing
5. Walking at varying speed on an course that is constantly changing
6. Walking in a pedestrian street.

We conducted a series of usability evaluations employing each of the six techniques above. In each evaluation, the user solved a number of specified tasks using a mobile system. The first five techniques were used in a usability laboratory. The sixth technique was used in the field. For configurations 4 and 5, the user had to walk a sequence of three different courses (depicted in figure 2).

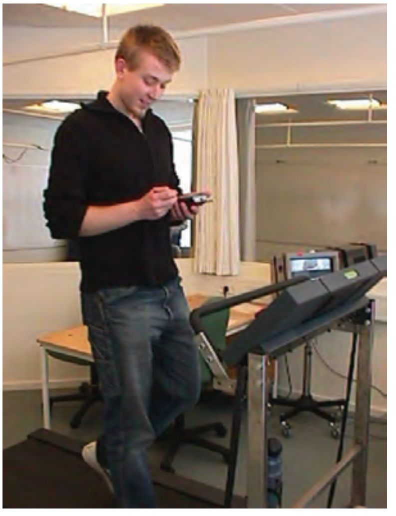


Figure 1. Test subject walking on a treadmill in the usability laboratory

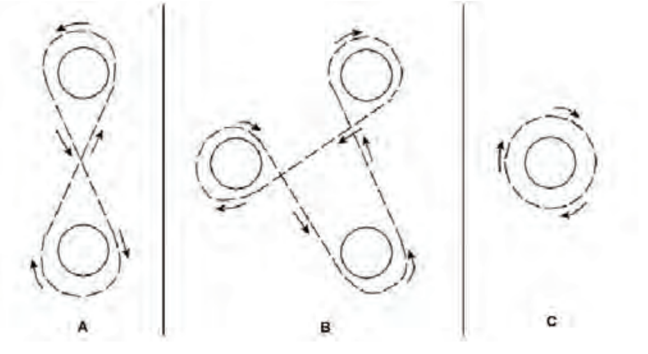


Figure 2. The sequence of the different courses to be walked in the laboratory (a, b, c)

The mobile system used for the experiment was an experimental short messaging service (SMS) application for the Compaq iPAQ personal digital assistant. This application provided the user with facilities for sending and receiving short text messages. In addition, the application complied with the specification of enhanced message service (EMS) (Ericsson, 2001), which enables exchange of small sound clips and pictures as part of a message. The application was specially designed for the experiment so that user actions and performance could be accurately measured through a dedicated monitoring application running in the background. We decided to use the short message service because it is widely used among mobile users and because it is highly interactive, involving both reading on the screen and typing in letters on a keyboard.

Each test subject was presented with five tasks involving sending and receiving short messages. While solving the tasks, the test subjects were required to think-aloud. In order to keep the time schedule, we decided to allocate ten minutes to each test subject. When the ten minutes had passed, the evaluation was stopped even if the test subject had not completed all tasks. For each evaluation, we collected three types of data:

- **Usability problems:** all evaluations in both the laboratory and the field were recorded on video. After the evaluation, the video recordings were analysed to produce a list of usability problems.
- **Performance:** a dedicated monitoring application automatically collected data about user interaction and time spent on each task.
- **Workload:** immediately after each evaluation, a NASA task load index (TLX) test was conducted with the test subject. This test assesses the user's subjective experience of the overall workload and the factors that contribute to it (Hart and Staveland, 1988; NASA).

The overall hypothesis was that, the ideal laboratory techniques would not differ from walking in a pedestrian street (technique 6) in terms of these three measures.

A test monitor managed each individual evaluation. Three experienced usability experts served as evaluators. In order to ensure that they were conducted consistently, the evaluators remained the same throughout all evaluations conducted using a specific technique. We were not able to carry out an unlimited number of evaluation sessions. Therefore, the number of test subjects used for each technique had to be a trade-off between the number of evaluations with each technique and the number of techniques we would be able to evaluate at all. Our aim was to explore a number of different techniques and not only evaluate one or two (cf. section 1). For that reason we had to minimize the number of test subjects used for evaluating by means of each technique. Key literature on usability evaluations suggests that it is possible to find around 80-85% of all usability problems if five test subjects are used (Nielsen, 2000; Virzi, 1992). Also, it has been argued that four to five test subjects are sufficient to gain an overall idea about the usability, but to avoid missing a critical problem, at least eight subjects should be used (Rubin, 1994). Based on this, we planned to use eight test subjects for each technique, amounting to a total of 48 test subjects.

The test subjects were all students of Informatics or Computer Science at the University of Aalborg, Denmark. They were demographically homogeneous, realistic users of the application, and easy to contact. We contacted students who were at the end of their first or second year on these two educations. They answered a questionnaire about personal characteristics, experiences with mobile phones and personal digital assistants, their knowledge about the short messaging service, and their prior involvement in usability evaluations. Based on their responses we distributed them on the different techniques with the intention of avoiding bias. Due to test subjects cancelling or not showing up, and our aim to conduct the same number of tests with each technique, we ended up conducting six tests with each of the six techniques, involving a total of 36 test subjects.

After the evaluation sessions, the three usability evaluators analysed the video recordings individually in random order and produced three lists of usability problems with severity ratings of critical, serious or cosmetic in accordance to the definition proposed by Molich (2000). Following this, the evaluators met and discussed their individual problem lists until consensus on one complete list was reached.

4.2. Experiment B

To complement experiment A, we designed and conducted a different experiment based on framework B (Jacobsen et al., 2002). The purpose of this experiment was to compare the extent to which laboratory and field evaluations supported the comparison of two different mobile phones. Experiment B involved two different techniques:

1. Using a mobile system while playing a computer game requiring the player to move around physically on a mat placed on the floor.
2. Using a mobile system while walking in a pedestrian street (a typical use situation) serving as a reference for the other technique.

The computer game used in the first technique was the "Jungle Book Groove Party" for Sony PlayStation 2. When playing this game, the user steps on different active areas on a "dance mat" according to sequences shown on a monitor. In all evaluations, the test subject played the game on the easiest level. The idea was to force the user into a situation with clearly divided attention between the use of the mobile system and playing the game. The second technique was similar to the pedestrian street technique mentioned in experiment A.

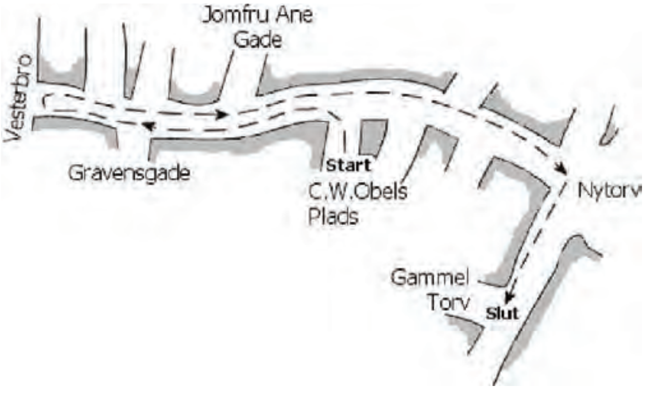


Figure 3. The route to be walked in the evaluations in the pedestrian street

We conducted a series of usability evaluations employing each of the two techniques above to evaluate the usability of two different mobile phones. In each evaluation, the user solved a number of tasks using one of the two mobile phones.

Figure 4. The Nokia 3310 and Nokia 5510

The two mobile phones were the Nokia 3310 and the Nokia 5510. They have comparable functionalities, but the keyboards are very different. The Nokia 3310 has a typical mobile phone keyboard with one digit and three or more letters assigned to each key. The Nokia 5510 has a full keyboard with only one character assigned to each key. Again, we decided to focus on the use of the short messaging service because this is a widely used service among mobile users and because it requires extensive text input as well as reading on the screen. In each evaluation, we collected two types of data:

- **Usability problems:** the evaluations in the laboratory were recorded on video. In the pedestrian street evaluations only audio was recorded and an observer took notes. After the evaluations, the video and audio recordings and notes were analysed to produce a list of usability problems.
- **Performance:** From the video and audio recordings, the time spent on solving each task was measured.

The overall hypothesis of the experiment was that the two techniques would come out with similar results in terms of the data collected.

Four persons took turns in serving as evaluators. The test subjects were a mixture of grammar school students, university students, and business employees. As we only had 12 test subjects available for experiment B, we decided to let them use both mobile phones in counterbalanced order. The test subjects were divided into two groups according to their frequency of mobile phone and short messaging service use. When the evaluations were carried out, one test subject did not show up, and two test subjects were only able to participate in one test.

The test subjects were presented with two tasks; a trial task and an evaluation task. The trial task was completed while standing still, without walking or using the dance mat. The purpose of the trial task was to make sure that all subjects had used the specific type of mobile phone being evaluated at least once. The evaluation task involved the same functionality as the trial task, but the sequence and data were different.

After the evaluations, the four usability evaluators analysed the video recordings individually in random order and produced four lists of usability problems with severity ratings of critical, serious or cosmetic (Molich, 2000). The individual problem lists were then merged into one complete list through discussions among the four evaluators until consensus was reached.

5. RESULTS

This section presents the key results from the two experiments described above.

5.1. Experiment A

In experiment A, we collected data about identification of usability problems, performance and workload.

Usability Problems

The primary basis for evaluating a technique should be the number of usability problems it helps identifying. Thus, we have analysed the collected data in order to evaluate the extent to which each technique supports identification of usability problems.

Table 3. Mean numbers and standard deviations of usability problems identified by each technique

	Technique					
	1	2	3	4	5	6
Mean	10.8	7.5	6.7	6.7	5.2	6.3
Std. deviation	1.6	1.5	2.0	2.4	2.4	2.1

Our hypothesis was that the best laboratory technique would identify a number of usability problems similar to the number identified in the pedestrian street condition (technique 6). Table 3 is based on the number of usability problems identified with the 36 test subjects. It shows the mean number of identified usability problems along with the standard deviation, distributed on technique. This is also illustrated in figure 5.

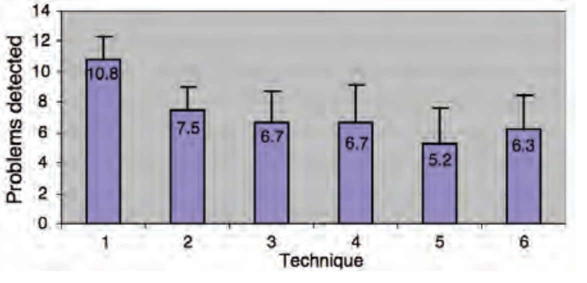


Figure 5. Number of usability problems detected with each of the six techniques

An analysis of variance of these numbers shows that the difference between the means are highly significant ($F_{5.30}$=2.53, p=0.001) and the use of Fisher's least significant difference test supports the conclusion that sitting at a table (technique 1) differs from the rest of the techniques. This means that the test subjects in the sitting condition supported identification of significantly more usability problems than with any of the other techniques.

Table 4. Number of identified usability problems

	Technique						
	1	2	3	4	5	6	**Combined**
Critical	4	4	3	4	3	3	4
Serious	11	11	9	9	9	8	17
Cosmetic	19	8	8	8	6	12	32
Sum	34	23	20	21	18	23	53

Table 4 shows the number of usability problems categorized from severity ratings and distributed on techniques. For severity rating, we used the criteria proposed by Molich (2000). In total, 53 unique usability problems were identified across all techniques, consisting of 32 cosmetic problems, 17 serious problems and 4 critical problems. No single technique supported the identification of all usability problems. In the sitting condition (technique 1), a total of 34 problems were identified. All of the others supported the identification of roughly half the usability problems. Looking at critical and serious problems, the six techniques supported the identification of nearly the same number. The main difference between the sitting condition and the other techniques relates to the number of cosmetic problems. Thus more than double the number of cosmetic problems was identified in the sitting condition than when using any of the other laboratory techniques.

Performance

We expected clear differences between the performances that test subjects would achieve with the different evaluation techniques. The main measure was the time spent on solving each of the five tasks. Our hypothesis was that the users who employed the best techniques would have similar performance to those walking in the pedestrian street (technique 6).

An analysis of variance of the time spent on each task with each technique did not enable us to identify any systematic differences between the techniques. The technique with the fastest task completion time changed from task to task. We also analysed the number of completed tasks, the number of wrong and undetected pressings of buttons, and the number of requests from the subject to have the task description repeated without finding any differences that could be clearly attributed to the different techniques.

Workload

The data on workload exhibited more difference between the techniques. Based on the theories behind the new techniques our hypothesis was that the workload would differ significantly as more body motion or attention was needed.

Table 5 shows the calculated workload numbers for the overall workload and three of the six contributing factors. The last three contributing factors; timing demands, own performance, and frustration did not provide any significant results.

Mental demands are described in the NASA task load index (TLX) as the amount of mental activity, e.g. thinking or use of memory, required to perform a piece of work. We were expecting to see an increase in these demands, as more attention was required in the evaluations. As can be seen in table 5, the numbers do seem very different, ranging from 29 to 204 and increasing as the techniques become more complicated. However, a variance analysis did not show any significant difference between the techniques ($F_{5.30}$=1.91, p=0.12). The reason for this is that the test subjects within each technique rated this factor very differently – for one technique the rating ranged from 15 to 85.

A pairwise comparison of technique 1-6 does show some significant differences though: compared to all but walking on a treadmill at constant speed (technique 2), the sitting condition (technique 1) demands significantly less mental activity, which seems to confirm that the new laboratory techniques and the pedestrian street evaluations require more attention from the test subjects. However further comparisons between the techniques in relation to physical demands, effort and overall workload do not reveal any significant results, aside from walking on a treadmill at constant speed (technique 2) compared to walking on a treadmill at varying speed (technique 3).

Table 5. Subjective experience of workload with the different techniques

	Technique					
	1	2	3	4	5	6
Mental demands	29	75	204	126	185	148
Physical demands	92	117	112	118	127	194
Physical demands	52	163	106	228	178	186
Overall workload	27	35	48	55	48	54

Physical demands mean how much physical activity was required, including both walking, dragging an icon or pushing a button. We expected to see an increase in this for the techniques, which involved more body motion. However, as before, a variance analysis shows no significant differences ($F_{5.30}$=0.48, p=0.79). Despite the seemingly big difference between e.g. walking on a treadmill at constant speed (technique 2) and walking in a pedestrian street (technique 6) a pairwise comparison of the techniques show no significant differences either.

Effort is explained as a combination of the mental and physical demands and for this our variance analysis shows significant difference between the techniques ($F_{5.30}$=3.27, p=0.02). Fisher's least significant difference test identifies that sitting requires significantly less effort than the other techniques. When we compare the figures for techniques with the same sort of attention but with constant versus varying speed, i.e. walking on a treadmill (technique 2 versus 3) or walking on a changing track (technique 4 versus 5), the test subjects seem to feel that constant speed requires more effort. However, the difference is not significant. The reason may be that the varying speed included some intervals with low speed where the test subjects got a chance to relax, before the speed increased again.

The overall workload exhibits a very significant difference between the techniques ($F_{5.30}$=4.14, p<0.01). Fisher's least significant difference test reveals no significant difference between sitting (technique 1) and walking on a treadmill at constant speed (technique 2), which indicates that the motion at constant speed is not experienced very different from sitting. The reason may be that walking at constant speed quickly becomes automatic and therefore not requiring much more attention than sitting at a table. On the other hand, walking at varied speed (technique 3) is significantly different from sitting. Walking on a changing course (technique 4 and 5) and in a pedestrian street (technique 6) are also significantly different from sitting and walking at constant speed, which implies that varied speed and conscious attention do in fact put a greater workload on the test subject. However there is no significant difference between the two techniques involving running or between the two techniques where the subjects are walking on a changing track.

It is impossible to pick a single technique that fully resembles the workload for the pedestrian street technique. For some factors having a need for both varying body motion and conscious attention seem to simulate the pedestrian street the best while in others, the less complicated techniques seem to be better.

5.2. Experiment B

The usability problems found from experiment B are shown in table 6. Again, severity ratings are based on the criteria proposed by Molich (2000).

Table 6. Number of usability problems detected with each technique

	Dance mat				Pedestrian street			
	Critical	Cosmetic	Cosmetic	Total	Critical	Serious	Cosmetic	Total
Nokia 3310	0	8	3	11	2	4	3	9
Nokia 5510	0	10	11	21	4	10	9	23

For both mobile phones, a comparable number of usability problems are identified across techniques (11 and 9 for the Nokia 3310 and 21 and 23 for the Nokia 5510). However, with the pedestrian street technique the evaluation identified some critical problems, which were not identified in the dance mat condition.

The measures of performance for this experiment are provided in table 7. Each number is the time taken for a test subject to complete the task in the given context. This table illustrates that the test subjects walking in the pedestrian street solved the tasks faster than the test subjects that used the dance mat. In the pedestrian street, the evaluation task was completed faster than the trial task. This improvement was expected as the subjects became familiar with the mobile phone. With the dance mat, we have the opposite result. The trial task was solved faster than the evaluation task. This indicates that the dance mat demands more attention from the user than walking in the pedestrian street.

Table 7. Average time spent solving the task with the mobile phones (minutes:seconds)

	Dance mat			Pedestrian street		
	Subject	Trial	Evaluation	Subject	Trial	Evaluation
Nokia 3310	S1	5:55	3:59	S7	3:10	3:38
	S2	2:48	3:52	S8	5:27	4:12
	S3	3:10	3:19	S9	4:20	2:44
	S4	6:52	7:44	S10	3:04	3:30
	S5	3:28	4:34	S11	3:05	2:40
	S6	–	–	S12	–	–
	Mean	4:27	4:42	Mean	3:49	3:21
Nokia 5510	S1	6:43	5:41	S7	6:50	3:57
	S2	4:27	4:30	S8	3:57	3:15
	S3	2:29	2:35	S9	3:49	2:44
	S4	7:09	7:02	S10	–	–
	S5	8:40	6:37	S11	5:49	4:01
	S6	7:52	5:46	S12	–	–
	Mean	6:13	5:22	Mean	5:06	3:29

6. DISCUSSION

This section discusses the experiments and issues from them that go beyond the results presented above.

6.1. The Sitting Technique

An interesting and surprising result of our experiments is that technique 1 (sitting at a table) supports the identification of more usability problems than any other techniques. In this sense, the traditional usability evaluation technique seems superior.

The data on workload indicates a potential reason of this result. Pairwise comparisons of the workload data across the techniques show that sitting (technique 1) compared to all other techniques, except walking on a treadmill at constant speed (technique 2), demands significantly less mental activity.

We have analysed the video recordings from experiment A for consequences of this reduced experience of mental demands. Generally, identification of usability problems is based on the test subjects thinking aloud. If the test subjects talk less, we may miss usability problems. Our video recordings indicate that the test subjects in the sitting condition (technique 1) spent more time and energy thinking aloud and commenting on what they observed compared to the test subjects in the other tests. The test subjects who were sitting down at the table also had energy to comment on all the small things they observed. The test subjects in the evaluations based on the five other techniques were mostly thinking aloud when they observed a larger usability problem.

Theory on human information processing can be used to explain this. Thinking aloud is a conscious action, which requires some amount of attention, much in the same way as motion and navigation. With sitting (technique 1) the users only needed to do one action, which was to solve the tasks. Therefore, these test subjects only had to divide their attention and effort between two actions: solving the task and thinking aloud. With the other five techniques, the users needed to solve the tasks, move physically and navigate. Therefore, these test subjects had to divide their attention and effort between three or more actions; solving the task, moving, navigating, and thinking aloud.

The number of usability problems found in the evaluations provides an indication of the consequences of this different demand. The results show that the techniques involving multiple actions support identification of less usability problems than the techniques with fewer actions. But when the usability problems are divided into the three categories of critical, serious and cosmetic problems, the results show that the major differences between the techniques reside in the amount of cosmetic problems found. This supports our assumption that the test subjects in the sitting condition (technique 1) point out every problem they find as opposed to the test subjects in the other five techniques, who only point out the serious and critical problems they encounter.

6.2. Usability Problems and Mobility

A detailed analysis of the performance results revealed that the test subjects in techniques with much motion and navigation are more likely to miss a button on the interface. This can happen if a user presses a button but moves the stylus out of the button before releasing it, or if a user unintentionally hits the wrong button.

In experiment A, the test subjects that were sitting at a table (technique 1) missed a button on average 2 times throughout the whole evaluation. In the techniques with motion but no navigation (technique 2 and 4) a button was missed about 3 times per test subject. In the techniques that involved both motion and navigation (technique 3, 5 and 6), a button was missed on average between 3.5 and 6 times per test subject. This difference between the six techniques is less significant ($F_{5.30}$=2.30, p=0.10), but it indicates that the techniques involving movement and navigation are better at finding problems concerning the interface layout and the sizes and placement of the individual interface elements.

Problems using the devices and programs can also be found in experiment B. In the pedestrian street evaluation, a total of 14 errors were made (divided on six of the 12 tests) when writing the short messages, whereas only five errors were made in the dance mat test (divided on only three of the 12 tests). Unfortunately, the data collected in this experiment do not allow further enquiry into the causes of this difference as only audio was recorded.

6.3. A Changing Track

One of the techniques involved walking on a track that was changing (technique 5). The idea of changing the track was to increase the need for attention. However, the pairwise comparisons between mental demands for each technique did not reveal such an effect.

One possible reason for this was discovered during the usability evaluations involving technique 4 and 5. We learned that most of the test subjects just followed the person ahead of them by keeping track at them out of the corner of their eyes. Rather than navigating between the obstructions the test subjects simply followed the person who set the speed and counted on him to avoid walking into anything, thereby reducing the attention needed. Thus the navigation did not appear to be as conscious as we wanted.

This problem may be solved by e.g. making the laboratory setup even more dynamic with more persons and moving objects.

6.4. Data Collection in the Field

We designed the pedestrian technique to include systematic data collection. In experiment A, all tests in the pedestrian street were video recorded using a video camcorder as shown in figure 6. In experiment B, we only recorded audio but also took written notes.

Collecting high-quality video data in the field turned out to be very difficult. It was not easy to record images of the screen of the iPAQ while walking. In addition, the users often moved their hands in a manner that covered the screen. Furthermore, it was difficult to experience “realistic” pedestrian motion, since the other pedestrians tended to move away from the three persons walking along the street; the test subject, the evaluator and the person operating the video camcorder (see figure 6). This problem may be solved by e.g. changing the role of the evaluator and mounting small cameras and microphones on the test subject. Recording only audio in the field resulted in a lack of detailed data about the interaction with the system. Consequently, we had to rely heavily on written notes during analysis. Thus while this approach was less obtrusive and easier to carry out in practice, it did not prove more valuable than recording the field evaluations on video.

Figure 6. Usability evaluation in the pedestrian street

6.5. Involving social context

When the places and environments where mobile systems are being used are compared to the theories that we have used for creating our testing techniques, there is a gap in the experiments. One of these gaps is the integration of social context. For instance, a user of a mobile system may be working with some colleagues while interacting with the mobile system, or the user may be working with colleagues through the mobile system. We have only used theories that enable us to see mobility as something that involves motion and navigation. None of the theories cover the social context. Therefore, this aspect has not been a part of our experiments. Future research should investigate further into this.

7. CONCLUSION

This paper has presented six techniques for evaluating the usability of a mobile system in a laboratory setting. The aim was to explore techniques that could facilitate evaluating mobile systems in a controlled environment while being as similar to a real use situation as possible.

Five techniques were developed from a framework that described mobility in terms of physical motion and the amount of attention needed to navigate while moving. A sixth technique was developed to divide the user's attention between conscious actions and the use of the mobile system.

The proposed techniques were evaluated through two experiments. In both experiments, walking in a pedestrian street while using the mobile system being evaluated was used as reference. There were no significant differences between the techniques in terms of user performance. On workload the techniques exhibited significant differences in terms of perceived effort and overall workload. However, there was no single technique that resulted in exactly the same workload as walking in the pedestrian street.

There was only one significant difference in terms of support to identification of usability problems. Sitting at a table, which was the simplest of the six new techniques, was clearly better than any other technique when focussing on identification of usability problems. However, the difference mainly related to cosmetic problems.

Both of the experiments have clear limitations. Each technique has been evaluated with six test subjects, except for three cases in the second experiment, where only four or five test subjects were used. More test subjects would have been desirable. Our aim was to facilitate comparison of several techniques with a limited number of test subjects. A follow-up experiment on selected techniques should increase the number of test subjects.

The proposed framework and subsequent experiments have focused on exploring techniques for recreating challenges of moving physically while interacting with a mobile computer system. Moving physically is, however, only one of many new factors involved in the use of mobile computer systems as opposed to traditional desktop applications. Other relevant factors are the social, physical and temporal context of mobile system use. Some studies exists, which explore how these factors influence usability evaluations and how they can be incorporated into laboratory based evaluation techniques (Lai et al. 2001, Salvucci, 2001, Kjeldskov and Skov 2003b) but further research and experiments are needed to develop new and refine existing ideas and techniques.

ACKNOWLEDGMENTS

The authors thank all test subjects and Elena T. Beck, Morten K. Christiansen, Nikolaj Kolbe, Charlotte S. Jacobsen, Pernille Langvad, Janne J. Sørensen and Heidi B. Thomsen who participated in the conduction of the two experiments.

REFERENCES

Bailey, R.W., Allan, R.W., Riello, P., 1992. Usability Testing vs. Heuristic Evaluation: A Head-to-Head Comparison. Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting, HFES, pp. 409-413.

Beck, E.T., Christiansen, M.K., Kolbe N., 2002. Metoder til Brugbarhedstest af Mobile Apparater. Department of Computer Science, Aalborg University (in Danish)

Brewster, S., 2002 Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing, 6, 188-205.

Brewster, S., Murray, R., 2000. Presenting Dynamic Information on Mobile Computers. Personal and Ubiquitous Computing, 4, 209-212.

Danesh, A., Inkpen, K., Lau, F., Shu, K., Booth, K., 2001. Geney: Designing a Collaborative Activity for the Palm Handheld Computer. Proceedings of CHI'2001, ACM, New York, pp. 388-395.

Dumas, J.S., Redish, J.C., 1999. A Practical Guide to Usability Testing. Exeter: Intellect.

Ericsson Mobile Communications, 2001. Enhanced Messaging Service, white paper, http://www.ericsson.com.au/about/media_center/white_papers/articles/EMS.pdf

Graham R., Carter C., 1999. Comparison of Speech Input and Manual Control of In-Car Devices while on-the-move. Proceedings of the Second Workshop on Human Computer Interaction with Mobile Devices, Mobile HCI'1999, Edinburgh, Scotland.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Hancock, P. A. and Meshkati, N. (Eds.), Human Mental Workload, Elsevier Science Publishers, Amsterdam, pp. 139-183.

Henderson, R., Podd, J., Smith, M., Varela-Alvarez, H., 1995. An Examination of Four User-Based Software Evaluation Methods. Interacting with Computers, 7(4), 412-432.

Jacobsen, C.S., Langvad, P., Sørensen J.J., Thomsen H.B., 2002. Udvikling og Vurdering af Brugbarhedstest til Mobile Apparater. Department of Computer Science, Aalborg University (in Danish)

Johnson, P., 1998. Usability and Mobility: Interactions on the move. Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices

Karat, C.M., Campbell, R., Fiegel, T., 1992. Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation. Proceedings of CHI'92, ACM, New York, pp. 397-404.

Kjeldskov J., Graham C., 2003 A Review of MobileHCI Research Methods. Proceedings of the 5th International Mobile HCI 2003 conference, Udine, Italy. Lecture Notes in Computer Science, Springer-Verlag, Berlin.

Kjeldskov J. and Skov M.B., 2003b. Creating a Realistic Laboratory Setting: A Comparative Study of Three Think-Aloud Usability Evaluations of a Mobile System. Proceedings of the 9th IFIP TC13 International Conference on Human Computer Interaction, Interact 2003. IOS Press, Zürich, pp. 663-670.

Kjeldskov, J., Skov, M.B., 2003a. Evaluating the Usability of a Mobile Collaborative System: Exploring Two Different Laboratory Approaches. Proceedings of the 4th International Symposium on Collaborative Technologies and Systems 2003, Orlando, Florida. SCS press, San Diego, pp. 134-141.

Koppinen A., 1999. Design Challenges of an In-Car Communication UI. Proceedings of the Second Workshop on Human Computer Interaction with Mobile Devices, Mobile HCI'1999, Edinburgh, Scotland.

Kristoffersen, S., Ljungberg, F., 1999. Mobile use of IT. In Käkölä, T. K. (Ed.), Proceedings of the 22nd Information Systems Research Seminar in Scandinavia, vol. 2, 271-284.

Lai, J., Cheng K., Green, P., Tsimhoni O., 2001. On the Road and on the Web? Comprehension of synthetic speech while driving. Proceedings of CHI'2001, ACM, New York, pp. 206-212.

Luff, P., Heath, C., 1998. Mobility in Collaboration.Proceedings of CSCW'98, ACM, New York, pp. 305-314.

Molich, R., 2000. Usable Web Design (In Danish). Ingeniøren|bøger

Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D., Kirakowski, J., 1998. Comparative Evaluation of Usability Tests. Proceedings of the Usability Professionals Association Conference, pp. 189-200.

NASA, Task Load Index, http://iac.dtic.mil/hsiac/Products.htm#TLX.

Nielsen, C., 1998. Testing in the Field. In Werner, B. (Ed.), Proceedings of the third Asia Pacific Computer Human Interaction Conference, IEEE Computer Society.

Nielsen, J., 1993. Usability Engineering. Morgan Kaufmann.

Nielsen, J., 2000. Why you only need to test with 5 users, Alertbox, http://www.useit.com/alertbox/20000319.html

Pascoe, J., Ryan, N., Morse, D., 2000. Using While Moving: HCI Issues in Fieldwork Environments. Transactions on Computer-Human Interaction, 7(3), 417-437.

Petrie, H., Johnson V., Furner, S., Strothotte, T. (1998) Design Lifecycles and Wearable Computers for Users with Disabilities. In Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices, Glasgow.

Pirhonen, A., Brewster, S.A., Holguin, C., 2002. Gestural and audio metaphors as a means of control for mobile devices. Proceedings of CHI'2002, ACM, New York.

Preece, J., Roger, Y., Sharp, H., Benyon, D., Holland, S., Carey T., 1994. Human-Computer Interaction. Addison-Wesley.

Rantanen, J., Impio, J., Karinsalo, T., Reho, A., Tasanen, M., Vanhala, J., 2002. Smart Clothing Prototype for the Artic Environment. Personal and Ubiquitous Computing, 6, 3-16.

Rubin J., 1994. Handbook of Usability Testing. Wiley.

Salvucci, D.A., 2001. Predicting the Effects of In-Car Interfaces on Driver Behavior using a Cognitive Architecture. Proceedings of CHI'2001, New York, pp 120-127.

Sawhney, N., Schmandt, C., 2000. Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments. Transactions on Computer-Human Interaction, 7(3), 353-383.

Sharples, M., Corlett, D., Westmancott, O., 2002. The Design and Implementation of a Mobile Learning Resource. Personal and Ubiquitous Computing, 6, 220-234.

Thomas, B., Grimmer, K., Zucco, J., Milanese, S., 2002. Where Does the Mouse Go? An Investigation into the Placement of a Body-Attached TouchPad Mouse for Wearable Computers. Personal and Ubiquitous Computing, 6, 97-112.

Virzi R.A., 1992. Refining the test phase of usability evaluation: how many subjects is enough? Human Factors, 34(4), 457-468.

Chapter 12

Simulating the domain

Jesper Kjeldskov and Mikael B. Skov

Abstract. Increased complexity of organizations and emerging technologies poses new and difficult challenges for the evaluation of software systems. Several years of research have proven that usability evaluations are invaluable tools for ensuring the quality of software technologies, but the increased complexity of technology requires new ways of understanding and evaluating the quality of software systems. This paper explores limitations, challenges and opportunities for studying mobile technologies "in use, in situ", in laboratories (in vitro), and in controlled high-fidelity simulations of the real world. We call the latter condition *in sitro*. We report from two different case studies of evaluating the usability of mobile systems within these three different conditions. Our results show that it is possible to recreate and simulate significant elements of intended future use situations in laboratory settings and thereby increase the level of realism while at the same time maintain a high level of control. In fact, the in sitro condition was able to identify most of the same usability problems as found in the other conditions. However, the in situ evaluation proved to provide a level of realism that is difficult to achieve in laboratory environments.

1. INTRODUCTION

As stated in the introduction to the 2005 In-Use, In-Situ: Extending Field Research Methods Workshop, "the increasing complexity of organizations and systems of communication, and the fast pace of technological change and adaptation, poses a challenge for researching the cognitive, social and cultural impact of technology that is in use in its natural settings, in situ" (Amaldi et al. 2005). One of the areas where this statement seems to be of particular importance is within the research field of mobile human-computer interaction and systems design, where the emergence of new mobile, pervasive and ubiquitous technologies continues to extend the scope of computer use in the workplace, in the home and in public, and consequently calls for research into people's use of technology going beyond our traditional laboratory approaches.

Originally published as Kjeldskov, J. and Skov, M. B. (2007) Studying Usability in Sitro: Simulating Real World Phenomena in Controlled Environments. *International Journal of Human-Computer Interaction,* 22(1), 7-37.

In the proceedings of the first workshop on Human-Computer Interaction for Mobile Devices in 1998, researchers and practitioners were encouraged to further investigate the criteria, methods, and data collection techniques for studying mobile system use (Johnson, 1998). Of specific concern to the development of such methods and techniques, it was speculated that traditional laboratory approaches would not adequately be able to simulate the context surrounding the use of mobile systems and that evaluation techniques and data collection methods such as think-aloud, video recording or observations would be extremely difficult in natural settings – in situ. These concerns have since been confirmed through a number of studies, for example, Graham and Carter (1999), Pascoe et al. (2000), Rantaten et al. (2002), Brewster (2002), Esbjörnsson et al. (2003), and Kjeldskov and Stage (2004).

A number of different techniques have been suggested for studying technology in use, in situ such as work place observations, contextual enquiries, interviews, focus groups, automatic logging of user actions, acting-out in context, and cultural and technology probes. While such techniques provide valuable insights into actual use of software technologies, they are often rather limited in their ability to assess the specific qualities of the technologies in use and weak in their ability to identify design problems and to inform redesign. Contrasting these methods, several years of HCI research have proven that usability evaluations are invaluable tools for measuring and improving the quality of software technologies, and hence usability engineering is today an established discipline within interaction design with widely acknowledged techniques and methods. With the emergence of mobile, pervasive and ubiquitous technologies and the fast speed of technological change and adaptation that these technologies involve, the field of usability engineering is now faced with challenges such as lack of realism and real world richness. In our view, this indicates an opportunity for combining the strengths of in situ empirical methods and usability evaluation techniques to overcome some of their individual shortcomings.

In 2003, a literature study on mobile HCI research methods revealed that 41% of the mobile HCI research and design reported in the main literature from 2000-2002 involved some sort of usability evaluation (Kjeldskov and Graham, 2003). However, even though evaluations of mobile systems are thus clearly prevalent, surprisingly little research had (and still has) been published concerning the methodological challenges described above. Exceptions include studies comparing two or more methods applied for evaluating mobile prototype systems in, for example, Brewster (2002), Graham and Carter (1999) and Pirhonen et al. (2002). Consequently, there is as yet no agreed set of appropriate usability evaluation methods and data collection techniques within the field of mobile HCI and we still have little knowledge about the relative strengths and weaknesses of laboratory-based and field-based usability evaluations of mobile systems. While the literature study (Kjeldskov and Graham, 2003) also revealed that 71% of mobile device evaluation was done through laboratory experiments and only 19% through field studies, it seems implicitly assumed that usability evaluations of mobile devices *should* be done in the field (Abowd and Mynatt, 2000; Brewster, 2002; Johnson, 1998). However, field-based usability studies are not easy to conduct. They are time consuming and the added value is questionable. For discussions of some of the challenges of evaluating mobile systems in the field see, for example, Pascoe et al. (2000), Rantanen et al. (2002), Esbjörnsson et al. (2003), Kjeldskov and Stage (2004). Partly motivated by these challenges, it has been suggested that instead of going into the field when evaluating the usability of mobile devices and services, adding mobility or other contextual features such as scenarios and context simulations to laboratory settings can contribute to the outcome of laboratory evaluations while maintaining the benefits of a controlled setting. For usability studies of mobile devices and services simulating mobility or other contextual factors in laboratory settings (see, for example, Salvucci 2001, Lai et al. 2001, Bohnenberger et al. 2002, Pirhonen et al. 2002, Kjeldskov and Skov 2003, Kjeldskov and Stage 2004).

The purpose of this paper is to contribute to the body of research on appropriate methods and techniques for evaluating mobile systems use by exploring the differences and similarities between studying such systems "in use, in situ", in the laboratory, and in controlled high-fidelity simulations of the real world. We do this on the basis of two case studies of mobile system evaluation for real world work tasks in highly challenging use contexts. These two case studies involve four empirical evaluations of mobile systems carried out in three different experimental conditions on the continuum from laboratory (in vitro) to field (in situ). On the basis of the two case studies, we outline limitations and challenges of evaluating mobile technologies in laboratory settings and in the real world. In response to these limitations and challenges, we have experimented with the use of a complementary approach, evaluating "in sitro", where real world phenomena is simulated in a controlled environment. Based on a comparison of the usability evaluation results produced from each of these three conditions (in situ, in vitro and in sitro) we explore the relative strengths and weaknesses of in sitro evaluations in comparison with in vitro and in situ evaluations.

The paper is structured as follows. First we briefly highlight and discuss the value of studying technology use through the lens of usability. We then take up the discussion of trade-offs between realism and control when evaluating in laboratory (in vitro) and field settings (in situ), and discuss the intermediate approach of evaluating in sitro. Based on this discussion, we map our two case studies of mobile systems evaluation onto a continuum of in situ, in sitro and in vitro evaluation approaches outlining the relationships between four different empirical usability evaluations of mobile systems that we have carried out over the last three years. Sections 3 and 4 describe our two case studies of mobile systems usability evaluation. Section 3 describes a comparative usability study of a mobile system for communication on board large container vessels, where we took up the challenge of increasing laboratory realism. We present the context for the study, the system developed, and two evaluations carried out in a traditional laboratory setting and in a high-fidelity ship simulator. Following this, we outline and compare the findings from these two evaluation approaches. Section 4 describes a comparative usability study of a mobile electronic patient record system for use in a hospital ward where we took up the challenge of going in to the field for the purpose of evaluating the system's usability. Again, we present the context for this study, the system developed, and two evaluations carried out in a simulated hospital ward and at a hospital during real work activities. We also outline and compare the findings from these two evaluation approaches. In section

5, we take a step back and highlight and explore the differences and similarities between our three experimental approaches, and discuss the implications of our findings in relation to the issue of evaluating technology in use in situ. Finally, section 6 concludes on our research and points out avenues for further work.

2. EVALUATING THE USABILITY OF MOBILE SYSTEMS

Several years of research have proven that usability evaluations are invaluable tools for ensuring the quality of software technologies. Therefore, usability evaluation of stationary computer systems is an established discipline within human-computer interaction with widely acknowledged techniques and methods. Several well-known textbooks on usability testing and engineering describe and illustrate how to plan, design, and conduct evaluations, (e.g. Nielsen 1993, Rubin 1994, Dumas and Reddish 1999). These have contributed to improved evaluations and have had industrial impact. Furthermore, several attempts have 'evaluated evaluations', that is, empirical evaluations of the relative strengths and weaknesses of the different approaches and techniques under different circumstances, e.g. differences between think-aloud testing and heuristic evaluations testing (Bailey et al. 1992, Karat et al. 1992), and different user-based evaluation methods (Henderson et al. 1995, Molich et al. 1998). So far, these kinds of comparative studies are only beginning to emerge in relation to the evaluation of mobile computer systems. A significant proportion of mobile technologies take many of the well known methodological challenges of evaluating usability to an extreme. Users are often ambulatory, typically highly mobile during their interaction with the system, and they are situated in a dynamic and sometimes unknown use setting (Vetere et al. 2003). The information presented to the users of mobile systems is closely related to their physical location, to objects in their immediate surroundings, or to their present as well as planned activities (e.g. Chincholle et al. 2002, Pospischil et al. 2002). Such challenges raise a number of interesting issues to consider when trying to understand the usefulness and usability of mobile systems. In particular, several discussions have been raised to determine when to evaluate mobile systems in vitro, e.g. laboratories, and when to evaluate mobile systems in situ, e.g. real use context (Kjeldskov and Graham 2003).

2.1. In Situ or In Vitro: The Trade-Offs between Realism and Control

In situ and in vitro evaluations inherently integrate a number of characteristics.

- **In situ:** "in its original place". This condition defines that it is in its original location. For experiments involving the evaluation of computer systems, this often means that the use of the system takes place in its natural environment.
- **In vitro:** "in glass". This condition is distinguished from conditions that actually apply in nature. For experiments involving the evaluation of computer systems, this often refers to experiments that take place in controlled environments, such as usability laboratories.

In situ experiments are often characterized by a high level of realism (as illustrated in figure 1). When dealing with evaluations of software systems, in situ experiments involve real users interacting with the system in a real situation and in the real context of intended use. Thus, the empirical basis for assessing the quality of the system is often very realistic. On the other hand, in situ evaluations are not easy to conduct (Nielsen 1998, Brewster 2002) and applying established evaluation techniques and data collection instrumentation, such as multi-camera video recording, think-aloud protocols or shadowing may be difficult in natural settings (Sawhney and Schmandt 2000). Further, in situ evaluations complicate data collection since users are moving physically in an environment over which we have little control (Johnson 1998, Petrie et al. 1998) and only partially comprehend. Also, for several mobile systems it is difficult to define and describe the original location, as location can be distributed in both time and space. Finally, some in situ evaluations may be impossible to conduct due to ethics or safety-critical issues (Kjeldskov and Skov 2003).

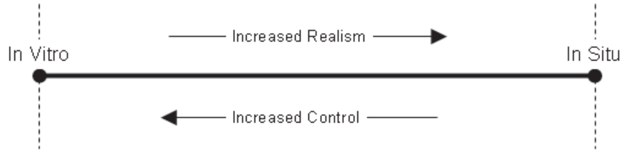


Figure 1. Trade-offs between a high level of control and a high level of realism.

In vitro experiments, on the other hand, are often characterized by a high level of control (as illustrated in figure 1). For interaction design or human-computer interaction, in vitro evaluations often refer to experiments that take place in controlled environments, such as usability laboratories. In vitro evaluations can often benefit from experimental control and high quality data collection when conducted in usability laboratories. Yet traditional usability laboratory setups may not adequately simulate the context surrounding the use of mobile systems. Thus, in vitro evaluations of mobile systems raise a number of challenges. First, the relation between the system and activities in the physical surroundings can be difficult to capture in expert evaluations such as heuristic evaluation or recreate realistically in a usability laboratory. Secondly, working with systems for highly specific domains (Kjeldskov and Skov 2003, Luff and Heath 1998), laboratory studies may be impeded by limited access to prospective users on which such studies rely. While benefiting from the advantages of a controlled experimental space, evaluating the usability of mobile systems without going into the field thus challenges established methods for usability evaluations in controlled environments.

2.2. Simulation: An Attempt to Bridge Realism and Control

The inherent challenges of in situ and in vitro experiments related to realism and control have facilitated the introduction of additional experimental conditions. It is quite obvious that several of the outlined challenges cannot be solved through simple means as they are inherently integrated into the nature of the experiments e.g. the lack of realism when using a computer system in a laboratory. As a consequence, such challenges are often rather difficult to address and solve. However, a number of attempts have been suggested to overcome some of these difficulties. One viable way is simulation where selected elements of the experimental condition are simulated using computers or simulators.

Several different types of simulations have been proposed and assessed and different terms are often used for these simulations. In the following, we will discuss two related but different types of simulations.

The first type of simulation is often referred to as *computational simulation* where computers fully simulate parts of an environment. Profoundly used in biology, such simulations serve to explore or investigate issues that are often difficult to do in vivo or in vitro, e.g. Roulet et al. (1998) state that computational molecular biology tools are becoming the method of choice for screening of certain DNA sequences. Computational simulations in biology have been coined *in silico* experiments (Wingender 1998). This experimental condition stems from the Latin phrases *in vivo* and *in vitro* which are commonly used in biology and refer to experiments done in living organisms and outside of living organisms respectively. Wingender (1998) states that in silico has been introduced into life sciences as a pendant to "in vivo" (in the living system) and "in vitro" (in the test tube) and implies the gain of insights by theoretical considerations, simulations, and experiments conducted on a silicon-based computer technology. Thus, simulation of real world phenomena is important for such experiments. In silico experiments have further been adapted in other disciplines, e.g. in computer science where Zhao et al. (2004) applied in silico experiments to simulate certain network behaviours. In summary, in silico experiments or computational simulations prove valuable when trying to understand effects of introducing new elements into a known environment that can be described (simulated) on a computer.

While computational simulations provide promising conditions for experiments that can be fully automated, we will bring attention to another kind of simulation referred to as *simulators* in which advanced high-fidelity, tailored environments provide a realistic context for human activity, for example training, system design, or personal assessment (Sanders 1991). For this particular stream of simulation, Harman (1961) defines it as "a technique of substituting a synthetic environment for a real one, so that it is possible to work under laboratory conditions of control". Hence, experimenters are able to obtain a significant high level of realism while maintaining control over the experimental condition. Simulations with simulators are widely adopted within ergonomics and human factors for primarily training purposes and secondarily system design and personal assessment purposes (Sanders 1991).

Simulators provide a very useful experimental approach but studies have stressed potential challenges characterising their use. One key problem with simulators is performance measurement. Vreuls and Obermayor (1985) found that system performance measurements in highly sophisticated simulators are virtually useless due to poor system design – part of the problem resides in the fact that it is not clear what should or could be measured. Another important issue to consider is validation. Sanders (1991) argues that validation of simulators is crucial in order to establish how well the simulation actually reflects reality. Not until the simulation has been satisfactorily validated, can it be used itself to evaluate the effect of deviation from the full physical fidelity. Alexander et al. (2005) also acknowledge the importance of fidelity and describe it as the extent to which the virtual environment emulates the real world. Different subcategories of fidelity have been proposed like physical, functional, cognitive fidelities (Allen et al. 1985, Hays and Singer 1985) and psychological fidelity (Mayor and Volanth 1985) where e.g. functional fidelity has been defined as the degree to which the simulation acts like operational equipment in reacting to the tasks executed by the trainee (Allen et al. 1986). Highly sophisticated simulators can almost truly simulate the different subcategories of fidelity but they are often rather expensive and not very lightweight (Alexander et al. 2005). So far, a lot of effort has been put into making the simulator as realistic as possible and Sanders (1991) states that full simulations should be a final test and demonstration of the suitability of a new design rather than an open-ended trail.

2.3. In Sitro: Striving for Mobile Usability Realism and Control

In this paper, we take a slightly different approach to simulation compared to the types illustrated above when trying to evaluate the usability of mobile systems. We are also concerned with simulating real world phenomena when trying to enhance a controlled laboratory setting, but none of the above outlined approaches for simulation fits our work properly. First, the in silico experiments require that the simulation is conducted on a computer (e.g. Wingender 1998). This is not the case for the evaluations we are interested in as we explore human activities with computer artefacts. Human activity is central in simulators, but full simulations as illustrated by (e.g. Sanders 1991, Hays and Singer 1985) tend to focus several aspects of the human activity and related challenges of measuring and tailoring the realism. For the usability evaluation, we are primarily concerned with the identification of usability problems that prohibit a successful and fruitful interaction with the mobile system. Therefore, we wish to create an environment that partly or fully simulates other activities found in the real use context.

As a consequence, we coin an analogous term for conditions simulating real world phenomena in controlled environments when evaluating computer systems: *in sitro*. In sitro is concatenated from *in situ* and *in vitro* and stresses the combinational nature of the two conditions and of simulation of context. We define it as follows:

- **In sitro:** “in simulated context”. This experimental condition describes a partially or fully simulated controlled laboratory-based evaluation where the intended future in use situation is being simulated.

The principle idea behind in sitro experiments is that part of the real world phenomena is simulated in the laboratory. As illustrated in figure 2, the aim of in sitro experiments is to increase the realism of in vitro evaluations while at the same time increasing the level of control of in situ evaluations.

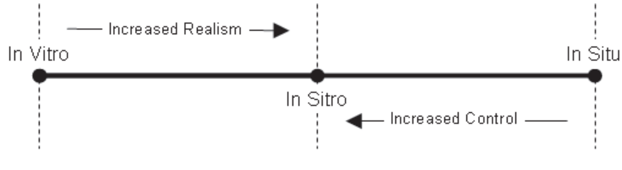


Figure 2. In sitro evaluations with increased levels of control and realism

2.4. In Sitro: Empirical Investigation

We present two independent cases involving four studies of usability evaluations of mobile systems involving 24 participants. These two cases serve to illustrate opportunities and limitations of our proposed experimental condition, in sitro. Our empirical investigation of the in sitro condition is illustrated in figure 3.

The investigation contains two cases (A and B) of evaluating usability of mobile devices; both cases contain two studies adopting different evaluation conditions. Case A focused on increasing laboratory realism for the evaluation of a mobile system for coordination and collaboration on a large container vessel contrasting the use of a traditional laboratory setup (in vitro) with a high-fidelity simulation of the intended use context (in sitro). Case B focused on a mobile system for healthcare contrasting the use of a high-fidelity simulation of the use context (in sitro) with going out into the real world (in situ). The two cases are presented and discussed in chapters 3 and 4.

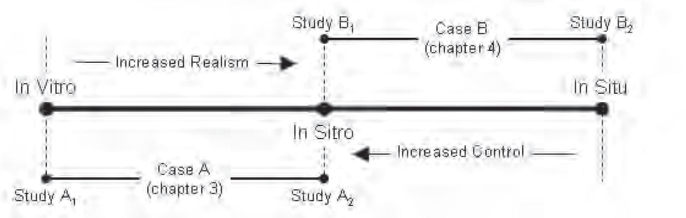


Figure 3. Illustration of our two studies investigating in sitro experiments.

3. INCREASING LABORATORY REALISM

Our first comparative usability case (Case A) focused on the opportunities and challenges of increasing laboratory realism for the evaluation of a mobile system contrasting the use of a traditional laboratory setup (in vitro) with a high-fidelity simulation of the intended use context (in sitro) (Kjeldskov and Skov, 2003). This study originated from our involvement in a large multidisciplinary research project involving ethnographic field studies of work activities in the maritime domain involving computerized process control and information systems (Andersen, 2000; Nielsen, 2000). As a part of this project, a mobile communication and coordination system, the Maritime Communicator, was developed for workers performing safety-critical collaborative work tasks on board very large container vessels. Evaluating the usability of this system was a particular challenge for several reasons. Firstly, the evaluation could not be done in situ for safety reasons but had to be done without going on board the container vessels. Secondly, the use of the system was closely related to highly contextualized work activities in a very specialized physical use context, which would be difficult to recreate realistically in vitro. Motivated by these challenges, we decided to explore a series of different opportunities for increasing evaluation realism in controllable and safe environments.

We briefly present the Maritime Communicator case study below and describe how the two evaluation studies were designed and carried out.

3.1. Case A: The Maritime Communicator

The Maritime Communicator system was developed for supporting work activities on board large container vessels (with sizes equivalent to 3½ soccer fields). The operation of such vessels requires workers to be highly mobile and physically distributed. Typically, the number of crew members is low and hence people are assigned to various tasks at different locations on the ship depending on the situation: cruising at sea, departing from the quay, etc. Work activities on large container vessels are typically safety-critical and involve high risks in the case of errors. Especially when manoeuvring inside a harbour when erroneous actions can cause serious material damage and possible injuries on personnel or loss of human life. Thus, systems for supporting these work activities have to be carefully evaluated.

Distributed work activities in the maritime domain

On the basis of ethnographic studies of work activities on a container vessel (Nielsen, 2000) the Maritime Communicator was developed to support the coordination of "letting go the lines" immediately before departing from a harbour. When departing from a harbour the first step is to let go the mooring lines holding the vessel in a fixed position. However, as physical space is restricted and means for precise manoeuvring are limited, all lines cannot simply be released simultaneously.

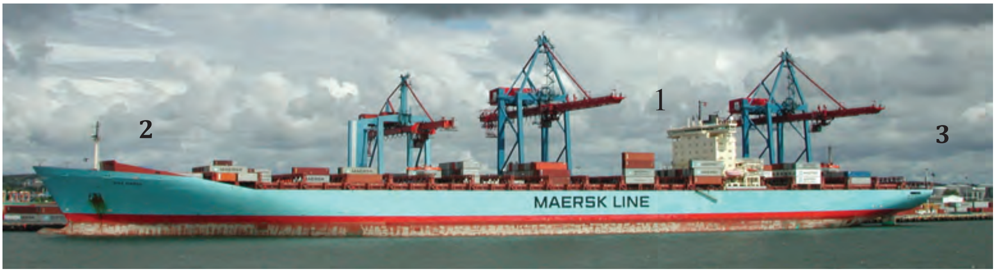


Figure 4. Sine Maersk in Gothenburg container terminal

Due to the huge size of the container vessel and the risk of lines getting sucked in and wrapped around the propeller or thrusters, leaving the vessel without any means of steering, the work tasks involved are distributed among a number of actors located at strategic positions (figure 4). These actors are all highly mobile throughout the whole operation. On the bridge (1), chief officers control the rudder, propeller and thrusters. At fore (2) and aft (3), the first and second officers control the winches for heaving in the lines. Ashore, two teams of assistants lift the lines off the bollards. The challenge of the operation consists of bringing the vessel clear of the quay sideways without running aground in shallow water or colliding with other ships. Because of wind, current, temporal lack of propulsion while lines are in the water, and poor visual view from the bridge, the operation of letting go the lines is not trivial and relies heavily on ongoing communication and careful coordination.

At present this coordination is primarily based on oral communication following a set of formalized procedures. While people on the bridge can see and hear each other, personnel on deck are out of direct visual and audio contact and have to communicate with the captain via walkie-talkies. In order to carry out the operation of departure, the captain needs an overview and total control over the propulsion, direction and mooring of the ship. While information about the rudder, propeller and thrusters are available on dedicated instruments no information about mooring is facilitated. At present this only exists as a mental model in the head of the captain based on his perception of the ongoing communication between bridge and deck. As this mental model is highly sensitive to errors or misunderstandings in the communication, and since disparity between the captain's mental model and the real world may cause wrong decisions, considerable cognitive resources are spent on establishing and maintaining common ground among the cooperating actors (Clark and Schaefer, 1989). Though flexible, radio-based communication suffers from limitations of technology as well as spoken language itself. Sound quality is often poor, utterances are not persistent, communication is time consuming and suffers from language barriers and bottlenecks (multiple parallel tracks). Furthermore, it cannot be automated or integrated with other systems.

The prototype system

Inspired by the potentials of text-based messaging as an asynchronous, flexible, ubiquitous and persistent communication channel requiring low cognitive overhead (see e.g. Churchill and Bly, 1999), it was the thesis of the research team that a text-based communication channel on mobile devices could eliminate or reduce some of limitations observed during the field studies. To investigate this potential further, a prototype of the Maritime Communicator was designed and implemented (Kjeldskov and Stage, 2003). The prototype setup consisted of three iPAQ 3630 connected through an IEEE 802.11b 11Mbit wireless TCP/IP network. One device was intended for the captain on the bridge while the other two were intended for first and second officers on the fore and aft deck respectively. The Maritime Communicator gives the distributed actors on the container vessel access to a mobile text-based communication channel and provides a graphical representation of the ship and its mooring lines.

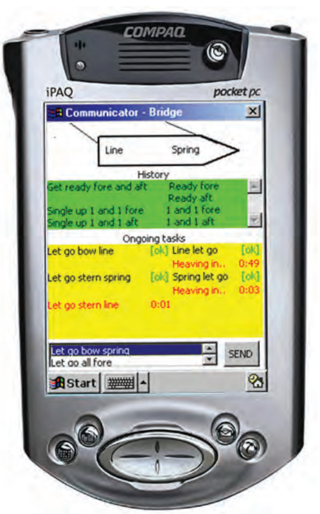


Figure 5. The Maritime Communicator

At the bottom of the screen, unexecuted commands and confirmations are displayed on a list. The order of the list corresponds to the standard sequence of the overall operation and commands appear only when appropriate. By default, the most likely next step of the operation is highlighted. Commands can be browsed and executed (send) with the five-way key on the device. Above this list, the workers can monitor ongoing threads of communication as they unfold during the operation synchronized with the graphical representation of the vessel and mooring lines.

In the following sections, we describe two evaluations of the Maritime Communicator carried out in vitro and in sitro.

3.2. Study A_1: Laboratory Evaluation (in vitro)

In our first evaluation study, we focused on evaluating the usability of the Maritime Communicator *in vitro:* through a "traditional" laboratory-based think-aloud evaluation with prospective users as described by, for example, Rubin (1994). In vitro evaluations are not by definition *non*-realistic just because they do not take place in the intended *situation* of use. In our first study, for example, some realism was provided through 1) the tasks to be solved using the system, 2) the physical separation of the communicating test subjects and 3) a simple cardboard mockup of the vessel, quay and mooring lines.

The first study was conducted in a standard usability laboratory consisting of two separate subject rooms (resembling the bridge and the fore deck respectively) and a control room. From the control room, both subject rooms could be surveyed through one-way mirrors and by means of remotely controlled motorized cameras mounted in the ceiling. Six test subjects participated in the study. They were divided into three teams of two and given the task of letting go the lines before departure of a large vessel coordinating the operation by means of the Maritime Communicator. All test subjects were educated and practically skilled sailors experienced with the operation of large vessels including hands-on experience with the task of letting go the lines. They were recruited from the nearby Skagen Maritime College. The test subjects received a 15-minute joint oral introduction to the specific use context of the prototype application and were presented with a use scenario. This was supported by a number of illustrations on a whiteboard (figure 6). The introduction and use scenario covered the overall operation of letting go the lines, the basic concepts and maritime notions involved, the distribution of work tasks, and present procedures of communication and coordination (as described above). Following this, one person was asked to take the role of captain on the bridge while the other took the role of officer on the fore mooring deck.

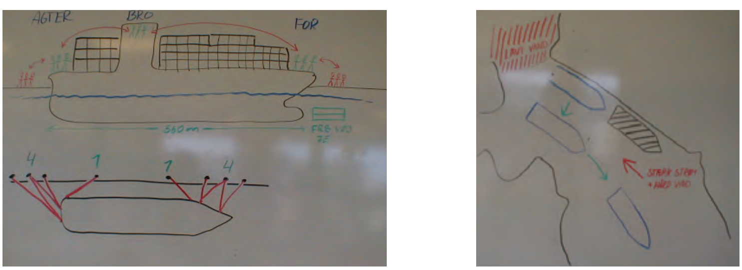


Figure 6. Introduction to use context and a possible use scenario drawn on whiteboard

The test subjects were seated at a desk with the mobile device located in front of them. During the evaluation, the test subjects were asked to think-aloud, explaining their comprehension of and interaction with the prototype. Supporting this, the captains were given a cardboard mock-up of central instruments on the bridge for controlling the thrusters, propellers and the rudder, as well as a model of the ship and mooring lines placed on a schematic drawing of the harbour (figure 7). The purpose of this mock-up was to supply the test subjects with a tool for explaining and illustrating their strategies and actions as the process of departing from the harbour developed over time. An evaluator located in each test room observed the test subjects and frequently asked them about their actions. On a video monitor facing away from the test subjects, the evaluators could see a close up view of the mobile device as well as the activities in the other subject room for the sake of overview. The evaluations lasted approximately 30 minutes and were followed by a 10 minute debriefing interview.

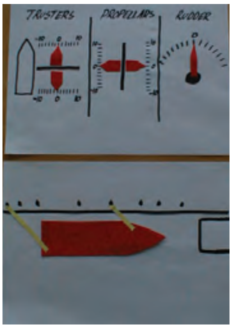


Figure 7. Cardboard mock-up of the bridge, ship and mooring

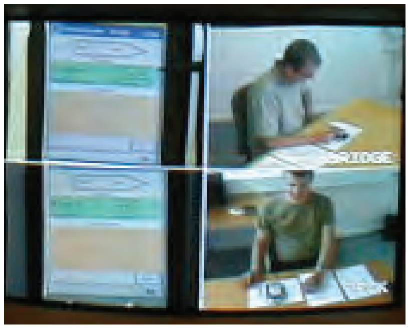


Figure 8. Video recording from evaluation in the usability laboratory

The laboratory setup consisted of two Compaq iPAQs and a PocketPC emulator on a laptop PC connected through a wireless network. The iPAQs displayed the interfaces for the officer on the fore mooring deck and the captain on the bridge respectively. The laptop displayed the interface for the officer on the aft mooring deck and was operated by one of the evaluators using a pre-defined script. Two A4 handouts depicted standard patterns of mooring and explained 10 basic concepts and notions of the maritime context for quick reference if necessary.

Remotely controlled video cameras mounted in the ceiling captured high quality video images of the evaluation sessions. Two cameras captured overall views of the captains and officers and two cameras captured close up views of the mobile devices. In order to ensure good video images of the displays, the test subjects were asked to keep the mobile devices within a delimited area, drawn on a white piece of paper taped to the desk. The four video signals were merged into one composite signal and recorded digitally (figure 8). Audio from the two subject rooms was recorded on separate tracks for later mixing and potential separation during analysis.

3.3. Study A_2: Simulating the Ship (in sitro)

In our second evaluation study, we aimed at evaluating the Maritime Communicator prototype in the hands of real users in a highly realistic but yet controllable and safe environment thus combining strengths and benefits from both in situ and in vitro studies. We define this approach as *in sitro*. Accomplishing this aim, we established a temporary usability laboratory at the simulation division of Svendborg International Maritime Academy (SIMA) and used their state-of-the art ship simulator for creating a realistic (but safe) setup simulating real world phenomena from the intended use context on a high level of fidelity. The ship simulator consisted of two separate rooms: a simulated bridge and a nearby control room. The bridge was fully equipped with controls for thrusters, propellers, rudder etc. as well as instruments such as dobler log, echo sounder, electronic maps, radars, VHF radio, etc. From the control room, simulator operators could see the bridge on a closed circuit video surveillance system. The computer application driving the simulation facilitated a high-fidelity interactive scenario of the operation of any computer-modeled vessel at any modeled physical location. Also weather and dynamic traffic conditions could be included into the scenario. For our specific study, the simulator was set up to imitate the operation of a large vessel in challenging weather and traffic conditions in Felixstowe harbour corresponding to a real world situation observed during our field studies (Nielsen, 2000)

Figure 9. The part of the simulated bridge at Svendborg International Maritime Academy

As in our first study, three captains and three officers, divided into teams of two, participated as test subjects in the study fulfilling their usual roles and were given the overall task of letting go the lines and departing from harbour using the Maritime Communicator for communication between bridge and deck. Again, all participants were educated and practically experienced prospective users fulfilling their usual roles in the use domain – this time recruited from the academy running the simulator facility. Carrying out the operation, the captain had to consider all aspects of manoeuvring the ship on the simulated bridge. This included controlling the rudder, propellers and thrusters as well as communicating with personnel on the ship, harbour traffic control, etc. and taking into consideration the movements of other vessels. The primary task of the first officer on deck (located in the neighbouring simulator control room) was to orally forward commands executed by the captain via the mobile device prototype to the operator of the simulation (impersonating the team of assistants carrying out the actual tasks) and report back to the captain. The operator would then enter the commands into the simulation (making the vessel respond differently to controls on the bridge as it would in the real world), and report to the first officer when the requested operations (such as letting go a line) had been carried out. For simplicity, commands targeted at the second officer on the aft deck were fed directly into the simulation and the simulation operator gave feedback.

Figure 10. Studying use in the ship simulator: the simulator and the video recording

During the evaluation, the captain and officer were asked to think-aloud, explaining their comprehension of and interaction with the prototype. Two evaluators located on bridge and deck respectively observed the test subjects and asked questions for clarification. On a video monitor facing away from the test subject, the evaluator on the deck could see a close up view of the mobile devices as well as an overview of the bridge. The evaluations lasted approximately 40 minutes and were followed by a 10 minute debriefing interview.

As in the traditional laboratory study, the prototype setup consisted of two Compaq iPAQs and a PocketPC emulator on a laptop PC connected through a wireless network. High quality video images were captured of the evaluation sessions. An already installed stationary surveillance camera captured an overall view of the simulated bridge while close-up views of the test subject's interaction with the prototype and other controls on the bridge was captured by the evaluator using a hand-held camera. In the room resembling the fore mooring deck, a camera captured an overall view of the test subject as well as the operators of the simulator. As in the traditional usability laboratory, the test subject acting as the officer on deck was seated at a desk with the mobile device located in front of him. Again, the device had to be kept within a delimited area drawn on a white piece of paper taped to the desk in order to ensure good video images of the display. The four video signals were merged into one composite signal and recorded digitally. Audio from the two rooms was recorded on separate audio tracks.

3.4. Analysis

The data analysis from studies 1 and 2 aimed at identifying, describing and classifying usability problems experienced during use of the Maritime Communicator prototype. The analysis of the data from the two studies was done in a collaborative effort between the two authors, allowing immediate discussions of identified problems, and involved two discrete steps: a *compilation* of findings for each study, and a *comparison* of findings across studies. In order to ensure a rigorous and credible process, we went through the following steps. Firstly, problems experienced by the test subject acting as first officer were identified by examining the video recordings while only listening to the audio track from the fore deck. Secondly, problems experienced by the test subject acting as captain were identified examining the video recordings while only listening to audio from the bridge. Thirdly, all video recordings were examined again while listening to a mix of the audio from both the fore deck and the bridge simultaneously in order to "get the whole picture" – confirming the problems already identified and identifying additional problems. The compilation process resulted in two lists of usability problems ranked as cosmetic, serious, or critical (Molich 2000).

Finally, the two lists of usability problems were merged into one complete list through extended discussion of each identified problem (member checking) among the authors until consensus had been reached. In case of different severity ratings of the same usability issue across techniques, the most severe rating was used in the merged list.

3.5. Findings

We identified a total of 53 different usability problems from the six in vitro and in sitro sessions. Eight problems were critical, 20 were serious, and 25 were cosmetic. The in vitro sessions identified 40 of the 53 problems whereas the in sitro experienced 36 of the 53 problems. Twelve of the problems were unique to the in sitro sessions whereas 17 problems were unique to the intro sessions. Most of the problems were experienced by many subjects. Some of the problems were interaction issues, for example, nearly all test subjects had problems about which elements to interact with on the screen whereas a few relate to the correlation between the representation of the ship on the system and real activities on the ship. As another example, many test subjects could not state the status of commands they had issued.

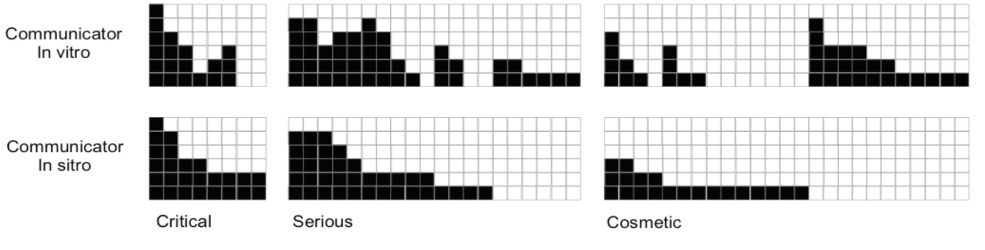


Figure 11. Distribution of identified usability problems

Figure 11 outlines the distribution of the identified 53 usability problems where each column represents one usability problem associated with the number of test subjects experiencing the problem (indicated by black boxes) for both settings. The distribution of problems on severity furthermore reveals that both conditions identified a large proportion of the critical and serious problems. However, the in sitro condition was able to reveal all eight critical problems whereas the in vitro condition only identified six of the eight critical problems.

As mentioned above, 12 usability problems were unique to the in sitro condition which constitute 1/3 of the total identified problems for that condition. These problems primarily concern the representation of the task in the system and lack of flexibility, e.g. more of the domain subjects wanted to specify in more details how they wanted to depart the harbour. However, this was not possible in the system. Furthermore, some of these problems relate to the lack of being able to cancel actions, e.g. one test subject lost complete overview of what was going on since he had to cancel one action. Finally, it should be noticed that both conditions identified several unique cosmetic problems only experienced by one participant.

4. GOING INTO THE FIELD

Fuelled by the challenges encountered and lessons learned from the comparative usability study of the Maritime Communicator, our second comparative usability case (Case B) focused on contrasting the use of a high-fidelity simulation of the use context (in sitro) with going out into the real world (in situ). Again, the motivation for doing so originated from our involvement in a larger research project – this time dealing with the use of computerized information systems in the healthcare domain with particular focus on the use of electronic patient record (EPR) systems at hospitals. As a part of this project, we developed MobileWARD – a mobile counterpart to the stationary EPR system of a large regional hospital supporting nurses' everyday collaborative and highly mobile work activities on a hospital ward. While in this project studying the use of the prototype system in situ was not ruled out, we were still faced with a series of challenges similar to those of the Maritime Communicator study. Firstly, hospital staff raised concerns that using the prototype EPR system could influence them taking care of their patients and potentially impact on their well-being. Secondly, significant ethical considerations were raised about involving real hospitalized patients (and potentially sensitive patient data) in a research study at all. Thirdly, allowing researchers access to the hospital ward during hectic work hours (which was when the system was intended to be used) was not popular among the nurses who were already very busy and under considerable work pressure due to understaffing. Motivated by these challenges, our experience with the use of simulated use contexts in laboratory settings, and the ongoing discussion about laboratory versus field evaluations of mobile usability, we decided to carry out a comparison between the results produced when simulating real world phenomena in a controlled environment (in sitro) and when studying usability *purely in situ;* observing real users doing real work in the real use context – with *no* researcher control or interference (which the hospital staff eventually agreed to).

Below we briefly present the MobileWARD case study and describe how the two evaluation studies were designed and carried out.

4.1. Case B: MobileWARD

MobileWARD was developed for supporting collaborative mobile work activities at a hospital ward through wireless access to electronic patient data on handheld computer terminals. Supporting work activities in healthcare is highly complex and challenging, and within the last 20 years considerable amounts of effort has been devoted to the development of computerized systems for this domain such as electronic patient records. An electronic patient record is a collection of information about a single patient's history in a hospital, which the hospital-staff use to diagnose diseases and to document and coordinate treatment. The primary motivation for this effort is that unlike paper-based patient records, electronic patient records will be accessible to all relevant persons independent of time and location. The design of electronic patient records is a huge challenge for the HCI community, raising a wide range of still unanswered questions related to issues such as screen layout, interaction design, and integration into work processes. However, while much research has studied the use of traditional paper-based patient records, suggesting electronic counterparts, little research has been published on studies inquiring into the use of EPR systems already out there.

Using electronic patient records in healthcare

Based on evaluations of EPR systems and field studies of mobile work activities in hospitals, we identified three key issues concerning the use of electronic patient records: 1) mobility, 2) complexity, and 3) relation to work activities.

Mobility. Most nurses expressed concerns about having to be mobile while at the same time working with the EPR system, which was stationary. Meeting this challenge, the use of laptop computers rather than desktop workstations had been suggested and discussed at the hospital. However, most of the nurses stated that they would find it impossible or unfeasible to carry a laptop computer around the ward every time they were to conduct work tasks away from their office. One problem was the size of the laptop, as they would also have to carry other instruments.

Complexity. Another overall concern reported and observed was the complexity and fragmentation of information. Most nurses found it difficult to locate the necessary patient information in the EPR system to carry out their work. This sometimes led to delays and incomplete task completions. Hence, the nurses would be unsure whether they had found the right information and whether they had succeeded in finding all relevant information.

Work Relation. Most nurses experienced problems with the EPR system due to difficulties with relating the data and structure of information in the system to real work activities and people. The problem was that they would typically use different kinds of information in context to determine how to solve a problem, for example, the visible condition of a patient. Another concern related to the fact that the system only partially reflected their current work tasks, making it difficult to the test subjects to find or store information.

The prototype system

Inspired by the potentials of context-aware mobile computing, it was our thesis that providing nurses with mobile access to electronic patient records automatically adapting to their current work situation could help overcoming some of the observed limitations of the stationary EPR system. In order to investigate this potential, a functional context-aware prototype system was designed and implemented to support the nurses' morning procedure (Skov and Høegh 2005), which 1) supported the highly mobile work activities of nurses by being handheld, 2) reduced complexity by adapting to its context and 3) eliminated double registering of information (first written down on paper and then entered into the PC later) by being integrated with the existing patient record. While facilitating access to patient information at the 'point of care' is not a new idea (Arshad et al. 2003, Kaplan and Fitzpatrick 1997, Urban and Kunath 2002), adapting information and functionality in a mobile EPR system to its context is a novel approach to improving the usability of such systems, which has not yet been investigated thoroughly.

MobileWARD runs on Microsoft PocketPC based Compaq iPAQ 3630 (or equivalents) connected to an IEEE 802.11b wireless TCP/IP network. In the intended setup all nurses on duty have their own personal device. The MobileWARD system is context-aware in the sense that the system presents information and functionality adapted to the location of the nurse, the time of the day, pending tasks, and nearby patients etc. Based on the classification by Barkhuus and Dey (2003), MobileWARD is an *active* context-aware system as it automatically presents information and adapts to its context. The system works as described below.

Before visiting assigned patients for morning procedure, nurses often want to get an overview of the specific information about each patient. As this typically takes place at the nurse's office or in the corridor, the system by default displays the overall patient list (figure 12a). Patients assigned for morning procedure are shown with a white background and the names of patients assigned to the nurse using the system are boldfaced (e.g. "Julie Madsen"). For each patient, the patient list provides information about previous tasks, upcoming tasks and upcoming operations. The indicators TP (temperature), BT (blood pressure) and P (pulse) show the measurements that the nurse has to perform. "O" indicates an upcoming operation (within 24 hours), which usually requires that the patient should fast and be prepared for operation. At the top of the screen, the nurse can see their current physical location (e.g. "in the corridor").

Figure 12. MobileWARD: Three different screens from the context-aware mobile EPR system

The window in figure 12b displays information related to one patient including name and personal identification number of the patient, previous sets of measured temperatures, blood pressures, and pulses as well as notes regarding the treatment of the patient. To enter new data into the system, the nurse must scan the barcode identification tag on the patient's wristband using the "scan" function in the bottom of the screen. When the nurse enters a ward, the system automatically displays information and functionality relevant to this location (figure 12c). Information about the patients on the current ward is presented, resembling the information available on the patient list displayed in the corridor, with the addition of a graphical representation of the physical location of the patient's respective beds. Data about patients is available by clicking on the names.

In the evaluated prototype of MobileWARD, some of the contextual sensing functionality was simulated by means of a "context control centre" application. The control centre runs on a separate iPAQ connected to the wireless network. Through this application, an operator can trigger "context events" in MobileWARD, e.g. instructing the system that the user has entered a specific room.

4.2. Study B_1: Simulating the Hospital Ward (in sitro)

The aim of our third study was to evaluate MobileWARD in a controlled simulated environment (similar to the ship simulator) where we could closely monitor the use of the system while at the same time simulate key real world phenomena such as mobility between rooms, work tasks and hospitalized patients. In order to achieve this, we modified and refurnished our usability laboratory to resemble a part of the physical space of a hospital department (figure 13). This included the use of two separate evaluation rooms connected by a hallway. Each of the evaluation rooms were furnished with beds and tables similar to real hospital wards. From a central control room, the evaluation rooms and the hallway could be observed through one-way mirrors and via remotely controlled motorized cameras mounted in the ceiling.

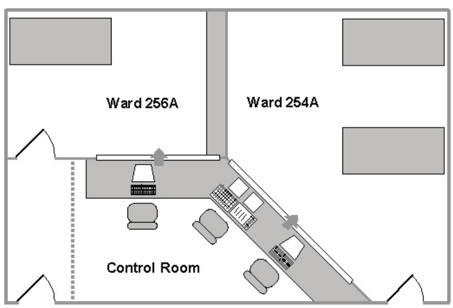


Figure 13. Physical layout of the usability laboratory simulating the hospital ward

Six test subjects (four females and two males) aged between 28 and 55 years participated in the study. All test subjects were trained nurses employed at a large regional hospital and had between 2 and 36 years of professional experience. Thus as in the maritime communicator evaluations, the test subjects were real prospective users fulfilling their usual roles. All test subjects were mobile phone users but only one had experience with the use of handheld computers. Everyone was also familiar with stationary electronic patient record systems and described themselves as experienced or semi-experienced IT users. All test subjects were given a series of tasks to solve while using the system. The tasks were derived from a field study at a hospital ward and were developed in collaboration with hospital staff. The tasks covered the duties involved in conducting standard morning work routines involving primarily 1) checking up on a number of assigned patients based on information in the system from the previous watch, 2) collecting and reporting scheduled measurements such as temperature, blood pressure, and pulse, and 3) reporting anything important for the ongoing treatment of the patients should be taken into consideration on the next shift.

Before the evaluation sessions, the test subjects were given a brief instruction to the system. This included the room-sensing functionality and the procedure for scanning patients' bar-code tags. The test subjects were also instructed on how to operate the available instruments for measuring temperature, blood pressure and pulse. The evaluation sessions were structured by the task assignments, which required the test subjects to interact with all three patients in the two simulated hospital wards, and move between the two rooms through the connecting hallway a number of times. The nurses were encouraged to think aloud throughout the evaluation explaining their comprehension of and interaction with the system. The evaluations lasted between 20 and 40 minutes and were followed by filling out a questionnaire.

Figure 14.
Wireless camera mounted on PDA

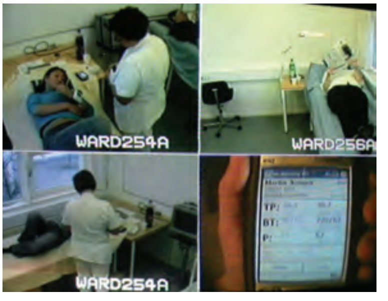


Figure 15.
Video images from simulated ward and PDA

Each evaluation session involved six people. One nurse used the system for carrying out the assigned tasks. Three students acted as hospitalized patients. One researcher acted as test monitor and asked questions for clarification. A second researcher operated the context-control centre and the video equipment. For data collection, high quality audio and video was recorded digitally from the ceiling-mounted cameras and a tiny wireless camera clipped on to the mobile device providing us with a close-up view of the screen and user-interaction (figures 14 and 15).

4.3. Study B_2: Studying Use at the Hospital (in situ)

The fourth evaluation study took place at a large regional hospital in Denmark. The aim of this evaluation was to study the usability of MobileWARD in situ for supporting *real work* activities at a hospital involving *real nurses*, *real patients,* and *real patient data*. In order to achieve this, we adopted an observational approach combined with questions for clarification while the nurses were not directly engaged in conducting their work.

The in situ evaluation was carried out at the Medical Department at the Hospital of Frederikshavn (figure 16). This included the physical area of seven hospital wards, an office with reception, a rinse room and a break-out area connected by a central hallway and involved nurses at work and patients committed to the hospital.

Six test subjects (all females) aged between 25 and 55 years participated in the in situ evaluation. All test subjects were trained nurses employed at the Hospital of Frederikshavn and had between 1 and 9 years of professional experience. They were all mobile phone users but novices with the use of handheld computers. All test subjects were frequent users of a stationary electronic patient record system and described themselves as experienced or semi-experienced users of IT.

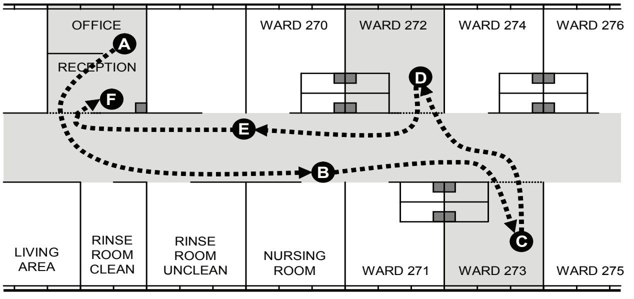


Figure 16. Physical layout of the hospital wards

The in situ evaluation did not involve any researcher control and interference in form of task assignments but was structured exclusively by the work activities of the nurses in relation to conducting their standard morning work routines. As in the task assignments of the laboratory evaluation this involved 1) checking up on a number of assigned patients in different wards and moving between different rooms through the connecting hallway a number of times, 2) collecting and reporting scheduled measurements, and 3) reporting anything important for the ongoing treatment of the patients. As in the laboratory evaluation, the test subjects were given a brief instruction to the MobileWARD system, including the room-sensing functionality and the procedure for scanning a patient's bar-code tag. The evaluations lasted 15-20 minutes on average and were followed by filling out a brief questionnaire. In order to be able to include a suitable number of nurses, the study took place over two days.

Each evaluation session involved six people. One nurse used the system for carrying out her work activities. One researcher observed the work and use of the mobile system from a distance and asked questions for clarification while in the hallway. A second researcher operated the context-control centre application and the portable audio/video equipment. In addition, each evaluation session involved three hospitalized patients in their beds. Due to the real-life nature of the study, each session involved different patients and the nurses did not think aloud.

Figure 17. Observer (left) carrying and operating portable audio/video equipment (right) for capturing high-quality data in the field.

Due to the challenges of capturing high-quality data during usability evaluations in natural settings (e.g. Pascoe et al. 2000, Rantanen 2002, Brewster 2002, Esbjörnsson et al. 2003, Kjeldskov and Stage 2004), we designed and purpose-build a portable configuration of audio and video equipment to be carried by the test subject and an observer, allowing a physical distance of up to 10 meters between the two. The configuration consisted of a tiny wireless camera (also used in the laboratory evaluation described above) clipped-on to the mobile device (figure 14) and a clip-on microphone worn by the test subject. Audio and video were transmitted by wireless to recording equipment carried by the observer (figure 17). In the test monitor's bag, the video signal from the clip-on camera was merged with the video signal from a handheld camcorder (Picture-in-Picture) and recorded digitally. This setup allowed us to record a high-quality close-up view of the screen and user-interaction as well as an overall view of user and context. During the evaluation, the observer viewed the user's interaction with the mobile device on a small LCD screen and monitored the sound through headphones. For ethical reasons, we were not permitted to film the hospitalized patients.

4.4. Analysis

The data from studies 3 and 4 amounted to approximately six hours of video recordings depicting the 12 test subjects' use of the MobileWARD system. On the basis of this data, the analysis aimed at identifying, describing and classifying two lists of usability problems experienced by the users in the two studies. All sessions were analyzed in random order by two teams of two trained usability researchers holding Ph.D. or Master Degrees in Human-Computer Interaction. As in the analysis of the Maritime Communicator studies, the data analysis involved a *compilation* of findings for each study and a *comparison* of findings across studies. Compiling the usability problems, each team first analyzed the videos recordings in a collaborative effort allowing immediate discussions of identified problems and their severity (cosmetic, serious or critical). As a guideline for the collaborative analysis, each identified usability problem would be discussed until consensus had been reached. The two teams of researchers produced two lists of usability problems indicating for each problem if it was experiences in sitro, in situ, or both. Subsequently, these two lists were merged into one complete list through a process of comparing each problem. Again, this was done in a collaborative effort, discussing each problem and its severity until consensus had been reached. In case of different severity ratings across techniques, the most severe rating was used.

4.5. Findings

We identified a total of 37 different usability problems from the 12 in sitro and in situ sessions. Eight problems were assessed to be critical, 19 problems were assessed to be serious, and ten problems were assessed to be cosmetic. Our case showed that the in sitro condition found more usability problems than the in situ condition. The six in sitro subjects experienced 36 of the 37 usability problems whereas the six in situ subjects experienced 23 of the 37 usability problems. Fourteen usability problems (1 critical, 9 serious, 4 cosmetic) were unique to the in vitro condition, whereas one serious usability problem was unique to the in situ condition.

Regarding the critical problems, the in sitro setting identified all eight critical problems and the in situ setting identified seven critical problems. Considering the serious problems, we find that the in sitro identified eight extra problems compared to the in situ evaluation.

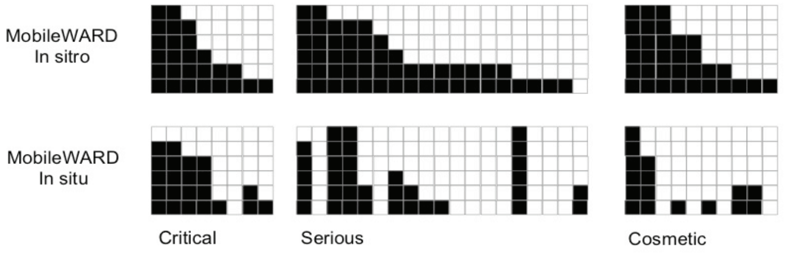


Figure 18. Distribution of identified usability problems

Figure 18 outlines the distribution of the identified 37 usability problems where each column represents one usability problem associated the number of test subjects experiencing the problem (indicated by black boxes) for both settings. Seven usability problems (two critical, two serious, three cosmetic) were experienced by all six subjects in the in sitro setting, whereas three usability problems (two serious, one cosmetic) were experienced by all six subjects in the in situ setting and one usability problem (cosmetic) was experienced by all 12 subjects.

Looking across the distribution of the usability problems, we find that while the critical problems have a roughly similar distribution, the serious and cosmetic problems have rather dissimilar distributions where some problems were identified by all or nearly all subjects in one setting, but only identified by a few or none in the other setting. For example, all subjects were informed to use either their fingers or the attached pen for device interaction, but only the in sitro subjects chose to use the pen and most of them experienced difficulties in placing the pen between tasks.

One problem was identified by only in situ participants. This problem concerned the validity of recorded data in the system. Two of the nurses only reported and recorded accurate data on patients' heart rate, temperature, and blood pressure. Occasionally the nurses would first measure these values and then perform other work activities. Later when recording the values into the system they had forgotten the exact measures and would then repeat the measures. This was not the case in the in sitro condition where the participants stressed the artificial condition and the lack of need for being accurate. The in-use and in-situ situation made the participants stress accuracy and validity.

5. DISCUSSION

Our motivation behind this paper was to contribute to the body of research on development of appropriate methods and techniques for evaluating mobile system use by systematically exploring the differences and similarities between studying such technologies "in use, in situ", in the laboratory, and in controlled high-fidelity simulations of the real world. Thus, our aim was to contribute to the general knowledge of usability evaluation and testing on how to set up an environment for testing (see e.g. Nielsen, 1993, Rubin 1994, Dumas and Reddish 1999) and more specifically to the extensive body of knowledge on comparative usability evaluations studies in general (see e.g. Bailey et al. 1992, Karat et al. 1992, Henderson et al. 1995, Molich et al. 1998) and especially for mobile systems evaluations (Kjeldskov and Stage 2004). Our comparative usability evaluation study explored three conditions for mobile usability evaluations.

We proposed an experimental condition for mobile usability evaluations called in sitro from the in situ and in vitro conditions and stressed the combinational nature of the two conditions and of simulation. The idea of in sitro is to simulate partial or full fidelity of a real intended use situation. In sitro is closely related to and takes inspiration from research on simulation (see e.g. Allen et al. 1985, Hays and Singer 1985, Mayor and Volanth 1985, Sanders 1991). Our results indicated that the in sitro condition was able to simulate significant elements of the intended future in use situation and thereby increase the level of realism in the evaluation while at the same time maintaining a high level of control.

Both investigated cases involved rather complex and dynamic use situations: 1) captain and crew members coordinating activities on a large container vessel during departure from harbour, and 2) nurses doing morning procedures at a hospital ward. While the former setup involved a rather expensive and sophisticated simulator used for training of future container vessel captains, the latter setup was rather simple taking place in our traditional usability laboratory including two wards, a number of beds, and some patients (student actors). While the container vessel setup resembles full simulations as described in Sanders (1981), the hospital setup was rather light weight and resembles some of the ideas for training using desktop games illustrated in Alexander et al. (2005).

Both conditions could potentially be difficult to simulate in a laboratory, however our studies confirmed that for usability evaluations this is possible. Our first case exhibited similar distribution of identified usability problems for the critical problems and partially similar distribution for the serious problems. In fact, the in sitro condition identified additional critical problems in the interface not found in the in vitro condition. Both these problems could be traced to the increased level of realism in the use situation. At the same time, we experienced no significant problems with experimental control when we introduced use situation elements, e.g. other artefacts, additional participants, and the participants still being able to think-aloud during the evaluation. Thus, it seemed possible to simulate parts of the environment in a controlled laboratory. Our second case confirmed this observation as the in sitro condition was able to identify nearly all the problems identified in situ. Several problems were further identified only in the in sitro condition, but most of these were due to lack of control in the in situ condition.

The real world condition (in situ) integrated levels of realism that were not achieved in the in sitro condition which primarily concerned the validity of the recorded data in the system. Thus, in our case it seemed impossible to recreate all levels of fidelity in the simulations. Some in situ participants would only report and record accurate (and thus realistic) data on patients' heart rate, temperature, or blood pressure. This observation was seen when some nurses first measured values and then performed other work activities. The nurses would record the values in the system later but had by that time forgotten the exact values and would therefore repeat the measures. This was not the case in the in sitro condition where several participants stressed the artificial condition and the lack of need for being accurate. As a consequence, they occasionally entered data into the system that was incorrect or simply wrong. We argue that this aspect has to do with psychological fidelity as illustrated by Alexander et al. (2005). They stress that many real-world environments evoke levels of stress and arousal that may not be directly replicable in virtual environments. This seemed to be confirmed in our study. Other recent studies have investigated the effects of adding contextually-relevant stress to training paradigms (Driskell et al. 2001) but this was not attempted in our approach. While both conditions lacked psychological fidelity, the container vessel setup did integrate the same kind of problem probably due to an established seriousness when using the simulator. Thus, even though stress could have been low as the participants were not manoeuvring a real container vessel, all of the participants acting as captain took the assignment very serious and never entered false data deliberately into the system.

Our in situ sessions with nurses confirmed the challenges of decreased control as none of the subjects used the note taking facility in our electronic patient record prototype. In the in situ study, we deliberately chose to give them no assignments as this would possibly increase level of control (and thereby perhaps decrease level of realism). As a direct consequence, we identified no usability problems in the note taking facility of the prototype from the in situ condition. Thus, control was definitely a challenge in our study.

The in vitro condition provided only a few extra findings compared the in sitro condition. The in vitro sessions in our first case were easier to plan and conduct. Utilizing our own usability laboratory, we could more easily set up and conduct the evaluations. Compared the in sitro sessions, this required less resources and less planning. Furthermore, the in vitro sessions identified a number of serious and cosmetic problems not identified in sitro. It is difficult to assess the value of these extra problems, but some of the cosmetic problems would probably be irrelevant when looking at the system in use, in situ. On the other hand, the in vitro sessions failed to identify two critical problems identified in sitro; both problems had to do with the simulated context of the evaluation. As an example, more of the in sitro participants needed to cancel issued commands, but this was not possible in the tested system. This turned out to be a critical problem in the evaluation since the captain had to apply different means of communication in order to cancel the command. Such issues never came up in the in vitro condition. But some of these issues from the in vitro condition can probably be explained from the low level of physical fidelity as means for extra communication, e.g. radio communication, were not present during the in vitro tests.

The general validity of the results of our study is limited in a number ways. First, the number of test subjects applied in each study implies that we can primarily explore qualitative issues of changing the condition for usability evaluations. Our study can hopefully set out avenues for further research. Also, general competences of the test subjects varied between the setups, which could have influenced some of the results.

This was especially true for the IT skills of some of the subjects, as they had never used a hand-held device before. While there may have been high variability within the groups of subjects, we tried to minimize variability between groups in an attempt to decrease the effects on our study. Thus, in all groups we had subjects that were not very familiar with mobile and hand-held devices. Secondly, conducting an in situ experiment is controversial in itself as it can be discussed whether it is possible to observe as closely as we did and still call the condition in situ. In our study, we tried to get as close as possible to a real use situation for the nurses on morning procedure. Our presence could possibly have influenced some of their behavior and interaction with the mobile devices and this would have influenced the collected data. This is difficult to avoid in the adapted setup, but ethnographic studies involving interviews and observations could address this issue even further. Thirdly, even though our cases are rather different, additional cases could verify the applicability of the in sitro condition. Again, we hope that this would inspire additional study of practical applicability of the condition.

With the increased levels of complexity of mobile technologies and use situations for these mobile technologies, we will probably need additional and innovative ways of evaluating such technologies in the future. Additionally, practitioners will eventually start to request methods, heuristics, and guidelines for testing the usability of these mobile technologies. Our research in this paper is an attempt to add to this body of knowledge. Possible avenues for further research on this topic could be determining what kind of fidelity is needed when evaluating different kinds of mobile technologies. Our study showed that high fidelity in controlled experiments increased the number and types of identified usability problems.

REFERENCES

Abowd, G. and Mynatt, E. (2000) Charting past, present and future research in ubiquitous computing. ACM Transactions on Computer-Human Interaction, Vol. 7(1), pp. 29-58

Alexander, A. L., Brunyé, T., Sidman, J., and Weil, S. A. (2005) From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-in and Their Effects on Transfer in PC-Based Simulations and Games. DARWARS Training Impact Group, pp. 1-14

Allen, J. A., Hays, R. T., & Buffordi, L. C. (1986). Maintenance training, simulator fidelity, and individual differences in transfer of training. Human Factors, 28, 497-509.

Amaldi, P., Satinder G.P., Fields, B. and Wong, W. (2005) In-Use, In-Situ: Extending Field Research Methods Workshop. Interaction Design Centre, Middlesex University, and the BCS HCI Education & Practice SubGroup, BCS HCI Group.

Andersen, P.B. (2000) Communication and work on maritime bridges, CHMI Research Report CHMI-1-2000

Arshad, U., Mascolo, C. And Mellor, M. (2003) Exploiting Mobile Computing in Health-care. In Demo Session of the 3rd International Workshop on Smart Appliances, ICDCS03

Bailey, R. W., Allan, R. W., and Raiello, P., 1992, Usability testing vs. heuristic evaluation: a head-to-head comparison. Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting, HFES, pp. 409-413

Barkhuus, L. and Dey, A. (2003) Is Context-Aware Computing Taking Control away from the User? Three Levels of Interactivity Examined. In Proceedings of the UbiComp2003 conference, LNCS 2864, pp. 149 – 156

Bohnenberger, T., Jameson, A., Krüger, A., and Butz, A. (2002) Location-Aware Shopping Assistance: Evaluation of a Decision-Theoretic Approach. In Proceedings of Mobile HCI 2002, Pisa, Italy, Springer-Verlag, LNCS

Brewster, S. (2002) Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing, 6, 188-205

Chincholle, D., Goldstein, M., Nyberg, M. and Erikson, M., 2002, Lost or found? a usability evaluation of a mobile navigation and location-based service. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 211-224

Churchill, E. F. and Bly, S. (1999) It's all in the words: Supporting work activities with lightweight tools. In Proceedings of ACM Siggroup'99, pp. 40 - 49

Clark, H.H. and Schaefer, E. F. (1989) Contributing to discourse. Cognitive Science, vol. 13, pp. 259 – 294

Driskell, J. E., Johnston, J. H., and Salas, E. (2001) Does stress training generalize to novel settings? Human Factors, Vol. 43(1), pp. 99-110

Dumas, J. S., and Redish, J. C., 1999, A Practical Guide to Usability Testing (Exeter: Intellect).

Esbjörnsson, M., Juhlin, O. and Östergren, M. (2003) Motorcyclists Using Hocman – Field Trials on Mobile Interaction. In Proceedings of the 5th International Mobile HCI 2003 conference, Udine, Italy, Springer-Verlag, LNCS

Graham R., Carter C. (1999) Comparison of Speech Input and Manual Control of In-Car Devices while on-the-move. In Proceedings of the Second Workshop on Human Computer Interaction with Mobile Devices, Mobile HCI 1999, Edinburgh, Scotland.

Harman, H. H. (1961) Simulation: a survey, Report SP-260, System Development Corporation, Santa Monica, CA

Hays, R. T., & Singer, M. J. (1989). Simulation fidelity in training system design. New York: Springer

Henderson, R., Podd, J., Smith, M., and Varela-Alvarez, H., 1995, An examination of four user-based software evaluation methods. Interacting with Computers, 7(4), 412-432

Johnson P. (1998) Usability and Mobility; Interactions on the move. In Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices, Glasgow, Scotland, GIST Technical Report G98-1

Kaplan, S. M. and Fitzpatrick, G. (1997) Designing Support for Remote Intensive-Care Telehealth using the Locales Framework. In Proceedings of DIS'97, ACM

Karat, C. M., Campbell, R., and Fiegel, T., 1992, Comparison of empirical testing and walkthrough methods in user interface evaluation. Proceedings of CHI'92 (New York: ACM) pp. 397-404

Kjeldskov J. and Graham C. (2003) A Review of Mobile HCI Research Methods. In Proceedings of the 5th International Mobile HCI 2003 conference, Udine, Italy, Springer-Verlag, LNCS

Kjeldskov, J. and Stage, J. (2004) New Techniques for Usability Evaluation of Mobile Systems. International Journal of Human-Computer Studies (IJHCS), 60, 599-620, Elsevier

Kjeldskov, J. and Skov, M. B. (2003) Creating a Realistic Laboratory Setting: A Comparative Study of Three Think-Aloud Usability Evaluations of a Mobile System. In Proceedings of Interact 2003, Zürich, Switzerland, IOS Press, pp. 663-670

Lai, J., Cheng, K., Green, P., Tsimhoni, O. (2001) On the Road and on the Web? Comprehension of synthetic speech while driving. In Proceedings of CHI'2001, ACM, New York, pp. 206-212

Luff, P. and Heath, C. (1998) Mobility in Collaboration. In Proceedings of CSCW'98. ACM, pp. 305 - 314

Mayer, J. D. & Volanth, A. J. (1985). Cognitive involvement in the mood response system. Motivation & Emotion, Vol. 9, pp. 261-275.

Molich, R. (2000) Usable Web Design (in Danish). Ingeniøren | bøger

Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D., and Kirakowski, J., 1998, Comparative evaluation of usability tests. Proceedings of the Usability Professionals Association Conference, pp. 189-200

Nielsen, J., 1993, Usability Engineering. (Morgan Kaufmann).

Nielsen, M. (2000) Letting go the lines: Departure from the Felixstowe harbour. CHMI Research report CHMI-4-2000

Pascoe, J., Ryan, N., Morse, D. (2000) Using While Moving: HCI Issues in Fieldwork Environments. Transactions on Computer-Human Interaction, Vol. 7(3), pp. 417-437.

Petrie, H., Johnson, V., Furner, S., and Strothotte, T. (1998) Design lifecycles and wearable computers for users with disabilities. Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices.

Pospischil, G., Umlauft, M. and Michlmayr, E. (2002) Designing LoL@, a mobile tourist guide for UMTS. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 140-154

Pirhonen A., Brewster S.A., Holguin C. (2002) Gestural and audio metaphors as a means of control for mobile devices. In Proceedings of CHI'2002, ACM, New York.

Rantanen, J., Impio, J., Karinsalo, T., Reho, A., Tasanen, M., Vanhala, J. (2002) Smart Cloth-ing Prototype for the Artic Environment. Personal and Ubiquitous Computing, 6, 3-16.

Rubin, J. (1994) Handbook of Usability Testing. Wiley

Salvucci D. D. (2001) Predicting the Effects of In-Car Interfaces on Driver Behaviour using a Cognitive Architecture. In Proceedings of CHI'2001, New York, pp 120-127

Sanders, A. F. (1991). Simulation as a tool in the measurement of human performance. Ergonomics, 34(8), 995-1025

Sawhney, N., and Schmandt, C. (2000) Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. Transactions on Computer-Human Interaction, 7(3), 353-383.

Skov, M. B. and Høegh, R. T. (2005) Supporting Information Access in a Hospital Ward by a Context-Aware Mobile Electronic Patient Record. Personal and Ubiquitous Computing (forthcoming)

Urban, M. and Kunath, M (2002) Design, Deployment and Evaluation of a Clinical Information System which uses Mobile Computers and Workstations. In Proceedings of the 2nd conference on Mobiles Computing in der Medizin, Heidelberg

Vetere, F., Howard, S., Pedell, S., and Balbo, S., 2003, Walking through mobile use: novel heuristics and their application. Proceedings of OzCHI 2003 (Brisbane, Australia: CHISIG)

Vreuls, D. and Obermayer, R. W. (1985) Human system performance measurement in training simulations. Human Factors, Vol. 27, pp. 214-250

Wingender, E. (1998) ISB: Just Another Journal? In Silico Biology. An International Journal on Computational Molecular Biology, Vol. 1

Zhao, J., Stevens, R. D., Wroe, C. J., Greenwood, M., and Goble, C. A. (2004) The Origin and History of in Silico Experiments. In Proceedings of the UK e-Science All Hands Meeting, Nottingham UK

Chapter 13

Bringing the system in to the field

Jesper Kjeldskov, Connor Graham, Sonja Pedell,
Frank Vetere, Steve Howard, Sandrine Balbo and Jessica Davies

Abstract. When designing a usability evaluation, choices must be made regarding methods and techniques for data collection and analysis. Mobile guides raise new concerns and challenges to established usability evaluation approaches. Not only are they typically closely related to objects and activities in the user's immediate surroundings, they are often used while the user is ambulating. This paper presents results from an extensive, multi-method evaluation of a mobile guide designed to support the use of public transport in Melbourne, Australia. In evaluating the guide, we applied four different techniques; field-evaluation, laboratory evaluation, heuristic walkthrough and rapid reflection. This paper describes these four approaches and their respective outcomes, and discusses their relative strengths and weaknesses for evaluating the usability of mobile guides.

1. INTRODUCTION

Mobile guides constitute a special class of mobile computer systems. Usually mobile guides are closely related to the user's physical location and objects in the user's immediate surroundings (e.g. Cheverst et al. 2000, Chincholle et al. 2002, Schmidt-Belz et al. 2002, Reid 2002, Umlauft et al. 2003). Also, they are often used while the user is ambulating, moving from one physical location to another. These properties make the design and evaluation of mobile guides challenging for human-computer interaction researchers and practitioners.

The design of mobile guides has received considerable attention over the last decade (see e.g. Abowd et al. 1996, Cheverst et al. 2000, Cheverst et al. 2002, Pospischil et al. 2002, Fithian et al. 2003). When authors consider the design of mobile guides, they

Originally published as Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S. and Davies, J. (2005) Evaluating the Usability of a Mobile Guide: The influence of Location, Participants and Resources. *Behaviour and Information Technology,* 24(1), 51-65.

also frequently report the results of evaluations. The reported usability evaluations involve the use of a wide range of methods and techniques borrowed from usability research into 'desk bound' computers and their use, then adapted to fit the special needs, opportunities and limitations of mobile guides. This includes, for example, formal and informal product presentations combined with questionnaires, expert evaluations (Andrade et al. 2002, Po et al. 2004), controlled laboratory experiments (Bohnenberger et al. 2002, Chincholle et al. 2002, Iacucci et al. 2004) and a variety of use studies in realistic field settings including direct observation of use (Cheverst et al. 2002, Schmidt-Belz and Poslad 2003, Laakso et al. 2003), indirect observation of use (Bornträger et al. 2003), field questionnaires (Rocchi et al. 2003), and longitudinal use studies combined with interviews (Kolari and Virtanen 2003, Iacucci et al. 2004). These evaluations all provide valuable insight into usability and usefulness and typically inform design refinements and/or inspire new design concepts. Such research will, one hopes, result in the development of more useful and usable mobile guides.

However, even though evaluations of mobile guides are prevalent, little research has been published on the particular challenges, to usability evaluation, posed by mobile guides; how should we evaluate mobile guides, what methodological challenges do we face, what are the pros and cons of different usability evaluation approaches? Exceptions include, for example, Bornträger and Cheverst (2003) who consider social and technical problems encountered during field evaluations of mobile guide systems, and Kray and Baus (2003) who review and compare nine mobile guide systems and touch upon the methods and techniques that were used in their evaluation. Examining the general literature on mobile HCI does not provide much additional support, with only a few authors considering different usability evaluation methods and techniques for mobile computer systems (see e.g. Brewster 2002, Pirhonen et al. 2002, Kjeldskov and Skov 2003, Kjeldskov and Stage 2004). As a result of our reluctance to 'evaluate evaluation', that is to understand how the utility of the techniques in our usability toolkit respond to the challenge of mobile guide evaluation, no agreed-upon set of usability evaluation methods and data collection techniques exist for mobile guides and little knowledge exists as to when and why one should choose one technique over another. Consequently, researchers and practitioners are provided with little support in making informed decisions about which methods and techniques to select and combine for mobile guide evaluation.

In this paper we report the evaluation of a mobile guide, following four different approaches: field-evaluation, laboratory evaluation, heuristic walkthrough and rapid reflection. The paper describes these four approaches, presents their respective outcomes and discusses their relative strengths and weaknesses from the perspective of the challenges of mobile guide evaluation. In the next section we present and discuss related research on evaluating the usability of mobile computer systems emphasising the special challenges related to the evaluation of mobile guides. Then we briefly describe a project in which we designed and implemented a mobile guide and evaluated it through four independent usability studies. Each of these usability studies are described in detail, followed by a comparison and discussion of the findings. Finally, we conclude with a number of recommendations for usability evaluation of mobile guides.

2. CHOOSING APPROPRIATE EVALUATION TECHNIQUES

Usability evaluation has proven to be an invaluable tool for ensuring the quality of computerised systems. Usability evaluation of stationary computer systems is an established discipline within human-computer interaction with widely acknowledged techniques and methods (e.g. Nielsen 1993, Rubin 1994, Dumas and Reddish 1999). This is complemented by a growing number of attempts to 'evaluate evaluation', empirical evaluations of the relative strengths and weaknesses of the different approaches and techniques, under different circumstances (e.g. Bailey et al. 1992, Karat et al. 1992, Henderson et al. 1995, Molich et al. 1998). So far, this kind of research is only beginning to emerge in relation to the evaluation of mobile computer systems.

Mobile guides take many of the well known methodological challenges of evaluating the usability of both stationary and mobile computer systems to an extreme. Users of mobile guides are ambulatory, typically highly mobile during their interaction with the system, and are situated in a dynamic and often unknown use setting (e.g. Schmidt-Belz et al. 2002, Tamminen et al. 2003, Vetere et al. 2003, Makimoto and Manners 1997). Furthermore, the information presented to the users of mobile guides is closely related or indexed to their physical location, objects in their immediate surroundings and to their present as well as planned activities (e.g. Chincholle et al. 2002, Pospischil et al. 2002, Kolari et al. 2003, Kray and Baus 2003, Kjeldskov et al. 2003). The questions and challenges related to choosing appropriate techniques for evaluating the usability of mobile guides are several. Should the evaluation be done in the lab or in the field? Should the evaluation be based on usability experts and/or involve users? How should the data be analyzed; using a thorough (but time consuming) qualitative and quantitative analysis or a 'discount' approach?

2.1. In-situ or in-vitro?

Since the use of mobile guides is so closely related to the user's context, evaluating in the field seems like an appealing, even indispensable, approach. Indeed most existing studies of mobile usability apply some type of field-based approach. Yet, as the relative strengths and weaknesses of laboratory and field-based methods and techniques for evaluating mobile devices become better understood, this assumption is challengeable (Kjeldskov et al. 2004, Po et al. 2004). Applying a laboratory-based approach, evaluations can benefit from experimental control and high quality data collection. Yet traditional usability laboratory setups may not adequately simulate the context surrounding the use of mobile systems. Using a field-based approach, it may be possible to obtain a higher level of 'realism'. However, field-based usability evaluations are not easy (Nielsen 1998, Brewster 2002) and applying established evaluation techniques and data collection instrumentation, such as multi-camera video recording, think-aloud protocols or shadowing may be difficult in natural settings (Sawhney and Schmandt 2000). Also, field evaluations complicate data collection since users are moving physically in an environment over which we have little control (Johnson 1998, Petrie et al. 1998) and only partially comprehend.

2.2. Users, surrogates or experts?

Usability evaluations in both laboratories and in-situ are problematic for mobile technology because they involve techniques that assume usage that is relatively fixed, tasks that endure over a reasonable period of time and (for laboratory evaluations) can be de-contextualised easily. Furthermore, laboratory and field based evaluations typically involve studying prospective users' interaction with the system being evaluated. This can be very time consuming and hampered by limited access to participants unfamiliar with the process. As an alternative, usability research has promoted a tranche of expert-based evaluation techniques, such as heuristic inspection (Nielsen and Molich 1990) and cognitive walkthrough (Wharton et al. 1994) which may offer benefits. These techniques typically benefit from providing evaluators with guidance (in the form of heuristics or a checklist) for identifying a prioritised list of usability flaws. However, inspection approaches are often criticised for finding proportionately fewer problems in total, and disproportionately more cosmetic problems (Karat et al. 1992). Further, inspection based approaches have been accused of *context immunity* (Po et al. 2004).

2.3. Exhaustive or discount data analysis?

One of the most resource-demanding activities in a usability evaluation is the analysis of collected empirical data, a stage vital to lessons learned, and yet difficult and time consuming to conduct. Whereas there is a strong body of research within human-computer interaction regarding the appropriate choices of data collection methods and techniques, data analysis is vaguely described by many authors, e.g. (Nielsen 1993, Preece et al. 1994, Rubin 1994). Many methods and techniques exists for analyzing the empirical data from usability evaluations like, for example, grounded analysis (Strauss and Corbin 1997), video data analysis (Sanderson and Fisher 1994, Nayak et al. 1995), cued-recall (Omodei et al. 2002), and expert analysis (Nielsen and Molich 1990), etc. However, approaches to instrumenting data analysis are often poorly discussed (Gray and Saltzman 1998) and the relative value of applying such exhaustive approaches to the analysis of usability data is still largely speculative. Of special note, it seems implicitly assumed by many authors that a thorough grounded analysis or video analysis with detailed log-files and transcriptions of usability evaluation sessions is the gold standard by which evaluation should be judged (Sanderson and Fisher 1994). However, the balance between the costs of spending large amounts of time on video analysis and the value added to the subsequent results has been questioned (Nielsen 1994) and is an open question in relation to the evaluation of mobile guides.

3. THE TRAMMATE PROJECT

Inspired by the challenges discussed above, during 2002 and 2003 we explored the issues surrounding the design and evaluation of a mobile guide.

We conducted a research project focusing on the potential of mobile guides for supporting the use of public transportation in Melbourne, Australia (Kjeldskov et al. 2003). The project was motivated by discussions among consultants and sales staff of a large IT company regarding alternatives to the use of cars for travelling in the city to meetings with clients. In large cities where traffic is often very dense, travelling by car can be time-consuming, necessitating much planning. Using Melbourne's tram-based public transport would not only have environmental benefits, but might also be more effective if supported by a mobile information service providing travellers with relevant information at the right time and in the right place.

From this study, we identified some key requirements for a mobile guide supporting the use of the public transportation system:

- Relating travel information directly to the users' unfolding schedule of formal and informal appointments;
- Providing route-planning information for the tram system based on the user's current location and time;
- Alerting the users when it is time to commence their journey in order to make it to the destination in time;
- Providing easy access to key information such as travel time, walking distance and number of route changes.

3.1. The prototype system

A functional mobile guide prototype for Melbourne's tram system was developed by researchers at the University of Melbourne's Department of Geomatics (Smith et al. 2004). The prototype provided route-planning facilities for the tram system based on the user's current location as a combination of textual instructions and annotated maps, satisfying some of the requirements described above. One of the overall screens in the prototype system is shown in figure 1.

Figure 1.
Entering a destination into the mobile guide

Figure 2.
Map view on the mobile guide

The prototype was designed for an iPAQ handheld computer equipped with a WAP browser. The device is connected to the Internet via a GPRS data connection and acquires its position via GPS. The application was designed to serve three functional processes with regard to public transport. These were accessible via the start-up screen.

1. Timetable Lookup: information about the tram timetable based on the input of stop numbers (origin and destination) and route numbers. This function was aimed at regular tram users who are very familiar with their route of travel. No maps are available within this section of the system.
2. Plan Trip: information about the whole route (containing route descriptions and maps) based on the input of suburb and street corners of origin and desired destination. Users were also presented with an option to enter an arrival time or departure time for their journey. From each screen within this function, it was possible to view a visual representation of the relevant portion of the journey on a map.
3. Determine Route: information about the whole route (containing route descriptions and maps) based on the input of the street corner of the destination and the suburb. The system determined the user's origin location via a GPS. Maps were also available for components of the journey in this function.

Upon entering all required input, the system computes a suitable travel plan for using the tram network between the desired origin and destination. The solution suggested by the system is optimal in terms of normative data on journey length (measured in number of stops), and the timing of tram vehicles. An example of the maps displayed by the system is shown in figure 2.

4. COMPARING THE FOUR APPROACHES

In order to investigate the advantages and disadvantages of different techniques for evaluating the usability of mobile guides, we conducted four different evaluation studies of the mobile guide prototype described above:

1. Field Evaluation: exhaustive analysis of user-based data; data collected in-situ but analysed in-vitro
2. Laboratory Evaluation: exhaustive analysis of user-based data; data collection and analysis conducted in-vitro
3. Heuristic Walkthrough: discount collection and analysis of usability problems by experts; data collection and analysis conducted in-vitro
4. Rapid Reflection: discount analysis of user-based data from field and laboratory studies; in-vitro data analysis. This analysis was done prior to the exhaustive analysis in studies 1 and 2

These four techniques illustrate some of the key issues of choosing an appropriate evaluation technique discussed earlier. The four evaluations are described in detail in the following sections.

4.1. Study 1: Field Evaluation

The field evaluation focused on guide use in realistic settings. It took place over two days in the city centre of Melbourne, Australia. The evaluation involved five test subjects between twenty one and forty two years of age similar to the profile of the participants involved in the earlier user studies of the TramMate project.

Figure 3. Field evaluation of the mobile guide.

The test subjects were all frequent computer users and had experience with the use of PDAs and mobile phones. The test subjects were all familiar with the tram system of Melbourne. The subjects had to complete four realistic tasks involving route planning while travelling to appointments in the city by tram. The tasks were derived from the earlier user studies in the TramMate project and were piloted prior to the evaluation, resulting in minor modifications in order to make them achievable within a feasible timeframe. In order to solve the tasks, the test subjects had to look up information available in the mobile guide and then perform the tasks 'for real' (e.g. catching a tram to a specific destination). An example task is shown below:

> You are going to catch a tram from the corner of Swanston and Queensberry Street in Carlton for a meeting at the corner of Little Collins and Exhibition Street in Melbourne. You have to be there in about 30 minutes from now.
>
> Using the plan trip option, find out:
>
> - Which tram route(s) to take
> - When the first possible tram is departing
> - The number of route changes (if any)
> - If there is a route change, where to board the second tram.
> - Which stop to get off the last tram.
> - How to get from the last stop to your final destination.
> - The estimated time of arrival.
> - Use this information to get to the meeting.

The prototype accessed live timetable information through a GPRS connection to the Internet. Due to technical problems with acquiring precise GPS positioning data in the city area and on the trams, positioning was simulated by the researchers by inputting predefined spatial data into the system 'behind the scenes' of the evaluation. Users were not aware of this.

The field evaluation involved four people for each evaluation session. One test subject used the mobile guide to solve the tasks. One researcher managed the evaluation sessions, encouraging the test subjects to think-aloud and asking questions for clarification similar to a contextual interview. Another researcher recorded the evaluation sessions on video switching between close-up views of the device and overall views of the surroundings. A third researcher took written notes (figure 3).

The data from the field evaluation was subject to a detailed grounded analysis (Strauss and Corbin 1997), producing a list of richly described usability problems. The problems were rated as critical, serious or cosmetic in accordance with Molich (2000).

Critical problem

- Recurred across all users
- Stopped users completing tasks

Serious problem

- Recurred frequently across users
- Inhibited /slowed down users completing tasks
- Users could (eventually) complete tasks

Cosmetic problem

- Did not recur frequently across users
- Did not inhibit users severely
- Users could complete tasks

The researchers involved in the analysis counted and grouped problems collaboratively. Then a qualitative judgment concerning each problem's severity was made. For example, the 'system vs. real world' problem category was rated as critical as this problem occurred across all users and severely impeded the user's ability to complete their work. The 'labelling' problem was rated as 'severe' because it occurred frequently across some users but did not inhibit completing the task. The 'social comfort' problem category was rated as cosmetic because only one user described this as a problem and it did not inhibit this user's task completion noticeably. The time spent on the field evaluation amounted to fifty six person-hours for data collection and twenty six person-hours for data analysis.

4.2. Study 2: laboratory evaluation

The laboratory evaluation focused on use in a controlled setting. It was conducted in a state-of-the-art usability laboratory at the University of Melbourne's Department of Information Systems. Due to less time required for logistics, we were able to conduct the laboratory evaluation in one day. We intentionally designed the laboratory evaluation to be similar to the field evaluation in a number of important ways as this allowed us to compare the results across techniques. However some differences were necessary if we were to 'play to the strengths' of each approach. The laboratory evaluation involved the same number and type of test subjects and the test subjects had to solve the same four tasks using the same mobile guide system. However, in the laboratory evaluation, the subjects were seated at a desk, with the mobile guide in their hand rather than being physically mobile. Also, they did not have to perform the tasks 'for real' as in the field – that is they were not required to board a tram and take the journey.

The laboratory setting allowed for high-quality audio and video recordings from multiple perspectives (figure 4). Three ceiling-mounted cameras captured overall views of the test subject and test monitor. A fourth camera on a tripod captured a close-up view of the mobile guide (figure 5). To ensure a good view of the screen and interaction, the test subjects were asked to hold the device within a limited area indicated on the table.

Figure 4. Laboratory evaluation of the mobile guide.

As in the field, the mobile guide accessed live timetable information while positioning was simulated. The laboratory evaluation involved four people: one test subject and three researchers; a test monitor or host, encouraging the test subject to think aloud and asking questions for clarification; and two data loggers, observing the evaluation through a one-way mirror respectively. The data from the laboratory evaluation was analyzed using the same method as for the field evaluation, resulting in a similar list of usability problems.

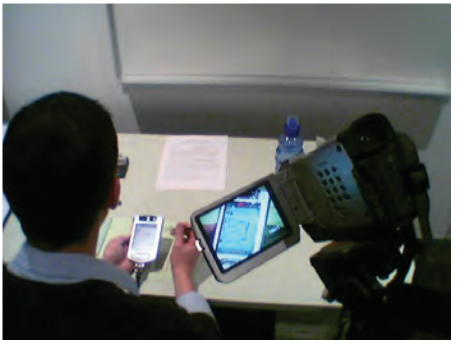


Figure 5. Close-up of interaction with the mobile guide.

The time spent on the laboratory evaluation amounted to thirty two person-hours for data collection and eighteen person-hours for data analysis.

4.3. Study 3: heuristic walkthrough

The third evaluation of the mobile guide focused on usability as perceived by experts in human-computer interaction. It was conducted in the same laboratory used for the laboratory study (figure 6) and consisted of a heuristic walkthrough guided by a set of heuristics developed specifically for the purpose of this evaluation, heuristics sensitive to the mobile challenge. For a detailed description see Vetere et al. (2003).

Four evaluators, all with expertise in HCI and usability, each independently performed a heuristic walkthrough of the mobile guide. The evaluators were given the mobile guide heuristics and a common set of tasks to contextualize the evaluation, thereby blending aspects of traditional heuristic evaluation and the cognitive walkthrough. The tasks were the same as used in the field and laboratory evaluations. Each evaluation lasted an average of one and one quarter hours. First, the evaluators were welcomed by the host (a representative from the design team), and given the opportunity to ask questions about the process. The evaluators then explored the device, without reference to either the heuristics or the task scenarios. Thereafter, the evaluators assessed the device against the heuristics and recorded their observations. Finally, the evaluators worked through each task, recording further observations against the heuristics. After all heuristic walkthroughs had been completed results were collated in a post session workshop, allowing the evaluators to discuss their identified usability problems. As in the field and laboratory evaluations, the mobile guide accessed live timetable information, while positioning was simulated.

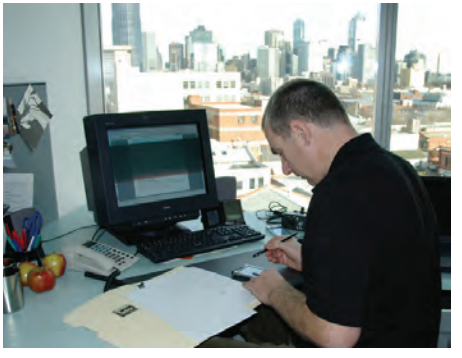


Figure 6. Heuristic walkthrough.

All but one of the evaluators completed all tasks, and all evaluators addressed the mobile guide heuristics. Additionally all evaluators drew broadly on their knowledge of usability, not confining themselves to ‘mobility issues’ or the mobile guide heuristics alone, and all reflected on the heuristic walkthrough process itself. The time spent on the heuristic walkthrough amounted to ten person-hours in total.

4.4. Study 4: rapid reflection

The fourth study had the purpose of investigating the potential for reducing the effort spent on data analysis by applying a ‘rapid reflection’ approach inspired by rapid ethnography (Millen 2000). The rapid reflection study of the mobile guide differed somewhat from the other three studies. Rather than being a completely separate study, the rapid reflection approach was based on the empirical data gathered through the field and laboratory evaluations. However, as an alternative to the rather time consuming grounded analysis of the video data, the rapid reflection approach applied a pragmatic discussion and consideration of the collected data by the involved evaluators. For a detailed description of this study see Pedell et al. (2003).

The rapid reflection sessions (figure 7) followed immediately after the field and laboratory evaluations and involved all participating researchers. On the basis of the observers’ written notes and experiences during the evaluations, the rapid reflection sessions had the purpose of discussing and agreeing on what main themes and usability problems had emerged on that specific day. Each session was restricted to one hour.

Figure 7. Rapid Reflection session.

The rapid reflection session was assisted by an observer, who was not present during the laboratory or field evaluations, asking questions for clarification. Furthermore, one of the researchers had the role of writing all identified usability problems and other issues on a whiteboard as they were presented, and keeping an overview of the discussed usability problems as the session progressed. After the reflection session, one of the researchers spent an hour on writing up the contents of the whiteboard into a richly described list of usability issues, which was then circulated among the team for validation and comments.

The time spent on the rapid reflection approach amounted to a total of fourteen person-hours for the field data and eight person-hours for the laboratory data. As the rapid reflection builds on the data already collected in the field (study 1) and lab (study 2) respectively, these numbers should be compared to the twenty six and eighteen hours spent on the exhaustive data analysis described above.

In this paper, we do not compare the outcomes of the rapid reflection sessions across field and laboratory data in this paper. For a more elaborated discussion on this issue, see (Pedell et al. 2003).

4.5. Analysis

The analysis of data from each of the four approaches described above focused on identifying and describing usability problems experienced with the use of the mobile guide prototype. In the case of the field and laboratory evaluation this was done through the use of grounded analysis (Strauss and Corbin, 1997). In the case of the heuristic walkthrough and the rapid reflection it was done through post-evaluation workshops. Two discrete steps were involved in the comparison of the results across the four approaches; a *compilation* of the results and a *comparison* of the results across techniques. In order to ensure that this process was rigorous and that both the compilation and comparison of results were credible, dependable and confirmable (Lincoln and Guba 1986) the following steps were taken.

Firstly, one researcher compiled the results for each of the four approaches into four lists of identified usability issues. This researcher was involved in data analysis for the field and laboratory evaluation, and data collection and analysis for the rapid reflection and heuristic walkthrough. Thus this researcher had proximity to the results from each of the four approaches, a prolonged engagement with the results and had engaged in persistent observation of the data (Guba and Lincoln 1989). Following the compilation of the results from the four different approaches, all participating researchers were required to revisit the list from each approach. In this way, the dependability (Guba and Lincoln 1989: 242) of the results for each of the four approaches was ensured. Secondly, another researcher (who had been involved in the data collection for the field and laboratory evaluation and rapid reflection) collaborated with the first researcher in the compilation of the results for each of the four approaches into one merged list. This collaboration involved extended discussions of the identified problems (member checking) and in the monitoring of the compilation of the results for each of the four approaches (progressive subjectivity) (Guba and Lincoln 1989). Due to the experience of the researchers across the four methods both with data collection and analysis, it was possible to ensure that the problems compared were on a similar level of abstraction. In case of different severity ratings of the same usability issue across techniques, the most severe rating was used in the merged list. To be able to identify disparities in severity ratings the original ratings were preserved as comments to each of the cells in the list. Finally, the merged list of usability issues was presented and discussed jointly by the full team of participating researchers (the authors of this paper) through a one-hour workshop. This was done to ensure that the comparisons across techniques were credible (through member checking and the involvement of the attendant researchers in the initial analysis), and dependable and confirmable (through an audit of the results and comparisons by two researchers). The resulting list of merged problems is shown in table 1.

In the next section we present our findings, and draw out some key differences between the four approaches as they apply to the task of evaluating a mobile guide. Differences between the approaches that are not germane to mobile guide evaluation are outside the scope of this paper.

It should be noted that, in presenting our results, we do not claim statistical power, but rather aim to present a rich, qualitative overview of the data, drawing out differences and similarities as they arise. This allows us to draw some overall conclusions concerning the pros and cons of different techniques for evaluating the usability of mobile guides.

5. FINDINGS

Jointly, the four usability studies generated a list of twenty two distinct usability problems. Of these twenty two problems, a total of five problems were classified as critical, eleven as serious, and six as cosmetic (see final column of table 2). Critical usability problems related to the interaction between the user/system and the surrounding environment, for instance the representation of map and textual information in the system and the way the system required the user to use this information. Another critical issue was caused by disparities in the relationship between information presented in the system and the context in which the user was situated. Critical problems were typically related to mapping issues arising from the use of the 'system in the world'. The distribution of usability problems in the 4 techniques is summarized in table 2.

Table 1. Merged problem list.

Critical problems

1	Maps. Issues related to how the user interprets and uses maps in conjunction with the textual information.
2	Navigation. Issues related to problems with navigating through the screens of the system.
3	Information. Issues related to lack of relevance and accuracy of information presented by the system.
4	System vs. World. Issues caused by disparities in the relationship between information in the system and in the world.
5	Use and usefulness. Issues related to a conception of use broader than usability including overall purpose of the device

Serious problems

6	Input and affordances. Issues with entering data into the system and the affordances offered by the system for doing so.
7	Help and recovery. Issues related to lack of support and assistance in helping the user recovering from errors.
8	Knowledge about city. Issues related to high requirements for user's knowledge about the city
9	Labelling. Issues caused by poor wording and use of abbreviations within the system.
10	Cognitive Load. Issues related to high requirements for cognitive resources (memory and attention) to use the system.
11	System. Issues caused by technical malfunctions in the prototype system.
12	Interface flexibility. Issues related to lack of support for variation from the predefined path of interaction.
13	Mental model. Issues related to disparities between how the system works and how the users think the system works
14	User Confidence. Issues related to lack of confidence in using the system or acting according to information provided.
15	Scope. Issues related to uncertainties regarding what functionalities the system offers to the user.
16	Value. Issues related to users experiencing limited value of the information presented by the system.

Cosmetic problems

17	Efficiency. Issues emerging from users experiencing the system being time consuming and cumbersome to use
18	Orientation. Issues emerging from lack of information in the system for supporting the user's orientation in the world.
19	Readability. Issues related to difficulties with reading small fonts on the screen of the device.
20	Dependency on the System. Issued related to the user being dependant on the system for making decisions.
21	Social comfort. Issues related to how comfortable the user is with using the system in public.
22	Emotional response. Issues causing strong emotional responses from the user while using the system.

Table 2. Distribution of the number of usability problems identified.

	Field evaluation	Lab evaluation	Heuristic walkthrough	Rapid reflection	Total
Critical	4	4	4	4	5
Serious	7	6	6	7	11
Cosmetic	2	3	4	4	6
Cosmetic	13	13	13	13	22

Regarding problem coverage, any individual technique identified little more than half of the total problem set (coincidently, thirteen from twenty two in each case).

Looking at the critical problems, all techniques identified four out of five critical problems, no technique identifying all problems. In the case of serious problems, more variation was observed across the four techniques, with the identification of between five and seven problems, from a total set of eleven. Again, no single technique was able to identify all eleven issues, and only the field evaluation identified more than half of the total number of serious problems. In the case of cosmetic problems, the rapid reflection technique was the most effective, identifying four out of six problems. While missing two of the five cosmetic problems identified through the video analysis, the rapid reflection was the only technique that reported the issue of problems with using the system causing strong emotional responses from the users. As an interesting aside, it should be noted that the heuristic walkthrough did not generate the usual level of 'cosmetic noise' that often characterizes expert evaluations based on general usability heuristics (Karat et al. 1992). It may be that tailoring the heuristics (see Vetere et al. 2003) to the mobile problem helped reduce such noise, especially false positives, in the data. The distribution of problems identified with the 4 techniques is illustrated in figure 8.

Figure 8. Distribution of usability problems. A black square indicates that a problem was identified using that specific technique. A white square indicates that a problem was not identified using that specific technique but was found using another technique.

Figure 8 shows twenty two usability problems (each column represents a specific problem), stratified as critical, serious or cosmetic, distributed across the four different techniques. A black square shows that a problem was identified using that technique. A white square indicates that a problem was not identified using that technique, but was found using another technique (see table 1 for a brief description of the problems). The distribution of problems in figure 8 is discussed below.

5.1. Critical problems

Three out of the total set of five critical problems were identified by all techniques, with a further problem identified by all but the heuristic walkthrough. Though comparing evaluation approaches is always challenging, due primarily to the lack of any independently established problem set, we can be confident that these four critical problems were indeed present in the evaluated mobile guide, rather than being false positives. On the other hand, the distribution of critical problems also indicates that the identification of critical problems depended little on the precise circumstances surrounding the deployment of a specific evaluation approach; it is encouraging that critical problems generally are uncovered regardless of approach. It is also noticeable that the field, lab and rapid reflection studies were consistent in the types of critical problems identified.

For the identification of the most severe issues in a mobile guide, discount data analysis appears to be adequate. The benefits of an exhaustive grounded analysis may not out weigh the associated costs.

Only one critical usability problem was unique to a specific approach. This 'problem', identified by the heuristic walkthrough, concerned the general purpose of the guide, and its alignment with broader lifestyle and use issues not evident in findings drawn from the other approaches. Issues raised here included the degree to which users could flexibly adapt the device to fit lifestyle activity (Vetere et al. 2003).

The critical problem not identified in the heuristic walkthrough was a problem related to disparities in the relationship between information in the system, and the users' context- the 'system in the world' problem referred to earlier. This problem was adjudged critical in both the field and rapid reflection studies (which in turn drew on the data collected in the field), but cosmetic in the laboratory study. Given the situated flavour of this problem, the different severity ratings are not surprising. However, it does highlight the fact that while contextually related problems may appear in laboratory settings, they can be experienced, and described, in very different ways to the field.

5.2. Serious problems

The distribution of serious problems shows a more varied picture across approaches. Of eleven serious problems, eight were identified by two or more of the techniques, four were found by three techniques or more, and only one problem was identified by all techniques. Three serious problems were uniquely identified by only one technique.

Whereas the critical problems reflected 'system in the world' issues, serious problems were more oriented to significant usability hurdles: difficulty in entering data into the system, difficulty in being able to recover from errors and poor labelling of interface elements. Additionally, the systems' implicit assumptions about the users' existing knowledge of the city in which the mobile guide was used also drew attention here. Other serious problems related to cognitive load demands, e.g. remembering data from one screen when interacting with another, and lack of flexibility to deviate from a predefined, by the system, path of interaction.

Looking at the clustering of problems, it is noticeable that there is a relatively large overlap between the findings from the field and laboratory studies. Five out of the total eleven serious problems were identified in both the lab and the field, with the field identifying only two additional unique problems and the laboratory only one further unique problem. The five serious problems identified in both the laboratory and the field included the four most prominent; input, recovery and labelling.

Whilst some of the more serious flaws were also identified by both the heuristic walkthrough and the rapid reflection, and both of these approaches contributed unique problems (one in each case), both the heuristic walkthrough and the rapid reflection missed four and five serious problems respectively, from those identified collectively in the field and in the lab.

5.3. Cosmetic problems

The picture is yet more confused when examining cosmetic problems. None of the cosmetic problems were identified by all techniques, and only two problems were identified by three of four approaches.

Looking at the clustering of problems, there was no overlap between the cosmetic problems found in the field and in the lab. The field approach drew attention to issues such as the real-world validity and precision of the data presented by the system and the 'social comfort' (e.g. whether it felt embarrassing to use the device in a public setting). In contrast, the lab-based approach drew attention to device oriented issues, such as the readability of text and efficiency of looking up information.

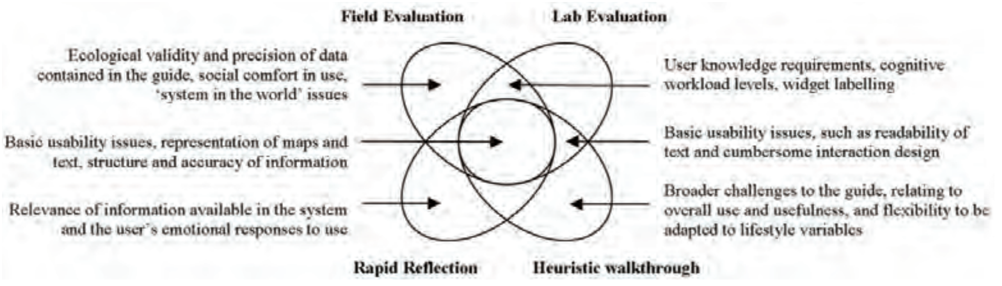


Figure 9. Overview of the types of usability issues identified in overlaps between techniques.

Interestingly, the lab and the heuristic walkthrough identified the same problem set, with the rapid reflection sitting somewhere in-between, identifying one unique problem related to the observation, that many users had a strong emotional response when encountering problems with the system.

In the next section we draw out general lessons learned, especially in relation to the similarities and differences between the four approaches.

6. DISCUSSION

Figure 9 outlines the overlap between the four approaches, in terms of the usability problems identified.

There are benefits to be gained from each approach in relation to the types of usability problems uncovered, but many strengths are shared by more than one technique. The cluster in the centre of figure 9 emphasises that many usability issues related to the representation, accuracy and structure of the map and textual information provided, and these issues are captured by all approaches.

All approaches, with the exception of the laboratory study, identified unique problems. The field evaluation uniquely identified issues of validity and precision of the data presented by the device, and the lack of social comfort when using the device in public. The heuristic walkthrough uniquely identified issues related to the overall use and usefulness of the mobile guide, and its flexibility in relation to different user activities. The rapid reflection approach, though based on the data from the lab and field studies, brought forward some issues related to the perceived relevance of available information and highlighted the users' strong emotional responses (ranging from frustration to sheer outrage!) to the hurdles presented by the design.

Examining the various pair-wise comparisons, it is interesting to note that the overlap between the laboratory evaluation and heuristic walkthrough contains basic usability problems, such as the readability of screen text, whereas the overlap between the field and laboratory studies contains the potentially more complex problems of the assumed extent of users' prior knowledge and the cognitive workload demands on the user.

Contrasting the laboratory and field studies, two differences in the problem sets are worthy of note. Whilst the laboratory problems were reported in great detail (often related to the artefact per se, for example, mislabelling of commands), the field study stressed problems of mobile 'use' rather than simply device usability, and typically those problems were expressed in the language of the situation. For example, spending too long inputting commands was made urgent through making explicit the pressing demands of the situation; the user might be stationary, reading the mobile display, and blocking a footpath in the situation of use.

The rapid reflection sessions briefly summarized the key issues from the field and laboratory user studies requiring considerably fewer person-hours for analysis. Generally, the problems reported through the rapid reflection were less specific and the list of problems was not complete compared to the joint outcome from the video analysis. On the other hand, the rapid reflection technique allowed the researchers to focus only on the top-most severe problems observed. Identifying four out of five critical problems in less than half the time required for the video analysis, the rapid reflection proved to be a very cost-effective usability analysis technique. This finding is consistent with a similar comparison done by Kjeldskov et al. (2004). The differences between problems reported through rapid reflection and exhaustive video analysis across the field and laboratory studies may be due to, among other factors, the people involved in the analysis having different views of and proximity to the data. For a more elaborated discussion of this issue see (Pedell et al. 2003).

Across the four approaches there is much similarity in the pictures that emerge of the mobile guide, but there are many compelling differences. We will now summarise some general lessons learned.

6.1. In-situ or in-vitro?

The development of electronic mobile guides remains a rather recent design challenge, and we cannot rely on established theory or rigorously tested examples of best practice to guide us. Collecting data in-situ prompted us with elements of the situation of use that we might have been ignorant of, or that might have passed un-remarked. Additionally, being in-situ provoked a very concrete consideration of how things might be changed; it is easy to be lazy when discussing the future, speculations turning from plausible fiction to science fiction. Being in-situ was our insurance policy against ignorance in the absence of a refined understanding of what 'the situation of use' was, or might become. For examples of problems identified in-situ but not in-vitro see table 1, problem 14, 20 and 21. Until we are able to supplement our meagre understanding of mobile use, and unless there are insurmountable practical or logistical hurdles to accessing the situation of use, we should continue to collect, at least as a part of a broader data collection protocol, data in the field.

6.2. Users, surrogates or experts?

The issue of expert versus user-based evaluation is part of a more general discourse (for example Dumas and Redish 1999, Nielsen 1994) that we will not cover here. With respect to mobile guides, a few comments are appropriate.

Due to the relative novelty of mobile guides, and the lack of a substantial relevant knowledge base, the perceived 'opinion free' flavour of user based tests, as compared to inspection based approaches, might strengthen the usability argument in the broader software development process. In contrast, the relative novelty of the mobile guide paradigm should drive us to ‘test early and often’; we saw some evidence in our data, though preliminary at best, to suggest experts are able to overcome the credibility hurdles involved in early paper-based prototypes more ably than end-users.

6.3. Exhaustive or discount?

Our activities in the development of mobile guides are thirsty for foundational concepts and theoretical insight. The motivation for exhaustive data collection and analysis extends beyond theory building, to practice as it relates to safety-critical or business-critical applications. We should continue to champion discount approaches for the fast cycle, discovery oriented phases of early development, whilst encouraging a concerted effort in building the theoretical foundations of an applied science of mobile use.

7. CONCLUDING COMMENTS

Whilst no individual approach to the usability testing and evaluation of mobile guides can be held to be the definitive approach, any testing and evaluation is much better than none at all. The level of agreement amongst the approaches was both significant and encouraging, but not complete and multi-method approaches to mobile guide evaluation are clearly useful, as implied in figure 9.

Mobile guides raise particular if not unique challenges, including the need to understand the users’ experience of the ‘system in the world’, establishing and designing for social comfort and evaluating the compatibility between the device and broader lifestyle considerations. These particular challenges provide new reasons to respect the unfolding nature of, and situated character of, the interactions between people and technology; challenges that, with time, will be met by advances in our theoretical apparatus, our methodological toolkit, and our sense of what is and what is not best practice in relation to the design of mobile guides.

The transferability of the findings presented in this paper to the evaluation of other mobile guides requires further investigation. Drawn from our experiences across four evaluation methods, we have presented three key issues pertinent to the selection of evaluation methods, which we believe are of interest to researchers and practitioners. By describing each method in detail and comparing the usability problems identified by each of them, we have presented a rich insight into the strengths and weaknesses of each method for evaluating a functional, prototypical mobile guide. Opportunities exist to attempt to apply these findings to mobile guides at different levels of fidelity residing in different contexts.

Regarding the transferability of the usability problems presented, some of them relate to the specific design of the evaluated system and may or may not apply to other guide systems. This includes, for example, some of the cosmetic problems such as the labelling of interface elements and readability. However, most of the critical and serious usability problems identified relate to more general issues, such as the design of maps, navigation in the system, relevance of information, the relation between the system and the real world, etc. These problems, we believe, are much more universal and will most likely apply to the usability of mobile guides in general.

ACKNOWLEDGEMENTS

Thanks to the consultants who acted as participants in the initial user studies, thanks to the lab and field test subjects, and thanks to those colleagues involved in the earlier phases of the TramMate project: Jennie Carroll, John Murphy, Jeni Paay and Daniel Tobin. This research was supported by the Danish Technical Research Council (project reference 26-03-0341).

REFERENCES

Abowd, D., Atkeson, C., Hong, J., Long, S. and Pinkerton, M., 1996, Cyberguide: a Mobile context-aware tour guide. Wireless Networks, 3(5), 421-433.

Andrade, M. T., Santos, E., Livaditi, J., and Tsakali, M., 2002, Managing multimedia content and delivering services across multiple client platforms using XML. Proceedings of Mobile Tourism Support (Pisa, Italy: in conjunction with Mobile HCI 2002).

Bailey, R. W., Allan, R. W., and Raiello, P., 1992, Usability testing vs. heuristic evaluation: a head-to-head comparison. Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting, HFES, pp. 409-413.

Bohnenberger, T., Jameson, A., Krüger, A. and Butz, A., 2002, Location-aware shopping assistance: evaluation of a decision-theoretic approach. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag).

Bornträger, C., Cheverst, K., Davies, N., Dix, A., Friday, A. and Seitz, J., 2003, Experiments with multi-modal interfaces in a context-aware city guide. Proceedings of Mobile HCI 2003 (Udine, Italy: LNCS, Springer-Verlag), pp. 116-130.

Bornträger, C., and Cheverst, K., 2003, Social and technical pitfalls designing a tourist guide system. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Brewster, S., 2002, Overcoming the lack of screen space on mobile computers. Personal and Ubiquitous Computing, 6, 188-205.

Cheverst, K., Davies, N., Mitchell, K., 2002, Exploring context-aware information push. Personal and Ubiquitous Computing, 6, 276-281.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C., 2000 Developing a context-aware electronic tourist guide: some issues and experiences. Proceedings of CHI'00 (The Hague, Netherlands: ACM), pp. 17-24.

Chincholle, D., Goldstein, M., Nyberg, M. and Erikson, M., 2002, Lost or found? a usability evaluation of a mobile navigation and location-based service. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 211-224.

Dumas, J. S., and Redish, J. C., 1999, A Practical Guide to Usability Testing (Exeter: Intellect).

Fithian, R., Iachello, G., Moghazy, J., Pousman, Z. and Stasko, J., 2003, The design and evaluation of a mobile location-aware handheld event planner. Proceedings of Mobile HCI 2003 (Udine, Italy: LNCS, Springer-Verlag), pp. 145-160.

Gray, W. D., and Saltzman M. C., 1998, Damaged merchandise? a review of experiments that compare usability evaluation methods. Human-Computer Interaction, 13(3), 203-261.

Guba, E. G., and Lincoln, Y. S., 1989, Fourth Generation Evaluation. (California: Sage Publications).

Henderson, R., Podd, J., Smith, M., and Varela-Alvarez, H., 1995, An examination of four user-based software evaluation methods. Interacting with Computers, 7(4), 412-432.

Iacucci, G., Kela, J. and Pehonen, P., 2004, Computational support to record and re-experience visits. Personal and Ubiquitous Computing, 8(2), 100-109.

Johnson, P., 1998, Usability and mobility: interactions on the move. Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices.

Karat, C. M., Campbell, R., and Fiegel, T., 1992, Comparison of empirical testing and walkthrough methods in user interface evaluation. Proceedings of CHI'92 (New York: ACM) pp. 397-404.

Kolari, J., and Virtanen, T., 2003, In the zone: views through a context-aware mobile portal. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Kjeldskov, J., Skov, M. B., and Stage, J., 2004, Instant data analysis: conducting usability evaluations in a day. Proceedings of NordiCHI 2004 (Tampere, Finland: ACM).

Kjeldskov, J., Howard, S., Murphy, J., Carroll, J., Vetere, F. and Graham, C., 2003, Designing TramMate - a context aware mobile system supporting use of public transportation. Proceedings of DUX 2003 (San Francisco, CA, USA: ACM).

Kjeldskov, J., and Stage, J., 2004, New techniques for usability evaluation of mobile systems. International Journal of Human-Computer Studies, 60, 599-620.

Kjeldskov, J., Skov, M. B., Als, B. S., and Høegh, R. T., 2004, Is it worth the hassle? Exploring the added value of evaluating the usability of context-aware mobile systems in the field. Proceedings of Mobile HCI 2004 (Glasgow, Scotland: LNCS, Springer-Verlag).

Kjeldskov, J. and Skov, M. B., 2003, Creating a realistic laboratory setting: a comparative study of three think-aloud usability evaluations of a mobile system. Proceedings of Interact 2003 (Zürich, Switzerland: IOS Press).

Kray, C., and Baus, J., 2003, A survey of mobile guides. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Laakso, K., Gjesdal, O., and Sulebak, J. R., 2003, Tourist information and navigation support by using 3D maps displayed on mobile devices. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Lincoln, Y. S., and Guba, E. G., 1986, But is it rigorous. Trustworthiness and authenticity in naturalistic evaluation. In Naturalistic Evalution, edited by D.D. Williams (San Francisco: Jossey-Bass).

Makimoto, T., and Manners, D., 1997, Digital Nomad (Chichester: John Wiley & Sons).

Molich, R., 2000, Usable Web Design (In Danish) (Ingeniøren | bøger)

Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D., and Kirakowski, J., 1998, Comparative evaluation of usability tests. Proceedings of the Usability Professionals Association Conference, pp. 189-200.

Nayak, N. P., Mrazek, D., and Smith, R.D., 1995, analysing and communicating usability Data. SIGCHI Bulletin, 27 (1).

Nielsen, C., 1998, Testing in the field. Proceedings of the third Asia Pacific Computer Human Interaction Conference, APCHI 1998 (IEEE Computer Society)

Nielsen, J., 1994, Guerrilla HCI: Using discount usability engineering to penetrate the intimidation barrier. In Cost-Justifying Usability, edited by R. G. Bias, and D. J. Mayhew (Academic Press Professional).

Nielsen, J., 1993, Usability Engineering. (Morgan Kaufmann).

Nielsen, J. and Molich, R., 1990, Heuristic evaluation of user interfaces. Proceedings of CHI'90 (Seattle, USA: ACM Press), pp. 249 – 256.

Omodei, M. M., Wearing, A. J., and McLennan, J., 2002, Head-mounted video and cued recall: a minimally reactive methodology for understanding, detecting and preventing error in the control of complex systems. Proceedings of the 21st European Annual Conference of Human Decision Making and Control., (Scotland: University of Glasgow).

Pedell, S., Graham, C., Kjeldskov, J. and Davies, J., 2003, Mobile evaluation: what the metadata and the data told us. Proceedings of OzCHI 2003 (Brisbane, Australia: CHISIG)

Petrie, H., Johnson, V., Furner, S., and Strothotte, T., 1998, Design lifecycles and wearable computers for users with disabilities. Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices (Glasgow).

Pirhonen, A., Brewster, S.A., and Holguin, C., 2002, Gestural and audio metaphors as a means of control for mobile devices. Proceedings of CHI 2002 (ACM).

Po, S., Howard, S., Vetere, F., and Skov, M., 2004, Heuristic evaluation and mobile usability: bridging the realism gap. Proceedings of Mobile HCI 2004 (Glasgow, Scotland: LNCS, Springer-Verlag).

Pospischil, G., Umlauft, M. and Michlmayr, E., 2002, Designing LoL@, a mobile tourist guide for UMTS. Proceedings of Mobile HCI 2002 (Pisa, Italy: LNCS, Springer-Verlag), pp. 140-154.

Preece, J., Rogers, H., Sharp, D., Benyon, D., Holland, S., and Carey, T., 1994, Human-Computer Interaction. (Workingham: Addison-Wesley).

Reid, H., 2002, Maps on your PDA - simple stuff #8 - Mapping the Trade Show Floor. Directions Magazine 2002.

Rocchi, C., Stock, O., and Zancanaro, M., 2003, Semantic-based multimedia representations for the museum experience. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Rubin, J., 1994, Handbook of Usability Testing (Wiley).

Sanderson, P.M., and Fisher, C., 1994, Exploratory sequential data analysis: foundations. Human-Computer Interaction, 9(3), 251-317.

Sawhney, N., and Schmandt, C., 2000, Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. Transactions on Computer-Human Interaction, 7(3), 353-383.

Schmidt-Belz, B. and Poslad, S., 2003, User validation of a mobile tourism service. Proceedings of HCI in Mobile Guides (Udine, Italy: in conjunction with Mobile HCI 2003).

Schmidt-Belz, B., Laamanen, H., Poslad, S., and Zipf, A., 2003, Location- based mobile tourist services - first user experiences, ENTER 2003 Conference, January 2003 (Helsinki).

Smith, J., Mackaness, W., Kealy, A., and Williamson, I.P., 2004, Spatial data infrastructure requirements for mobile location based journey planning. Transactions in GIS, 8(1).

Strauss, A., and Corbin, J., 1997, Grounded Theory in Practice (Thousand Oaks: Sage Publications).

Umlauft, M., Pospischil, G., Niklfeld, G., and Michlmayr, E., 2003, LoL@, a mobile tourist guide for UMTS, Information Technology & Tourism, 5(3), 151-164.

Vetere, F., Howard, S., Pedell, S., and Balbo, S., 2003, Walking through mobile use: novel heuristics and their application. Proceedings of OzCHI 2003 (Brisbane, Australia: CHISIG).

Wharton, C., Rieman, J., Lewis, C., and Polson, P., 1994, The cognitive walkthrough: a practitioner's guide. In Usability Inspection Methods, edited by J. Nielsen, and R. L. Mack (John Wiley and Sons, Inc,) pp. 105-140.

Chapter 14

Taking the lab with you

Rune T. Høegh, Jesper Kjeldskov, Mikael B. Skov and Jan Stage

Abstract. Evaluating mobile technologies "in the real world" is hard. It is challenging to capture key situations of use, hard to apply established techniques such as observation and "thinking aloud", and it is complicated to collect data of an acceptable quality. In response to these challenges, we have developed a "field laboratory" for evaluating mobile technologies in situ. Facilitating high-quality data collection as well as unobstructed user interaction, the field laboratory allows a small wireless camera to be attached to a mobile device, capturing a close-up image of the screen and buttons. This chapter describes the iterative development of our field laboratory over 4 years of evaluating several mobile systems in field settings. It leads to a description of the current setup and how it is used, and explains the rationales for key decisions on technology and form factors made throughout its development.

1. INTRODUCTION

Studying peoples' use of technology is a key activity within the research field of Human-Computer Interaction (HCI) providing software developers with invaluable information about the usability and usefulness of their systems at different stages of the process from conceptual design to a final implementation. Traditionally, such studies have taken place in dedicated "usability laboratories" where users' interaction with computer systems can be observed in a controlled experimental setting providing video and audio data of very high quality. Studying the usability of mobile technologies, however, raises new questions and concerns. Mobile systems are typically used in highly dynamic contexts involving a close interaction between people, systems and their surroundings. Therefore, studying mobile technology use in situ seems like an appealing or even indispensable approach – rather than trying to recreate the use situation realistically in a laboratory.

Originally published as Høegh, R. T, Kjeldskov, J., Skov, M. B. and Stage J. (2008) A Field Laboratory for Evaluating In Situ. In J. Lumsden (Ed.), *Handbook of Research on User Interface Design and Evaluation for Mobile Technology* (pp. 982-996). PA, USA: Idea Group Inc (IGI).

However, studying mobile technology usability "in the real world" is difficult. It is difficult to capture key situations of use, apply established usability techniques such as observation and "thinking aloud" without interfering with the situation, and it is complicated to collect data of an acceptable quality.

In response to some of these challenges, we have extended our stationary usability laboratory at Aalborg University's Department of Computer Science with a mobile counterpart, the field laboratory, which can be taken into the field when studying mobile system use and usability. Facilitating high-quality data collection as well as unobstructed user interaction, the field laboratory allows a small wireless camera to be attached to the mobile device, capturing a close-up image of the screen and buttons while a third-person view is captured by a handheld camcorder.

The purpose of this chapter is to communicate our experiences with developing and using the field laboratory for evaluating mobile technology use and usability in situ by taking the readers through four years of major iterations leading to its current configuration. By doing this, it is our aim to make practitioners, researchers and designers of mobile technologies able to set up and use their own field laboratories for evaluating mobile systems in situ. It is also our aim to inspire further development of even better field laboratory setups facilitating better, easier, faster, and cheaper use and usability data collection in the field. It is not the purpose of this chapter to discuss the relation between evaluating in the field or in the lab. We take our point of departure in the assumption that you have decided to evaluate in the field and focus on how you can collect high quality data while out there. It is also not our aim to present or discuss findings about the usability of the specific systems we have evaluated with our field laboratory (these can be found elsewhere). Instead, the purpose of mentioning these studies here is to illustrate how they functioned as vehicles for iterating on the field laboratory's configuration.

The chapter begins with a short summary of related work motivating the development of techniques for improving evaluation data collection in the field. We then describe three iterations of developing our own field laboratory. For each of these iterations, we describe our initial motivations and aims, the corresponding configuration of equipment, an example evaluation where it was used, and the pros and cons identified. The next iteration then describes how we modified the field laboratory configuration accordingly, and what we learned from using it in practice. Finally, we describe the current setup, outline some future trends within this area of research, and conclude on the work presented in the chapter.

2. BACKGROUND

In the proceedings of the first workshop on Human-Computer Interaction (HCI) for Mobile Devices in 1998, researchers and practitioners were encouraged to investigate further into the criteria, methods, and data collection techniques for usability evaluation of mobile systems (Johnson 1998). Of specific concerns it was stated that traditional usability laboratory setups would not adequately be able to simulate the context surrounding the use of mobile systems and that evaluation techniques and data collection methods such as think-aloud, video recording or observations would be extremely difficult in natural settings. These concerns have since been confirmed through a number of studies such as (Brewster 2002, Esbjörnsson et al. 2003, Pascoe et al. 2000).

In 2003, a literature study revealed that 41% of mobile HCI research involved evaluation (Kjeldskov & Graham 2003). However, even though evaluations of mobile systems were clearly prevalent, only 19% of these evaluations were carried out in the field while 71% were carried out in laboratory settings. Although the issue of how to study and evaluate mobile technology use and usability in the field has since received increased attention, no established set of usability evaluation methods and data collection techniques yet exists for field evaluations.

The research into field-based evaluations of user interfaces for mobile technologies can be divided into two overall categories of equal importance. The first category focuses on the methodological challenges of adapting traditional usability evaluation methods such as the use of the think-aloud protocol, as well as developing new ones, to suit the challenges and prospects of evaluating mobile user interfaces in the field. The second category focuses on the practical challenges of improving existing techniques for data collection in field settings and developing new ones. In this chapter, we focus on the latter: how to facilitate data collection better when evaluating user interface design for mobile technologies in the field.

One of the primary sources of data when evaluating the usability of an IT system is video and audio recordings of use depicting the system, the users' interaction with it, and the context in which this takes place. When evaluating in the field, the primary challenge of data collection is that these recordings can be very hard to make at a sufficient level of quality. Video filming evaluation sessions in the field with a handheld camcorder is seemingly an attractive approach because it is cheap and easy (figure 1 left). However, while suitable for capturing the overall use context of a field evaluation, capturing good close-up views of mobile device screen, buttons and user interaction can be quite difficult while moving (Kjeldskov et al. 2005). Furthermore, filming a good overview of a use situation with a handheld camcorder require a bit of distance while obtaining good close-ups and good sound requires that the cameraman stay relatively close to the test subject and interviewer. The latter often results in the so-called "bodyguard effect" (figure 1 right) where the test subject is practically isolated from other people in their surroundings, hence questioning the value of going into the field in the first place (Kjeldskov & Stage 2004).

Figure 1. Usability evaluations of mobile technologies in the field using a handheld camcorder and note taking for data collection.

Within the "practical" category of improving data collection techniques for evaluations in the field, three specific approaches are particularly worth mentioning.

One approach has aimed at obtaining field data in a non-intrusive way through automatic logging of user interaction for later analysis. Through logging, researchers can accurately record a user's interaction with a system, such as clicks or keyboard entries, or even record the entire graphical user interface of the software being evaluated. One of the advantages of logging is that it does not necessarily require the presence of a test monitor, and involves a minimum of interference with the user's context. This makes logging particular useful for longitudinal studies of mobile technology usability. Logging is also an efficient method to obtain data in a cost effective way from a large population of users. One of the drawbacks of logging is that it does not usually record any information outside the mobile device. It provides no record of, for example, the physical surroundings of the user, and it does not record so-called "near-interactions" where, for example, the user fails to interact with the system (Waterson et al., 2002). Another quite significant limitation is that logging usually requires installation of dedicated software on the device being evaluated. This is not only cumbersome but also sometimes simply not possible. While highly suitable for generating large amounts of data for quantitative studies, logging does usually not provide good data for qualitative studies. A way of overcoming this limitation could be to combine automatic logging with, for example, video and audio recordings, interviews etc.

Another approach has aimed at bringing traditional laboratory setups into the field by means of a "portable usability laboratory" or "lab-in-a-box" (Kimber et al., 2005; Winters et al., 2001). The advantage of a portable laboratory is that it allows rich data to be collected using high quality equipment. Not being truly mobile, portable usability laboratories are, however, best used in field settings where the user remains semi-mobile within a delimited spatial area for a period of time – for example in a restaurant or on the bridge of a ship. Other drawbacks for this approach are that the equipment is often cumbersome to transport and setup and may be intrusive in the context (Rowley, 1994). Setting up large amounts of video and audio recording equipment in the field may also cause users and surrounding people to act differently, which, in essence, stand diametrically opposed to the purpose of evaluating usability in the field. As a final downside, it may be difficult to record video of users' interaction with a mobile device with standard camera equipment.

Taking its offset in the challenges of using portable laboratories in the field, a third approach has been aimed at developing more compact and mobile usability laboratory facilities that are able to record high quality video data from various sources in an un-intrusive way. Different configurations of such mobile usability laboratories have been described in recent literature (e.g. Betiol & Cybis, 2005; Kaikkonen et al., 2005; Roto et al., 2004) and demonstrated at leading conferences within the field (e.g. Nyyssönen et al., 2002). The typical setup of a mobile laboratory makes use of a mini camera that can be attached to a mobile device for a good close up of the screen and user interaction. In some setups, such as the one proposed by Roto et al. (2004), additional cameras are used to capture views of the evaluation context. Images from these cameras are then mixed and recorded for later playback during analysis. While mini-camera approaches like this are highly promising – not only in field evaluations but also in laboratory settings – experiences from the deployment of mobile usability laboratories in the field also point out a series of challenges. Some of the issues relates to the quality of video and audio recordings when using wireless equipment, and how to best record multiple video sources and audio in sync. Other issues relate to battery lifetime and the weight of the equipment having to be carried around during the evaluation sessions.

In the following sections, we outline how we have dealt with these and other challenges through three iterations of setting up and using a field laboratory for in situ evaluations of mobile technologies.

3. CLOSE-UP VIDEO AND IMPROVED SOUND

Motivated by the challenges of capturing high-quality video data during usability evaluations in the field described in the literature and experienced in a series of evaluations carried out between 2002 and 2003, we decided to develop a portable configuration of audio and video equipment that could be carried by the test subject and an observer during a field evaluation.

Our primary focus for the first version of our “field laboratory” was to enable close-up recording of the mobile device screen and user interaction. Inspired by commercially available products, such as the “mobile device camera” from Noldus (Noldus, 2005), we constructed a small camera-mount on which a mobile phone or PDA could be mounted with Velcro (figure 2 left). The camera-mount contained a wireless camera mounted on a flexible “gooseneck” as well as a 9v battery-supply. This allowed us to capture a detailed close-up view of the mobile device in colour (figure 2 right) and record this throughout the whole evaluation. Apart from recording close-up video of the mobile device, we also wanted to improve the sound quality of our data recordings to minimize ambient noise and ensure capturing all utterances made by the test subject and the interviewer. For this, we combined the camera on the mobile device with an off-the-shelf professional wireless microphone from Sennheiser; a lapel microphone with a belt-pack transmitter worn by the test subject and a belt-pack receiver carried by the observer.

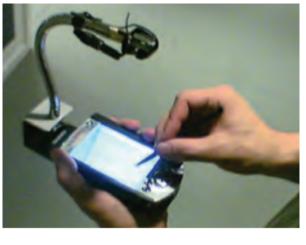


Figure 2. PDA on camera-mount allowing for close-up view of screen and user interaction.

Video from the camera on the mobile device and audio from the lapel microphone is transmitted wirelessly to receivers and recording equipment carried by an observer (figure 3). In the observers bag, the video and audio signals are recorded on a portable DV recorder, for example a camcorder, set up to record from an external source. During the evaluation, the observer can monitor the user‘s interaction with the mobile device on a small LCD screen and monitor the sound through earphones.

Figure 3. Observer (left) carrying and operating portable audio/video equipment (right) for capturing close-up view of screen and user interaction.

3.1. Using The First Field Laboratory in Practice

We used the first version of our field laboratory described above for an evaluation of a mobile information system in situ in 2003/04 (Kjeldskov et al. 2004). The evaluation focused on the use of a mobile, context-aware, electronic patient record system by nurses and doctors at a large regional hospital in Denmark. Six test subjects (all females) aged between 25 and 55 years participated in the field evaluation. They were all trained nurses with 1-9 years of professional experience.

Due to the real-life nature of the study, the field evaluation did not involve any researcher control in form of task assignments but was structured exclusively by the work activities of the nurses. The studied work activities were highly mobile, and involved interaction with assigned patients in different wards (i.e. collecting and reporting scheduled measurements), and moving back and forth between different rooms and hallways. As in a standard usability evaluation, the test subjects were given a brief instruction to the mobile system being evaluated and were encouraged to think aloud when possible. Each evaluation session lasted 15 minutes on average and involved three people. One nurse used the system for carrying out her work activities. One researcher acted as interviewer and asked questions for clarification while in the hallway. A second researcher operated the field laboratory. In addition, each session involved a number of hospitalized patients in their beds. For ethical reasons, we did not film the hospitalized patients. In order to be able to include a suitable number of different nurses as test subjects, the field evaluation took place over two days.

3.2. Lessons Learned From Using Field Lab #1

The field evaluation at the hospital highlighted a series of the challenges related to evaluating mobile technologies in situ. It was highly time consuming and complex to plan and execute the study, and it was difficult to capture key situations of use. However, in relation to data collection, the camera on the mobile device provided us with high-quality close-up views of the nurses' interaction with the system being evaluated, while at the same time allowing them to move around freely in the environment and focus on their work. The use of a professional wireless microphone supplemented the video close-up recordings with a clear audio track capturing all the nurses' utterances as well as enough ambient sounds to give a sense of context. During the later analysis phase, these video and audio recordings were invaluable sources of data for identifying usability problems and suggesting opportunities for redesign. The video track allowed us to see exactly which parts of the system were perceived as problematic and where the nurses had problems with operating the interface. The audio track allowed us to hear the nurses' comments about their interaction with the system and provided us with context of use. When evaluating mobile technologies in a laboratory, this kind of data is very much standard. The first version of our field laboratory made it possible to capture this kind of data in situ as well. It was lightweight, and it was relatively easy to operate.

On the downside, the first version of our field laboratory also had a number of limitations. First of all, the video recording only contained the close-up view of the mobile device and the user interactions taking place within 5-6 centimetres of the screen. It did not capture the users or their surroundings. During the data analysis phase, this proved to be very problematic at times where the use context was significant for understanding what the user was trying to do with the system. It was also hard to tell from the video track when the users were looking at the screen of the mobile device and when they focused elsewhere during the evaluation. Although the audio track did provide some information about context and the focus of the users, this information was often partial, ambiguous, and not conclusive. Secondly, the audio track only captured the voice of the interviewer if he or she was standing close to the test subject (who was wearing the microphone). In a stationary evaluation setup, this would usually not be a problem because the interviewer and test subject will be seated close to each other. However, when evaluating in the field it is most likely that interviewer and test subject will sometimes be physically separated by enough distance for directional microphones not to be able to pick up the voice of them both. In the field evaluation at the hospital this was often the case simply because the nurses were sometimes hard to keep up with by the interviewer and because the interviewer would sometimes have to stand back a bit in order not to interfere with the nurses' work tasks (i.e. attending to patients in bed). Thirdly, the mini camera was far from perfect. Although it was considerably smaller than commercially available alternatives, the gooseneck camera-mount clearly influenced the form factor of the mobile device being evaluated. It was too heavy, and made it impossible for users to hold the device the way they would usually do.

4. SMALL CAMERAS AND MULTIPLE VIDEO SOURCES

On basis of the lessons learned from the field evaluation at the hospital, we set out to improve our field laboratory in three ways. Firstly, we wanted to reduce the influence of the wireless camera attached to the mobile device being evaluated. We wanted to minimize the size and weight of the camera, and make it more flexible for use with different types and sizes of mobile devices. Secondly, we wanted to facilitate data recording from multiple sources of video allowing us to capture close-up views of the mobile device, close-up views of the user, 3rd person views of the user in context, and 1st person views of the surroundings as seen in, for example, Roto et al. (2004). Thirdly, we wanted to be able to capture audio from multiple sources (e.g. test subject and interviewer).

Minimizing the size and weight of the camera on the mobile device turned out to be surprisingly simple while at the same time also increasing its flexibility. Our solution was to simply strap the camera house on to a small plastic clamp with a flexible piece of plastic and a few cable strips. All items necessary to produce the wireless "camera-clamp" were purchased from a local hardware store for less than 20 USD. The clamp made it possible to mount the camera on almost any mobile device without interfering with its form factor (figure 4). The 9v battery powering the camera was simply attached to the mobile device with double-sided tape, wherever it would interfere the least with the user's grip of it. Using the same approach, we created other variations of the camera-clamp. One was also clipped-on to the mobile device but faced the camera towards the user (figure 4 right). Another one was designed to sit on the user's ear (like a Bluetooth headset) capturing a first-person view of the surroundings. These additional wireless cameras allowed us to capture video data from multiple sources in parallel.

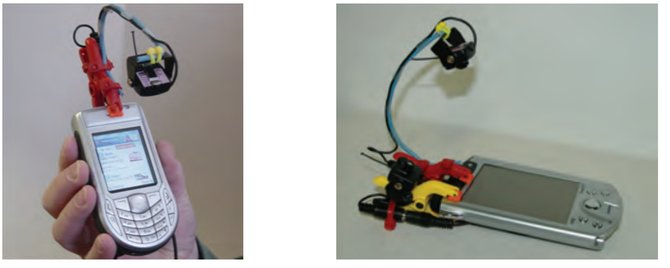


Figure 4. Lightweight camera-clamps attached to mobile devices

In order to capture a third-person view of the evaluation session, we decided to equip the observer with a handheld video camcorder. For better audio capture, we added a second wireless lapel microphone for the interviewer.

Figure 5. Equipment used for the second field laboratory (configured for two wireless cameras). Batteries and power regulators are not shown

While reducing the size and weight of equipment carried by the test subjects (even though we added more cameras), the addition of more cameras and microphones significantly increased the equipment necessary to be carried and operated by the observer (figure 5). The additional lapel microphone required an additional belt-pack receiver. For each additional wireless camera we had to add another video receiver and 12v battery. In order to include the video signal from 3-4 different sources in one composite video recording, we had to include some sort of battery driven video mixing as well. For this purpose, we modified a stationary Panasonic WJ-MS 424 Quad display unit to run on batteries. In order to minimize the number of different batteries in use and avoid batteries running flat at different times, we custom-built a power supply, which could power all equipment from the same 12v battery source (apart from the camcorders which ran on their own batteries).

4.1. Using The Second Field Laboratory in Practice

The second field laboratory setup described above was used in pilot studies preparing for a large-scale evaluation of a mobile information system in situ in 2005. The aim was to facilitate data collection about the use and usability of a context-aware mobile web site used by pairs of friends while socializing "out on the town". Hence, it was important to document both peoples' interactions with the device, with each other, and with their physical surroundings.

Figure 6. Example recording with multiple video sources

4.2. Lessons Learned From Using Field Lab #2

The pilot field evaluations in the city centre of Melbourne once again highlighted the complexity of evaluating mobile technologies in situ. However, this time we clearly got more out of our efforts to move from the laboratory into the field. The second version of our field laboratory made it possible to capture multiple video and audio sources in situ. As in the evaluations at the hospital, test subjects could move around relatively freely, and were undisturbed by the cameraman who could easily keep a distance of 5-8 meters while still capturing good images and sound. As we had aimed for, the second field laboratory provided rich data of high quality capturing both detailed views of the users, their interaction with the device, and their surroundings from several perspectives (figure 6). During the later analysis phase, especially the third-person view of the users in context provided an invaluable resource for contextualizing peoples' verbal utterances and their interaction with the system. Unlike our early evaluations in the field, where we were only using a handheld camcorder, the field laboratory allowed the cameraman to remain focused on the surroundings rather than having to zoom back and forwards between a third-person view and a close-up view of the mobile device. The use of two microphones resulted in a stereo audio track, which very clearly captured all utterances by test subjects and interviewer. Recording two separate audio tracks made it easy to separate between utterances made by different people during playback for analysis.

It also made it possible to make post-evaluation adjustments of the relative levels of peoples' voices.

On the downside, however, it only took us two pilot sessions in the field before we realized that the current setup of the field laboratory had a series of fundamental problems and needed to be modified. While we were able to capture great data like never before, the cost of this was very high in terms of battery life, weight, and complexity of operating the equipment needed. We had been able to fit all the field laboratory equipment depicted on figure 5, as well as the necessary batteries and power supply regulators, into a large laptop bag with internal cabling. However, the total weight of the bag exceeded 10 kg, which turned out to be physically challenging for the cameraman to carry for more than a few hours. At the same time, the modified Quad display splitter and the video receivers ran the battery-pack of four 12v motorcycle batteries flat in less than 1.5 hours. In effect this made back-to-back evaluation sessions impossible without recharging or carrying extra batteries with us into the field as well! While running all equipment on the same 12v power supply reduced the task of monitoring and replacing a lot of individual batteries for, for example, the audio receivers, we also found that the power regulators needed for doing this introduced noticeable noise to the audio recordings. Finally, the amount of equipment and the number of different video and audio sources made it highly complex for one person to operate the field laboratory in the (already) stressful conditions of an evaluation in situ.

On top of these problems, the number of different wireless technologies involved at this stage also resulted in problems with radio interference between equipment operating on the same or close frequencies. While we had no problems whatsoever with the professional wireless microphones, wireless video from multiple cameras turned out to be problematic. Camera signals sometimes interfered with each other, as well as with the wireless capabilities of the mobile device being evaluated (WLAN and Bluetooth). In fact using more than one wireless camera at a time sometimes completely disrupted the PDAs WLAN connection making parts of the evaluation impossible to carry out. At other times, the use of Bluetooth significantly distorted the images from some wireless cameras. Dealing with the problem of radio interference was quite a challenge. While we were to some extend able to modify our own use of wireless technologies during the evaluations to avoid problematic combinations of Bluetooth, WLAN and the wireless cameras, evaluating in the field of course made it impossible for us to control *other peoples'* nearby use of wireless technologies, which sometimes interfered with our equipment.

On the bright side, however, revisiting the field recordings quickly made it evident that collecting data from four independent video sources was not necessary in order to get a sufficient view of users, use, and context. The only sources we made any significant use of during the analysis of the evaluation sessions were the close-up view of the device and the third-person view of users and context. Hence, we could reduce our equipment.

5. MINIMIZING EQUIPMENT AND INCREASING BATTERY LIFETIME

Informed by the lessons learned from the pilot field studies described above, we made some significant changes to the field laboratory with the aim of minimizing equipment, reducing weight and complexity, and increasing battery life.

Our first major decision was to reduce the number of video sources to two: a wireless camera attached to the mobile device and a handheld camcorder operated by an observer. Reducing the number of wireless cameras limited the issue of radio interference and allowed us to make some significant reductions in the equipment to be carried by the observer. Firstly, the number of video receivers could be reduced correspondingly. Secondly, we were able to replace the battery-hungry Quad display unit with a much smaller Picture-in-Picture unit running on 12v (drivedata DPIP1). In return, these reductions made it possible to phase out a few heavy power regulators and run the field laboratory for almost 4 times longer on half the batteries. Replacing the wireless audio receivers with newer and more lightweight models (Sennheiser ew100 G2), we were also able to phase out an audio preamplifier and noise generating power regulators while at the same time improving the sound quality. We also replaced our portable tape-based DV recorder with a smaller and more lightweight 100GB AV hard disk recorder (Archos AV400). The third generation of our field laboratory is configured as schematically depicted in figure 7.

Video signals from the wireless camera attached to the mobile device is sent to a receiver in a small bag carried by an observer where they are mixed on the fly with a third-person view of the users captured by the handheld camcorder. Ensuring high-quality sound, users and interviewer are wearing small directional wireless lapel microphones. Mixed video and sound is recorded digitally on a hard disk recorder in the observer's bag. This configuration of the field laboratory weights approximately 4 kg, measures 26x18x30 cm, and has a battery time of approximately 5 hours on two 12v batteries (figure 8).

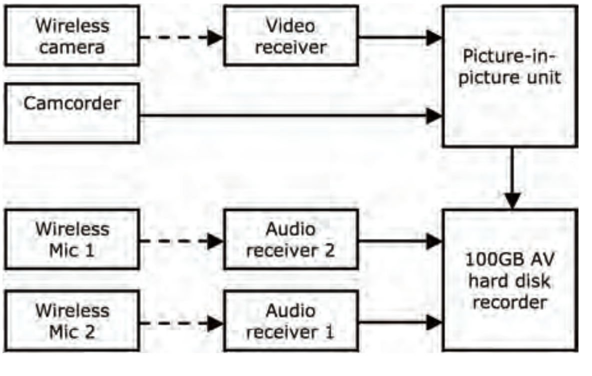


Figure 7. Schematic configuration of the current version of the field laboratory with two video sources and two audio sources recorded in one composite digital file.

5.1. Using The Third Field Laboratory in Practice

The third field laboratory setup described above was used in a large-scale evaluation focusing on the use and usability of a context-aware mobile web site facilitating sociality in the city centre of Melbourne, Australia (Kjeldskov & Paay, 2005). The field evaluation involved 20 people (grouped in pairs). All pairs of users were familiar with the location at which the evaluation took place and frequently socialized there together (figure 9).

Figure 8. The third field laboratory in a medium-sized light-weight camera bag

Figure 9. The field laboratory in action at Federation Square, Melbourne, Australia

With the purpose of being true to the real-life qualities of studying mobile technology use in situ, the field evaluations were not structured by tasks in a traditional usability evaluation sense of the term. Instead, the evaluations were structured by a set of overall prompts for use of different parts of the system and a list of corresponding interview questions. The socializing activities studied were highly mobile and involved the users moving between several physical locations in the city; bars, cafés, museums, etc. Prior to the evaluation, the users were given a 10-minute introduction to the system and were allowed to familiarize themselves with it for 5-10 minutes. Inspired by the constructive interaction approach to thinking-aloud studies with more than one user, the groups were asked to talk among themselves about their perception of and interaction with the system interrupted only with questions for clarification. The evaluation sessions each lasted between 45 and 70 minutes and took place over several days.

5.2. Lessons Learned From Using Field Lab #3

The field evaluation of the mobile web site was very successful. With the third iteration of our field laboratory, we had reached a very useable and stable solution with a good trade-off between supported data sources, weight and battery lifetime. We were able to capture audio and video sources needed for studying the use and usability of mobile technologies in situ, and were able to do so in a quality that matched (and even sometimes superseded) our stationary usability laboratory.

Figure 10. Video recording with third-person view of participants and close-up view of PDA. Note that the camera focused on the device screen is turned 90 degrees to optimize use of the Picture-in-Picture view.

Figure 11. Our most recent version of the field laboratory weighing only 2 kg and measuring just 18x14x25 cm - containing video and audio receivers, Picture-in-Picture unit, hard disk recorder, and battery.

The field laboratory was small, light-weight, relatively simple to operate and had a battery lifetime allowing for 2-3 evaluation sessions in a row without worrying about recharging (at this point the weakest link was in fact the battery lifetime of the PDAs used to run the prototype system). It allowed the observer to effortlessly follow the participants and interviewer from a bit of a distance while filming them and their surroundings with the handheld camcorder. In turn, this allowed the interviewer to focus on the participants' use of the mobile system being evaluated without having to worry about data collection. Figure 10 shows an example of the video data recorded in the field.

While the third version of our field laboratory was already considerably smaller and lighter than any of our earlier ones, we have since been able to reduce the weight and physical size further through a fourth iteration of reducing cabling, battery supply, and optimizing the use of bag-space (figure 11). In our most recent design (version 4), the field laboratory has the same specifications for data capture as described above, but now weights only 2 kg and measures only 18x14x25 cm, making it highly mobile and very easy to bring into the field for longer periods of time. Powered by only one 12v battery, this configuration can operate for approximately 2.5 hours.

6. FUTURE TRENDS

The future trends for developing field laboratories for evaluating mobile technology use and usability in situ focus primarily on improving the quality, reliability, and size of the cameras attached to the mobile device. As wireless video technology matures and becomes more widespread, we are likely to see an emergence of cheap high-end wireless video cameras matching the professional standard of the wireless microphones used in our current version of the field laboratory. Broadcast quality interference-free wireless video technologies exist today, but are still rather expensive and not sufficiently lightweight for our purposes.

Coming from another area of application, new camera technologies are also emerging within the field of video surveillance, which would allow video signals to be transferred digitally via wireless network connections rather than over an analogue radio link. Apart from offering much higher quality and stability, this approach is particularly interesting because it bypasses the use of any analogue video equipment, which is typically quite battery intensive. It also enables the development of field laboratories where all video sources are recorded digitally in separate, time stamped tracks avoiding the down-sampling of Picture-in-Picture and allowing for synchronised playback of multiple camera angles without any loss of quality.

A third emerging way of dealing with the camera problem is to replace it with a software solution that logs screen images from the mobile devices, or replicates them on a laptop or stationary computer via a network connection and then grabs the images from there. However, as discussed earlier in the section about automatic data logging, this approach does not capture the user-interaction with the physical device and situations where, for example, input is not registered by the system. Nevertheless, parallel data logging of the mobile device screen could be a very interesting way of complementing video and audio data captured through wireless cameras and microphones and should be investigated further. In a similar way, capturing video and audio data of user interaction could be an interesting way of enhancing the use of data logging when evaluating mobile technologies in the field.

7. CONCLUSIONS

In this chapter, we have described the iterative development of a field laboratory facilitating in situ evaluations of mobile technology use and usability. We have described a series of initial motivations, how we responded to these, and the lessons learned from deploying our field laboratory to a series of evaluations.

It is hard to evaluate mobile technologies in situ. It is difficult to capture key situations of use and it is complicated to collect data of an acceptable quality. However, by means of a field laboratory with small wireless cameras and wireless microphones, we have shown that it is possible to capture field data about the use and usability of mobile technologies in a quality that matches that of a stationary usability laboratory. Furthermore, we have shown that field laboratories can be made small, lightweight, and operational for hours before having to recharge batteries. Equipped with a field laboratory as the one described in this chapter, we believe that researchers and designers will be able to make more and better evaluations of user interfaces for mobile technology in the field.

REFERENCES

Betiol, A. H., & Cybis, W. D. A. (2005). Usability Testing of Mobile Devices: A Comparison of Three Approaches. In *Proceedings of Interact 2005* (pp. 470-481). LNCS: Springer-Verlag.

Brewster, S. (2002). Overcoming the Lack of Screen Space on Mobile Computers. *Personal and Ubiquitous Computing, 6*(3), 188-205.

Esbjörnsson, M., Juhlin, O., & Östergren, M. (2003). Motorcyclists Using Hocman – Field Trials on Mobile Interaction. In *Proceedings of Mobile HCI 2003* (32-44). LNCS: Springer-Verlag.

Johnson, P. (1998). Usability and Mobility; Interactions on the move. In *Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices*. Glasgow, Scotland: GIST Technical Report G98-1.

Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., & Cankar, M. (2005). Usability Testing of Mobile Applications: A Comparison between Laboratory and Field Testing. *Journal of Usability Studies, 1*(1), 4-16.

Kimber, J., Georgievski, M., & Sharda, N. (2005). Developing Usability Testing Systems and Procedures for Mobile Tourism Services. In *Proceedings of Annual Conference on Information Technology in the Hospital Industry*. Los Angeles: Informational Technology Association.

Kjeldskov, J., & Paay, J. (2005). Augmenting the City: The Design of a Context-Aware Mobile Web Site. In *Proceedings of DUX 2005* (pp. 1-7). San Francisco, CA, USA: ACM.

Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S., & Davies, J. (2005). Evaluating the Usability of a Mobile Guide: The influence of Location, Participants and Resources. *Behaviour and Information Technology, 24*(1), 51-65.

Kjeldskov J., Skov M. B., Als B. S., & Høegh R. T. (2004). Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In *Proceedings of Mobile HCI 2004* (pp. 61-73). LNCS: Springer-Verlag.

Kjeldskov, J., & Stage, J. (2004). New Techniques for Usability Evaluation of Mobile Systems. *International Journal of Human-Computer Studies*, *60*, 599-620.

Kjeldskov, J., & Graham, C. (2003). A Review of Mobile HCI Research Methods. In *Proceedings of Mobile HCI 2003* (pp. 317-335). LNCS: Springer-Verlag.

Noldus (2005) Mobile Device Camera. Retrieved March 2, 2007 from http://www.noldus.com/site/doc200402054

Nyyssönen, T., Roto, V., & Kaikkonen, A. (2002). Mini-camera for usability tests and demonstrations. Paper presented at demo sessions of Mobile HCI 2002, Pisa, Italy, September 2002.

Pascoe, J., Ryan, N., & Morse, D. (2000). Using While Moving: HCI Issues in Fieldwork Environments. *Transactions on Computer-Human Interaction, 7*(3), 417-437.

Roto, V., Oulasvirta, A., Haikarainen, T., Kuorelahti, J., Lehmuskallio, H., & Nyyssönen, T. (2004). *Examining Mobile Phone Use in the Wild with Quasi-Experimentation* (HIIT Technical Report 2004-1). Helsinki, Finland: Helsinki Institute for Information Technology.

Rowley, D. (1994). Usability Testing in the Field: Bring the Laboratory to the User. In *Proceedings of CHI 1994* (pp. 252-257). ACM Press.

Waterson, S., Landay, J. A., & Matthews, T. (2002). In the Lab or Out in the Wild: Remote Web Usability Testing for Mobile Devices. In *Proceedings of CHI 2002* (pp. 796-797). ACM Press.

Winters, J. M., Story, M. F., Campbell, S., Lemke, M., Danturthi, S., Barr, A., & Rempel, D. M. (2001). Mobile Usability Lab: A Tool for Studying Medical Device Accessibility for Users with Diverse Abilities. In *Proceedings of RESNA 27th Annual Conference.*

KEY TERMS AND THEIR DEFINITIONS

Field laboratory – a configuration of laboratory equipment, such as video and audio recording devices, put together so that it can be taken into to field for data collection about the use and usability of mobile technologies in situ.

AV hard disk recorder – a video unit that records an external video and audio source directly onto a hard disk in a digital format that can be played back on a computer. The video recording is typically compressed when it is recorded resulting in manageable file sizes.

Picture-in-Picture unit – a video unit that inserts a video image over a part of another one. The inserted video image is rescaled and thus loses a bit of quality in the process.

Quad display unit – a video unit that merges four different video signals into one composite signal. All four video images are rescaled and thus lose a bit of quality in the process.

Lapel microphone – a small microphone that can be clipped on to a person's collar or revere. The microphone is usually connected to a small transmitter that can be carried in a pocket or clipped on to the belt.

Camera-clamp – a tiny camera that can be clipped on to a mobile device such as a PDA or a mobile phone. Camera-clamps can be either cabled or wireless. The latter require a battery supply and a video receiver.

Third person view – a video recording of the user(s) of a mobile device and their immediate surroundings during an evaluation of use or usability from the perspective of a third person observing from a distance.

Close-up view – a video recording of the screen and buttons o a mobile device, such as a PDA or mobile phone, during an evaluation of use or usability. Usually captured with a mobile devices camera attached to the device.

Part IV

Artefacts

ARTEFACTS

Part IV addresses the question *how can we make use of context in the implementation of concrete interactive mobile systems?* In order to create actual interaction design artefacts, we need to know what is technologically possible now and in the near future, and we need to know how emerging technologies can be used for pushing this frontier further. Five of my own contributions to this are included in chapters 15-19. These chapters present our experiences with the technical implementation of five prototype systems.

MobileWARD

Chapter 15 presents MobileWARD. In this chapter we explore the construction of an interactive context-aware mobile prototype system for the healthcare domain. The system deploys a combination of active and adaptive context-awareness by automatically pushing information and functionality to the user filtered on the basis of their location, current work activities and people nearby. The chapter presents the details of the prototype system and results from an empirical study of its use. It concludes that context-aware mobile information systems hold potential value within the healthcare domain as a component of a ubiquitous computing environment supporting the mobile and distributed nature of work activities in this domain. However, the implementation of interaction designs for such systems is highly complex and must be carefully thought out and evaluated in order to ensure a good fit between systems, users, and their context.

Just-for-Us

Chapter 16 presents Just-for-Us. In this chapter we explore the construction of an interactive context-aware mobile prototype system for socialising in the city. The system primarily deploys passive adaptive context-awareness by filtering user requested content and functionality on the basis of their context. Just-for-Us was an early attempt at making mobile context-aware systems web-based. The aim of this technical approach was to explore system alternative architectures that would allow such mobile systems to benefit from the fast paced development of new programming facilities for the mobile web. The chapter concludes that the web-based approach holds interesting potentials for the creation of highly dynamic and graphical mobile context-aware applications, but that mobile web browsers and programming environments of the time lacked a series of capabilities for handling dynamic exchange of information between clients and servers.

GeoHealth

Chapter 17 presents GeoHealth. In this chapter we explore the construction of a web-based location-based service for home healthcare workers distributed over a large geographical area. The system combines active and passive context-awareness depending on importance of information, and combines adaptation and mediation of context through a mesh-up of information from various sources on an interactive map. Extending directly from the work presented in chapter 16, the GeoHealth system explores the powers of Web 2.0 technologies in combination with GPS positioning, Google Maps and the Mobile Internet. The chapter presents the prototype system in detail followed by results from an empirical study of its use. It shows that mobile location-based services built around Web 2.0 technologies and interactive maps have unexploited potentials for mobile interaction design within the domain of home healthcare.

ArchiLens

Chapter 18 presents ArchiLens. In this chapter we explore the construction of a mobile augmented reality system for architectural visualization. The system deploys passive context-awareness and mediates contextual information by allowing the user to explore the visual and spatial characteristics of their future house in context. This is done by overlaying the 3D scene onto the live images from the phone's built-in camera. The ArchiLens system was implemented as an application for the Android operating system, making use of its powerful graphics engine mainly designed for interactive 3D games. The chapter presents the prototype in detail followed by results from a study of use with 40 participants. It show that the 3D capabilities of modern mobile devices in combination with their contextual sensors and built-in high quality video cameras have strong potentials as a platform for the design of mobile interactions.

Power Advisor

Chapter 19 presents Power Advisor. In this chapter we explore the construction of an interactive mobile system to promote sustainability by allowing people to monitor their domestic electricity consumption and adjust usage behaviour accordingly. The system deploys active and mediated context-awareness by pushing information about the household's electricity consumption to the user as a resource for interpretation and exploration. This information is collected wirelessly from a "smart" power meter unit. The Power Advisor system was implemented as a mobile web application allowing it to be used on Android and iOS enabled devices. The chapter presents the prototype system in detail followed by results from a study of use in 10 households over a period of 7 weeks. The findings provide insight into peoples awareness of electricity consumption in their home and how this may be influenced through interaction design of mobile systems.

Chapter 15

MobileWARD

Jesper Kjeldskov and Mikael B. Skov

Abstract. Ubiquitous technologies have potentials to serve major roles in different real world organizational settings. One of the areas where applying ubiquitous technologies has been given a lot of attention is in the healthcare domain. Here, users are frequently on the move while at the same time relying increasingly on centralized computerized information. In this paper, we explore ubiquitous technologies in the real world through two studies in the healthcare domain. Firstly, we look at the use and usability of a ubiquitous Electronic Patient Record (EPR) system distributed on desktop and laptop computers throughout a large hospital. Secondly, we present an extension to this ubiquitous computing environment in the form of a context-aware mobile computer terminal prototype. The usability of the mobile EPR prototype was evaluated in both laboratory and field settings. Our results indicate that the usefulness of a ubiquitous computing environment supporting work activities in healthcare can benefit from context-aware mobile information access. However, interaction design for such systems must be carefully thought out and thoroughly evaluated. Also, while the use of mobile and stationary computers complement each other very well, we found that the usefulness of ubiquitous computing environments in healthcare may benefit from additional elements such as situated displays at key locations and on key objects, and from seamless integration between the different devices of the system as a whole.

1. INTRODUCTION

Over the past years, emerging computer technologies have drawn enormous attention as they often yield new and innovative use in work as well as in leisure. We are currently on the move away from traditional desktop-based computer technologies towards ubiquitous computing environments that will potentially enfold us in almost everyday situation and activity. We encounter these computing environments everywhere: in our

Originally published as Kjeldskov, J. and Skov, M. B. (2007) Exploring Context-Awareness for Ubiquitous Computing in the Healthcare Domain. *Personal and Ubiquitous Computing,* 11(7), 549-562.

homes, cars, work places, shops, restaurants, cinemas etc., and thus such computing environments have to accommodate several different use situations and user groups. Consequently, ubiquitous computing environments have received immense attention from both academia and industry in order to explore their promising opportunities, apparent limitations, and experienced implications for interaction design. Because the use of ubiquitous computing environments is often closely related their contextual settings, one of the avenues of research which has concerned human-computer interaction researchers and practitioners is the ability of ubiquitous computing environments to explore context-awareness in interaction design.

In this paper, we present our experiences with two ubicomp environments to support nurses and doctors in conducting their work activities within a healthcare domain. Firstly, we explore the use of an existing ubiquitous computing environment at a large hospital in the form of a commercial Electronic Patient Record (EPR) system distributed on a series of desktop and laptop based computers throughout the organization. Secondly, we explore the design and use of and an experimental ubiquitous computing environment extending the existing ensemble of technologies with a mobile context-aware component facilitating a higher degree of pervasive and nomadic use.

The paper is structured in the following way. First, we discus a series of issues concerning emerging ubiquitous technologies and present some of the current experiences with information technologies in healthcare reported in the literature. Secondly, we present our initial usability study of a ubiquitous computing environment in healthcare. The findings from this study are then outlined as a series of high-level themes describing problems and advantages encountered during use. Thirdly, we present the design of an experimental context-aware mobile Electronic Patient Record system, supplementing the existing ubiquitous computing environment. Fourthly, we show how this prototype system was evaluated in a laboratory as well as in field settings at the hospital, and outline the primary findings from these evaluations. Finally, we discuss the findings from the evaluations of the mobile EPR component prototype in the light of the themes identified in our initial study and discuss context-aware interaction design for ubiquitous computing environments.

2. RELATED WORK

The diversity of users and use situations makes it challenging and difficult to design ubiquitous computing environments. Designers have to pay attention to several issues in the use domain if the computing environment is to become useful and successful. Furthermore, only limited practical experiences with the design and use of ubiquitous computing environments for the real world are reported in the literature.

2.1. Ubiquitous computing environments and usability

As instances of ubiquitous technologies, distributed terminals, mobile and handheld technologies etc. have the potentials to serve major roles in different organizational contexts in the future because they will provide users the ability to access information and services when away from their desktop (Green et al. 2001). Within a multitude of work domains, this may potentially lead to new ways of working, as people will be able to conduct or perform work activities different from what they are presently able to. However, designing and implementing these types of information systems will also be highly challenging and difficult. Research studies show that ubiquitous and mobile technologies continue to challenge our existing body of knowledge on analysis, design, implementation and evaluation of information systems (Barkhuus and Dey 2003, Crabtree et al. 2003). While such challenges are not unique for any specific emerging technologies (Skov 2002), ubiquitous, and mobile technologies may exhibit novel and unprecedented complexity for interaction design as user interaction with such technologies will be continuous and pervasive (Barkhuus and Dey 2003).

It is generally considered of great importance for designers to consider the future use situation when designing and implementing software information systems. Naturally, this is also true for ubiquitous computing environments. However, for ubiquitous computing environments, the traditional focus of effectiveness and efficiency of software use may not be applicable, suitable, or desirable in the same way as we are used to. Instead, designers have to look broader at the ensemble of activities in the physical world (e.g. work activities), interactions between people in the physical world, and their use of technology. Thus, some of the key research problems in relation to the design of ubiquitous computing environments have become to understand the everyday character of the environment (Crabtree et al. 2003) and to design systems that conform, not disrupt, the natural workflow of the user (Green et al. 2001). Few studies provide suggestions on how to achieve a smooth interaction between the user and a ubiquitous computing environment. Exceptions count, for example, Barkhuus and Dey (2003), who examines three levels of interactivity for context-aware mobile systems.

Rubin (1994) argues that usability evaluation of software systems is an efficient and well-documented approach to understanding and classifying the interaction between a user and a software system. Usability evaluations can be utilized to identify problems in the interaction design, and can potentially inform designers about extent to which their software product is useful (Molich 2000). In recent years, evaluating the usability of *emerging* technologies has also become vital from a business perspective as indicators of potential success of failure of new technologies within an area typically associated with considerable financial investments and risk for technology providers and manufacturers (Nielsen 1993, Rubin 1994). Combining this with the growing complexity of new technology, identifying fundamental usability problems with, for example, ubiquitous computing environments may prove significant for informing successful design and implementation of such systems. Furthermore, as users become more diverse in terms of, for example, skills, motivation, and experience, obtaining a high level of usability of new technologies becomes a substantial challenge to designers and businesses, as their they have to accommodate this great diversity.

2.2. Ubiquitous computing in healthcare

Ubiquitous computing environments can potentially influence and change work practices within a multitude of different settings in the healthcare domain dramatically. Healthcare work, for example in hospitals, is typically characterized by very complex and specialized procedures in which information technology may contribute to improved performance, reduction of errors made in treatment of patients, reduction of economical costs etc. Different types of software systems are currently being introduced into many hospitals and other parts of the healthcare domain. Typically these systems are connected through various types of computer networks and are widely dispersed throughout the physical organization. Hence, they can be classified as ubiquitous computing environments.

Of particular interest, a lot of resources are being put into the implementation and use of Electronic Patient Record (EPR) systems. EPR systems collect information about the history of treatment performed on the patients admitted to a hospital. The hospital personnel use the patient record to diagnose diseases, and to document and coordinate treatment. Within the last 20 years, a considerable amount of effort has been devoted to the development of Electronic Patient Record systems. The primary motivation for this effort is that unlike paper-based patient records, electronic patient records will be accessible to all relevant persons independent of time and location.

The design of Electronic Patient Records systems is a huge challenge for our community, raising a wide range of still unanswered questions related to issues such as screen layout, interaction design, and integration into work processes. Where should the systems be located and who should enter the data? How do we make sure that input is complete and accurate? How are the different work processes in healthcare structured and coordinated? What is the most useful way of displaying and accessing the vast quantity of patient data? (Consolvo et al. 2002). In the light of these questions, a lot of research has been published in the literature about EPR systems and how to meet challenges related to design and use of computer system in healthcare. Specifically, much attention has been given to issues such as information sharing (Grimson and Grimson 2000), support for cooperation (Kaplan and Fitzpatrick 1997) and privacy (Rindfleish 1997). While much of this research is based on studies on the use of traditional paper-based patient records, suggesting viable electronic counterparts, however, little research has been published based on studies that inquire into the use of the mass of EPR systems already in use.

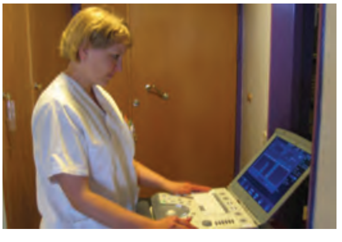



Figure 1. Ubiquitous EPR use in the healthcare domain

In 2002, the Danish government decided that all Danish hospitals must have replaced the traditional paper-based patient records with electronic patient records by 2005. However, it was up the regional authorities to decide on the details of deployment. Thus, a number of projects were conducted within this period of time with the aim of developing and evaluating electronic patient record systems (see for example the work of Bardram et al. 2003). In relation to a regional Danish research program entitled "The Digital Hospital" we have studied the use and usability of a commercial EPR system currently constituting the backbone of a ubiquitous computing environment at a large regional hospital (IBM IPJ 2.3). In addition to this, we have designed, implemented and evaluated an experimental mobile EPR prototype terminal extending the current system's functionality and scope. Driving this study, we were concerned with the following research questions:

1. What challenges and potentials characterize the use and usability of the ubiquitous EPR system currently in use at the hospital?
2. How can the use and usability of the ubiquitous EPR system be improved through context-aware mobile information access?

The following sections describe the details of our initial usability evaluations of the EPR system currently in use and subsequent experimental design process of a mobile context-aware counterpart.

3. STUDY A: UBIQUITOUS EPR IN THE REAL WORLD

As our point of departure, we conducted a longitudinal usability evaluation of IBM's electronic patient record system IPJ 2.3 currently in use throughout the Hospital of Frederikshavn, Denmark. The aim of study A was to enquire into the use and usability of ubiquitous Electronic Patient Records in relation to carrying out typical work activities at a regional hospital prior to national implementation of such systems. The two main screens of the evaluated EPR system are illustrated in figure 2 and 3.

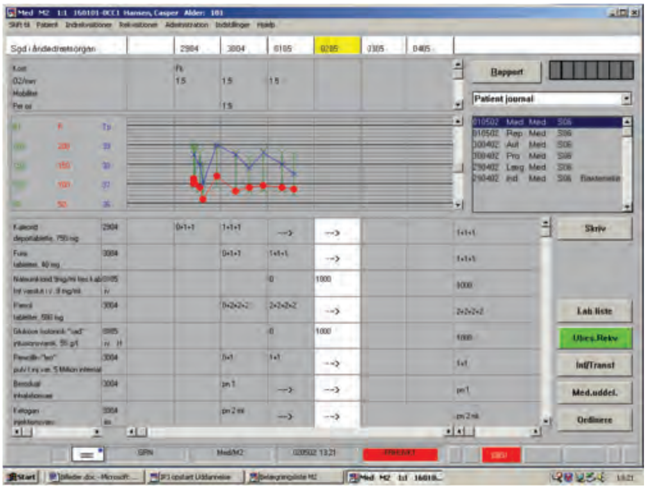


Figure 2. Screen shot from IPJ 2.3 illustrating the primary information, for example, previous measured temperatures, heart rates, on a specific patient

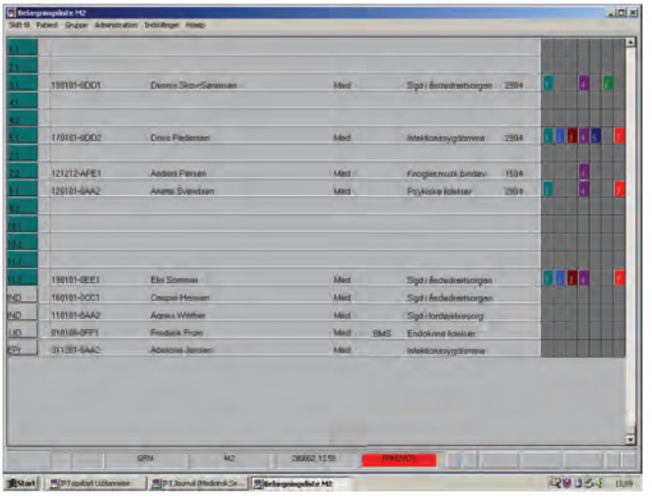


Figure 3. Screen shot from IPJ 2.3 showing the central list of patients on the ward with overall information about diagnosis and treatment.

3.1. Method

The evaluation of the EPR system was carried out in a dedicated usability laboratory at Aalborg University over a period of little more than one year. In order to avoid just getting a snapshot of the use and usability of the system, we conducted a longitudinal study involving two separate evaluations with the same users. The first evaluation was conducted over two days in May 2002 shortly after the system had been taken into use at the hospital. The second evaluation was conducted over two days in August 2003 after the system had been used extensively in daily work at the hospital for more than a year. In total, 16 evaluation sessions of approximately 1 hour were carried out.

The evaluation involved two parts:

1. A think-aloud evaluation with the purpose of identifying usability problems in the electronic patient record system.
2. A series of interviews with the purpose of gaining insight into the integration of the electronic patient record into the work of the nurses.

The details of the evaluation are described below.

Preparation

Prior to the first evaluation, the research team inquired thoroughly into the work activities at the hospital related to the use of patient records. This was done through observations at the hospital and interviews with key personnel. At the same time, the research team also investigated into the specific functionality of the Electronic Patient Record system to be evaluated. Based on this, the research team and key personnel at the hospital assigned to the implementation of the Electronic Patient Record System in the organization developed a series of use scenarios and a number of tasks for the evaluation in collaboration. The use scenarios and tasks went trough several iterations of pilot testing and modification before the evaluation took place.

Test Subjects
The evaluation involved eight trained nurses from the Hospital of Frederikshavn. All nurses were women, aged between 31 and 54 years with professional work experience ranging from 2 to 31 years (at the time of the first evaluation). Prior to the first evaluation, all nurses had attended a course on the IPJ 2.3 system amounting to between 14 and 30 hours and were just beginning to use the system in their daily work. They characterized themselves as novices or beginners in relation to the use of IT in general. At the time of the second evaluation, they had all used the system extensively in their daily work for more than a year and now characterized themselves as system experts. All test subjects participated in both the 2002 and the 2003 evaluations except for one person who had left the hospital in the meantime. As a replacement, the hospital found another nurse who matched the characteristics of the original participant.

Tasks
The purpose of the usability evaluations was to inquire how the EPR system supports typical work activities at the hospital. Based on our scenarios we designed three tasks with a number of sub tasks centered on the core purpose of the system – such as retrieving information about patients, registering information about treatments, making treatment notes, and entering measurements.

Procedure
The evaluation sessions consisted of two parts: 1) a hands-on part where the nurses used the system to solve the assigned tasks, and 2) a semi-structured interview. The hands-on part of the evaluation was based on the think-aloud protocol as described in (Molich 2000, Nielsen 1993, Rubin 1994). If a test subject had problems with a task and had tried all the options she could identify, the test monitor provided her with help to find the solution. If a test subject was completely unable to solve a task, the test monitor asked her to go on to the next task. The hands-on sessions lasted approximately 45 minutes. After the hands-on sessions, four randomly selected nurses were interviewed about their work and their opinions about the system as well as its integration in and influence on their work. The interviews were semi-structured based on a list of questions and a number of issues that could be raised. The evaluation sessions were conducted over two days. One of the authors of this article was test monitor throughout all test sessions. The other author conducted the interviews in collaboration with a third researcher.

Setting
The evaluations were all conducted in a state-of-the-art usability laboratory at Aalborg University facilitating close-up observation of the test subject's interaction with the evaluated system. This involved the use of three different rooms: a subject room, an observation room and a control room. For the EPR evaluation, the subject room was equipped with a standard PC with a 19" screen and a standard mouse and keyboard matching the hardware used at the hospital. During the hands-on sessions, the test monitor observed the test subjects, encouraged them to think aloud and asked questions for clarification. Two additional researchers operated the video equipment and took notes in the separate control and observation rooms respectively.

Data collection
All evaluation sessions were recorded on digital video. Two remotely controlled motorized cameras captured overall and close-up views of the test subject and test monitor. The image on the PC screen was converted to composite video and mixed with the camera images to one composite video signal. The video recordings also contained audio tracks with the verbal utterances of the test subject and test monitor.

Data analysis
Following the evaluations a log file was produced from each test. Each of the three individual researchers then used the video recordings and these log files as empirical foundation for producing three individual lists of usability problems. These lists were then merged to one. The severity of the identified problems was rated as critical, serious or cosmetic based on the guidelines proposed by Molich (2000). The rating was done individually and followed by negotiation in cases of disagreement.

3.2. Findings from the evaluation of IPJ 2.3

The evaluation identified a substantial amount of usability problems with the current Electronic Patient Record system. The total number of identified usability problems was 75, distributed on 9 critical, 39 serious and 27 cosmetic problems. The 75 usability problems were related to various concerns of human-computer interaction, e.g. problems with finding and storing information, problems related to high complexity of screen layouts and problems with inconsistent interaction design. The 75 identified usability problems are described in detail in (Kjeldskov et al. 2002). At the same time, the usability test also identified a number of strengths in the interface of the Electronic Patient Record system, for example, easy registration of values, good integration with some existing work activities and good overview of the patients in the ward.

While the identified strengths and weaknesses provided valuable information for redesign and focused training, the usability problems in themselves provided little input for breaking out of the boundaries of the existing system. In order to do so, we took a step back from the list of usability problems and conducted a grounded analysis of observed interactions with the system and the statements from the interviews. From the grounded analysis, we identified three themes reflecting more abstract concerns of the usability of Electronic Patient Records:

1. Mobility
2. Complexity
3. Relation to work activities

These are described in more detail below.

Mobility
During the usability evaluations and in the interview afterwards, most nurses would stress concerns about being mobile while working with the system. Meeting this challenge, the use of laptop computers rather than desktop workstations had been suggested and discussed at the hospital. However, most of the nurses stated that they would find it impossible or unfeasible to carry a laptop computer around the ward

every time they were to conduct work tasks away from their office. One problem was the size and weight of the laptop, as they would also have to carry other instruments. As it turned out later, the issue of mobility was so immense that the hospital implemented a semi-mobile way of accessing the EPR system through laptops connected to a wireless network being wheeled around the hospital wards on trolleys.

Complexity
Another overall concern or problem was complexity and fragmentation of information. Most nurses found it difficult to locate the necessary patient information as they were solving tasks. This sometimes led to inadequate or incomplete task completion. Hence, the nurses would be insecure whether they had found the right information and whether they had succeeded in finding all relevant information.

Relation to Work Activities
Most nurses experienced problems with the use of the EPR system because they had difficulties relating the information in the system to their work activities. The problem was that they would typically use different kinds of information in real life to decide how to solve a problem, for example, visible conditions of a patient. Another concern related to the fact that the system only partially reflected the current work task, making it difficult to the test subjects to find or store information.

4. STUDY B: AN EXPERIMENTAL MOBILE EPR SYSTEM

Motivated by the findings from our evaluation of the Electronic Patient Record system, we carried out a field study into the work activities at the Hospital of Frederikshavn related to the use of Electronic Patient Records in practice. In concert with the findings from the usability evaluation outlined above, this suggested a design solution with the following characteristics.

1. Supporting the highly mobile work activities of nurses by being mobile /handheld.
2. Reducing complexity by adapting to its context: location, time, tasks
3. Eliminating double registering of information (first written down on paper and then entered into the desktop or laptop PC later) by being integrated with the existing patient record system.

While facilitating access to patient information at the 'point of care' is not a new idea (Arshad et al. 2003, Morton and Bukhres 1997, Urban and Kunath 2002), adapting information and functionality in a mobile EPR system to its context is a novel approach to improving the interaction design of such systems that has not yet been investigated thoroughly. On the basis of the findings from the usability evaluation and subsequent field study, an experimental prototype of a handheld context-aware EPR system for supporting the morning procedure, MobileWARD, was designed and implemented (Hansen et al. 2003, Høegh and Skov 2004). The system is described briefly below; a more detailed description can be found in (Skov and Høegh 2006).

4.1. Architecture

MobileWARD was designed for a series of Compaq iPAQ 3630 running the Microsoft® PocketPC operating system. The system uses a Wireless Local Area Network (wLAN) for network communication. The system was implemented in Microsoft embedded Visual Basic 3.0. For the first experimental prototype of MobileWARD, context-awareness was simulated by means of a "context control center" application. The control center runs on a separate iPAQ connected to the wireless network. Through this application, an operator can trigger "context events" in MobileWARD simulating that the user has entered a specific room, scanned the barcode on a specific patient etc. This approach was chosen to facilitate early evaluation of the experimental design solution without having to worry about the technical challenges of context sensing at the hospital before this had proven to be a viable approach from the user's perspective. In later versions of the system, real sensing of the environment can be implemented where found promising.

For discussions on how to sense environments see, for example, the work of Schilit and Theimer (1994).

4.2. Interface design

MobileWARD is designed to support work tasks during morning procedure at the hospital ward. The design is based on two basic concepts. First, the system is designed to reflect the context of the user in the sense that it is able to sense and react to a number of changes in the environment. Secondly, as the use of a pen for typing in information would sometimes be inappropriate because the nurses would often use the system while being mobile or engaged in other activities, the interface design incorporates a visual layout with large-scale buttons that enables finger-based interaction through the touch-screen of the iPAQ.

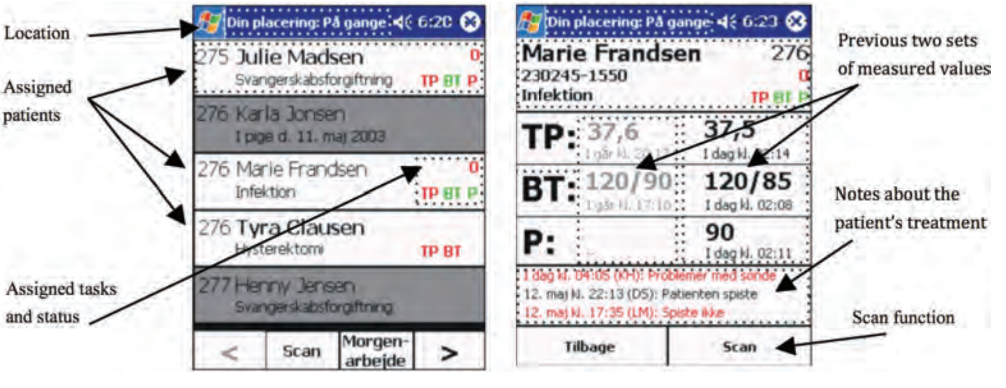


Figure 4. Interface layout from MobileWARD illustrating patients admitted to the hospital (left) and information on a selected patient (right).

MobileWARD is context-aware in the sense that the system recognizes the location of the nurse and presents information and functionality accordingly. Before visiting assigned patients, the nurses often want to get an overview of the specific information about each patient, for example, previous measured values. This typical takes place at the nurse's office or in the corridor. The windows related to these locations are shown in figure 5.

When located in the corridor, the system by default displays all the patients admitted to the ward. The patient list is ordered by ward number. Patients assigned for morning procedure are shown with a white background and the names of patients assigned to the nurse using the system are boldfaced (e.g. "Julie Madsen" and "Tyra Clausen" on figure 4). At the top of all windows, the nurse can see their current physical location as interpreted by the system. In the example on figure 4, the nurse is in the corridor. For each patient, MobileWARD provides information about previous tasks, upcoming tasks and upcoming operations. The indicators TP (temperature), BT (blood pressure) and P (pulse) show the measurements that the nurse has to perform. The indicators are either presented with red text (value still to be measured) or green text (value already measured). Above the three indicators, an "O" indicates an upcoming operation (within 24 hours), which usually requires that the patient should fast and be prepared for operation.

If the nurse wants to view data about a specific patient, she can click on one of the patients on the list. This will open the window shown on figure 4 (right), displaying the name and personal identification number of the patient, the previous two sets of temperature, blood pressure, and pulse measurements taken as well as written notes regarding the treatment of the patient. This window is accessible at any time and location. Thus the nurse can choose to look up more specific details about each patient while located in the corridor or in the office. In order to enter new data into the system, the nurse has to scan the barcode identification tag on the patient's wrist-band using the "scan" function in the bottom of the screen. This is described further below.

Figure 5. Screens displayed in the ward in relation to the tasks of measuring temperature, blood pressure, and pulse

The aim of this design is to provide the nurse with information that helps her plan the scheduled morning procedure. The system presents information and functionality adapted to the location of the nurse and the time of the day. Furthermore, the system knows the status of each patient and represents already measured values and values yet to be measured by simple color codes.

When the nurse enters a ward, the system automatically displays a different set of information. This is illustrated in figure 5. At the top of the screen, the nurse can see her physical location as interpreted by the system (e.g. ward 276). Below this, information about the patients on the current ward is presented, resembling the information available on the patient list displayed in the corridor, with the addition of a graphical representation of the physical location of the patient's respective beds. In this way, MobileWARD aims at presenting only relevant information to the nurse, e.g. by excluding patients from other wards. Like in the corridor, data about the patients is available by clicking on their names (or on their bed-icon). At the bottom of the screen, the nurse can activate the barcode scanner ("scan") used to identify a patient prior to entering data into the system.

After having scanned a patient, the nurse can type in measured values (figure 5, center). This window shows previous measurements of values and provides functionality for typing in new values. By clicking the new value button ("ny"), the system displays a window for entering new values (figure 5, right). Below the personal information (name and personal identification number), date and time is shown. In the gray box, the nurse can input the measured value by editing the shown value. This is done by pressing the large sized buttons on the screen with a finger. The number shown by default is the latest measurement. The reason for this is that the latest measure is most likely to be close to the newest one. If there is a huge difference, this is clearly signaled to the nurse, as she will have to perform more button presses than usual, providing an implicit opportunity to see whether, e.g. the temperature for a given patient is rising or falling. The Save button stores the value in the database along with the date, time, and user identification. Furthermore, the system updates the status of the task as having been carried out.

4.3. Usability evaluation of MobileWARD

We evaluated the studied the use of the MobileWARD prototype through two separate studies conducted in the laboratory and in the field respectively. The first study was conducted in our usability laboratory with the objective to evaluate MobileWARD in a controlled environment where we could assign the nurses to specific tasks and closely observe their use of the system. In addition to this, we also wanted to the laboratory evaluation to involve mobility and context. In order to achieve this, we modified the standard laboratory setup in a number of ways. The second study took place at the Hospital of Frederikshavn. The aim of this evaluation was to study the usability of MobileWARD for supporting real work activities at a hospital setting involving real nurses and real hospitalized patients. In order to achieve this, we adopted an observational approach combined with questions for clarification while the nurses were not directly engaged in conducting their work. The details of the two studies are described below.

Setting

The usability laboratory was set up to resemble a part of the physical space of a hospital department (figure 6). This included the use of two separate evaluation rooms connected by a hallway. Each of the evaluation rooms was furnished with beds and tables similar to real hospital wards. From a central control room, the evaluation rooms and the hallway could be observed through one-way mirrors and via remotely controlled motorized cameras mounted in the ceiling.

The field evaluation was carried out at the Medical Department at the Hospital of Frederikshavn (figure 7). This included the physical area of seven hospital wards, an office with reception, a rinse room and a break-out area connected by a central hallway and involved nurses at work and patients committed to the hospital.

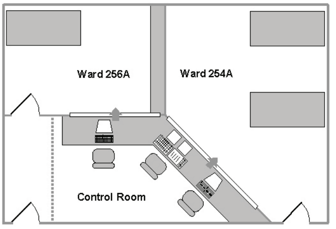


Figure 6. Physical layout of the lab set up to emulate a section of the hospital ward

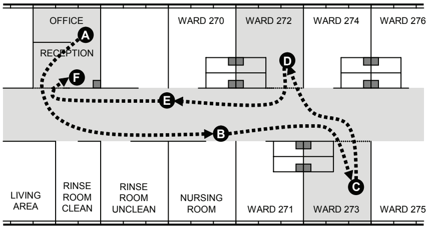


Figure 7. Physical layout of the hospital ward used for the field evaluation

Data collection

High quality audio and video data from the laboratory and field evaluations were recorded digitally. In the laboratory, a tiny wireless camera was clipped on to the mobile device (figure 8), providing us with a close-up view of the screen and user-interaction. This was then merged with video signals from the ceiling-mounted cameras (figure 9).

Figure 8. Wireless miniature camera mounted on PDA

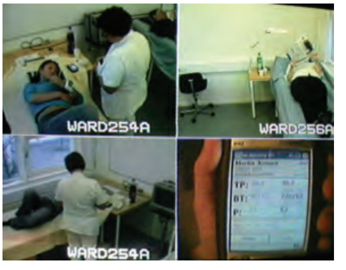


Figure 9. Video recording from usability evaluation

Motivated by the challenges of capturing high-quality video data during usability evaluations in the field, we designed and built a portable configuration of audio and video equipment to be carried by the test subject and an observer, allowing a physical distance of up to 10 meters between the two. The configuration consists of a tiny wireless camera (also used in the laboratory evaluation described above) clipped-on to the mobile device (figure 8) and a clip-on microphone worn by the test subject. Audio and video is transmitted wireless to recording equipment carried by the observer (figure 10).

Figure 10. Observer (left) carrying and operating portable audio/video equipment (right) for capturing high-quality data in the field.

In the test monitor's bag, the video signal from the clip-on camera can be merged with the video signal from a handheld camcorder (Picture-in-Picture) and recorded digitally. This allows us to record a high-quality close-up view of the screen and user-interaction as well as an overall view of user and context. During the evaluation, the observer can view the user‘s interaction with the mobile device on a small LCD screen and monitor the sound through earphones. For ethical reasons, we were not permitted to film the hospitalized patients.

Test subjects

12 test subjects participated in the evaluations. Six test subjects (four females and two males) aged between 28-55 years participated in the laboratory study whereas six test subjects (all females) aged between 25-55 years participated in the field evaluation. All test subjects were trained nurses employed at a large regional hospital and had between 2 and 36 years of professional experience. All subjects were mobile phone users but novices with the use of handheld computers. All test subjects were frequent users of a stationary electronic patient record system and described themselves as experienced or semi-experienced users of IT.

Tasks

All test subjects in the laboratory were given tasks to solve while using the system. The tasks were derived from an ethnographic study at a hospital ward and covered the duties involved in conducting standard morning work routines. This involved 1) checking up on a number of assigned patients based on information in the system from the previous watch, 2) collecting and reporting scheduled measurements such as temperature, blood pressure, and pulse, and 3) reporting anything important for the ongoing treatment of the patients should be taken into consideration on the next shift. The field evaluation did not involve any researcher control in form of task assignments but was structured by the work activities of the nurses in relation to conducting standard morning work routines.

Procedure

Before the evaluation sessions, the test subjects were given a brief instruction to the system. This included the room-sensing functionality and the procedure for scanning patients' bar-code tags. The test subjects were also instructed on how to operate the available instruments for measuring temperature, blood pressure and pulse. In the lab sessions, the evaluation sessions were structured by the task assignments. The tasks required the test subjects to interact with all three patients in the two hospital wards, and to move between the two rooms through the connecting hallway a number of times. The nurses were encouraged to think aloud throughout the evaluation explaining their comprehension of and interaction with the system. The lab evaluations lasted between 20 and 40 minutes. In the field sessions, the evaluation sessions were not structured by tasks but by the actual work activities of the nurses. This involved interaction with three patients in different wards and moving between different rooms through the connecting hallway a number of times. The nurses were encouraged to think aloud when possible. The evaluations lasted 15 minutes on average. In order to be able to include a suitable number of nurses, the field evaluation took place over two days. After all evaluation sessions, the test subjects filled out a brief questionnaire.

Roles

Each evaluation session involved six people. One nurse used the system for carrying out the assigned tasks or real work activities depending on the experimental setting (laboratory or field). One researcher acted as test monitor and asked questions for clarification. A second researcher operated the context-control centre and the video equipment. In addition, each evaluation session involved either three students acting as hospitalized patients (in the laboratory) or three real hospitalized patients in their beds (in the field). Due to the real-life nature of the study that took place at the hospital, each field evaluation session involved different patients.

Data analysis

The data analysis aimed at creating two independent lists of usability problems identified from the two experimental settings. The usability problems were classified as cosmetic, serious or critical according to Molich (2000). The two usability evaluations amounted to approximately 6 hours of video recordings depicting the 12 test subject's use of the system. All sessions were analyzed in random order by two teams of two evaluators where each team analyzed the videos in a collaborative effort allowing immediate discussions of identified problems and their severity. As a guideline for the collaborative analysis, each identified usability problem would be discussed until consensus had been reached. The two teams produced two lists of usability problems. Subsequently, these were merged into one complete list. Again, this was done in a collaborative effort, discussing each problem and severity until consensus had been reached.

5. FINDINGS AND DISCUSSION

We identified a total of 37 different usability problems from the 12 laboratory and field sessions. 8 problems were assessed to be critical, 19 problems were assessed to be serious, and 10 problems were assessed to be cosmetic. Our study showed that the laboratory setting revealed more usability problems than the field setting. The six test subjects in the lab experienced 36 of the 37 usability problems whereas the six test subjects in the field setting experienced 23 of the 37 usability problems. 14 usability problems (1 critical, 9 serious, 4 cosmetic) were unique to the lab setting, whereas one serious usability problem was unique to the field. Regarding the identified critical problems, the lab setting identified all 8 critical problems and the field setting identified 7 critical problems. Considering the serious problems, the lab identified eight additional problems compared to the field.

Looking qualitatively at the usability problems encountered, these can be divided into three categories, concerning interaction, mobility, and context-awareness.

Firstly, usability problems arose from nurses having problems with interacting with the system or understanding the interface. For example, one nurse did not know the semantics of the keyboard for typing in textual notes about patients. Specifically, she felt insecure about the buttons "Tab", "Caps" and "Shift" as she would expect them to be "tablets", "capsules" and "shift medication". Secondly, usability problems arose from aspects of mobility and working conditions. For example, one nurse was concerned about putting the mobile device in her pocket. She was afraid that she would accidentally click some buttons while walking and she stated that it would be impossible to carry the device in her hand at all times. Another problem related to mobility and working conditions was the fact that one nurse feared that the device could spread bacteria from patient to patient. Thus, she did not want to place the device on the patient's bedside table or on the bed. Finally, the studies revealed seven usability problems related to the context-aware element (all encountered in both conditions). These problems were primarily related to confusion among the nurses when the interface "suddenly" changed contents when, for example, walking into a ward. Typically, this would make the users either confused or annoyed - especially if reading information on the screen at the time of the automatic update. Surprisingly, however, all six field test subjects (but only one lab subject) did not understand *why* the system would automatically update information and functionality according to the physical location. So even though their use situation was in-situ and closely related to the context, they still got confused about the system being actively context-aware. Analyzing this result, we find that their reluctance towards the automatic-update element in the mobile device may stem from the consequently decreased lack of control. Operating and working in a safety-critical environment like healthcare, the decreased level of control may not appear to support systematic work practices, but merely to compromise the work activities. The feeling of lack of control is well-known to active context-aware mobile system and should probably be investigated further.

In the following sub sections, we revisit the three issues of mobility, complexity and relation to work activities encountered in the study of the commercial ubiquitous EPR system in the light of the findings from our context-aware mobile counterpart. We then take a step back and discuss some general implications for ubiquitous computing emerging from our study.

5.1. Mobility revisited

Issues of mobility are crucial in many activities for nurses acting in a professional environment. Nurses would normally find themselves visiting patients in different physical locations and they often require different kinds of information for dependent and independent work tasks. The aspects of mobility in our study can be considered local mobility, as described in (Bardram et al. 2003), and therefore the nurses would normally not require directional guidance from the system. Thus, we attempted to support the local mobility through a relatively small, handheld device that could be carried around by the nurses (potentially in their pockets) while visiting patients or conducting other work tasks. The idea of having a mobile device was appreciated by all nurses in our evaluation. However, we found that the nurses would continuously switch between reading or storing information on the device and conducting work tasks without the device, for example, taking measurements from patients. Thus, holding the device in their hands all the time would be impossible and therefore they would occasionally need to put it away or lay it down. This caused problems to most of the test subjects as they did not know where to put the handheld device. As a consequence, some of them requested functionalities allowing them to lock the screen. Others questioned the general usefulness of handheld devices.

5.2. Complexity revisited

The first study identified another important issue with electronic patient records namely issues concerned with complexity and fragmentation of information. Most subjects experienced problems in locating relevant and adequate information in the traditional electronic patient record. This could be a result of many different circumstances, but one of the problems was the extensive amount of different types of information on each screen (figure 2). The nurses would occasionally fail to notice relevant or even critical information on, for example, patients and scheduled operations. As a result, more subjects failed to solve all assigned tasks in the study. To address this problem we aimed at presenting much less information at a time on the mobile device by exploiting context-awareness to, for example, only presenting information about patients close by. Validating this approach, the nurses encountered no severe complexity problems when using the mobile device. However, they would occasionally request more information than could be fitted into the screen at one time.

5.3. Relation to work activities revisited

As a final issue from the two usability evaluations, we discovered that nurses would typically require very specific information based on current work tasks and activities. The traditional electronic patient record did not fully support this but presented too much, too little, or too fragmented information. In the mobile EPR prototype, we utilized context-awareness in different ways as a mean for determining the work task of the nurses. However, this also introduced some pitfalls as nurses would sometimes miss reminders presented on the screen because their focus was engaged elsewhere. Furthermore, some nurses became confused or even annoyed by the automatic adaptation of information on the screen to their physical location. Thus, the use of context-awareness was not experienced as universally useful and further research into issues such as user control in interaction design with such systems is clearly needed.

5.4. General implications for ubiquitous computing in the real world

Taking a step back from the specific findings from our empirical studies at the hospital wards, a series of general implications for ubiquitous computing in the real world emerge. Overall we find that mobile computing in the healthcare domain – whether context-aware or not – is not an *alternative* to the use of ubiquitous computing environment consisting of networked desktop and laptop terminals situated throughout the environment. Rather, mobile access to such ubiquitous computing system has potentials to *supplement* information access from stationary (and semi-mobile) terminals strategically situated in the working environment. In the real world, mobile systems are only one of many components of a truly useful ubiquitous computing system. While based on a study of specific work activities in a specific real world domain, we believe this will also apply generally to other organizations in which workers are required to be mobile within a (relatively) limited physical area while at the same time dependent on access to a large amount of shared information. Mobile access to patient record information at a hospital ward is useful for nurses in many situations because their work often require them to move between different physical locations. In these situations, the nurses usually only require very specific information related to their current physical location, task at hand or patient under treatment. Hence, handheld device automatically adapting to these contextual factors have great potentials for adding to the usefulness of the ubiquitous computing environment – provided that specific and usability issues related to context-awareness (such as user control) are carefully taken into consideration in their interaction design. At the same time, however, stationary access to patient record information via conventional computer terminals situated throughout the hospital wards is also very important. To use Kristoffersen and Ljungberg's notion of mobility (1999), the mobility of the work activities in the hospital wards does not just include "wandering" from place to place but also "visiting" specific key location for longer periods of time (such as offices, consultation rooms, etc.). When accessing patient data from one of these key locations, more detailed information is often sought for than when standing at the patient's bed – including looking through the patient's history of treatment, medication, treatment notes etc. Also, this is very often where more detailed notes and reports on treatments are entered into the system. Hence, traditional PC terminals with larger screens and better input devices for browsing information and entering text than offered by both portable (laptops) and mobile (PDAs) is by far a preferable approach.

In summary, we found that the use of both stationary and mobile terminals at the hospital complemented each other very well in response to the three identified issues of mobility, complexity, and work relation.

Having said that the use of both stationary and mobile terminals complemented each other very well, however, we still believe that ubiquitous computing environments in the healthcare domain (and in similar domains) could be improved much further. While the combination of mobile context-aware and stationary context-independent access to the shared resource of patient information accommodated for mobile, nomadic, and stationary work, reduced complexity of information access without removing the ability to access complex information, and related patient information more closely to work activities, it is our impression that there is still a huge potential and relevance for additional technologies in between the two. In the case of the hospital, nurses and doctors are not the only ones who are mobile. So are patients, beds, and medical equipment. In fact, some of these are already often associated with highly specific situated information, which is currently not linked in with the Electronic Patient Record systems. Patients have wristbands with written information and sometimes carry printouts of subset of their patient record allowing not only medical staff, but also administrative staff, quick overview of, for example, upcoming medicals. Beds are equipped with printouts of information about blood pressure, temperature etc. for their associated patients, again allowing for easy (and implicitly context-related) access to key information without having to interact with a mobile device. Equipment sometimes has written notes attached to it about how to operate it and who to contact in case of malfunction. Whereas the traditional PC terminals provide centralized access to information about all patients and the mobile context-aware terminals provide access to a subset of this information adapted to the user's context (location etc.), information could also be provided through the environment itself in form of for example, situated displays located on locations, objects and people of importance.

Furthermore, we find that the use of PDAs and PCs (as well as situated displays etc.) as points of access to a ubiquitous computing system should not be seen in isolation from each other. Rather it should be acknowledged that users of a ubiquitous computing system are most likely to sometimes use different points of access such as PDAs, PCs, etc. in combination to solve a given task at hand, and that they will frequently shift backwards and forwards between these. In response to this, real world ubiquitous computing systems in the healthcare domain (and domains alike) should strive for seamless integration of their different elements, and allow users to apply and appropriate their combined functionality in a highly flexible manner – as also described in (Bardram et al. 2003). Information must be easily portable between devices, and it should be easy to shift from one device to another in the middle of a task without having to start over from scratch. As a simple example of this, time spent on browsing complex information hierarchies in the Electronic Patient Record system could be significantly limited if when having accessed information about a specific patient on your PDA, you could immediately direct other terminals, such as a PC in the office or a laptop in the ward, to the same place in the records and vise-versa.

6. CONCLUSIONS

We have explored the use of ubiquitous computing in the real world through a series of studies in the healthcare domain. Supporting work activities in healthcare is a highly complex and challenging task and the healthcare domain is a potential candidate for advanced ubiquitous computer systems. In response to this, we have conducted a study over two overall phases. First, we identified important challenging for supporting work tasks in healthcare through the evaluation of a ubiquitous electronic patient record system at use at a large hospital. Secondly, we designed, implemented, and evaluated a mobile extension of this ubiquitous EPR system addressing identified challenges of mobility, complexity and relation to work activities by utilizing context-awareness as a key means for supporting the nurses' interaction with the EPR system. Our results show that workers in the healthcare domain can benefit from ubiquitous computing environments and that ubiquitous computer environments in the healthcare domain may be improved through mobile and context-aware points of access. However, our studies also confirm that the design of ubiquitous computing systems for the real world needs a lot of further investigations. Also, even though our findings showed that context-awareness can be applied as a useful means of exploring mobility of healthcare workers to reduce complexity of information and improve the relation between information in the EPR system and the nurses' work tasks, we found that context-awareness is a very difficult style of interaction to master, raising serious new challenges in relation to, for example, user control. Context-awareness has huge potentials for ubiquitous computing environments but should not be seen as a universally useful paradigm of interaction design.

The use of mobile and stationary terminals in a ubiquitous computing environment compliments each other very well for the types of work activities studied in this research. At the same time, however, we speculate that the useful of such ubiquitous computing system in the real world would benefit further from additional means of in formation access. This could, for example, be in the form of small, situated displays located at key locations and on key objects in the working environment. In line with related research, we also speculate that the usefulness of ubiquitous computing environments comprising of an ensemble of different devices such as mobile terminals, desktop PCs, laptops and situated displays would benefit from seamless integration between these devices and services, including easy and flexible exchange of files, pointers to files, user identities etc, allowing for unforeseen user appropriation of ubiquitous technologies over time.

ACKNOWLEDGEMENTS

The authors thank all test subjects and the Hospital of Frederikshavn for kind collaboration throughout the project. The usability evaluations were conducted in collaboration with Jan Stage, Benedikte Skibsted Als and Rune Thaarup Høegh. MobileWARD was designed and implemented by Rune Thaarup Høegh, Karsten Kryger Hansen and Søren Lauritsen. This research is supported by the Danish Technical Research Council (projects 26-04-0026 and 2106-04-0022).

REFERENCES

Arshad, U., Mascolo, C. and Mellor, M. (2003) Exploiting Mobile Computing in Healthcare. In Proceedings of Demo Session of the 3rd International Workshop on Smart Appliances, ICDCS03

Bardram, J. Kjær, T. and Nielsen, C., (2003) Supporting Local Mobility in Healthcare by Application Roaming among Heterogeneous Devices. In Proceedings of Mobile HCI'03, Springer-Verlag (LNCS), pp 161-176

Barkhuus, L., and Dey, A. (2003) Is Context-Aware Computing Taking Control away from the User? Three Levels of Interactivity Examined. In Proceedings of the 5th International Conference on Ubiquitous Computing, Springer-Verlag (LNCS), pp. 149-156

Bohnenberger, T., Jameson, A., Krüger, A. and Butz, A. (2002) Location-Aware Shopping Assistance: Evaluation of a Decision-Theoretic Approach. In Proceedings of Mobile HCI'02, Springer-Verlag (LNCS), Berlin, 2002, pp. 155-169

Brinck, T. and York, G. (1998) User Interfaces for Computer-Based Patient Records. In Proceedings of CHI'98, NewYork, ACM

Consolvo, S., Arnstein, L. and Robert-Franza, B. (2002) User Study Techniques in the Design and Evaluation of a Ubicomp Environment. In Proceedings of the 4th International Conference on Ubiquitous Computing, Springer-Verlag (LNCS), pp. 73-90

Crabtree, A., Rodden, T., Hemmings, T. and Benford, B. (2003) Finding a Place for UbiComp in the Home. In Proceedings of the 5th International Conference on Ubiquitous Computing, Springer-Verlag (LNCS), pp. 208-226

Green, N., Harper, R., Murtagh, G. and Cooper, G. (2001) Configuring the Mobile User: Sociological and Industry Views. Personal and Ubiquitous Computing, 5(2) (2001) 146-156

Grimson, J. and Grimson, W. (2000) The SI Challenge in Health Care, Communications of the ACM, 43(6) (2000) 49-55

Hansen, K.K., Høegh, R.T., and Lauritsen, S. (2003) Making Context-Awareness (at) Work, Aalborg University, Aalborg

Høegh, R.T. and Skov, M.B. (2004) Exploring Context-Awareness as Mean for Supporting Mobile Work at a Hospital Ward. In Proceedings of the 7th International Conference on Work with Computing Systems (WWCS 2004)

Kaplan, S.M. and Fitzpatrick, G. (1997) Designing Support for Remote Intensive-Care Telehealth using the Locales Framework. In Proceedings of DIS'97, ACM

Kjeldskov, J., Skov, M.B. and Stage, J. (2002) Usability Evaluation of IBM IPJ 2.3 Electronic Patient Record, Aalborg University, Aalborg

Kristoffersen S. and Ljungberg F. (1999) Mobile Use of IT. In Proceedings of the 19th Information Systems Research Seminar in Scandinavia. Finland

Molich, R. (2000) Usable Web Design (In Danish). Ingeniøren | bøger

Morton, S. and Bukhres, O. (1997) Utilizing mobile computing in the Wishard Memorial Hospital ambulatory service. In Proceedings of SAC'97, ACM

Nielsen, J. (1993) Usability Engineering, Morgan Kaufmann

Rindfleish, T.C. (1997) Privacy, Information Technology, and Health Care. Communications of the ACM, 40(8) (1997) 93-100

Rubin, J. (1994) Handbook of Usability Testing, Wiley

Schilit, B.N. and Theimer, M.M. (1994) Disseminating Active Map information to Mobile Hosts. In IEEE Network, 8(5) (1994) 22-32

Skov, M.B. (2002) Design of Interactive Narratives: Concepts, Methods, and Architectures. Aalborg University, Aalborg

Skov, M. B. and Høegh, R. T. (2006) Supporting Information Access in a Hospital Ward by a Context-Aware Mobile Electronic Patient Record. In Crestani, F. et al. (Eds.) Personal and Ubiquitous Computing, Springer London, vol. 10, no. 4, pp. 205 – 214

Urban, M. and Kunath, M. (2002) Design, Deployment and Evaluation of a Clinical Information System which uses Mobile Computers and Workstations. In Proceedings of the 2nd conference on Mobiles Computing in der Medizin, Heidelberg.

Chapter 16

Just-for-Us

Jesper Kjeldskov and Jeni Paay

Abstract. Mobile computer technologies are increasingly being appropriated and used to facilitate people's social life outside the work domain. Addressing this emerging domain of use, we present the design of a context-aware mobile information system prototype facilitating sociality in public places: Just-for-Us. The design of the prototype system was informed by two empirical studies: an architectural analysis of a recently built public space in Melbourne, Australia and a field study of small groups socialising there. We describe these two studies and illustrate how findings informed our prototype design. Finally, we outline an ongoing field study of the use of the Just-for-Us prototype.

1. INTRODUCTION

Mobile computer technologies are increasingly being appropriated to facilitate people's social life outside the work domain. Mobile phones, and especially SMS texting, have changed the way people communicate, interact in the physical world, and coordinate their social activities (Grinter and Eldridge 2001, Rheingold 2003). Smart Phones and Personal Digital Assistants (PDAs) connected to the Internet bring access to web-based communities to the mobile user and extend the potentials of SMS through Internet-chat capabilities and facilities for video-based communication. By embedding networked sensors into the built environment, adding advanced positioning technology and short range network capabilities (such as Bluetooth, RFID tags, etc.), mobile services are emerging that adapt their content to both the user's physical and social contexts. For example, mobile dating services exist which alert the user when they are in the proximity of a potential partner who matches their own pattern of attributes (CNN 1998). As another example, swiping electronic membership cards at the entrances and exits of cafés, discotheques, music clubs, etc. in some Danish cities makes it possible for members of a social group to identify the whereabouts of their friends and other people

Originally published as Kjeldskov, J. and Paay, J. (2005) Just-for-Us: A Context-Aware Mobile Information System Facilitating Sociality. In *Proceedings of Mobile HCI 2005,* Salzburg, Austria (pp. 23-30). New York: ACM. (A condensed version appeared in IEEE Computer: Kjeldskov J. and Paay J. (2006) Public Pervasive Computing in the City: Making the Invisible Visible. *IEEE Computer,* 39(9), 60-65.)

through the mobile Internet, and to see which places in the city are currently busy and which are not (Hvem er i byen 2004). In the more experimental domain, context-aware mobile systems show the location of friends in vicinity (Fithian et al. 2003), take into consideration the user's social context when presenting event and tourist information (Kolari and Virtanen 2003), and create couplings between physical and virtual spaces by enabling people to attach text and media content to physical locations for others to find (Persson et al. 2002).

The emergence of systems like these represents a new trend of huge interest to the mobile HCI community: facilitating sociality through context-aware mobile devices. However, adapting mobile information systems to people's social context is not trivial and further research is needed into a series of fundamental questions. What should such systems do and how is the user's social context taken into consideration in interaction design in a way that makes sense and is useful to a user currently engaged in a situated social activity? Addressing some of these questions, this paper presents a case study example of how sociality can be facilitated by a context-aware mobile information system adapting to the user's physical and social context. On the basis of a field study, it is shown how people's situated social interactions in public places are complex and how the physical and social affordances of a place influence the situated interactions that occur there. Informed by these studies, the presented prototype exemplifies how sociality can be facilitated on mobile devices through information that is "Just-for-Us".

The paper is structured in the following way. Section 2 briefly introduces the concepts of sociality, social indexicality and Just-for-Us Information. Section 3 describes two empirical studies of physical and social context and outline key findings informing our prototype design. In section 4, we describe the interaction design and technical implementation of a mobile prototype just-for-us system, which adapts to the user's physical and social context. Section 5 briefly describes how we are currently studying the use of the prototype system through a series of field visits. Section 6 concludes and indicates further work.

2. SOCIALITY, INDEXICALITY AND JUST-FOR-US INFORMATION

Sociality is a term traditionally used by sociologists for describing the support for social interaction between people. In relation to interaction design, sociality is strongly associated with on-line interaction in virtual worlds but has also started to appear in relation to the use of mobile information and communication technologies in context. Gaver (1996) relates sociality and the physical world by defining "affordances for sociality" as being "possibilities offered by the physical environment for social interaction". Crabtree and Hemmings (2001) use the terms sociality and social interaction interchangeably while exploring the relationship between architecture, technology and social interaction in domestic space. In discussing the integration of technology and our interactions in urban spaces, Wittel (2001) correlates the terms of face-to-face sociality with face-to-face interaction.

Indexicality is a concept drawn from semiotics, applied to the design of mobile device interfaces to streamline and reduce the amount of information delivered to the user (Kjeldskov 2002, Paay and Kjeldskov 2005). Indexicality is a property of a representation that gives it a context-specific meaning and thus only makes sense in a particular setting. In applying indexicality to mobile human-computer interaction, information in the interface is indexed to information present in the user's immediate surroundings, such as, for example, signposts, buildings and roads. Social indexicality takes the idea of indexicality a step further, and makes reference to, for example, the user's current social group and their history of shared experience, such as where people usually meet or what they usually do when they go out together. Combining physical and social indexicality, we create information that is "just-for-us". It is not a compilation of all available knowledge about a place, delivered in a hierarchical or location based format, but a selection of that information that is relevant to the time, place, and the people involved in a situated social interaction.

3. EMPIRICAL STUDIES: PHYSICAL AND SOCIAL CONTEXT

With the purpose of investigating the role of physical and social contexts of people's situated social interactions we conducted two empirical studies at the newly opened public space of Federation Square, Melbourne, Australia. The first study was an expert inspection of the architectural and informational elements constituting the physical context at Federation Square. The second study was a series of contextual interviews made with people socialising in this space, studying their social interactions to understand the social context of the space. Federation Square was chosen for this study because it is a multi-modal public space with a mixture of distinct architectural features and embedded digital elements that provide a variety of activities to visitors. Federation Square has cafes, restaurants, shops, cinemas, galleries, gathering places and open spaces and functions as a small-scale city district (figure 1).

Figure 1. Federation Square, Melbourne, Australia

In the first study, methods were adapted from Urban Planner Kevin Lynch (1960) and Architect Christopher Alexander (1977). Lynch and Alexander both modelled built environments, specifically cities, with regard to the people that inhabit them, incorporating social theories and user needs into analyses of physical environments. Lynch (1960) developed a method for visual analysis of city precincts through descriptions of key aspects of physical space held by people as they navigate and orient themselves within cities, extracting an environmental image of a place. Alexander et al. (1977) empirically investigated the interplay between architectural space and its inhabitants and identified architectural design problems in context and their impact on inhabitants of that environment. Although situated within the field of architecture, the works of Lynch (1960) and Alexander et al. (1977) have also been explored within the field of human-computer interaction. For example, Dieberger and Frank (1998) use the work of Lynch to create a city metaphor for navigating complex information spaces. Crabtree et al. (2002) use the concept of Alexandrian patterns to inform the design of new technologies for use in domestic settings. Informing the design of web based systems for supporting social interaction, Erickson and Kellogg (2002) use Alexander and Lynch for exploring the relationship between physical spaces and social interaction. Within the field of mobile human-computer interaction, Lynch has inspired the design of several mobile guide systems such as (Goodman and Gray 2003, Kulju and Kaasinen 2002). The works of Alexander and Lynch also inspired the study of "familiar strangers" in urban settings reported by Paulos and Goodman (2004) and the design of the "Jabberwocky" personal mobile device facilitating social interaction between familiar strangers in public places.

Inspired by Lynch and Alexander and the use of their works in HCI, the aim of our first empirical study at Federation Square was to inquire into how architectural and informational elements of the built environment contribute to the visitor's experience of a public place. This involved observational expert audits in the field by a trained architectural observer recording the relationship between elements of the environment. Identifying the main characteristics of the space, content analysis (Neuman 1994) and affinity diagramming (Beyer and Holtzblatt 1998) were used to derive, group and refine categories from the collected data. The findings from the first study are described in detail in (Paay and Kjeldskov 2005a).

The aim of the second study was to enquire into the "use" of Federation Square by regular visitors as a place for socialising. In this study, McCullough's typology of everyday situations (McCullough 2001) was used as a starting point for a classification of the social activities of people associated with being out on the town: eating, drinking, talking; gathering; cruising; belonging; shopping; sporting; attending; and commemorating.

Figure 2. Contextual interview at Federation Square

On the basis of McCullough's typology, field observations were carried out using contextual interviews (Beyer and Holtzblatt 1998) and observational ethnographic techniques (Blomberg and Burrell 2003) with established social groups on location at Federation Square (figure 2). The participants in the second study were three different established social groups. Each group consisted of three young urban people, mixed gender, between the ages of 20 and 35, who had a shared history of socializing at Federation Square together. Prior to the field visits each group received a 10 minute introduction to the study followed by a 20 minute interview about their socializing experiences and preferences.

The participants were then taken to the nearby Federation Square, where they were asked to involve themselves in the kinds of interactions and activities they would usually do together when socialising out on the town. The participants were not given specific tasks but were asked to verbalize their actions and interactions and to respond to the interviewer's questions for clarification about things being said, and implicit decisions and interactions being made. An observer recorded the field visits on digital video. Three field visits were carried out, lasting approximately three hours for each group. The video recordings of the participants' situated social interactions were subsequently used at the foundation for a thorough grounded theory analysis (Strauss and Corbin 1990) of transcripts and affinity diagramming of themes to draw successively higher levels of abstraction. The findings from the second study are described in detail in (Paay and Kjeldskov 2005b).

4. THE DESIGN AND IMPLEMENTATION OF JUST-FOR-US

On the basis of the findings from the two empirical studies, we designed and implemented a high-fidelity context-aware mobile prototype system, Just-for-Us, facilitating sociality when out on the town. Just-for-Us keeps track of the user's location, current activity, friends within close proximity, the location and activities of other people, and the current environmental conditions. It also keeps a history of the user's visits to places in the city.

4.1. Technical Implementation

Just-for-Us was implemented as a web application running in Microsoft Pocket Internet Explorer on HP iPAQ h5550 connected either directly to the Internet through WLAN or via GPRS on a Bluetooth enabled mobile phone. This approach facilitated delivering all content by means of simple HTML pages and graphics files over a standard HTTP connection. Since, however, HTTP does not support pushing information to the user this was implemented by maintaining an open HTTP connection using pushlets (Pushlets 2004)[2]. The content of Just-for-Us is powered by a relational mySQL database divided into three parts. Firstly, it contains information about the physical layout of Federation Square, the location and accessibility of places, descriptions and photographs of landmarks and transition points, and simple way finding descriptions for getting around the space. As the physical layout of the space is continuously modified, the information in the database is designed to be easily updateable.

2) Just-for-Us system was one of the first systems for mobile devices exploring context-aware content delivery through web browsers rather than through dedicated client applications. This approach has since gained a lot of momentum and new programming tools have been developed that overcomes some of the limitations of HTTP in relation to, for example, information push in form of the AJAX framework (Garrett 2005). The AJAX framework allows the development of more advanced dynamic web applications for mobile computer systems (Kjeldskov et al. 2010).

Secondly, the database contains information about the different establishments at Federation Square and about the space itself. This includes descriptions and photographs of places, descriptions of special events, menus, lists of special offers, programmes, logos, etc. For maintenance of information, vendors can modify their own data through a simple web interface. Thirdly, the database contains dynamic information about a user's current context (location, activity, social group, etc.) and their history of visits.

The web interface of Just-for-Us was implemented using PHP for server-side generation of HTML pages on an Apache web server on the basis of the content and context information in the mySQL database. Dynamic client-side interaction and handling of pushed information was implemented using JavaScript. In this sense, Just-for-Us works very much like any other dynamic website, and is very easy to extend and modify. Making the application primarily server-side has the benefit of requiring only the exchange of small pieces of HTML code, images and URL requests over the network rather than the exchange of heavy program code. Furthermore, it makes use of the server's processing power for quick program execution rather than relying on slow mobile device processors.

Supporting the functioning of the web application, a number of server-side programs were implemented to perform specific sub-tasks. One such application monitors the contextual information in the database and pushes information to the user when appropriate. Another server-side application uses the database to dynamically generate PNG maps and annotated photographs. The overall architecture is outlined in figure 3.

Figure 3. Just-for-Us prototype system architecture.

Just-for-Us allows several ways of resolving the physical location of the user: through GPS, WLAN or by Bluetooth beacons embedded in the built environment. Using Bluetooth beacons for positioning, the system does not know the exact coordinates of the user but only has a rough idea of his position (for example, if he is in a specific café or in the main square). The presence of friends in the vicinity of the user is resolved by scanning for other Bluetooth devices with network ID's matching the user's list of friends. Positioning and friends in vicinity is resolved locally on the user's device and relayed to the mySQL database through a small client application running in the background.

4.2. Interaction Design

The Just-for-Us prototype was designed for Microsoft Pocket Internet Explorer running in full screen mode. The interaction design of the system to a large extent resembles a normal web page being accessed from a handheld device. The user can use the stylus to select links on the touch screen and use the on-screen keyboard or digitizer for entering text. To avoid scrolling, long pages were cut into a sequence of sub-pages. The design of the Just-for-Us prototype is based on four ideas emerging from the empirical studies:

1. Making the Invisible Visible: Augmenting the User's Physical Surroundings
2. Supporting Ad-Hoc Communication about Places, Activities and Time
3. Indexing Recommendations and Content to History and Context
4. Representing Activities within Proximity and Indexing to Familiar Places

Making the Invisible Visible: Augmenting the User's Physical Surroundings

The empirical studies of Federation Square revealed that the space is divided into four overall districts each with its own distinct characteristics and landmarks. This division into four districts was used as a starting point for our prototype design.

When logging on to the Just-for-Us system while physically located at Federation Square (or when entering Federation Square while logged on to the system) a "home screen" is pushed to the device displaying information corresponding to the district where the user is presently located (figure 4).

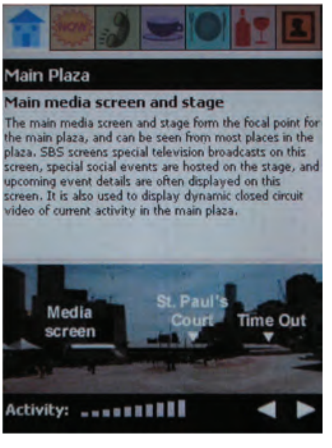


Figure 4. Home screen: augmentation of the user's built surroundings

The home screen consists of four elements; 1) the name of the district, 2) textual descriptions of places in that district, 3) an activity meter showing the current patronage and primary activity at a selected place, and 4) a 360° annotated panoramic view divided into a series of sequential photographs. The activity-meter indicates how many people are currently present at a place, and shows what they are doing there. The annotations on the panoramic photograph show what is located behind the physical structures surrounding the user, thus making the invisible visible through a form of indirect augmented reality. Clicking on an annotation, a short description of that place and a list of what's currently happening there, such as specials or upcoming events, movies etc., are displayed. Furthermore, the activity meter below the panoramic photograph is updated to reflect the patronage and primary activity of the selected place. Using the arrow icons below the photograph, the user can pan left and right, exploring the district and accessing more information about places in it. By default, the panoramic photograph is focused on the landmark with a corresponding textual description of the district. When entering a new district, the corresponding home screen is automatically pushed to the device.

Supporting Ad-Hoc Communication about Places, Activities and Time

Another significant finding from the field studies was that people typically coordinate meeting up with their friends in a highly ad-hoc manner. Typically, this involves a lot of communication back and forth, negotiating who, why, where and when to meet. The preferences for these factors are highly dependent on context and final decisions are seldom made until the very last minute. Activities depend on who you are meeting and what the others want to do. Places to go to depend on what activity you want to do and your shared history of going out. Places to meet depend on people's physical familiarity with a place, how long you have to wait, and the presence and activity of others. When to meet depends on people's physical distance from potential meeting places, who you are meeting up with, and why you are meeting up. These findings informed the design of the "contact" screen of Just-for-Us, providing a text-based communication channel to one's friends with a set of shared representations for negotiating a rendezvous (such as mutual location information).

When the user selects the "Contact" option on the top menu bar of the screen, the system displays a list of friends similar to the contacts list in e.g. MSN Messenger. The list is divided into three parts: 1) friends who are online and within proximity of the user; 2) friends who are online but further away; and 3) friends who are offline. If two or more friends are currently together (within close proximity of each other) they are displayed as a group. When the user selects a friend or a group of friends, an Internet chat session is established (figure 5).

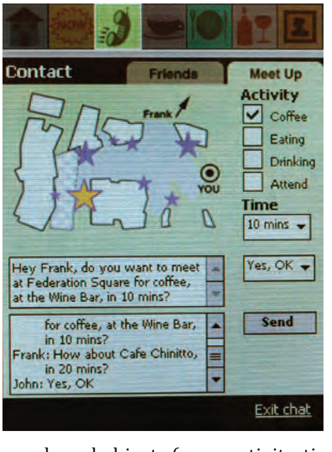


Figure 5. Contact screen: shared objects (map, activity, time) and chat window

At the receiving end this causes a brief ringing tone and a flashing telephone icon on the top of the screen. Apart from supporting free text input, the chat screen, more importantly, also supports automatic generation of small pieces of text with the purpose of supporting communication about people, places, activities and time. At the top of the screen, a small map shows the user's immediate surroundings and the location of the participants in the chat (absolute if within the map area and relative if not). Next to the map, the user can choose between the activities supported by the places on the map. When selecting an activity, recommended places generated on the basis of the user's history and current context are shown on the map by means of different sized coloured stars. Interacting with the map, activity checkboxes and a time drop-down menu generate auto text in the outgoing message window such as, for example, "Hey Frank, do you want to meet at Federation Square for coffee at The Wine Bar in 10 minutes?". When an automatically generated text message is sent, it causes the selected place, activity and time to be synchronized among the participants in the chat, who can now modify the original suggestion, causing a counter suggestion such as "No, but what about a drink at Transport Hotel in 25 minutes?". As in a traditional Internet chat, people can leave the conversation and new people can be invited.

Indexing Recommendations and Content to History and Context

Another finding from the empirical studies, which had impact on the design of Just-for-Us, was that places and spaces are dynamic and that setting matters immensely for the quality of socializing – especially in relation to its physicality, the presence and activities of other people and its convenience in terms of closeness. It was notable that people like to go back to places they know well, have been to before with friends in their current social group, or that have been recommended by a friend. People seldom go to completely new places if a familiar place exists within convenient distance. On the basis of these findings, Just-for-Us attempts to support the ongoing negotiation of where to go by indexing to the social group's shared knowledge of familiar places and providing information for the group to be able to size-up the situation before committing to entering a place. When the user clicks on one of the four activity-icons at the top of the screen, the system presents a list of recommendations of providing this activity (figure 6).

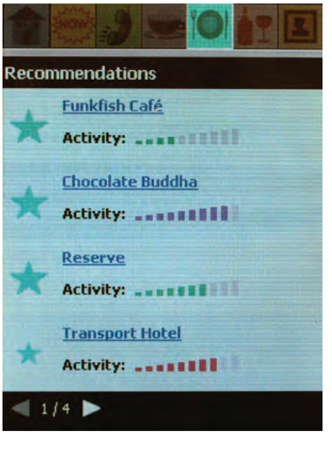


Figure 6. Recommendations screen: ranked list of places to go for food

Rather than simply sorting this list, for example, alphabetically, it is sorted on the basis of the systems knowledge about user's familiar places (history of prior visits), current physical and social setting (where the user is and who he is with) and the current weather conditions. Firstly, the list contains places where the user has been to before with the people that he is currently socializing with. This is followed by places that all people in the social group have been to before, but have not been to together. Thirdly, the list contains places where the user has been before, that none of the other people in the social group have visited (the user's own past experience). This is followed by places that the user has not been to, but that other members of the current social group have (implicit recommendations from friends). Finally, the remaining places in the vicinity of the social group, not familiar to anyone, are displayed. Within these listings, places are ranked in consideration to the frequency of past visits, the proximity of places, the current activity in places, and how well the weather situation of past visits to a place fits the current conditions. The highest scoring places are highlighted with a star next to the place name in the list. Furthermore, each place has an associated "activity-meter" displaying the current patronage and primary activity (like the one used on the "home" screen). From the list of recommendations, the user can access more information about a place such as menus or programmes. If the user is physically present at a place supporting the selected activity, Just-for-Us will assume that he is primarily interested in information about that specific place and thus takes him directly to the menu or programme. If he is interested in places other than his current situation, he can access the list of recommendations through a "Show other places" link on the bottom of the screen.

Representing Activities within Proximity and Indexing to Familiar Places

A final finding from the field studies of socialising at Federation Square, which had impact on the design of Just-for-Us, was that people make sense of a place through the social affordances provided by other people; where they are going and what they are doing there. People often use this information as important cues for where to go and what to do themselves. It also accommodates people's desire for interaction by proximity between their own social group and others. This finding informed the design of the "now" screen.

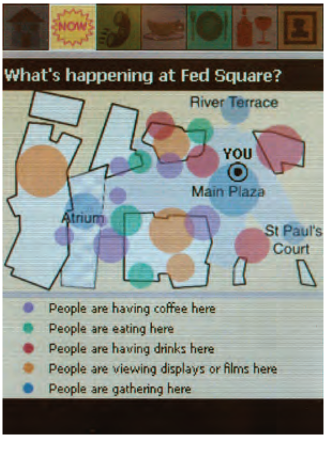


Figure 7. Now screen: showing clustering and activities of close-by people

When the user clicks on the "now" icon on the main menu bar, the system displays a small map of the user's immediate surroundings with superimposed, dynamically updated, coloured circles indicating the clustering and activities of people in proximity (fig. 7).

The radius of the circles indicates the number of people at a place while the colour represents their prevalent activity (e.g. "having coffee", "having a drink", "eating" or "attending a cultural event") using the system's general colour coding of these activities. The map also shows the user's (approximate) location. Clicking on the coloured circles on the map, the user can access more information about a place: detailed descriptions, photographs, menus, programmes, and directions on how to get there from their present location (fig 8). The maps are generated continuously on the server on the basis of the context information in the mySQL database.

Figure 8. Now screen: details about selected place and small activity map

The field studies also revealed that people seldom navigate by means of detailed maps and route descriptions when making their way around a built environment such as Federation Square as a part of a social group out on the town. Neither do they need or want to. Instead they make use of indexes to familiar places and landmarks in their surroundings. This finding was used to inform the design of the "Getting There" tab in Just-for-Us (fig 9). Selecting this tab, presents a photograph of the selected place and a small set of cues about where it is located based on the user's current location and indexes to places where the user has been to before. If the place is in another district than the user, directions are divided into a series of sub steps on separate pages guiding the user to that district through references to familiar places, landmarks and transition points. In this way, Just-for-Us takes into consideration what people already know about the environment they are situated in and makes use of people's ability to make sense of an unfamiliar place on the basis of a few simple cues. The "Getting There" tab is only available where it makes sense to provide this information. This excludes the home screen, where the user can see the location of places on the panoramic photograph, and recommendations to familiar places.

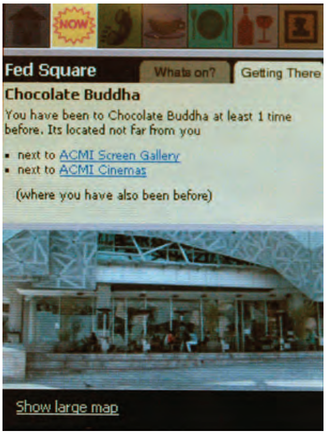


Figure 9. Now screen: way-finding instructions based on history of visits

5. EVALUATION

The presented prototype has been evaluated through a series of heuristic inspections by the authors throughout its creation but has not yet been subjected to real users. At the time of writing, we are preparing a large-scale study of use and usability in the field at Federation Square to be conducted in March-April 2005. The study will include 16 social groups consisting of two or more people. The groups will be asked to use the system for approximately 1.5 hours while socializing at Federation Square. As previous research has stressed the value of researcher control in field evaluations (Kjeldskov et al. 2004), the users will be given a number of overall tasks to prompt use of the different parts of the system. Supporting this approach, the users will also be asked to validate the relevance and realism of these tasks in relation to the activity of socialising in public. Inspired by the constructive interaction approach to thinking-aloud studies with more than one user, the social groups will be asked to talk among themselves about their perception of, and interaction with, the system with the researcher only asking questions for clarification.

The use of the system will be documented on digital video by means of a newly developed, state-of-the-art mobile data collection facility developed as part of this project. This facility allows miniature wireless cameras to be attached to the mobile devices, capturing high-quality images of the screens and the users. Video signals are transmitted to a small bag carried by the test monitor, where they are mixed on the fly with a third-person view capturing the user's context. Ensuring high-quality sound, all users are wearing directional wireless microphones transmitting to an audio mixer in the test monitor's bag. Video and audio is recorded digitally on a 100GB AV recorder.

At present time, we have conducted one full pilot study, causing minor modifications to the tasks and validating the reliability of the data collection equipment and the robustness of the prototype system. Findings from the planned study will be forthcoming.

6. CONCLUSIONS AND FURTHER WORK

This paper has addressed the challenge of facilitating sociality through mobile system design. On the basis of field studies of physical space and people socializing there, we have presented a mobile prototype system, Just-for-Us, which facilitates sociality by adapting not only to location but to the ensemble of physical and social context. The system exemplifies how "just-for-us" information and functionality can be tailored to the user's physical and social setting by indexing to information implicitly present in the user's surroundings, the user's existing knowledge and the user's current social setting.

Further work should investigate the use of the current prototype by identifying strengths and opportunities for improvements. Of special interest, we are exploring opportunities for allowing the users to contribute with their own content creating a rich layer of digital media overlaying the city. It would, for example, be interesting to create virtual entities of social groups and to blur the boundary between physical and virtual spaces by allowing for socializing by virtual proximity. On the technical side, we are developing a more flexible platform for extending the system to cover a much wider physical area, making the interface adaptable to even smaller mobile devices such as mobile phones, and extending the mobile, context-aware interface with omnipresent, traditional web-based access.

ACKNOWLEDGMENTS

This research is supported by the Danish Technical Research Council's talent project "Indexical Interaction Design for Context-Aware Mobile Computer Systems" (project reference 26-04-0026) and The University of Melbourne's PORES program. The authors thank everyone participating in the empirical study and Federation Square Management for their support. The authors also thank Steve Howard and Bharat Dave for valuable input to the design of the empirical studies.

REFERENCES

Alexander, C., Ishikawa, S., and Silverstein, M. A Pattern Language. Oxford University Press, New York, 1977.

Beyer, H., and Holtzblatt, K. Contextual Design - Defining Customer Centred Systems. Morgan Kaufmann, San Francisco, 1998.

Blomberg, J., and Burrell, M. An Ethnographic Approach to Design. In J. Jacko, and A. Sears (eds.). Handbook of Human-Computer Interaction. Lawrence Erlbaum Associates Inc., Mahwah, New Jersey, 2003, 964-986.

CNN.com. Japan's Lonely Hearts Find Each Other with "lovegety". CNN.com, 7 June 1998, http://www.cnn.com/WORLD/asiapcf/9806/07/fringe/japan.lovegety/ (1998) (accessed 16 January 2005).

Crabtree, A., and Hemmings, T. Shaping the Home: Architecture, Technology, and Social Interaction. Report for Equator IRC, Document D1.2, The Sociality of Domestic Environments, The University of Nottingham, UK, 2001.

Crabtree, A., Hemmings, T., and Rodden, T. Pattern-based support for interactive design in domestic settings. In Proceedings of the 2002 Symposium on Designing Interactive Systems (DIS2002) (London, UK, 25-28 June, 2002). ACM Press, London, 2002, 265-275.

Dieberger, A., and Frank, A. A City Metaphor to Support Navigation in Complex Information Spaces. In Journal of Visual Languages and Computing, 9 (1998), 597-622.

Erickson, T., and Kellogg, W. Social Translucence: An Approach to Designing Systems that Support Social Processes. In ACM Transactions on Computer-Human Interaction, 7, 1 (2002), 59-83.

Fithian, R., Iachello, G., Moghazy, J., Pousman, Z., and Stasko, J. The Design and Evaluation of a Mobile Location-Aware Handheld Event Planner. In Proceedings of Mobile HCI 2003 (Udine, Italy, 8-11 September, 2003). Springer-Verlag, Berlin, 2003,145-160.

Garrett, J. J. (2005) Ajax: A New Approach to Web Applications. Adaptive Path. http://www.adaptive path.com/ideas/essays/archives/000385.php

Gaver, B. Affordances for Interaction – The Social is Material for Design. In Ecological Psychology, 8 (1996), 111-129.

Goodman, J., and Gray, P. A Design Space for Location-Sensitive Aids for Older Users. In Proceedings of Workshop on HCI in Mobile Guides (Udine, Italy, 8 September, 2003).

Grinter, R. E., and Eldridge, M. y do tngrs luv 2 txt msg? In Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work (ECSCW'01) (Bonn, Germany, 16-20 September, 2001). Kluwer Academic Publishers, Dordrecht, Netherlands, 2001, 219-238.

Hvem er i byen (who is out on the town). http://www.hvemeribyen.dk/ (2004).

Kjeldskov, J., Christensen C.M., and Rasmussen K.K. (2010) GeoHealth: a location-based service for home healthcare workers. Journal of Location-Based Services, 4(1): 3-27

Kjeldskov, J., Skov, M.B., Als, B., and Høegh, R.T. Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In Proceedings of Mobile HCI 2004 (Glasgow, Scotland, 13-16 September, 2004). Springer-Verlag, Berlin, 2004, 61-73.

Kjeldskov, J. Just-In-Place Information for Mobile Device Interfaces. In Proceedings of Mobile HCI 2002 (Pisa, Italy, 18-20 September, 2002). Springer-Verlag, Berlin, 2002, 271-275.

Kolari, J., and Virtanen, T. In The Zone: views through a context-aware mobile portal. In Proceedings of Workshop on HCI in Mobile Guides (Udine, Italy, 8 September, 2003).

Kulju, M., and Kaasinen, E. Route Guidance Using a 3D City Model on a Mobile Device. In Proceedings of Workshop on Mobile Tourism Support Systems (Pisa, Italy, 17 September, 2002).

Lynch, K. The Image of the City. MIT Press, Cambridge, MA, 1960.

McCullough, M. On Typologies of Situated Interactions. In Human-Computer Interaction, 16 (2001), 337-347.

Neuman, W. Social Research Methods: Qualitative and Quantitative Approaches. Allyn and Bacon, Boston, 1994.

Paay, J. and Kjeldskov J. (2005b) Understanding Situated Social Interactions in Public Places. Proceedings of INTERACT 2005 (Rome, Italy, 12-16 September, 2005).

Paay, J., and Kjeldskov, J. (2005a) Understanding and Modelling the Built Environment for Mobile Guide Interface Design. In Behaviour and Information Technology, 24 (2005), 21-35.

Paulos, E., and Goodman, E. The Familiar Stranger: Anxiety, Comfort, and Play in Public Places. In Proceedings of Conference on Human Factors in Computing Systems (CHI 2004) (Vienna, Austria, 24-29 April, 2004). ACM Press, London, 2002, 223-230.

Persson, P., Espinoza, F., Sandin, A., and Coster, R. GeoNotes: A Location-based Information System for Public Spaces. In K. Hook, D. Benyon, and A. Munro, (eds.). Designing Information Spaces: The Social Navigation Approach. Springer-Verlag, London, 2002, 151-173.

Rheingold H. Smart Mobs - The Next Social Revolution. Perseus Publishing, Cambridge, 2003.

Strauss, A., and Corbin, J. Basics of Qualitative Research. Sage Publications, Newbury Park, California, 1990.

van den Broecke, J. Pushlets – Whitepaper. http://www.pushlets.com (2002)

Wittel, A. Toward a Network Sociality. In Theory, Culture and Society, 18 (2001), 51-76.

Chapter 17

GeoHealth

Jesper Kjeldskov, Claus M. Christensen, and Klaus K. Rasmussen

Abstract. We describe a map-based location-based service "GeoHealth" for home healthcare workers who attend patients at home within a large geographical area. Informed by field studies of work activities and interviews with care providers, we have designed a mobile location-based service prototype supporting collaboration through information sharing and distributed electronic patient records. The GeoHealth prototype gives the users live contextual information about patients, co-workers, current and scheduled work activities, and alarms adapted to their geographical location. The application is web-based and uses Google Maps, GPS positioning, and Web 2.0 technology to provide a lightweight, dynamic and interactive representation of the work domain supporting distributed collaboration, communication, and peripheral awareness among nomadic workers. Through a user-based evaluation, we found that the healthcare workers were positive towards the use of location-based services in their work, and that the dynamic and interactive geospatial representation of the work domain provided by GeoHealth supported distributed collaboration, communication, and peripheral awareness. We also identified areas for improvements.

1. INTRODUCTION

Location-Based services constitute an emerging paradigm of computing where users on the move are provided with information and functionality that is particular relevant at a specific geographical location or within a specific distance. Within this broad scope, many systems have been developed and described in the research literature over the last decade, envisaging "blue sky" innovations, exploring user experiences, and tackling technical challenges of implementation (Raper et al. 2007). The research field of location-based services has developed significantly in its relatively short lifetime and core technologies are now maturing for deployment outside the scope of research

Originally published as Kjeldskov, J., Christensen, C. M. and Rasmussen, K. K. (2010) GeoHealth: a location-based service for home healthcare workers. *Journal of Location-Based Services,* 4(1), 3-27.

prototypes. Facilitated by this, increasing research is reported into the user side of the equation exploring new potential “real-world” application areas and studying the use of location-based services in these settings. This is, indeed, an interesting time to be doing LBS research.

As a specific case for our own research within the area of location-based service development and use, we have studied work activities in regional home healthcare as a potential application area of this emerging technology. Home healthcare is a subset of the healthcare domain where patients are treated or attended to in their own homes rather than at a hospital or a clinic. Like in most areas of the healthcare domain, home healthcare workers are under increasing pressure to provide improved care to more people while using fewer resources. At the same time it is often speculated that ubiquitous and mobile computing may provide some means for improving quality of service through, for example, better communication between patients and healthcare workers, by providing access to centralised medical data at the point of care, supporting effective emergency response, etc. In response to this, we have been exploring the development and use of a *location-based mobile healthcare system* supporting the nomadic collaborative work activities of physically distributed home healthcare workers by providing mobile and location-based access to electronic patient records as well as information about co-workers and means for communication and shared work task management.

Recent developments in web-based technologies and services, availability of GPS positioning on mobile and portable devices, and widespread access to broadband mobile Internet has enabled the creation of a new class of location-based services. Combining positioning capabilities with the powers of “Web 2.0” technologies for user-generated content, open APIs for geographical interfaces such as Google Maps, and mobile or portable devices with full-blooded web browsers now makes it easier than ever to implement location-based services and applications for users on the move that take into consideration their location and the surrounding environment. Hence, Raper et al. (2007: 29) reports “an explosion of activity and interest” in using Web 2.0 technologies for the development of “the next wave” of location-based services.

Inspired by these new technological possibilities, our latest location-based prototype application “GeoHealth” represents an early instance of this potential next wave of Web 2.0 based location based services described by Raper et al. (2007). GeoHealth is a web-based location-based service running within the web-browser on a laptop or tablet PC. It graphically augments a home healthcare district with a meshup of context information (e.g. about patients, co-workers, current and scheduled work activities, and alarms), user-generated information (e.g. text messages and treatment notes), and information stored centrally (e.g. medical patient records and work schedules). The aims and objectives of the GeoHealth prototype were to explore the user experience of a mobile location-based service within the healthcare domain exploring the powers of Web 2.0 technologies in combination with GPS positioning a geographical interface, and the Mobile Internet.

Reporting on this work in the Journal of Location Based Services, we contribute to two of the “urgently needed” areas of research outlined in Raper et al.’s (2007) editorial lead paper of the first issue of the journal: interaction design for LBS and LBS architectures and platforms. We also respond to the lead paper's recommendation for more user-centred system conception of location-based services (Raper et al. 2007: 22).

The paper is structured as follows. First, we present related work within context-awareness, location-based services and mobile healthcare applications. Secondly, we introduce our case of home healthcare workers. Thirdly, we present the GeoHealth prototype followed by findings from a user-based evaluation. Finally, we conclude on our research and point out avenues for further work.

2. RELATED WORK

In this section we outline related work within context-awareness, location-based services and mobile healthcare applications.

2.1. Context awareness and location-based services

There has been a lot of research on context-awareness in recent years. Much of this has enabled computers to collect and formalize information about contextual factors describing, for example, the geographical, environmental, biometrical, and social setting and stage of a user. In parallel this has inspired the use of machine intelligence to make mobile systems transform themselves accordingly and minimize need for explicit user interaction. Experiences with developing and using such systems are many and context-awareness has proven possible within specialised domains such as industrial process control, healthcare, transportation, and tourism (e.g. Paay et al 2009, Skov and Høegh 2006, Bardram 2006, Cheverst et al. 2000). However, research has also identified many fundamental problems with this approach to human-computer interaction. Even simple adaptations to context, such as the mobile phone that knows when not to ring, are extremely complex to implement, and any context-aware technology is most likely to make mistakes and take control away from the user (Brown and Randell 2004, Barkhuus and Dey 2003). In contrast, it has been stressed that while computational systems are good at gathering and aggregating data, humans are good at noticing, integrating and interpreting obvious as well as subtle cues and determining appropriate actions (e.g. Kjeldskov and Paay 2006). In response, some researchers have advocated that another way to deal with complex and ambiguous context information is to actively involve the user (Dey and Mankoff 2005) and make humans, not machines, the consumers of context information by representing it in the interface rather than automatically adapting to it (Edwards 2005, Aaltonen and Lehikoinen 2005). As a fictive example, a representational context-aware service for a hospital ward would display collected context-information about patients, healthcare workers, and current work activities, but leave interpretation and decision-making based on this information to the user rather than to the computer.

Real examples of context-aware systems embodying this *representational* approach are still very few and mostly research prototypes rather than commercial systems. For example, Paulos and Goodman's jabberwocky prototype (2004) use detected patterns of nearby Bluetooth devices over time for visualizing “familiar strangers” in the users' vicinity. Erickson et al.'s Babble prototype (2002) visualize “social proxies” based on the presence and activities of nearby people. In a more recent study, Lemmelä and Korhonen (2007) use geographical context information, gathered from salient keywords within public "geo-postings", to visualise physical clusters of similar places on a digital map, such as areas with a high occurrence of the key word "restaurant", shopping", "art" or museum". Common for these systems is that they are context-aware but use their awareness to generate new information rather than trying to reduce it. It is also common that they are particularly aware of their *location*.

One of the most promising aspects of user context for a mobile or portable computer system to respond to is *location* (e.g. Raper et al. 2007, Jones et al. 2004, Kaasinen 2003). As a subset of context-aware systems, location-based services represent an emerging class of computer systems providing mobile device users with information and functionality that is particular relevant at a specific geographical location or within a specific distance. Within the last decade, this class of mobile or portable computer systems have received increasing attention from researchers within areas such as computer science, human-computer interaction, and interaction design, as well as from software industry. Commercially available location-based services are still relatively few, but an increasing number of services are now emerging that integrate wide-area broadband wireless Internet access, web resources and geographical information for PDAs and "smart phones" with GPS and other positioning capabilities. As pointed out in Raper et al.'s opening article of the first issue of the Journal of Location Based Services (2007: 6), one of the reasons for the less than expected user take-up of location based services is that many of the potential application areas for useful location based services have simply not yet been explored by mobile operators. One of these under explored domains is healthcare. Within the healthcare domain, geographical positioning technologies have a history of use in emergency response systems as means of pinpointing people in distress and directing information to relevant personnel. However, the use of location-based *services* as such is still a rarity despite the huge potentials of, for example, relating the vast amount of available digital information to geographical locations in the creation of new and effective medical services.

Yet, developing successful location-based services for a potential application domain such as healthcare is not trivial either. It inherits challenges of context-awareness and other location-based services described in the literature, such as issues of user control, privacy, and determining people's location in physical space, as well as introducing new ones. Little is known about people's use of such of services, how to design them well, and what is (and isn't) useful. It is unknown how users perceive and use information provided through a location-based service, what content is considered relevant (and what is not), and how people will adopt and appropriate information services that react to their location and combine, for example, web content, satellite imaging, and cartography. Hence, more research is needed into the development and real world use of location-based services. Through the case study presented below, our aim is to contribute to this area with 1) understanding of a specific potential use domain for location-based services and technologies, 2) concrete design ideas for location-based service for home healthcare workers, and 3) a Web 2.0 approach to the technical implementation location-based services for exploring latest trends in web-based technology and geographical information systems.

2.2. Mobile healthcare applications and location-based services

Mobile healthcare applications is an evolving area within health informatics where recent advances in ubiquitous and mobile computing are exploited for the creation of new, flexible, and effective medical services (e.g. Varshney 2003, Woodward et al 2001, Tachakra et al. 2003). The area covers a wide range of healthcare applications and services where patients and/or healthcare workers are free to move around physically while still connected to the service (for a review of systems see, for example, Pattichis et al. 2002). One of the most prevalent mobile healthcare applications reported in the literature is the use of biometrical sensors and mobile phones for monitoring patient data by oneself (Mattila et al. 2008) or remotely in real time (e.g. Figueredo and Dias 2004, Tachakra et al. 2003) allowing, for example, non-critical patients to remain at home as much as possible. Other applications involve the use of high-bandwidth mobile data networks for live transmission of video data from, for example, ambulances on dispatch to health professionals at the hospital (e.g. Mandellos et al. 2004), pervasive access to electronic patient data for healthcare workers on the move (e.g. Reponen et al. 2000), and patient telemonitoring in emergency situations (Maglogiannis et al. 2009, Maglogiannis and Hadjiefthymiades 2007). Some mobile healthcare applications also support "telecooperation" amongst physically distributed and mobile healthcare workers through information sharing, team communication, and coordination of team activities via mobile terminals and wireless data networks (e.g. Pitsillides et al. 1999).

As a specific area within healthcare applications, location-based services and interactive map applications have received increasing attention within the last coupe of years. In (Boulos 2003b), location-based health services are promoted as a new paradigm for personalised information delivery responding to the potential information overload potentially created by unfiltered delivery of health information to various mobile devices. The main goal of this paradigm is to create better presentation of health and healthcare needs and Internet resources across a geographical area in order to facilitate better support for decision-making. As a specific example of such location-based service, "HealthCyberMap" is an interactive geographical map for browsing medicine and health information on the Internet (Boulos 2003a). Enabling location-based browsing of health information and collection of spatial data, Rainham et al. (2008) presents the design and evaluation of a wearable GPS system and data logger developed to assist in deriving a more complete picture of the different places that influence people's health and wellbeing. Using similar technologies, Boulos et al. (2007) describe a location-based healthcare service aimed at increasing older people's autonomy through the use of a wearable monitoring device with GPS connected to a central care service and monitoring system. Exploring the use of location-based services as a part of a ubiquitous computing environment at a large hospital, Kjeldskov and Skov (2007) presented findings on the usability of the "MobileWard" system giving nurses and doctors access to electronic patient records and work tasks based on their physical location.

In a user-based study of requirements for GIS based decision support in public health, Driedger et al. (2007) found that maps and mapping tools were powerful means of "converting locally collected data into information" and that desired functionality for such systems included the ability to geo-code patient data and push-pin patients and services on a shared map base of the region. Extending the use of such shared map-based representations, Chang et al. (2009) presented a combined Google Earth and GIS mapping system designed to assist managing of dengue fever in a developing country. This system makes use of several interactive information layers overlaid onto a map of a specific region, allowing public health workers to collaboratively interpret spatial relationships between data. Most of the applications reviewed here use GPS as the primary positioning technology. For a detailed review of technologies for detecting user location in healthcare see (Boulos 2003b).

In the study presented in this paper, we focus on a combination of technologies where a map-based location-based service provides the foundation for supporting GPS tracked home healthcare workers in their work activities through mobile assess to electronic patient record data, channels for group communication, tools for coordination, and simple patient monitoring.

3. FIELD STUDY: HOME HEALTHCARE

Home healthcare is a subset within the healthcare domain where patients are treated or attended to in their own homes rather than at a hospital or a clinic. Home healthcare nursing and assistance is mostly used for providing care to elderly people, or people with disabilities, who with this nursing or assistance are then capable of taking care of themselves in their everyday lives. It is also often used as a means of follow-up treatment subsequent to periods of hospitalisation. Making house calls to patients at home obviously requires nurses and healthcare assistants to be mobile within a specific geographical area "nomadically" (Sawhney and Schmandt 2000) moving their location of work from place to place throughout the day. It also requires them to be physically distributed from their co-workers, while still collaborating on shared tasks coordinated through division of labour, shared plans, and ongoing communication.

With the purpose of informing the design of a location-based service for nomadic home healthcare workers, we conducted an empirical field study of work activities in the municipality of Aars in Denmark in 2006. The study was done using a rapid ethnography approach (Millen 2000) involving a combination of observations of current practice, contextual interviews, and semi-structured interviews. The study involved 10 healthcare assistants and 11 nurses with different levels of experience with the work domain, and resulted in 17.5 hours of audio data accompanied with field notes and digital photographs. Findings from the study were verified through subsequent meetings and focus groups with the home healthcare staff.

The focus of the study was to enquire into the characteristics of the distributed work activities and the relationships between work activities, patients, healthcare workers, information, information needs, and geographical space. For example, we were interested in knowing who needs *what* information *when* and *where*, how division of labour is negotiated, assigned, represented, and managed, how individual workdays are structured (and restructured) in relation so geographical space, what happens in case of events out of the ordinary, what information is kept as a permanent record for later use, what communication goes on between the distributed workers throughout the day, etc. Furthermore, we wanted to identify specific opportunities for supporting work activities, collaboration, and communication by means of a location-based service representing key contextual data in a geographical interface.

Data from the field study was analysed through a process of grounded coding and affinity diagramming of higher-level themes (Strauss 1987, Beyer and Holtzblatt 1998). On the basis of the field data, we also created four different personas (Pruitt and Adlin 2006) describing key characteristics of prospective users, their work activities, goals and habits with the purpose of serving as hypothetical archetype users for the design team. Finally, we created a number of physical models (Beyer and Holtzblatt 1998) of the workspaces of nurses and healthcare assistants illustrating how they operate during a typical workday in the municipality.

3.1. Findings from field study

In this section, we present a concentrate of findings from our find studies. These findings serve as a foundation for our prototype design, but also illustrate the type of data, insight, and richness that is typically gained from this methodological approach within the field of human-computer interaction. In conjunction with our prototype system, we wish to illustrate how this field study approach is also of value in relation to design and development of location-based services.

Home healthcare in the observed municipality involves two types of workers: nurses and healthcare assistants. The nurses have overall responsibility for the treatment of patients and also administer medicine to patients at their homes. The healthcare assistants are responsible for daily treatment and serve as "daily observers" reporting back to the nurses in case of, for example, side effects from medication. Currently, the nurses and healthcare assistants use a series of paper-based information artefacts and mobile and landline phones for coordinating and communicating with each other throughout the day and between shifts.

The healthcare assistants have a pre-planned workday with schedules for visiting assigned patients at particular times and in a particular order.

The division of work assignments happens in central meetings, but is also altered by the healthcare assistants during a day if something out of the ordinary comes up. Supplementing the day schedule, they carry specific task lists for the care taking of each patient including, for example, shopping lists for groceries. In addition to this, they carry a paper notebook used to jot down needs for changes to the services or schedule, which are then entered in to the permanent record at the central office by a manager after the shift. Finally, the healthcare assistants carry sheets of paper for documenting mileage, which is used for compensating the use of private cars for work purposes.

The nurses' workdays are more ad-hoc than the healthcare assistants'. Rather than a pre-planned schedule of visits, each nurse carry reference sheets for all his or her assigned patients in the district. These sheets contain personal information, phone numbers for the general doctor or practitioner, overview of health conditions, remarks on treatment and things to be aware of, diagnosis, phone numbers for relatives, lists of nursing tasks, and practical information such as the location of the key the patients' homes. They also indicate what times the patients are usually visited by nurses or healthcare assistants during the week. Every morning the nurses look through their reference sheets, pick out the patients they are supposed to visit that day, and manually transfer this information to the corresponding page in their calendar. However, rather than making a precise timed schedule for the day, the "day-view" functions more as a loosely scheduled "to-do list" with a few notes. In addition to their reference sheets and calendars, the nurses carry paper notebooks in which they register changes to the treatment and/or medication of patients, which they then enter in to the permanent record later at the central office. Finally, the nurses also register how many kilometres they drive in their private car for work purposes for economical compensation.

In order to facilitate collaboration the healthcare assistants and nurses share a series of paper-based records located either at the individual patients or at the central office. These records function partly as repositories of information on treatment for documentation purposes, and partly as media for communication between workers and between shifts. Firstly, each patient have an individual paper-based "book" located at his or her home and used for general "all-to-all" information and communication between healthcare workers, the patient, and their relatives. In addition to this, medicated patients have a paper-based medical prescription located with their medicine. Secondly, the healthcare assistant and nurses collaborate and communicate through paper-based records at the central office containing more detailed information not suitable for public view (i.e. relatives). Thirdly, the nurses collaborate through the official medical records, which are increasingly in electronic form.

In addition to coordinating their collaborative efforts through shared written notes and records, the nurses and healthcare assistants make extensive use of direct synchronous communication through mobile and landline phones. All nurses are equipped with mobile phones, making them immediately reachable in case of an emergency situation. The healthcare assistants do not have mobile phones but communicate by means of their patients' landline phones, and thus have to estimate the whereabouts of each other before making calls. Three kinds of synchronous communication were observed. Firstly, co-workers (healthcare assistants, nurses, physiotherapists, janitors, staff at the central kitchen, etc.) are informed about progress or delays in their work schedule, which may influence the work assignments of others. This could be, for example, when a healthcare assistant has finished bathing a patient, who is then ready for the nurse. It could also be in the event that a patient is suddenly hospitalised and thus does not need food delivery or assistance at home. Secondly, the healthcare assistants sometimes call each other to negotiate ad-hoc changes of work tasks in the case of something out of the ordinary happening. Thirdly, and perhaps most importantly, the healthcare workers communicate orally in emergency situations, which are usually triggered by patient alarms. When a patient triggers their alarm button (situated centrally in their home), the alarm is directed to a dedicated "emergency phone" carried by one of the healthcare assistants. This person is then responsible for delegating dispatch to the closest healthcare worker. This is done by making a series of phone calls from the closest landline; first to the place of the alarm in order to assess the situation, and then to the colleagues who is believed to be closest to the patient in need. Eventually, a worker is dispatched to the alarm.

From our field studies, it is clear that current work practice have enablers as well as challenges for providing optimal care. On the positive side, the lose mechanisms for ad-hoc exchange of work tasks makes it easy for the workers to respond to events out of the ordinary. Access to shared repositories of information makes it possible to coordinate treatment among co-workers, and communicate relatively easily between shifts. Finally, self-structuring of the workday allows for flexibility and for assisting each other if needed. On the negative side, firstly, the use of paper-based notes and records requires double registration of information and results in a delay between time of registration and time of availability to co-workers. Having several parallel records with overlapping content also introduces redundancy and uncertainty about what information is the newest. Secondly, relying on non-mobile technology for communication between mobile workers introduces inefficiency and risks. Important messages may not reach their recipients in time or may never be communicated in the first place because of doubt about where to call. In terms of alarms, the risk of losing valuable time in the process of locating and dispatching the closest healthcare assistant or nurse is obvious.

As a case for location-based services, the field study reveals a number of properties, which makes this a potentially interesting application area. Home healthcare workers are highly mobile and geographically distributed throughout a workday, and their work tasks are highly tied to specific geographical locations. Not all information is relevant at all times of the workday, but can to a large extend be coupled to specific locations that the healthcare workers will visit during the day. Hence, location could be used as a powerful filtering parameter for this particular work domain, and the functionality needed by home healthcare workers could be considered and presented as location-based information services. In addition to this, the dynamic locations of co-workers in relation to one self constitute a central factor for ad-hoc management of situations out of the ordinary, and could also be considered a location-based service candidate.

4. THE GEOHEALTH PROTOTYPE

Informed by our field study, we designed and implemented a functional prototype, GeoHealth, which supports distributed and mobile collaboration through representation of live contextual information about patients, co-workers, current and scheduled work activities, and alarms adapted to the user's location (figure 1). Using the terminology of Raper et al. (2007: 20), GeoHealth is a "map-based location based service".

The prototype explores new technical opportunities for the creation of multi-user location-based web-applications through a combination of Web 2.0 technology, GPS positioning, interactive map/satellite image overlays and mobile Internet access. GeoHealth was purposely designed for a sub-notebook size laptop or tablet computer that the healthcare workers can easily carry with them in a small bag. We deliberately refrained from designing for a smaller mobile form factor such as PDA handheld devices because of the limitations of these in terms of screen size. We have previously designed and evaluated other systems for healthcare using context-aware handheld devices (Kjeldskov and Skov 2007), and found from empirical studies that this form factor was often too limiting for healthcare workers, and that a small laptop or tablet was often a more suitable form factor for this kind of work. Aiming at a small laptop or tablet also allowed us to design GeoHealth to integrate well with other relevant applications (such as Skype), and to fully explore the use of Web 2.0 functionality. Based on findings from our field study, GeoHealth was developed through a semi-structured iterative process for moving from ethnographic data towards user interface design, involving activities of design sketching, paper prototyping and technology exploration (Paay 2008).

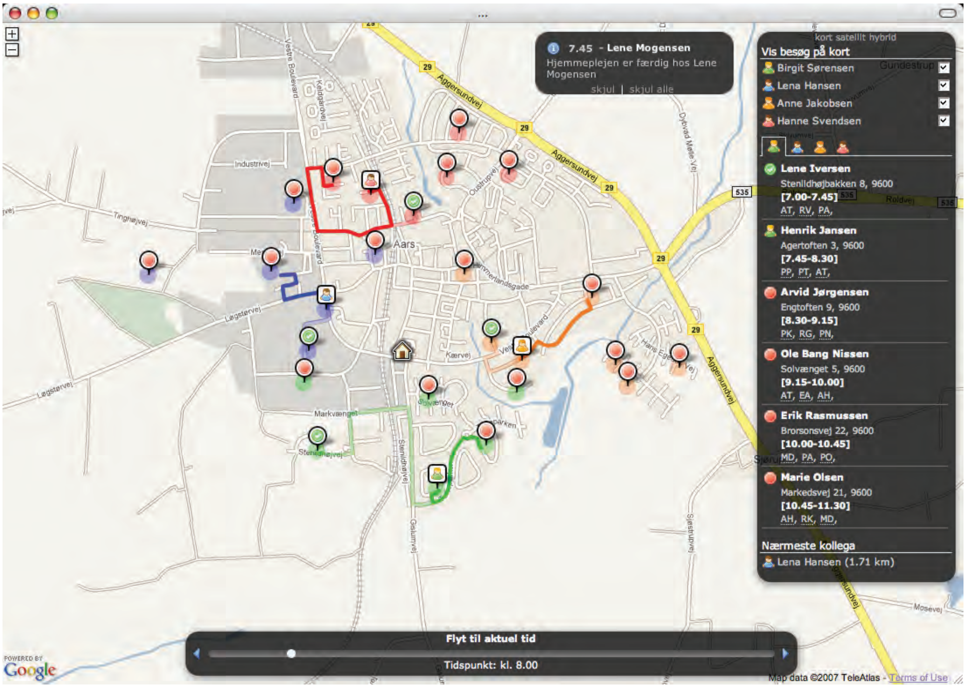


Figure 1. The main screen of the GeoHealth prototype system with the layers of four healthcare workers' patients, routes and current location enabled. In the palette on the right, the user can see her own scheduled work tasks and keep an eye on which colleague is closest to her. Using the timeline, the user can "fast forward" into the future and see how the schedules for the day will unfold according to plan, or "rewind into the past" to see where people have already been today.

The basic idea behind the GeoHealth system was to provide the home healthcare workers with an interactive graphical representation of their shared and individual work domains on a small laptop or tablet computer. This representation should be embedded into a digital map of their physical surroundings in different layers (figure 2). Within this representation, the workers should be able to directly access key information about their patients, scheduled tasks, and the location and activities of their colleagues. At the same time, the system should be able to function as a medium for text-based messaging, facilitate easy ad-hoc exchange of work tasks, and it should function as a graphical interface to the existing patient alarm system replacing needs for making phone calls. We also wanted to facilitate spoken communication among the workers through VoIP, as well as automate mileage registration for private car use.

to the users what to do, we wanted to represent context information in the interface and leave the intelligence to the users. Hence, GeoHealth deploys *active* context-awareness (Barkhuus and Dey 2003) in the sense that it pushes information to the user when, for example, he or she enters the vicinity of an assigned patient, or if a patient alarm is triggered nearby. However, leaving the user in control, the system also deploys *passive* context-awareness in the sense that pushed context information appears discreetly within the interface, and goes away again automatically if the user does not act on it. We call this combination of active and passive awareness *discreet context-awareness.*

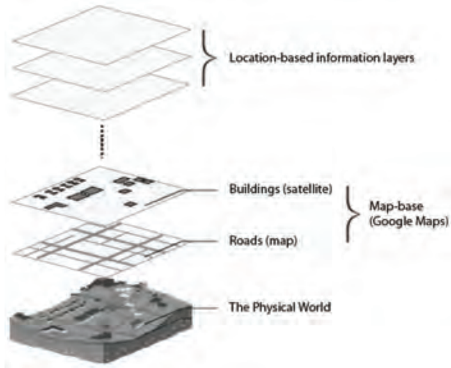


Figure 2. The geographical information-layering model of GeoHealth

We intentionally made the system context-aware in the sense that it would know the whereabouts and activities of its users, the patients, and scheduled work assignments. However, rather than using this information to make the system "intelligently" suggest

The functionality of GeoHealth can be divided into four overall themes: 1) location based spatial view of the work domain, 2) location based information push, 3) location-based ad-hoc exchange of work tasks, and 4) location based alarms. These are described and illustrated in details in the subsections below.

4.1. Location-based spatial view of the work domain

The primary functionality of GeoHealth is to provide the distributed healthcare workers with a spatial view of their work domain. This is done by plotting the locations of patients, healthcare workers (tracked via GPS), and their planned tasks/routes for the day on a full-screen map of the municipality (figure 1). The map can be navigated by means of dragging it left/right and up/down. Zooming and toggling between map view, satellite view, and hybrid view (satellite images with superimposed map information) is done through on-screen controls. By clicking on a patient on the map, a call-out box appears with access to all the information available about this particular person (which used to be distributed on different paper-based records) and with functionality for taking notes (figure 3). Through this box it is also possible for the nurses to make changes to the medical records, treatment etc. Changes and new information are immediately available to co-workers. Clicking on a telephone number in the patient records initiates a SkypeOut call. By clicking on a co-worker on the map, it is possible to initiate a Skype call to this person's laptop.

Figure 3. Four different views of patient information: 1) medical record, 2) contact information, 3) notes on treatment, and 4) satellite image and map details of home location.

4.2 Location-based information push

In addition to facilitating information "pull" about the work domain through the map representation and call-out boxes, the system also "pushes" information on the basis of location and assigned work tasks. However, rather than prompting the user with intrusive information screens, information push is done discreetly through gentle changes to the map representation in the form of appearance (and disappearance) of information. Firstly, the map view moves automatically when the user comes close to the edges (as in a sat-nav system), and icons change automatically in accordance to co-workers progress throughout the day. Secondly, when a healthcare worker drives into the vicinity of an assigned patient, the call-out box for this patient automatically pops up on the map, making information about the upcoming visit ready at hand (see figure 4). In addition to electronic patient information, one of the tabs contains a close-up satellite image of the patient's address facilitating way finding through reference to landmarks and physical attributes of the environment (e.g. the colour of roofs, density of houses, presence of trees, etc.). When the healthcare worker leaves vicinity of a patient, the box automatically disappears again.

Figure 4. Information automatically appearing when driving into the vicinity of an assigned patient's home.

4.3 Location-based ad-hoc exchange of work tasks

Apart from facilitating reshuffling of ones' own work tasks, the GeoHealth prototype system also facilitates easy ad-hoc exchange of work tasks between healthcare workers during a workday based on the location of co-workers. Negotiation of changes to the work plan is done through the systems' built-in text-messaging functionality or verbally through a Skype call, and the actual change to the work plan is done by simply dragging a patient icon onto the location of another healthcare worker on the map and then dropping it there. By doing this, the visit to that particular patient will disappear from one healthcare worker's list of tasks and appear on the other's. Moving a work task on to a co-worker also triggers a predefined text notification to appear on the receiving person's screen. Hence, rather than forcing the users through tedious formalized procedures for negotiation and confirmation of changes to the work schedule (which happens all the time), this design relies on trust, professionalism and social conventions for collaboration. The representation of the location of patients and co-workers on the map assists the healthcare workers in assessing whom it would be best to hand over a particular task to.

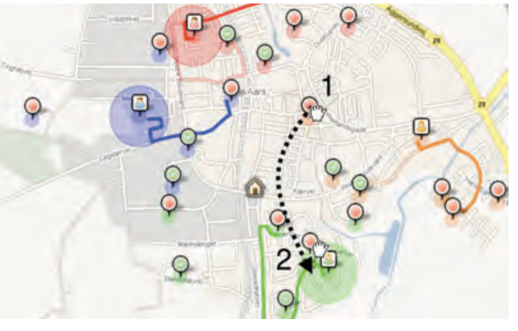


Figure 5. Location-based ad-hoc exchange of work tasks by dragging a patient icon onto the location of another healthcare worker.

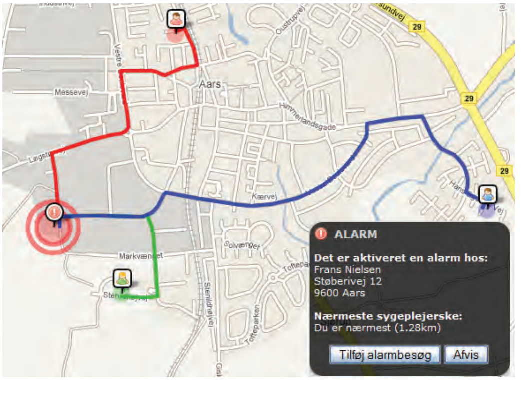


Figure 6. Patient alarm appearing as a pop-up message on the closest healthcare worker's screen along with information about location of the two most nearby colleagues.

4.4 Location-based alarms

For the purpose of replacing the use of mobile and landline phones for negotiating responds to alarms, and for supporting better decisions regarding such responses, GeoHealth embeds alarms directly into the spatial view of the work domain, and uses the continuous GPS positioning of the healthcare workers to automatically direct alarms to the nearest person. Hence when an alarm is triggered, the closest healthcare assistant or nurse receives a pop-up message on their screen, and the location of the alarm, as well as the location of two nearby colleagues, is highlighted on the map (figure 6). If assistance is needed, the nearby colleagues can then be contacted via Skype. In the event that the closest healthcare worker is not able to respond, he or she can reject the alarm by which it is directed to the second closest person, and so forth. Again, this implementation represents a design decision to rely on the ability of the spatial representation of the shared work domain, in conjunction with the healthcare workers' professionalism, mutual trust, and well-established social conventions, to inform well-functioning collaboration rather than enforcing more rules and formalizations on peoples' workday.

4.5 Implementation

GeoHealth was implemented using web technologies and languages such as XHTML (see W3C 2002), CSS (see W3C 2009), JavaScript (see Eich 2005), and the Ajax framework (see Garrett 2005), thus allowing the healthcare workers to interact with the system through a standard web browser. The server-side component was implemented on an experimental set-up consisting of a Debian Linux system, an Apache web server, a MySQL database server, and the PHP scripting language engine (see Lerdorf et al. 2002). For the evaluation studies, the GeoHealth prototype was accessed through a Firefox browser on an Apple MacBook connected to the Internet via GPRS on a Bluetooth-enabled mobile phone. The user's position was tracked with a GPS device connected to the MacBook's USB port.

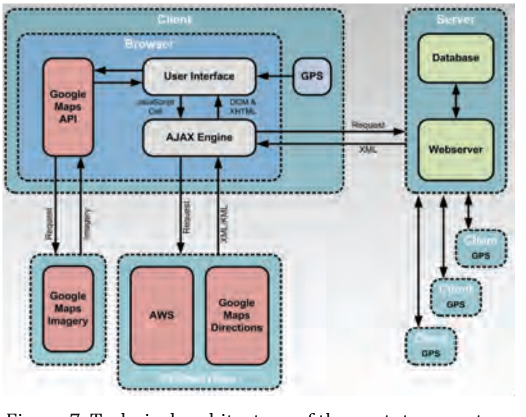


Figure 7. Technical architecture of the prototype system

GeoHealth's user interface is built on top of Google Maps embedded into a web page in full screen. The floating palettes are made from layered XHTML dynamically updated using JavaScript. The engine used to run the interface was created using object-oriented

JavaScript and asynchronous XML requests, which makes the user interface independent of page reloading and improve possibilities for interactivity and system responsiveness to user interaction both locally and remote. In popular terms, Ajax provides a framework that allows developers to create desktop-like applications for the web by relying on XML for data transport "behind the scenes" of the graphical interface. Thereby the user is able to interact with the application while it is requesting or processing data in the background even though the application is run in a web browser. The prototype architecture is illustrated in figure 7.

The information layers superimposed onto the geographical map are constructed from markers and polylines. Locations of patients on the map are calculated on the basis of their addresses' latitude and longitude coordinates provided through a look-up in the Address Web Service (AWS) provided by The National Survey and Cadastre of Denmark, which contain the geographical coordinates of all addresses and in the country. Routes between addresses on the map are generated on the basis of a look-up in the Google Maps directions service, which return a KML file with a sequence of coordinates making up a correct path following the road network. All requests to AWS and Google Maps directions are done through the Ajax engine. The flow is illustrated in figure 8.

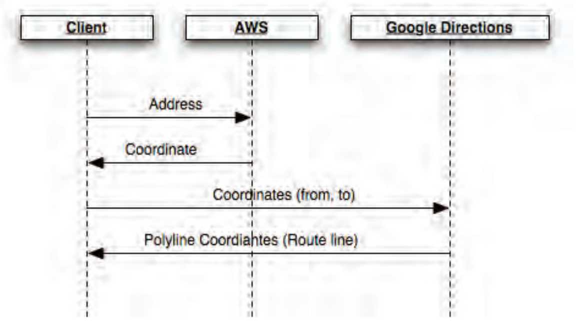


Figure 8. Information flow in the prototype system

Below is an example extract from a KML file defining a point on the route (one of several) and the geometry of the complete route from start to end point.

```
<Placemark>
      <name>Head west on Solvænget</name>
      <description><![CDATA[go 60 m]]></description>
      <Point>
                                <coordinates>9.514260,56.796510,0</coordinates>
      </Point>
      <LookAt>
                                <longitude>9.514260</longitude>
                                <latitude>56.796510</latitude>
                                <range>100.000000</range>
                                <tilt>45.000000</tilt>
                                <heading>350.060608</heading>
      </LookAt>
</Placemark>

(…)
```

```
<Placemark>
      <name>Route</name>
      <description><![CDATA[Distance: 1.1 km (about 2 mins]]></description>
      <GeometryCollection>
      <LineString>
      <coordinates>9.514260,56.796510,0.000000 9.514260,56.796510,0.000000 9.513280,56.796500,0.000000
9.513280,56.796500,0.000000 9.513090,56.797420,0.000000 9.513110,56.797790,0.000000
9.513250,56.798380,0.000000 9.513250,56.798380,0.000000 9.515990,56.798170,0.000000
9.517050,56.798120,0.000000 9.518940,56.798180,0.000000 9.520890,56.798440,0.000000
9.520890,56.798440,0.000000 9.521490,56.797410,0.000000</coordinates>
      </LineString>
      </GeometryCollection>
</Placemark>
```


The individual client parts of the application running on the healthcare workers' laptops communicate with each other via a central database and a web server. However, unlike a conventional web site or web application, all requests to the web server are done through the Ajax engine and are requests for XML data rather than pages of XHTML. Hence, the content displayed in the browser is dynamically updated whenever content changes, and pages are not reloaded every time the user interacts with the system. As a result, the user experience of GeoHealth is much more smooth and continuous than that of conventional web-applications.

Applications based on Google Maps are quite demanding on download bandwidth for graphics. For a map of 1700 x 1200 pixels, each zoom level requires 200Kb-1Mb of data depending on view mode (map, satellite or hybrid). Since bandwidth on the mobile Internet is still a rather scarce and expensive resource in areas covered only by GPRS or 3G networks, the GeoHealth prototype system caches all data from the Google Maps imagery server locally on the client through a proxy server set-up. Thereby, maps and satellite images are only downloaded to the client once (until updated on the server), and the response time of the application is improved significantly. Using this approach, a GPRS connection of 56kbit/s proved sufficient enough for running the prototype application smoothly once maps and satellite images have been cached. Although areas with cheap high speed wireless broadband access will unquestionably grow, we believe that it is important to construct mobile Internet services and applications in a way that is still as bandwidth efficient as possible.

As GeoHealth is a network application with distributed clients, it relies on network connectivity to function optimally. In case of a network failure some of the functionalities of the system would be lost, notably the ability to communicate and the exchange of real-time information. In relation to this, we subscribe to the principle of *volatility* promoted by Kindberg and Fox (2002). It is important to avoid designing systems on the misleading assumption that such failures are only rare. Instead, system designs should assume failure, and explicitly allow graceful degrading of functionality in response without expensive processes of recovery. In the case of GeoHealth, using a caching-approach as described earlier, the work tasks of each individual worker, and associated information from the central database, can be downloaded and stored locally on the client while within in network coverage, thus minimizing the implications of a network outage on the work activities. In case of a positioning failure, all information remains accessible by clicking on the map.

5. EVALUATIONS

We evaluated GeoHealth through qualitative field and laboratory studies involving nine healthcare workers. The evaluations took place in the municipality of Aars, Denmark. The evaluation sessions were structured by task assignments prompting the users to interact with particular parts of the prototype. During the evaluation sessions, the users were asked to think-aloud and respond to a series of interview questions about their perception, interaction, and use of the system.

The aim of the field evaluation was to investigate into the use of the prototype system in realistic surroundings. For this purpose, we installed a laptop computer with GPS in a minibus equipped with video and audio recording equipment (figure 9), and then let the healthcare workers use this vehicle to visit locations on their normal routes. As we were not at this point interested in issues related to interacting with the system while driving, most interaction took place while the vehicle was not in motion. The time used to drive from location to location was spent on follow-up questions and general discussions about the use of the system. The aim of the laboratory evaluations was to supplement findings from the field evaluation with more empirical data about the healthcare workers' perception and user experience of the prototype system. For this purpose, we set up a temporary laboratory in a dedicated room at the healthcare workers' central office, in which they could interact with the prototype system while seated at a desk.

Figure 9. In-car field evaluation of GeoHealth prototype.

The nine participating healthcare workers were all women and either employed as nurses or healthcare assistants. They were between 32 and 53 years old, and had 6 to 28 years of experience with their job. Most were experienced users of IT but one user characterized herself as a novice and one as an expert. All evaluations were recorded on digital video showing the users and their interactions with the system. In addition, the prototype screen was captured to a movie file using a screen logging utility running on the laptop. The evaluation sessions in the field took 30-40 minutes while the sessions in the lab took 25-30 minutes. For ethical reasons none of the evaluation sessions involved real patients or real patient data but was based on fictive data and scenarios informed by our ethnographic studies of the domain.

Video data from the evaluation sessions underwent qualitative content analysis (Strauss 1987) carried out by two researchers with experience in human-computer interaction and qualitative analysis methods. Following the prescribed method, video data was first analysed and coded by the two researchers independently. Secondly the two lists of codes were merged into one coherent list. The final list amounted to 25 unique codes. These codes were subsequently affinity diagrammed (Beyer and Holtzblatt 1998) into themes on higher levels of abstractions leading to the creation of 8 overall categories. Below, we present and discuss highlights from these.

6. FINDINGS AND DISCUSSION

From our evaluations it was clear that the users understood the basic design of GeoHealth and were able to interact with its basic functionalities after very short time. The users understood the map representations and the floating control panels (icons, colour coding etc.), as well as the close relationship between them. They also expressed that the spatial and temporal/sequential representations of work tasks complemented each other very well, and that the spatial representation gave them a good and useful overview of their own work activities and supported awareness about their co-workers without requiring too much attention. On the down side, some concerns were voiced in relation to difficulty in separation of similar coloured routes and tasks on the map (i.e. the ones coloured red and the ones coloured orange). Obviously, this problem would increase if even more healthcare workers were to be represented at the same time, and great care should be taken when selecting a palette of possible colours. Some users also had problems with navigating and zooming the map.

Making information and note taking facilities about patients available electronically directly through call-outs on the map was received very positively, and it was confirmed that the use of satellite imagery aided navigation by means of prominent physical properties of the surroundings. However, it was also observed that the strong colours of satellite imagery sometimes interfered with the colour coding of work tasks and routes making the information overlays difficult to read (Denmark is a very green-looking country).

In relation to the location-based information push, the healthcare workers were very happy with the way GeoHealth automatically presented information about assigned patients when entering vicinity of their home. It was expressed that the change of visual appearance on the screen caught attention without being intrusive, and that it was "natural" that the call-out box automatically disappeared when leaving vicinity. Nobody expressed, and it was not observed, that this use of context-awareness information push took control away from the user. However, concern was expressed that the status of a patient might change simply by driving past their house (i.e. from "pending" to "done"), and that such automatic change should require a certain "dwell-time" at the physical location of the house. This concern confirms the issue of partitioning streams of location data highlighted by Raper et al. (2007), and addressed by, for example, Krimm and Horvitz (2006) and Liu et al. (2006).

In term of support for location-based ad-hoc exchange of work tasks, all healthcare workers were able to use the drag-and-drop technique with icons on the map. The simplicity of the procedure was appreciated, and users explicitly expressed that doing this within a spatial representation of the work domain helped them consider which colleague to negotiate exchanging tasks with based on vicinity to planned routes. On the negative side, however, some healthcare workers expressed that although they were able to understand and use the specific implemented interaction technique, it felt odd to be moving icons of patients to the healthcare workers on the map, since this did not match the corresponding effect in the real world (where the healthcare workers are the ones who go to the patients). As a solution, it was suggested to "reverse the technique", so that one should drag the healthcare worker icons to the locations of the patients on the map instead in order to pair the two.

In relation to alarms, the healthcare workers were very positive towards the visualisation and automatic prompt of the closest person. Combined with the representation of nearby colleagues, it was expressed that this would significantly ease the amount of time as well as cognitive efforts spent on coordination. In extension of the new way of communicating in relation to alarms, the healthcare workers also found that the use of text-based messages constituted a significant advantage over current practice because it eased their need for making phone calls about every little detail. However, it was also stressed that text messages should not replace verbal communication, and that access to a verbal communication channel (e.g. Skype) would be essential for complex coordination as well as for the social purposes.

Extending the scope of patient monitoring beyond simple user-triggered alarms, it was discussed to embed live patient data into the GeoHealth system relayed from mobile wireless sensors over the Internet, as described in for example Figueredo and Dias (2004), Tachakra et al. (2003), Varshney (2003).

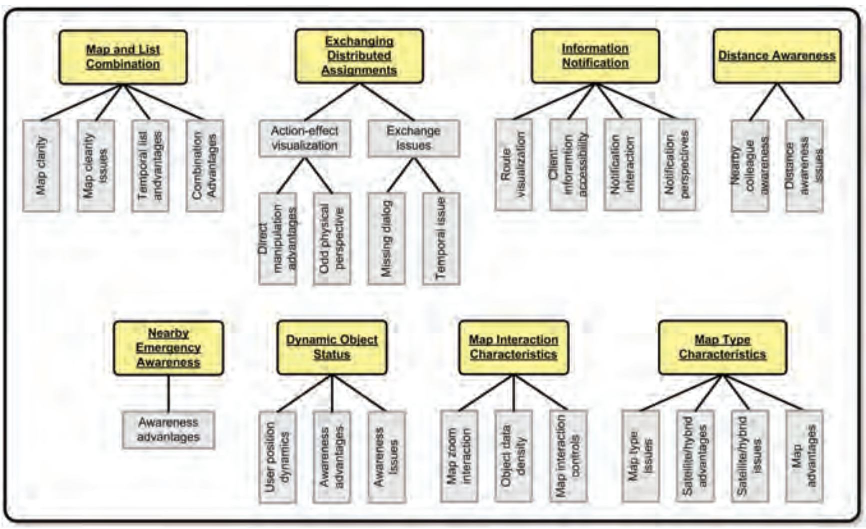


Figure 10. Thematic clustering of findings from the evaluations.

As discussed in Raper et al. (2007, 32-33), location-based services tracking the location of its users represent an ethical challenge of privacy and disclosure. In relation to this potential ethical issue of the GeoHealth application, none of the users expressed concerns about their movements during the workday being tracked via GPS. However, it should be noted that given the short time frame of the evaluations, this is not sufficient enough ground to conclude that tracking the location of mobile healthcare workers throughout their workday is acceptable as such. The ethical issues of this functionality should be investigated further, and mechanisms for ensuring privacy in this class of location-based services should be developed and evaluated with users.

7. CONCLUSIONS AND FURTHER WORK

We have described a location-based service, GeoHealth, for physically distributed and nomadic home healthcare workers that combine Google Maps, GPS positioning, and Web 2.0 technology. The system provides information based on the user's location, represents live contextual data about a shared work domain on an interactive map displayed in a web browser, and has facilities for communicating and coordinating work tasks among distributed mobile users. Through user-based evaluations, we have verified the overall design and functionality of the prototype system as well as identified a number of areas for improvements to its functionality and interaction design.

From the user evaluations, we learned that the healthcare workers were positive towards the use of a location-based service in their work activities while distributed throughout the municipality. Regarding the user experience of GeoHealth specifically, we learned that the dynamic and interactive map-based representation of the work domain was able to support collaboration, communication, coordination, and peripheral awareness amongst the distributed co-workers. The users appreciated the overview of their shared work-domain, the way information was gently pushed to them based on their location, the access to electronic patient data directly in the map view, the integration with voice and text based communication, and the location-based handling of alarms. On the basis of the feedback from our test-users, we conclude that location-based services such as GeoHealth, built around Web 2.0 technologies, have unexploited potentials for the creation of new and positive user experiences within the domain of home healthcare. We speculate that similar location-based services would be applicable and valuable to other domains.

The presented study leaves several areas for further work. Firstly, the prototype should be extended into a fully functional and stable system with improved interaction design, which can be deployed into the home healthcare organization for a longer period of time (e.g. 6-12 months). During this trial period, the prototype design should continuously be improved through iterative cycles of user feedback and changed to the interface design. Facilitated by this process, longitudinal studies of real world use should be done in order to gain knowledge about long-term usage and potential changes happening to the way work is carried out in the use organization triggered by the implementation of this new technology. As a part of this research, quantitative and qualitative measures should be done of improvements (or degradations) to the quality of service provided. Enquiries should also be made into the patients' perception of and opinion about the use of such technologies within home healthcare. Finally, the healthcare workers perception of being tracked by GPS throughout their workday should be investigated.

For the first functional prototype iteration described in this paper, our focus has been on the user experience of the functionality envisioned for GeoHealth. Hence, there are a number of technical aspects that we have not yet implemented, which should be dealt with in future iterations. Firstly, the electronic patient record part was not integrated with the real national back-end database. Instead, a database was set up with temporary patient information for testing purposes. In order to do a realistic longitudinal study of GeoHealth, the information in the system should be integrated with the real electronic patient record for the municipality. Secondly, no mechanisms for dealing with issues of data security and privacy were implemented yet. Since GeoHealth handles sensitive and personal data, a fully implemented system, or even a prototype system deployed in the use organization over a longer period of time, should, of course, protect data from unauthorized access, for example by using encryption or secure protocols. Thirdly, future implementation work should deal with the challenges of system scalability. Finally, in terms of network failure, it is important that the current network status and a potentially degraded state of operation (i.e. no real-time information updates) are communicated clearly to the user. The same goes for potential positioning failure.

GeoHealth was implemented in 2006 at a time where handheld computing platforms imposed several limitations on interaction designers and systems developers. One year later Apple released their iPod Touch and iPhone. These devices have significantly changed the landscape of mobile computing, and are much more powerful both in terms of system performance, mobile usability and flexibility to implement well-functioning mobile applications and, in particular, location-based services. Extending the work presented in this article, it would be highly interesting to design, implement, and study the use of an iPhone-powered GeoHealth System. While much of the functionality of the presented prototype can be achieved technically on an iPhone, the question remains if the handheld form factor and limited screen real estate is sufficient or not for this type of application and in this particular use domain.

ACKNOWLEDGEMENTS

The authors thank the staff at the home healthcare unit of the municipality of Aars, Denmark for their participation throughout the project.

REFERENCES

Aaltonen, A. and Lehikoinen, J. (2005) Refining visualization reference model for context information. Personal and Ubiquitous Computing, 9(6), 381-394.

Bardram, J. (2004) Applications of Context-Aware Computing in Hospital Work – Examples and Design Principles. Proc. SAC'04, pp. 1544-1579.

Barkhuus, L. and Dey, A. K. (2003) Is Context-Awareness Computing Taking Control away from the User? Three Levels of Interactivity Examined. Proc. UbiComp 2003.

Beyer H, and Holtzblatt, K. (1998) Contextual Design. Morgan Kaufmann, San Francisco.

Boulos, M.N.K, Rocha, A., Martins, A., Escriche., Vicente, M.E., Bolz, A., Feld, R., Tchoudovski, I., Braecklein, M., Nelson, J., Laighin, G.Ó. Sdogati, C., Cesaroni, F., Antomarini, M., Jobes, A., and Kinirons, M. (2007) CAALYX: a new generation of location-based services in healthcare. International Journal of Health Geographics, 6:9.

Boulos, N.M.K (2003b) Location-based information services: a new paradigm in personalised information delivery. International Journal of Health Geographics, 2:2.

Boulos, N.M.K (2003a) The use of interactive graphical maps for browsing medical health Internet information ressources. International Journal of Health Geographics, 2:1.

Brown, B. and Randell, R. (2004) Building a Context-Sensitive Telephone: Some Hopes and Pitfalls for Context Sensitive Computing. CSCW, 2004(13), 329-345.

Chang, A.Y., Parrales, M., Jimenez, J, Sobieszczyk, M.E., Hammer, S.M. Copenhaver, D.J., and Kulkarni, R.P. (2009) Combining Google Earth and GIS mapping technologies in a dengue surveillance system for developing contries. International Journal of Health Geographics, 8:49.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C. (2000) Developing a Context-Aware Electronic Tourist Guide: Issues and Experiences. Proc. CHI 2000, pp. 20-31.

Christensen C.M., Kjeldskov, J., and Rasmussen, K.K. (2007) Geohealth: A location-based service for nomadic home healthcare workers. Proc. OzCHI 2007, ACM and CHISIG.

Dey, A. and Mankoff, J. (2005) Designing Mediation for Context-Aware Applications. ACM Transactions on Computer-Human Interaction, 12(1), 53-80.

Driedger, S.M., Kothari, A., Morrison, J., Sawada, M., Crighton, E.J., and Graham, I.D. (2007) Correction: Using participatory design to develop (public) health decision system support through GIS. International Journal of Health Geographics, 6:53.

Edwards, W.K. (2005) Putting Computing in Context: An Infrastructure to Support Extensible Context-Enhanced Collaborative Applications. ACM Transactions on Computer-Human Interaction, 12(4), 446-474.

Eich, B. (2005) JavaScript at ten years. In Proceedings of the tenth ACM SIGPLAN international conference on Functional programming. ACM, p 129.

Erickson, T., Halverson, C., Kellogg, W. Laff, M. and Wolf, T. (2002) Social Translucence: Designing social infrastructures that make collective activity visible. Communications, 45(4), 40-44.

Figueredo M.V.M. and Dias J.S. (2004) Mobile Telemedicine System for Home Care and Patient Monitoring. Proc. IEEE Engineering in Medicine and Biology Society (EMBS 2004), pp. 3387-3390.

Garrett, J. J. (2005 Ajax: A New Approach to Web Applications. Adaptive Path. http://www.adaptivepath.com/ideas/essays/archives/000385.php (accessed 6 September 2009)

Jones, Q., Grandhi, S.A., Terveen, L. and Whittaker, S. (2004) People-to-People-to-Geographical-Places: The P3 Framework for Location-Based Community Systems. CSCW, 2004(13), 249-282.

Kaasinen, E. (2003) User Needs for Location-Aware Mobile Services. Personal and Ubiquitous Computing, 7(1), 70-79.

Kindberg, T., and Fox, A. (2002) System Software for Ubiquitous Computing. IEEE Pervasive Computing, vol. 1, no. 1, pp. 70-81

Kjeldskov J. and Skov M.B. (2007) Exploring Context-Awareness for Ubiquitous Computing in the Healthcare Domain. Personal and Ubiquitous Computing, 11(7): 549-562.

Kjeldskov J. and Paay J. (2006) Public Pervasive Computing: Making the Invisible Visible. IEEE Computer, 39(9), 60-65.

Krimm, J. and Horvitz. E. (2006) Predestination: inferring destinations from partial trajectories. UbiComp 2006.

Lemmelä, S, and Korhonen, H. (2007) Finding Communication Hot Spots of Location-Based Postings. Proc. CHI 2007 work-in-progress, pp 2549-2554.

Lerdorf, R., Tatroe, K., Kaehms, B., and McGredy, R. (2002) Programming PHP. O'Reilly & Associates Inc

Liu, J., Wolfson, O., and Yin, H. (2006) Extracting semantic location from outdoor positioning systems. Proceedings of the 7th International Conference on Mobile Data Management (MDM'06). New York: IEEE Computer Society.

Mandellos, G.J., Lymperopoulos, D K., Koukias, M.N., Tzes, A., Lazarou, N., and Vagianos, C. (2004) A Novel Mobile Telemedicine System for Ambulance Transport. Design and Evaluation. Proc. IEEE Engineering in Medicine and Biology Society (EMBS 2004), pp. 3080-3083.

Mattila E, Parkka J, Hermersdorf M, Kaasinen J, Vainio J, Samposalo K, Merilahti J, Kolari J, Kulju M, Lappalainen R, and Korhonen I (2008) Mobile Diary for Wellness Management—Results on Usage and Usability in Two User Studies, IEEE Transactions on Information Technology in BioMedicine, 12(4): 500-512

Maglogiannis I., Kazatzopoulos L., Delakouridis K., Hadjiefthymiades S. (2000) Enabling Location Privacy and Medical Data Encryption in Patient Telemonitoring Systems. IEEE Transactions on Information Technology in Biomedicine, 13(6): 946-954

Maglogiannis, S. Hadjiefthymiades (2007) EmerLoc: Location-based services for emergency medical incidents. International Journal of Medical Informatics, 76(10): 747-759

Millen, D.R. (2000) Rapid ethnography: time deepening strategies for HCI field research. Proc. DIS 2000, pp. 280-286.

Paay J., Kjeldskov J., Howard S, and Dave, B. (2009) Out on the town: a socio-physical approach to the design of a context aware urban guide. ACM Transactions on Computer-Human Interaction (TOCHI) 16(2).

Paay, J. (2008) From Ethnography to Interface Design. In J. Lumsden (Ed.), Handbook of Research on Mobile User Interface Design and Evaluation. Idea Group, Inc. Chapter 1.

Pattichis, C.S., Kyriacou, E., Voskaride, S., Pattichis, M.S., Istepanian, R., and Schizas, C.N. (2002) Wireless Telemedicine Systems: An Overview. IEEE Antennas and Propagation Magazine, 44(2), 143-153.

Paulos, E. and Goodman, E. (2004) The Familiar Stranger: Anxiety, Comfort and Play in Public Places. Proc. CHI 2004, pp. 223-230.

Pitsillides, A., Samaras, G., Dikaiakos, M., Olympios, K., Christodoulou, E. (1999) DITIS: Collaborative Virtual Medical Team for Home Healthcare of Cancer Patients. Proc. Information Society and Telematics Applications.

Pruitt, J. and Adlin, T. (2006) The Persona Lifecycle: Keeping People in Mind Throughout Product Design. Morgan Kaufmann.

Raper, J., Gartner, G., Karimi, H., and Rizos, C. (2007) A critical evaluation of location based services and their potential. Journal of Location Based Services, 1(1), 5-45.

Rainham, D., Krewski, D., McDowell, I., Sawada, M., and Liekens, B. (2008) Development of a wearable global positioning system forr place and health research. International Journal of Health Geographics, 7:59.

Reponen, J., Niinimaki, J., Holopainen, A., Jartti, P., Ilkko, E., Karttunen, A., Kumpulainen, T., Tervonen, O, Paakko, E. (2000) MOMEDA - A Mobile Smartphone Terminal for DICOM Images and Web-Based Electronic Patient Data. Journal of Telemedicine and Telecare, 6(1), 45-49.

Sawhney, N. and Schmandt, C. (2000) Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. ACM Transactions on Computer-Human Interaction (TOCHI) 7(3): 353-383

Strauss, A. (1987) Qualitative analysis for social scientists. Cambridge University Press.

Skov, M.B. and Høegh, R.T. (2006) Supporting Information Access in a Hospital Ward by a Context-Aware Mobile Electronic Patient Record. Personal and Ubiquitous Computing, 10(4), 205-214.

Tachakra, S., Wang, X.H., Istepanian, R.S.H., and Song Y.H. (2003) Mobile e-Health: The Unwired Evolution of Telemedicine. Telemedicine Journal and e-Health, 9(3), 247-257

Varshney, U. (2003) Pervasive Healthcare. IEEE Computer, 36(12), 138-140.

Woodward, B., Istepanian, R.S.H., and Richards, C.I. (2001) Design of a Telemedicine System Using a Mobile Telephone. IEEE Transactions on Information Technology in Biomedicine, 5(1), 13-15.

W3C (2009) Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification http://www.w3.org/TR/CSS2/ (accessed 6 September 2009)

W3C (2002) XHTML™ 1.0 The Extensible HyperText Markup Language (Second Edition) http://www.w3.org/TR/xhtml1/ (accessed 6 September 2009)

Chapter 18

ArchiLens

Mikael B. Skov, Jesper Kjeldskov, Jeni Paay, Niels Husted,
Jacob Nørskov and Kenneth Pedersen

Abstract. Some movements within modern architecture particularly emphasize the importance of matching buildings to their surroundings. However, practicing such "contextual architecture" is highly challenging and typically not something the future inhabitants of a building are well equipped for participating in. This paper explores the potentials of using mobile phone technology for facilitating such client participation in the parts of an architecture process that takes place on the building site. For this we introduce ArchiLens, a mobile system for interactive on-site 3D visualization of houses, and findings from a field study with 40 participants in the process of building or modifying their home. The study showed that using the system helped evoke people's imagination of the look and feel of their future house, and envision it in context. This enabled them to participate more closely in the design process on-site by iteratively reviewing design alternatives and exploring, for example, other placements and materials.

1. INTRODUCTION

As articulated by the famous Finnish architect Eliel Saarinan *"one should always design a thing by considering it in its next larger context - a chair in a room, a room in a house, a house in an environment, an environment in a city plan".* This view represents a movement within modern architecture that particularly emphasizes the importance of working closely with the context in which a building is going to be located. This design philosophy has given rise to several highly acclaimed buildings around the world praised by their inhabitants for the way they fit naturally with their surroundings. However, practicing such "contextual architecture" (Brolin 1980, McGillick and Carlstrom 2002, Ray 1980), is a challenging undertaking for even the best architects, and is typically not something

Originally published as Skov, M. B, Kjeldskov, J., Paay, J., Husted, N., Nørskov, J. and Pedersen, K. (2012) Designing on-site: Facilitating Participatory Contextual Architecture with Mobile Phones. *Journal of Pervasive and Mobile Computing.* (in press)

that architects' clients – the future house owners or inhabitants – can easily participate in. One of the reasons for this is that exploring different architectural possibilities on the site where a building is going to be built requires highly developed skills of visual imagination often exceeding those of a non-architecturally trained person. For ordinary people to participate in such activity we need tools that can help stimulate imagination.

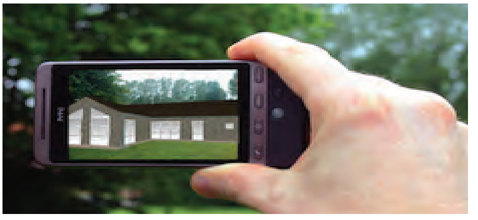


Figure 1. Visualizing architecture in context on a mobile phone while standing at the building site

Augmented reality (AR) allows superimposing digital graphics on the real world and thereby creating a visual blend between the real and the virtual (Piekarski and Thomas 2003). Key promises from this technology are that such composite scenes on mobile devices can enrich the experience of a physical space (Morrison et al. 2009). Until now real-time processing of 3D models and camera feeds have been problematic on mobile devices due to limited performance. However, with recent leaps in smart phone technology, such as higher CPU power, better displays and cameras, and spatial sensors like GPS and compasses, designers are now able to make advanced mobile AR applications available to ordinary people without requiring special and expensive hardware.

As an interesting area of application, we have been exploring how mobile AR can be used to facilitate ordinary people's participation in architectural design by evoking their imagination. For this purpose we have developed a mobile AR system, ArchiLens, which supports on-site interactive 3D visualization of buildings on a mobile phone (figure 1). The use of the system was investigated through a field study with 40 participants in the process of building or modifying their home, allowing them to explore and review design alternatives while present in its physical context.

2. RELATED WORK

Our research is grounded in two main bodies of related work. The first is Contextual architecture. Contextual architecture (Brolin 1980, Ray 1980, Summerfield and Hayman 2006), also referred to as contextualism (Shane 1976), represents a design philosophy where the interplay between form and context is given particular priority. Apart from having a notable effect on the produced outcomes, working within this design philosophy has some profound impacts on the process of design. Architects working closely with the context of their buildings spend significant time developing and assessing their design *on the building site* rather than at the drawing board in their studio. As an example, it is a well-known fact that the Danish architect Jørn Utzon, who is probably best known for the Sydney Opera House, spent considerable time on building lots exploring their contextual properties before and during the development of his building designs. In a rare interview he even described how he would sometimes map out the possible location of walls and windows by placing lines of small rocks on the ground, and then walk around the lot imagining the view of the surrounding environment from these as yet un-built rooms. An account of the works of Sydney contextual architect Alex Popov (a former associate of Utzon) describes how the result of buildings created with such sensitivity to the way they engage with their environment is that they do not just fit their context well, they themselves become parts of that evolving context (McGillick and Carlstrom 2002).

In architectural theory, contextual architecture is described as a matter of pursuing the notion of *genius loci* (the protective spirit of a place in classical Roman religion) by responding to the topographical, geographical, social and cultural context of a building site (McGillick and Carlstrom 2002). This concept is most notably explored by Norwegian architect and theorist Christian Norberg-Schulz in his phenomenology of architecture (Norberg-Schulz 1980) arguing that genius loci – or sensitivity to context – has profound implications for place making. Echoing this line of thought, American architect Christopher Alexander argues that architecture exhibiting a quality of timelessness always evolves through a series of “wholeness preserving transformations” in which the designer has not only focused on the creation of new form but done this with deep understanding of and respect for the existing context (Alexander 2002-05). Alexander’s Pattern Language (Alexander et al. 1977) contains a collection of problem-solution pairs composed to help evoking the reader’s imagination of the elements of a future design in context - enabling architects, as well as ordinary people building or modifying their home, to achieve such wholeness preserving contextual architecture.

The other body of related work is Augmented Reality. Augmented reality (AR) was coined as a term in the early 1990s and has its origins in virtual reality, where computer graphics completely immersed users in synthetic worlds or environments (Höllerer and Feiner 2004). However, where virtual reality disconnects the user from the physical world, AR superimposes virtual objects onto it (Azuma 1997).

Research in AR has explored various application domains and different technologies. There are several examples of tour guide applications such as in the ARCO project (Wojciechowski et al. 2004) where objects in a museum were augmented digitally, transforming the visitors from passive viewers and readers into active actors and players. Other examples include restaurant guides (Feiner et al. 197, Feiner 2002), other museums guides (Miyashita et al. 2008), and general guides for tourists (Reitmayr and Schmalstieg 2003, 2004). Another common application domain for AR is visualization of hidden or invisible information. Examples of this include the SiteLens system, which visualizes carbon monoxide levels in the user’s immediate surroundings (White and Feiner 2009), and the use of AR for visualizing hidden wires inside Boeing aircrafts in support of the manual manufacturing process (Caudell and Mitzell 1992, Mitzell 1994). Finally, and of particular relevance to the work presented here, AR has been used for visualizing architecture in-situ, with a number of studies investigating the strengths and limitations of this application domain. In an early piece of AR research, Guo et al. (2008) describe the presentation of buildings as one of the main features of this technology. As their example they describe the ARCHEOGUIDE system where Greek researches and the government used AR to visualize the Grecian Olympia. In an example of working with present construction projects, Schall et al. (2009) present an AR system using 3D models of underground infrastructure superimposed on the construction site.

While these and other research projects indicate that AR holds promising potentials for visualizing architecture, technology shortcomings have so far limited use in practice (Gleue and Dähne 2001; Reitmayr and Schmalstieg 2003). Two issues of particular importance are *accuracy of positioning* and *perception of depth.* As examples, during outdoor testing, Schall et al. (2009) experienced several problems related to shortcomings of GPS tracking resulting in poor accuracy for workers in street canyons. The same problem was reported by Reitmayr and Schmalstieg (2004). The issue of depth perception is reported by, amongst others, Schall et al. (2009) who experienced that some users had problems with perceiving the overlaid graphical representation of infrastructure buried under ground correctly. This is an issue related to registration, which is a cornerstone for achieving successful utilization of AR generally (Azuma 1997), and in particular when visualizing architecture. The challenge of registration is that objects in the real and virtual worlds must be accurately aligned; otherwise the illusion that the two worlds coexist will be compromised. Key sources of registration errors include optical distortion, poor positioning, mechanical misalignment and incorrect viewing parameters. This obviously motivates technology improvements, but the effect of suboptimal registration can also be reduced if users have good spatial understanding or perception (Azuma 1997).

As a consequence of the technological challenges of visualizing architecture in AR, previous research studies are dominated by the use of relatively advanced hardware setups during trials (Feiner et al. 1997, Feiner 2002, Schall et all 2009), while the use of off-the-shelf mobile devices is rarely seen. This means that although promising results may have been achieved in these trials, real world use of the developed applications by non-experts has not been realistic. However, in recent years smart phones have begun to meet the basic requirements for AR in terms of processing power, positioning, display size and camera resolution. With AR technically within reach of the general population there is an opportunity for developing applications for real world use as seen with Wikitude (2011) and Layar (2011). One such application, we speculate, is the use of AR for enabling private homeowners, or people in the process of building their future house, to participate more actively in a contextualized design process through an interactive visualization on their mobile phones. Exploring this idea, we have developed and deployed ArchiLens for evoking people's visual imagination and enabling them to explore and make decisions about placement, appearance, shape, and size of the building in context.

3. ARCHILENS

ArchiLens is a mobile application that integrates AR and 3D visualization. The system places a 3D house model in context, overlaid on the camera feed on a smart phone, making it possible for users to examine future house designs from different angles by moving around physically on the site. ArchiLens was implemented for a HTC Hero running Android 1.5. It was implemented in Java using Eclipse IDE paired with the official Android API and plug-in. The 3D house models were created in Google SketchUp and exported to OBJ format. In order to make the models faster, a file parser algorithm was constructed to import the OBJ information into OpenGL. In order to keep the 3D model from jumping around the screen when the user was holding the device still, we implemented smoothing algorithms for the GPS and orientation sensors data.

Figure 2 depicts the key views in ArchiLens, and in the use scenario we illustrate a typical situation. After this, we describe three central parts of ArchiLens' functionality: two-dimensional placement, spatial layout and characteristics, and visual appearance.

Figure 2. ArchiLens - A: Screen capture of a virtual house positioned on a lawn. B) Screen capture of the Blueprint View, showing a general overview of the floor plan. The red square represents the user's position on the blueprint. C) Placement View for house placement and rotation. The red shape represents the house, and touching the screen makes it is possible to rotate the house. D) View from inside the house. Notice the added dashed dividing lines between the roof and walls moving inwards and thereby creating perspective. E: Future view through the windows of the living room. Notice how the physical surroundings are visible through the windows in the virtual walls.

Use Scenario: John and Kate want to build a house where they can live with their children. They visit a lot for sale. Arriving at the lot, they start ArchiLens and choose a house model. They now place the house on the virtual representation of the lot. ArchiLens visualizes the house based on their preferences.

While walking around the lot, John gets different perspectives on how their choices fit into the surroundings. He notices all the neighbours have red bricks, and that their house is yellow. To see how their house would fit into the surroundings with red bricks, he changes materials through ArchiLens. He is surprised to see how the house changes with the new bricks. It nearly vanishes into the surroundings. John confers with Kate about the chosen colours. In cooperation, they decide that their initial idea about the yellow bricks is the best solution.

Kate notices that a neighbour's house is quite close to their panorama windows in the living room. She uses the blueprint drawing in ArchiLens to position herself at the panorama inside the virtual living room. She starts to check how the neighbour's house impacts their view from the living room and to her disappointment she realizes that the neighbour's house totally blocks the view. They are therefore forced to rethink the position and orientation of their house.

3.1. Two-Dimensional Placement

When the user has chosen a house model, the first step is to place it on the lot. ArchiLens supports this through a separate two-dimensional placement view using Google Maps. In this view, a visual link is created between the ArchiLens system and the user's physical surroundings by overlaying a proportionally correct outline of the house model on a satellite image (see figure 2C). Two main interactions are then available to the user. They can change the position of the house by moving it to any location on the Google map - the map adjusts and the house remains centred in the screen view. Using an on-screen control the user can also rotate the house around its centre point. When the user is happy with the placement the application changes to 3D view, but at any point in time the user can go back to the placement view and re-position or rotate the house model.

3.2. Spatial Layout and Characteristics

ArchiLens uses GPS positioning and orientation sensors to create an illusion of moving around the house and viewing it from different angles (see figure 2A). The GPS coordinates for the house and for the user are used to calculate the position of the house relative to the user. The calculation involves the distance to and bearing between these two points. In order to determine camera orientation, the built-in sensors provide information about how the phone is rotated according to azimuth, pitch, and roll.

The visualization of the virtual house has two different views that the user can alternate between by holding the phone horizontally with the camera facing the ground or vertically with the camera facing the horizon. We call the view presented when held vertical "3D View" (shown in figure 2A) and the view presented when held horizontally "Blueprint View" (figure 2B). In Blueprint View, the red square in the centre of the screen marks the users position in relation to the house.

3.3. The Visual Appearance

ArchiLens allows the user to change the visual appearance of the house while standing on the lot and to explore the use of alternative building materials, including brick type, roof type, colour of the windows, doors and woodwork. This is done by activating a menu on the screen, selecting the element to modify, and browsing through the available options by swiping the screen from left to right. When a different material is selected, it is passed along the rendering pipeline and the specific segment of the house model will change texture in the 3D view.

4. FIELD STUDIES

We studied the use of ArchiLens through a 5-week field study in spring 2010. The purpose of the study was to explore how ArchiLens evoked the imagination of ordinary homebuilders or renovators when given an interactive AR representation of their future house on the building site. In order to achieve diversity in feedback on the use of ArchiLens for different user groups, the field study involved participants from three different phases of building and modifying a private home: 1) open lot 2) signed contract, and 3) house extension. *Open lot* is the phase where people want to build a new house, but have not yet initiated the actual design or construction process. Instead they are searching for a suitable piece of land and to get overall ideas about possible house locations and positioning. *Signed contract* is the phase where people have bought the lot and made a contract with a construction company to build their house. Hence, they are now in the process of having their house built, and might like to be actively involved in the process. Finally, *house extension* is the phase where current homeowners wish to extend their existing house. They are therefore interested in how the new extension will align with the old part of the house, as well as the surroundings. This group may or may not have a contract with a construction company.

4.1. Participants

Forty people (including 11 children) from 14 families participated in the field study over a combined period of 5 weeks. The adults were aged from 28 to 41 with the majority in their mid thirties. Participants were recruited and grouped in respect to the three phases of building.

Open lot

Twelve adults and two children from six different families participated in this group. During Open House or Open Lot events in the vicinity of Aalborg, families were asked to be a part of the study. They were approached when leaving these events held by local house development companies. Participants in this category usually only had vague ideas of their building criteria. Therefore they sought inspiration on house designs and most of them did not yet own land for the house. These participants were presented with variations of tract houses made by the house development company "Vendia-Huset".

Signed contract

Twelve adults and six children from six families participated in this group. These participants all owned a vacant lot and had either signed a contract with Vendia-Huset for a specific house, or were in the process of signing. All families had seen floor plans of their houses, and three of them had also seen front elevation drawings. One family even had a 3D drawing of their future house. All participants were in their mid thirties with young children. For this group of participants we received individual house blueprints from Vendia Huset and created corresponding 3D models with possible choices of materials for the families to explore through ArchiLens.

House extension

Five adults and three children, from two families were recruited into this group, based on their wish to make a house extension. One of the adults was the designated architect for both families. Both families had ground plans and different drawings of their planned house extensions. These plans, models, and possible choices of materials were turned into corresponding 3D models viewable in ArchiLens.

4.2. Procedure and Data Collection

All field study sessions were conducted in-situ at the users own lot or at a lot that was for sale (if the participant was part of the Open Lot group). In total, we conducted 16 field sessions, as two families from the signed contract group participated in two sessions each. Each session in the field lasted between 10 and 65 minutes (most around 45 minutes) with participants from the signed contract group generally spending more time with the system than others.

The participants were given a short presentation of the system and the purpose of the study. They were then asked to fill out a questionnaire and take part in a semi-structured interview. We then encouraged them to explore their house through ArchiLens by walking around the 3D house placed on the lot. A facilitator and an observer followed each group of participants, at a distance. The facilitator's role was to help with any problems and to probe for thoughts and comments from the participants. The observer noted down comments, actions, and interactions during the trial. When the participants had finished using the system, another questionnaire was handed out and a final interview was conducted to explore their experience and use of the system. Two families from the signed contract group wished to participate in a second session with a different 3D model altered on the basis of their input and requests from the first session. These sessions followed the same procedure.

Four months after the field trials, we conducted follow-up interviews with the families from the signed contract group to investigate their experiences of the built house. At this stage all six houses for this group were fully completed and most of the families had recently moved in. We conducted the post-construction interview to put the families' initial experiences of ArchiLens into perspective. In particular, we sought to understand how well the completed house matched their envisioned house, and what role ArchiLens had played in evoking and shaping their visualization.

For the field study as a whole, we collected multiple types of data: questionnaires, written notes, video recordings and voice recordings. The first questionnaire asked for demographic data and responses to four general assertions using a five point Likert scale. Interviews were recorded in audio files, and families using the system were recorded on digital video. Researcher observations were described in notes. The second questionnaire consisted of the same four general assertions and specific statements regarding the use of ArchiLens, with a five point Likert scale for responses.

4.3. Data Analysis

The data analysis followed a Grounded Theory approach (Strauss and Corbin 1998) to identify and classify identities and relations. Themes were generated by systematic use of techniques and procedures to divide the qualitative data into controllable elements before using this foundation to create higher-level concepts. Transcripts of the recordings and written researcher notes were analyzed concurrently. First, open coding was used to discover 362 different properties, which identified 41 phenomena. Secondly, axial coding was used to create structure in the data and make categories based on the phenomena. Fourteen categories were derived from the phenomena. Thirdly, selective coding was used to relate the categories to each other, with the purpose of gaining an understanding of how the categories were interrelated and to find the main themes. This resulted in four themes: 1) enhancing visual perception, 2) priming imagination in context, 3) collective social understanding, and 4) dimensional insight.

Questionnaire answers were then compared with the four themes, distilled from the notes and interviews, looking for convergence between the two.

5. FINDINGS

The following sections report the findings from our field study. This section is organised according to the four themes emerging from the data analysis.

5.1. Visual Appearance

Our participants were generally very interested in exploring the visual appearance of their future house. Several of them spent considerable time viewing the house from different angles when it was projected onto the building site through ArchiLens.

Furthermore, some of the participants stated that they had problems in perceiving more traditional blueprints and imagining what different shapes and materials would look like on-site: *"I simply don't like blueprints, I don't understand them ... while this is much easier to understand" [A].*

Figure 3. People achieved a better understanding of the visual appearance of their future house when using ArchiLens.

Exploring the possibilities of ArchiLens made them appreciate that seeing their future house on the building site and being able to move around it in context made a huge difference on their visualization of it (as illustrated in figure 3). As a related finding we observed that ArchiLens helped uncover areas where the participants' conception of the current design was incorrect or incomplete. This can, for example, be seen in the following comment made by a participant, who prior to trying ArchiLens had expressed that she had a very good visualization of her future house: *"I actually thought I had a good idea of how my house was going to look, but there were things that surprised me and helped me understand a few things. All because I could see the house from different angles"* [B]. When we conducted the post-construction interview with her four months later, just after she had moved into the house, she followed up on this statement and expressed that in contrast to the blueprints the visualization evoked by ArchiLens was very well aligned with how the house eventually looked.

Overall, most participants felt that they had achieved a better visualization from using the system. In fact, many participants were surprised to realize that what they saw on the display when walking around the lot did not really match their imagination. One of the things they learned was that they didn't have a correct sense of the scale of the building as a whole. This was clearly reflected in the second questionnaire where they were asked, for a second time, to grade their visualization of the house. Here, several participants expressed that they had scored this item too high in the first questionnaire. Consequently average scores for the question "It is easy for me to visualize the facades and appearance of the house on the lot" moved from 3 before using ArchiLens to 4.5 after.

Visualizing the house in context also had a huge effect on how the participants bonded with their future home. As one participant stated while standing at the location of her future living room: *"This is so cool, that I can stand here in my living room and imagine what views I will have" [C].* Such bonding and understanding had reportedly been difficult using traditional blueprints or elevation drawings as people found it difficult to create mental images of the house in context on the basis of such drawings. Being able to see the house visually, and in its real physical surroundings, from the exact spot of interest supported this experience. The importance of the surroundings was also highlighted in a session where the camera feed suddenly turned off. This immediately caused the participant to comment that the experience was spoiled, and ask us to fix the device because they weren't finished exploring the house.

Two of the families participated in two sessions on their lot. One of these families used the second session to review the redesign of a carport, which had emerged from the first session, as well as other changes. This active participation in the design process and single iteration of changes helped them develop their understanding and imagination of physical movement around the house: *"We have talked about how quick we would be able to drive into the carport from the road - we were afraid it would be too 'racing like'. But now that I can see it from here, I actually don't think it is going to be a problem" [D].*

Somewhat surprisingly, 3D registration (the accuracy of the visualization) was not a major concern for most participants. However, for the house extension group, this did cause some problems, and some participants were not satisfied with ArchiLens' ability to display this. In terms of extensions it was obvious that the system's current 3D registration accuracy simply wasn't good enough for seamlessly integrating a 3D model alongside an existing physical structure. This resulted in some major difficulties maintaining the illusion of an augmented reality along the walls and corners of existing buildings (figure 4).

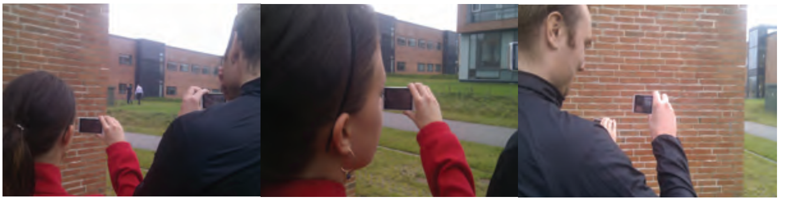


Figure 4. Participants experiencing problems with visualizing a house extension related to inaccurate 3D registration.

5.2. Relationship to Context

Visualizing the house on the actual lot served as a primer for the participants to consider the interplay between their house and its surroundings. This was evident in the participants' intentions and behaviour when investigating the appearance of their house in its physical surroundings, and the appearance of these surroundings from the house.

Building and moving into a new house involves not only getting to know a new location and house, but often also getting acquainted with new neighbours. During observations and interviews, we observed several times that participants focused on challenges concerning the views of and from their neighbour's house as illustrated by this statement: "When we chose the lot and position of the house, we thought a lot about how the house should be orientated - Will the view be directly into the neighbouring bedroom?" [D].

Hence, several participants used ArchiLens to investigate and understand the orientation and placement of their house, and especially its windows, in relation to their neighbours (see figure 5). As an example, the use of ArchiLens made one family became aware of how close their house was going to be to their neighbours. Exploring this further, the father realized that their bathroom window was directly in front of a big window in the living room of a neighbour. To better understand the extent of these two issues, he first used ArchiLens to check the view from their bathroom window from inside the house. He then walked over to the adjoining lot and checked out how his house would appear from the neighbour's living room. Such detailed investigation in context enabled the future inhabitants to consider how position and orientation impacted not only the surroundings but also how the surroundings made an impact on their house, thereby helping then to make more informed design decisions.

Figure 5. Participants trying to understand contextual influences of nearby houses.

For good reasons, the detailed interplay with neighbouring houses was not observed for participants where the neighbouring lots had not yet been sold, or did not have a planned house on them. It was also not an issue for any of the open lot participants, who were at this stage more interested in the visual appearance of the different basic house options, and their overall relation to surroundings.

Changing the setting of decision-making and visualization, from being in an office to being on-site influenced the process. Participants obtained new understandings and got new ideas for alternative solutions, sometimes even overriding previous design decisions. One of the couples from the signed contract group used the ArchiLens session to consider their selection of materials: *"I actually think I was confirmed in my choices, when I tried changing the bricks to some of the colours, we had talked about. I compared them to the surroundings and they simply did not work, so clearly our yellow bricks fit better than black" [B]*. Here the participant used ArchiLens as a tool for verifying decisions in context. She later confirmed this decision in the post-construction interview. It also showed that the couple did not want to stand out from their neighbours. As newcomers they wanted to fit in with the architectural style of the area.

Other couples started questioning some of their earlier decisions when they noticed different possible solutions. In most cases this happened when one of them was looking at a specific detail of the house, mostly about how the roof or some section of windows could be designed. One of the couples had the following considerations while using the system: *"Come take a look at this. Should we maybe have used some more money and removed that wall? The others' [neighbouring houses] solutions look really smart" [B].* Later, this couple confirmed this change in the post-interview after their house was completed, expressing happiness that they had changed their original design. This illustrates how ArchiLens served as a primer for discussing and investigating the contextual influence, and in some cases made participants change their earlier decisions. As a female participant said: *"It makes you think about a lot of things, when you can walk around and see the house" [D].* This included everything from how the garden should be laid out, where and how large the windows should be, to where plugs for the TV should be placed in the living room.

Somewhat surprisingly, none of the participants from the signed contract group attempted to move the house to a new position on the lot. As a possible explanation we learned from the interviews that the families in this group were under restrictions in this respect, as most of them had chosen a house of the maximum size allowed by local building regulations, greatly reducing possible placements within allowed distances from boundaries and the road.

5.3. Collective Understanding

Building a new house affects the whole family, but several participants stated that it was difficult to share opinions and negotiate details. We experienced that most participants got a refined mutual understanding while using ArchiLens. In fact, the visualization in context and the dynamics of using ArchiLens gave them a shared point of reference. It facilitated a universal understanding between the involved parties: parents, children, friends, families, and architects. We discovered several interesting social interactions during our sessions. Couples interacted with each other to exchange information, to get feedback on material or placement changes, or to co-discover the house using ArchiLens (as in figure 6 where a couple is discussing their experience).

Figure 6. A couple obtaining a collective understanding of their future home through ArchiLens

It seemed quite clear that couples strived to obtain a shared understanding of the house – as expressed by participants: *"I could imagine it would make it easier to reach an agreement about details, because you discuss based on the same foundation ... the only thing I could have wanted, was to have had this system earlier in the process, as there at that point was many disputed points" [E]*

ArchiLens also primed social interaction between parents and children. More of the participating families had children and as soon as children were present parents involved them by showing their future bedroom and what view they would have (as illustrated in figure 7).

Figure 7. Parents showing their future house to the children and pointing out their bedrooms

The children indicated that they thought it was amazing to see this visual representation of their bedrooms. Their parents were also quite pleased to be able to involve the children in the process – this gave them an explanation tool to directly illustrate matters that they had discussed earlier.

Another social aspect was the need to present the house to other people for affirmation or acknowledgment of decisions. It became evident that a lot of couples had presented their ideas and blueprints to friends and families. But there was one major problem when doing this, as explained by one participant: *"We have been out with friends to look at their building project. They proudly presented their blueprints at the lot, but it's simply impossible to get more than a vague idea on how it's going to look from that" [D]*. ArchiLens was mentioned as a great tool in such situations and several participants asked for the possibility to borrow the system either for decision-making or presentation to friends.

5.4. Dimensional Insight

The one-to-one mapping of physical movements of the user while using ArchiLens gave them a more genuine understanding of the dimensions of their house by requiring them to explore it in its actual size. Most participants expressed that prior to trying ArchiLens they would usually sit around a table or in front of a computer looking at blueprints or 3D models. In such settings participants reportedly often experienced a poor sense of scale due to either difficulties in transforming sizes like 10 centimetres on a drawing to its actual size of 10 meters in the real world, or difficulties in imagining the real world size of a 3D model on a computer screen. As one participant revealed: *"A 3D house model on a computer screen appear smaller than in the real world" [F].*

This poor sense of scale led to a general wish for using ArchiLens to gain a better understanding of the physical size of the building in reality. This led to a common behaviour (among the male participants) of pacing out the size of the house, and how close it was going be to boundaries, while looking at the blueprint view (figure 8).

Figure 8. Walking around the lot using the position on the blueprint to understand dimensions

Many participants indicated that the practice of physically walking around the lot was the most important aspect of using the system for improving their dimensional understanding of the house. This was detailed with comments like: *"... it was a lot easier to understand when you were in it [the physical surroundings]" [E] or "... it is great to 'touch' the house with my body" [F]*.

Understanding dimensions and size was difficult to many of our participants. Several stated that the blueprints only gave a partial understanding of the sizes of, for example, bedrooms, and the post-construction interviews revealed that one of the participants had imagined the bedrooms smaller than they actually were when the house was built. In this case, the system had only given her limited support.

6. DISCUSSION

Our field studies of ArchiLens led to a number of interesting observations about the use of mobile technology in architectural design and development. The participating families were generally quite enthusiastic and positive towards using the system and towards the idea of being able to see and explore their future house on their mobile phone in its real future surroundings. In the field sessions most of them engaged quite extensively in the interaction, and in the follow-up interviews they applauded the practical value of the system as a tool for evoking their imagination and enabling them to participate in the architectural design process. In the following we will discuss these overall observations in a bit more detail.

6.1. Evoking Imagination

Through the use of ArchiLens, we tried to facilitate people visualizing images of their future house in their minds. We have referred to this as evoking their imagination as a part of forming new ideas or adjusting existing ones. Evoking imagination serves an important role in house design and development. This usually involves envisaging what the house, or a part of it, is going to look like, where it should be placed, how it will integrate with its surroundings, what materials it should be made from etc. To achieve this, contextual architects suggest that some of the design process is carried out on the site where the house is going to be built (Alexander 2002-05; Brolin 1980; Lynch 1960; McGillick and Calrstrom 2002; Ray 1980; Summerfield and Hayman 2006). As a specific technique for this, Alexander simply suggests that the architects place themselves in a particular spot of interest, close their eyes, and create a picture in their mind of what things should look like from there (Alexander et al 1977).

Evoking and facilitating imagination is a particularly important factor to consider if one wishes to enable ordinary homebuilders or renovators to participate in an architectural design process alongside architects, developers and builders with expert skills in working with physical building design. However, as this is challenging even for the best architects, it can be a significant obstacle for their clients. Especially when it comes to *contextual* architecture, exploring design alternatives through visual imagination can be very difficult for non-architecturally trained people. Therefore it is often not something that the future house owners or inhabitants of a building actively partake in – simply because they do not have the visual skills and experience of their professional counterparts.

The participants in our study strongly confirmed this obstacle for homebuilders and renovators when trying to participate actively in the process. Some participants found it difficult to imagine the visual exterior of the house, and many found it difficult to imagine the appearance of the house in the context of its surroundings. But most importantly, everyone found it difficult to shape their visual imagination of either of these solely on the basis of 2D drawings and blueprints.

In response to this challenge, we experienced that ArchiLens was able to evoke visual imagination by bringing peoples un-built houses “alive” so to speak. This experience of aliveness stemmed from two things, namely that the visualization took place on site with a one-to-one mapping between the physical and the virtual world, and that the visualization was interactive and enabled people to consider alternative choices and share and discuss opinions right where the those designs were going to be situated. In several cases, this led to our participants achieving an improved understanding of the spatial layout and visual characteristics of their future house. They were also able to imagine better what it would look like in reality and how it would interrelate with its context. While our participants generally had some imagination about their un-built house beforehand, exploring it through ArchiLens helped extending this imagination and making it more concrete and complete.

Another challenge for active participation in the design process that is related to visual imagination is the coordination and adjustment of this imagination among the different people involved. This can be described as a challenge of achieving common ground through co-experience. As co-experience relies heavily on communication based in shared experiences (Battarbee 2003), this is particularly difficult to achieve if the shared object of imagination does not have a physical form, or a representation understood, at least partly, by everyone involved. In relation to this, our participants expressed that they would like to share and develop ideas with the architect, the builder, their partner, and their family, but that they often found it difficult to express their imagination clearly enough for a shared understanding to be reached. From this perspective, providing visual 3D representations in context, like the ones in ArchiLens, becomes a means for aligning different people's imagination by enabling co-experience of something virtual and providing a boundary object (Star and Griesemer 1989) for communication.

6.2. Enabling Participation in Contextual Architecture

One of our guiding motivations for developing ArchiLens was a belief in the value of contextual architecture as a design philosophy combined with a wish to enable a larger degree of "end user" participation in it. Just like we believe in participatory design of computer systems (e.g. Kensing and Blomberg 1998), we believe that there is important value to be gained from enabling people to participate actively in the design and development of the physical spaces that they inhabit.

In this particular study we have explored the use of augmented reality as a specific candidate emerging technology for enabling such participation. Based on our findings, this approach appears to have some merit, both technologically and in terms of the user experience created.

The ArchiLens system illustrates how the design philosophy of contextual architecture can be supported through the use of a new but widely available technology by enabling a highly iterative exploration of the interplay between buildings and their surroundings. From the perspectives of architectural theory described earlier, what a system like ArchiLens does is allowing the users to develop and exercise a higher level of *genius loci* (Norberg-Schulz 1980; McGillick and Carlstrom 2002) in their response to the topographical, geographical, social and cultural context of the building site. Not only does it *require* the users to be present in the context of the building site, and allow them to see what the future building will look like, it also allows them to make immediate changes on site in response to what they experience, thereby making the response to the environment much more iterative and direct. From the perspective of Alexander's

(2002-05) principle of *wholeness preserving transformations*, which describes good designs as those that with deep respect for the existing context transform an existing space into a new one while preserving its original quality of wholeness, the important thing about systems like ArchiLens is that they allow for an iterative exploration and assessment of a selection of possible transformations of a space by situating form in context without materialising it in a structure that is difficult, expensive, or impossible to modify – or remove.

The study of ArchiLens in use illustrates that putting such technology in the hands of ordinary homebuilders or renovators not only facilitated them practicing aspects of contextual architecture. It also enabled them to participate more actively in the architectural design process in collaboration with the architects and builders as a result of this. Without any sod being cut, and without a single brick being laid in mortar, they were able to explore genius loci in the design of their future homes, and they were able to identify which design options would transform the space without violating its existing qualities of wholeness.

7. CONCLUSIONS

We have explored the potentials of using mobile phones for enabling ordinary homebuilders or renovators to participate more actively in the parts of an architectural design process that takes place on the building site. For this purpose we have described a mobile AR system called ArchiLens, which was designed and implemented for on-site interactive 3D visualization of private homes. We have also described a field study with 40 participants in the process of building or modifying their home who used ArchiLens for exploring their design, and making changes to it, on site in its future physical surroundings.

The study showed that interacting with the system helped evoke people's imagination and envisage the look and feel of their future house. This enabled them to participate more closely in the design process on-site by interactively reviewing design alternatives and exploring, for example, other placements and materials. Through ArchiLens non-architecturally trained people were able to envisage and better understand the interplay between their house and its context. The system gave them the ability to change plans, locations orientations, materials and facades, and thereby exert a degree of genius loci in iteratively responding to the topographical, geographical, social and cultural context of their building site. As a result of this they were able to allow that context to influence the evolution of harmonious and compatible design features and shaping a house design suiting that particular environment.

Using the example of facilitating participatory contextual architecture with mobile phones, the study illustrates how new technology can be used to enable people's participation in important activities and decision making processes, that would otherwise be difficult or out of reach. We hope this will inspire other similar efforts of facilitating participation within other domains.

8. FURTHER WORK

Several avenues for further research seem apparent after our study. First we acknowledge the possible value of a bigger display for interactive visualization of future architecture in context. Especially when looking at the use of a system like ArchiLens as a social act, sometimes involving whole families, a larger display would be preferred simply to allow more people a better view at a time. With the emergence of tablet computers with rear-facing cameras, these would be an obvious choice of platform to explore for a follow-up study. In order to further investigate the impact of systems like ArchiLens in relation to decision-making in the architectural design process, a longitudinal study of use would be beneficial. While several indicators of impact on crucial design decisions were observed in our field study, the relative short duration of the study, in light of the typical time-span of a house building or renovation project, made it difficult to make definitive conclusions about this. Finally, further investigation of the social co-experience facilitated by systems like ArchiLens would be fruitful for informing future designs. On the technical side, we encourage the investigation of ways for improving the accuracy of the system, such as the use of physical markers, however with particular attention to the trade-off in user experience between precise graphics and unrestricted user mobility.

ACKNOWLEDGMENTS

We would like to thank all participating families in the field studies. Further, we would like to thank Vendia-Huset for their collaboration in the project. Finally, we would like to thank several reviewers and colleagues for comments on earlier versions of the paper.

REFERENCES

Alexander, C., Ishikawa, S., and Silverstein, M. (1977) A Pattern Language. Oxford University Press, New York.

Alexander, C. (2002-05) The Nature of Order. Vol. 1-4. CES Publishing. Berkeley, California

Azuma, R. T. (1997) A survey of augmented reality. Presence: Teleoperators and Virtual Environments. Vol 6 (4), pp. 355–385

Battarbee, K. (2003) Co-experience: the social user experience. In Proceedings of the Conference on Human Factors in Computing Systems (CHI'03), ACM Press, pp. 730–731

Brolin, B.C. (1980) Architecture in context: fitting new buildings with old. Van Nostrand Reinhold, New York .

Caudell, T. P. and Mizell, D. W. (1992) Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes. In Proceedings of 1992 IEEE Hawaii International Conference on Systems Sciences. IEEE Press.

Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. ISWC '97, pp. 74-81.

Feiner, S. K. (2002) Augmented Reality: A New Way of Seeing. Scientific American. Vol. 286(4), pp. 48-55

Gleue, T., and Dähne, P. Design and implementation of a mobile device for outdoor augmented reality in the archeoguide project. In VAST '01 (New York, NY, USA, 2001), ACM, pp. 161–168.

Guo, Y., Du, Q., Luo, Y., Zhang, W., and Xu, L. Application of augmented reality gis in architecture. ISPRS VIII (2008).

Höllerer, T. H. and Feiner, S. K. (2004) Mobile Augmented Reality. In Karami and Hammad (Eds.) Telegeoinformatics: Location-Based Computing and Services. Tayler and Francis Books Ltd

Kensing, F. and Blomberg, J. Participatory Design: Issues and Concerns In Computer Supported Cooperative Work. 7 (1998), 167–185.

Layar. www.layar.com, 2010. Accessed on 22nd September 2011.

Lynch, K. (1960) The Image of the City. The MIT Press, Cambridge, Massachusetts

Miyashita, T., Meier, P., Tachikawa, T., Orlic, S., Eble, T., Scholz, V., Gapel, A., Gerl, O., Arnaudov, S., and Lieberknecht, S. An augmented reality museum guide. In ISMAR '08 (Washington, DC, USA, 2008), IEEE Computer Society, pp. 103–106.

Mizell, D. W. Virtual reality and augmented reality for aircraft design and manufacturing. Boeing Information and Support Services (1994).

McGillick P. and Carlstrom K. (2002) Alex Popov: Buildings and Projects. Axel Menges.

Morrison, A., Oulasvirta, A., Peltonen, P., Lemmelä, S., Jacucci, G., Reitmayr, J., and Juustila, A. (2009) Like Bees Around the Hive: A Comparative Study of a Mobile Augmented Reality Map. Proc CHI'09, Boston, USA, ACM Press, pp. 1889-1898

Norberg-Schulz C. (1980) Genius Loci: Towards a Phenomenology of Architecture. Rizzoli, New York.

Piekarski, W., and Thomas, B. H. Interactive augmented reality techniques for construction at a distance of 3d geometry. In EGVE '03 (New York, NY, USA, 2003), ACM, pp. 19–28.

Ray K. (1980) Contextual Architecture: Responding to existing style. New York, McGraw-Hill.

Reitmayr, G., and Schmalstieg, D. Location based applications for mobile augmented reality. In AUIC '03 (Darlinghurst, Australia, Australia, 2003), Australian Computer Society, Inc., pp. 65–73.

Reitmayr, G. and Schmalstieg, D. (2004) Collaborative Augmented Reality for Outdoor Navigation and Information Browsing, Symposium on Location-Based Services and TeleCartography, pp. 31-41

Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B., and Schmalstieg, D. Handheld augmented reality for underground infrastructure visualization. Personal Ubiquitous Comput. 13, 4 (2009), 281–291.

Shane, G. Contextualism. Architectural Design, 46, 11 (1976), 676-679.

Star S.L. and Griesemer J.R. Institutional Ecology, 'Translations' and Boundary Objects. Social Studies of Science 19, 3 (1989), 387-420.

Strauss, A. and Corbin, J. (1998) Basics of Qualitative Research – Techniques and Procedures for Developing Grounded Theory. SAGE Publications, Thousand Oaks

Summerfield, A. J., and Hayman, S. On capturing context in architecture. Dept. of Architectual Design Science - University of Sydney (2006).

White, S., and Feiner, S. Sitelens: situated visualization techniques for urban site visits. Proc CHI '09 , Boston, USA, ACM Press, pp. 1117–1120.

Wikitude. www.wikitude.org, 2010. Accessed on 22nd September 2011.

Wojciechowski, R., Walczak, K., White, M., and Cellary, W. Building virtual and augmented reality museum exhibitions. In Web3D '04 (New York, NY, USA, 2004), ACM, 135–144

Chapter 19

Power Advisor

Jesper Kjeldskov, Mikael B. Skov, Jeni Paay and Rahuvaran Pathmanathan

Abstract. Recent focus on sustainability has made consumers more aware of our joint responsibility for conserving energy resources such as electricity. However, reducing electricity use can be difficult with only a meter and a monthly or annual electricity bill. With the emergence of "smart" power meters units, detailed information about a households' electricity consumption is now available digitally and wirelessly. This enables the design and deployment of a new class of persuasive systems giving consumers insight into their use of energy resources and means for reducing it. In this paper, we explore the design and use of one such system, Power Advisor, promoting electricity conservation through tailored information on a mobile phone or tablet. The use of the system in 10 households was studied over 7 weeks. Findings provide insight into peoples awareness of electricity consumption in their home and how this may be influenced through design.

1. INTRODUCTION

In recent years, we have seen a significant increase in people's awareness and interest in sustainability and environmental impact of resource use (Froehlich 2009). Hence, we are now witnessing a strong focus on people's responsibility for, and ability to, save energy (DiSalvo et al. 2010). However, people are often unaware about their own, or their household's, consumption of resources such as water, gas and electricity because they are being metered out of sight, and details about patterns of consumption are not available. Research has shown that consumers mainly rely on their monthly or annual bills, which typically reports limited or irrelevant consumption information and such feedback information is insufficient for efficient energy management (Darby 2006). This prevents people from successfully reducing, for example, the amount of electricity that

Originally published as Kjeldskov, J., Skov, M. B., Paay, J. and Pathmanathan, R. (2012) Using Mobile Phones to Support Sustainability: A Field Study of Residential Electricity Consumption. In *Proceedings of CHI 2012,* Austin, Texas, USA. New York: ACM.

they use at home (Darby 2006, Fisher 2008). Research has also shown that, in order to raise awareness about electricity consumption, timely feedback and guidance is required to stimulate conservation and enable users to change their behaviour in a way that decreases their power usage. For example, by providing daily feedback, consumers can potentially save between 5% and 15% of the electrical household energy consumption (Darby 2006).

Emerging digital "smart" power meter units provide new opportunities for collecting and storing data about electricity consumption in households. Such units, often referred to as eco-feedback technology appliances, can provide households with detailed information about power consumption wirelessly and regularly with the goal of reducing environmental impact (Foth et al. 2009). The development of eco-feedback appliances is based on the hypothesis that the majority of consumers are unaware of how their everyday activities impact the environment, and that better awareness about this will help preserve the environment. The growing availability of such sensing systems for environmentally related activities and interactive feedback displays provides a great opportunity for exploring new types of eco-feedback solutions (Kirman et al. 2010, Petersen et al. 2009, Shiraishi et al. 2009, Weiss et al. 2009). However, with the growth in eco-feedback technologies, the amount of data available to the user is also increasing, and ways of using it to make people more aware remains largely unexplored (Darby 2006, Froehlich et al. 2009.

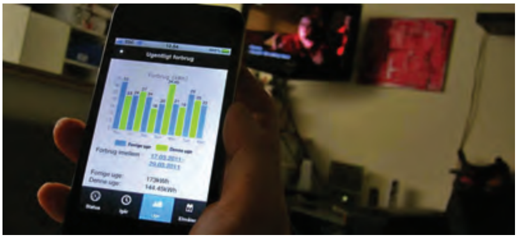


Figure 1. Power Advisor in use at home while watching TV

In this paper we present Power Advisor and findings from a case study of electricity consumption and awareness in 10 households. Our findings provide insight into how people view different types of feedback on their household electricity consumption and how they use this knowledge to reduce electricity use. This insight is then used to inspire ideas for designing mobile applications that can support people in understanding their own consumption, reduce it, and contribute to the sustainability of energy resources.

2. RELATED WORK

Sustainability has received increased attention over the last few years in HCI research (DiSalvo et al. 2010, Pierce et al. 2010). Numerous studies have investigated and experimented with the use of energy consumption feedback technologies (e.g. DiSalvo et al. 2010, Kirman et al. 2010, Yann et al. 2010, Weiss et al. 2009, Schultz et al. 2007). Froehlich et al. (2009) conducted a comparative survey on 133 HCI and environmental psychology papers. They identified key motivational techniques that HCI-designers should be aware of when promoting pro-environmental behaviour. In particular, they outline persuasive information as means for changing people's environmental attitudes and behaviour. Such persuasive information should be easy to understand, trusted, and presented in a way that attracts attention and is remembered (Froehlich 2009).

Domestic interactions and energy consumption habits were investigated by Pierce et al. (2010) in their study of energy consumption in 12 residential households. They developed a unique vocabulary for analyzing and designing energy-conserving interactions. This vocabulary consists of operational terms that capture actions and strategies of energy conservation, including: cutting, trimming, switching, upgrading, and shifting. They concluded that everyday interactions with home technologies are mostly performed without conscious consideration of energy consumption, stating that interactions tend to be unconscious, habitual, and irrational (Pierce et al. 2010). However, Shiraishi et al. (2009) show that people tend to increase their energy consumption awareness and knowledge by simply viewing a list of advice.

Some research show that consumers need better and more frequent information in order to reduce consumption of energy. Yann et al. (2010) investigated electricity usage by looking at requirements for always-on feedback electricity in private homes. They outline a three-stage approach for supporting electricity conservation routines. These stages are (i) raise awareness, (ii) inform complex changes, and (iii) maintain sustainable routines. Several of the participants in their study expressed that they would become more aware of their consumption if they could get detailed information about their past history of use – for example, one day ago or a comparison between the previous week and the current week. Weiss et al. (2009) identified a similar result where consumption history information raised awareness about the consumption patterns of electricity consumers.

Future technologies for consumption awareness can present information in different ways and normative information can influence people's behaviour and attitudes. Schultz et al. (2007) conducted a study with 290 households over a period of 7 weeks. The households were divided into two groups who received two types of feedback messages: half of their participants received descriptive-norm-only feedback messages while the other half received descriptive-plus-injunction messages. The descriptive-norm-only message condition contained information about how much energy (kilowatt hours per day) they had used in the previous week. The descriptive-plus-injunction-messages included the same information but also injunction information. The message included a happy face (☺) if the household had consumed less energy than the average in their community, and a sad face (☹) if their consumption was above average (Schultz et al. 2007). Their findings showed that households consuming less than average (thus receiving descriptive-norm-information) actually increased their consumption, creating an unintended boomerang effect. On the other hand, by adding a happy face (injunctive), households consuming less than average continued to consume at a desirable low level.

One of the pioneers of persuasive technology B.J. Fogg has conducted several studies on computers as persuasive social actors. In his book on persuasive technology, he outlines five types of social cues: physical, psychological, language, social dynamics, and social roles (Fogg 2003). He states that praise such as words, images, symbols, or sound in computing technology can lead users to be more open to persuasion. He further stresses the use of reciprocity because consumers like to reciprocate actions and favours. According to Fogg (2003), using roles of authority in computing technology also enhances its power for persuasion.

To understand and motivate behavioural change, Kirman et al. (2010) explored persuasive technologies in many ways. They claim that many technologies fail to take advantage of the established body of empirical research within behavioural science, for example, in the way that current persuasive technologies rely too much on positive reinforcement (Fogg 2003). They highlight that existing persuasive technology products fail to take advantage of negative reinforcement such as sad faces or negative tone texts (Kirman et al. 2010). They argue that one can achieve positive changes in behaviour using negative reinforcement, and it is therefore important to make use of this to promote behaviour change.

Within motivational psychology, Helen et al. conducted an analysis synthesizing a wide range of studies to develop a motivational framework of different stages of readiness and motivation to change (Helen et al. 2010). The Trans Theoretical Model splits behavioural change of individuals into several stages (Prochaska and Velicer 1997) and lists recommendations to motivate individuals at different stages. One way to make consumers more aware of their electricity consumption is to provide personalized feedback that acknowledges benefits and consequences of the individual's non-sustainable energy-behaviour in a neutral non-biased way. Another recommendation is to use injunctive normative messages and provide understandable feedback to consumers with already known symbols and signs. For example, one could utilize a smiley or thumbs-up sign. The third recommendation is to use personal self-set goals, which have the possibility of leading to higher performance and commitment (Helen et al. 2010).

As argued by DiSalvo et al. (2010), the majority of research on sustainability target users as individual consumers, e.g. to understand them or to change their behaviour. This is a result of the fact that many studies see people behaviour as causing environmental problems. Foth et al. (2009) have argued for research on sustainability at the group or the national levels. In this paper we investigate electricity consumption, but while maintaining focus on the individual; we also perceive them as community members. Therefore, we integrate two different information sources as feedback to residential members for creating awareness of their own power usage. These are (i) information on own consumption (individual), and (ii) information on electricity usage of other consumers (community). Whilst changing behaviour and maintaining sustained behaviour is important for electricity consumption (Yann et al. 2010), it is not the primary focus of our study. Rather we aim to raise consumer awareness by providing different kinds of feedback on own consumption as related to the surrounding communities.

3. POWER ADVISOR DESIGN

We designed a mobile application called Power Advisor to explore different kinds of information and feedback on power consumption in residential households. Inspired by previous research (e.g. DiSalvo et al. 2010, Weiss et al. 2009, Schultz et al. 2007), we designed Power Advisor to include descriptive and injunctive information. Descriptive information gives information about power consumption as historical data, for example, power usage during the last week, whereas injunctive information includes assessments or judgements of average consumption over a period compared to user's own goals or other consumer averages, for example, smiley's. When comparing consumption to averages of other residential households, we used data from the Danish Energy Saving Trust (both regional and national usage information), but also data from the participants in our study. When using data from the Danish Energy Saving Trust, we used calculated usage averages based on similar residential type (e.g. house, apartment), size (e.g. m2, number of bedrooms), and home inhabitants (e.g. number of adults and children).

The Power Advisor application integrates four menu items: My Consumption, Inbox, Enoks Guide, and Tip of the day. The first two are illustrated in the following sections. The other two provide general information and advice about power consumption in private residential households in Denmark. This advice included information about lighting, household appliances, IT and home office settings, and indoor climate and originates from The Danish Energy Saving Trust (2011).

3.1. Consumption Views

As suggested in existing research (Fogg 2003, Weiss et al. 2009), self-monitoring can lead to changed or adjusted behaviour as consumers become more aware of their own behaviour and actions. Under the menu item My Consumption, the Power Advisor provides self-monitoring through personalized information about the user's power consumption through four different views (three views are illustrated in figure 2).

The first view visualizes the total household power usage (measured as kilowatt hour) for the last week compared to the average consumption rate for a similar household in the northern part of Denmark (figure 2, left). The assessment shows whether their consumption is low, average, or high and the inclusion of a smiley supports the assessment – as suggested by Schultz et al. (2007). The second view shows household consumption over the last 24 hours (with one measurement every hour). This is visualized as a graph (see figure 2, middle). The third view shows consumption per day for the last week compared to the week before (inspired by Weiss et al. (2009) and shown in figure 2, right). The fourth view (not shown in figure 2) displays the last meter reading. Thus, Power Advisor seeks to provide different kinds of information for assessing one's own power consumption.

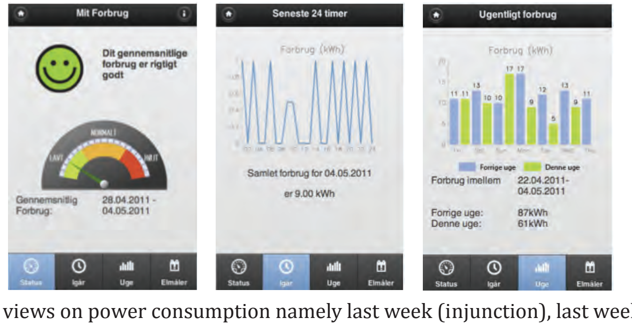


Figure 2. Three views on power consumption namely last week (injunction), last week consumption (descriptive), and last week compared to the week before (descriptive).

3.2. Consumption Messages

Power Advisor integrates personalized information for the user through a messaging service. All messages are placed in the Inbox and we use three different types of persuasive messages– (i) expert advice, (ii) community behaviour, and (iii) personal consumption performance.

Figure 3. Messages in Power Advisor

Expert advice messages give expert advice on power consumption. Based on the user's power consumption, Power Advisor compares the consumption with information from knowledge databases of The Danish Energy Saving Trust (2011) and generates a message with information about whether users should change their power-consumption behaviour. We named the expert Enok (an animated Eskimo) based on a TV-advert from the Danish Energy Saving Trust aired during the same period as our study (illustrated in figure 3, left including a positive smiley).

Community messages consist of information about what other consumers are planning, or are already doing, in their residential households. This information includes whether the majority of the community achieved their goals to reduce power consumption for the week, and how the user is doing in relation to the wider community.

Personal consumption performance messages contain information about the user's own personal power consumption. The information provided in these messages is objective information about the user's consumption and is detailed, compared to the other sources mentioned above. This information is also provided to have a diversity of information in the provided messages, from using smiley's to graphs and bars.

3.3. Technical Implementation

The system was implemented as a mobile web application to avoid issues of platform compatibility. It was developed using the open-source framework jQuery Mobile, which is touch-optimized for smart phones and tablets. The system communicates with a MySQL database in real-time using PHP, in order to ensure that user actions are logged.

4. POWER ADVISOR DEPLOYMENT

We conducted a case study over 7 weeks with Power Advisor deployed in 10 different households for 3 of those weeks. Our aim was to study the power consumption in these residential households and to explore how users responded to the different kinds of information feedback on power consumption provided by Power Advisor.

4.1. Apparatus

We collected usage data from the participating households through a product called Automatic Meter Reader designed by the utility company “Modstrøm” (2011) (a play on the Danish word for electricity indicating an anti-establishment view). The automatic meter reader is mounted on the existing power meter in the household and takes a picture of the readout every hour. This is then submitted to Modstrøm’s server and passed on to the Power Advisor application.

Figure 4. Modstrøm’s Automatic Meter Reader mounted on the readout of an old mechanical meter

One Automatic Meter Reader (AMR) was installed at each of the participating households, enabling us to collect power usage data automatically and unobtrusively throughout the entire study. The collected usage data from the households was used in the Power Advisor application as described in the previous section.

4.2. Household Recruitment

We recruited the participating residential households through the Modstrøm company’s customer database. The requirements for participating in the study were that (i) they should have a mechanical power meter unit, (ii) at least one member of the household should have a Smartphone or Tablet PC, (iii) they should have an active customer plan with a phone company, and (iv) they should be a customer of an energy provider company.

Of the ten different households recruited, five already had an Automatic Meter Reader installed in their home and were aware of and users of Modstrøm’s website to track their own power usage. The five other households had never used AMR before. Eight of the households owned their house while two households were renting. All households were in the northern part of Jutland, Denmark.

Most of our participants had standard household appliances, like fridges, freezers, washing machines (except for F and G), ovens, microwave ovens etc., and also a wide collection of other power-consuming devices, e.g., laptops, TVs, stereos, gaming consoles. Three households indicated high awareness of their own power consumption and 5 households indicated limited awareness. Two households expressed low or no awareness and they did not really care about their consumption.

- *Household A:* 2 parents (36 and 34 years old) and 3 children. Expressed high awareness of their consumption and used energy saving bulbs if possible.
- *Household B:* 2 adults (51 and 49 years old) and 2 children. Were only slightly aware of their power consumption and still used incandescent bulb.

- *Household C:* 2 parents (38 and 38 years old) and 1 child. Were not very aware of power consumption but mostly used A-level light bulbs.
- *Household D:* 1 adult (23 years old). Stated low (almost no) interest in saving power energy.
- *Household E:* 2 adults (45 and 50 years old). Indicated limited awareness of their own consumption, but had chosen to use only energy saving bulbs.
- *Household F:* 2 adults (25 and 28 years old). Were somewhat aware of the consumption and had a mixture of different light bulbs.
- *Household G:* 2 adults (24 and 28 years old). Stated only limited awareness of power usage and had only a few energy saving bulbs.
- *Household H:* 2 adults (37 and 37 years old). Were highly aware of their energy usage, and they used some energy saving light bulbs.
- *Household I:* 2 adults (40 and 40 years old). Were very aware of power consumption, and they used different kinds of energy saving light bulbs.
- *Household J:* 2 adults (40 and 41 years old). Indicated low interest in energy consumption and had a few different kinds of light bulbs.

The participants were requested to use Power Advisor every day to achieve a day-to-day experience with the application and their power consumption. Furthermore they were asked during the case study to check for incoming SMS-messages on their mobile phones as well as to read and respond to the messages in the system.

4.3. Study Protocol

Over 7 weeks of study, we conducted interviews, power usage monitoring, and deployment. Our households were visited at least three times where each visit lasted between 45 minutes and 2 hours.

During the first visit, we initiated the study by installing automatic meter readers (AMR) at the five households that had no reader in advance, and we introduced and explained the study to all household members. We further conducted a semi-structured interview to profile the participants on power consumption awareness and attitudes. We did this to get a sense of their behaviour when it comes to energy consumption and saving of energy, and we wanted to introduce them to the study. We also collected household demographics about inhabitants, age, income, appliances, etc. Interviews were audio recorded for later transcription. After the first visit, we logged data about their consumption using the installed AMR over one week. This was to achieve an understanding of the different households' power consumption and usage patterns before deployment of Power Advisor.

The second visit was carried out in the third week of the study, one week before the actual deployment of Power Advisor. The purpose of this meeting was to introduce them to system. We guided them through the system and explained how the different views and messages worked. With this introduction, participants also received a small manual that described the system and a video tutorial on how to use the system to minimize use and interaction problems during the deployment. They were also told that they would receive an SMS on their mobile phone if they had unread messages in their Inbox. On the day before the deployment, they received an email with instructions on how to log in to the system.

During the deployment period (3 weeks), we continuously and remotely monitored the use of the system as well as logged power consumption data through the AMR. During these three weeks, each participant received nine different messages in their Inbox, containing information about their power usage. Each participant received exactly three persuasive messages about personal power consumption, three messages about the community, and three messages with expert advice. After ten days of deployment, we sent a message to all participants Inbox with information about their own power consumption for the previous two weeks. They were requested to set a goal for power consumption for the forthcoming week. After the deployment period, the participants stopped using the system.

The last visit was conducted after the three-week period of deployment. We conducted a semi-structured interview where the aim was to explore how the participants found the different information sources they received in the incoming messages and the different views of consumption and what information sources they liked and disliked. The second interview started with questions reflecting on using a mobile device as a supportive tool in household. The incoming messages were discussed one by one. We used laminated cards with a physical illustration of each message to help the participants to remember the messages. Again, the interviews were audio recorded for later transcription.

4.4. Data Analysis

We analyzed our data (primarily the transcribed interviews) using techniques from qualitative research (Strauss and Corbin 1990). During the transcription, we strived to identify interesting and relevant topics and themes. Hence, while in this process, whenever an interesting topic was mentioned, the direct transcription of the speech was coloured-coded and given a specific number representing the properties in open coding. We identified 601 properties during this process. They were subsequently categorized into 22 different phenomena. By using axial coding, relations between the different phenomena were made and refined into 12 categories. These categories were then split among 3 themes. Finally, logged data on interaction with Power Advisor and logged usage data were analyzed to determine behaviour during the deployment period and to support qualitative findings from the interviews. The entire data analysis process including conducting the interviews, transcribing the audio recording, and using selective coding to find themes, took around 87 hours.

5. FINDINGS

Our findings stem from the two interviews with the households, their interaction with Power Advisor, and their consumption data for the period of use. Participants from the 10 households made a total of 347 logins in Power Advisor during the 3 weeks of deployment, thus households had on average 1.65 log-ins per day. We logged a total of 1851 interactions, the highest number per household being 299 entries (Household H) and the lowest number being 121 entries (Household C).

5.1. Using Power Advisor

Motivation for using the system varied quite significantly. As instructed, participants used Power Advisor continuously during the entire deployment. However, we noticed major differences in how often they used the application, and how and when they used the application. For example, Household H accessed the second view in the menu, My Consumption (shows consumption for the last 24 hours), 35 times. Whereas, Household C only accessed this view once during the entire period of deployment and accessed any type of consumption data only 13 times: *"I only used the application when I received a reminder"* (C).

Eight participant households used mobile phones as their primary access to the application. They preferred to use the system on a Smartphone because of the flexibility to receive messages on the phone and the ability to check the system on the same platform quickly and easily.

Some used the Smartphone access to the system for convenience: *"It is easier to monitor the usage on the phone than going to the basement to check the power meter"* (C). Most participants stated that the implemented SMS service was very useful as a prompt for starting to use the system. They liked to be notified and reminded when they had unread messages in the Inbox: *"When you have an unread message, you will be notified immediately. This raises your awareness of the application, and it works pretty well"* (F). For those who did not routinely check their power usage, the SMS triggering service influenced their behavior, effectively raising their awareness of their power consumption.

Quite interestingly, a few participants used the application during idle time, for example, while waiting for the bus or metro. They said that they often used mobile phones for "killing time" by playing games and that Power Advisor had some of the same characteristics as mobile games as it was both quick to use and fun. This actually led to an extension of the domestic space beyond the walls of their house where inhabitants could follow consumption while away from home (which was made possible by the system). We anticipated that people would prefer to use the application at home (and thus be able to change power consuming behaviour), but the short-sprint usage sessions with Power Advisor seemed well suited for idle time.

Two participant households mainly used Power Advisor on their iPads. They received reminders on their mobile phones and then used the application on the iPad. However, they did comment on the overhead involved in using a multi-platform setup. One iPad participant stated: *"I see some advantages on using the application on a mobile phone because I would still get an SMS on my mobile. Therefore it is easier to check the application on it than to pull my iPad up for that purpose. Furthermore, it is not always I have my iPad with me, but I always keep my phone with me"* (C).

5.2. Raising Awareness of Power Consumption

After the three weeks of system deployment all participants had increased their awareness of power consumption. The seven households who had little or no interest in their own power consumption found the application very useful as they suddenly achieved an understanding and awareness of their power usage: *"I think the study has* *been very good as I had no idea on how much I used before. It was also nice to get to know which group I belong to, so I can relate my consumption to others"* (D).

Participants from households A, H, and I indicated high awareness of power consumption before the deployment and they did not achieve the same benefits from the application even though they appreciated the opportunities for visualization of usage: *"The fact that you can have different views on your consumption makes this application useful"* (A). In fact, all participants found that the power consumption information provided the best informative illustration. They stated that the messages containing information about their own power consumption had a higher chance of persuading them to be more conscious of their own behaviour.

All ten participating households appreciated the personal and tailored information and messages in the system. This led to an increase in perceived applicability and credibility, which in turn were seen as necessary for persuasion: *"It is all linked to my personal consumption and provides with an opportunity to act and react"* (J). As a related point, the general advice in the application (Enoks Guide and Tip of the Day) was only used 18 times during the deployment. Two participant households (C and J) never used the advice component while three others (A, D, and E) only used the advice once. Surprisingly, almost all participants rated the general advice functionality as good value, easy to use and rather useful, even though when it came to changing behaviour and attitude, it was not particularly successful. Participants also added that other forms of general information, such as brochures or TV campaigns, should be more closely linked to users' own consumption in order to persuade them to change behaviour.

The participant households received three messages consisting only of information about their personal power consumption. The first message contained a bar showing the participant's highest consumption, their lowest consumption and the average consumption for a household of equivalent size. The participants expressed that this made them aware of their average consumption rate compared to others. The second and third messages consisted of information about their prior consumption and made the participants set up a goal for their own consumption for the forthcoming week. These messages were rated as most useful out of all messages received during the case study. As one participant enthused: *"With these messages, I have become more conscious of how much we consume and then you can maybe try to work with it, if you want to bring it down"* (E).

While general awareness of electricity consumption was raised during the study, it also became obvious that some participants were unfamiliar with the kWh (kilowatt/hour) unit of electricity use. This was a particular challenge for the participants in households where the initial awareness of power consumption was low. Participants from A, H, and I found it particularly difficult to perceive current and prior usage measured in kWh: *"You have to know something about the unit kilowatt-hour to able to assess your consumption and to decide whether you are satisfied with your current consumption rate … kilowatt/hour is an arbitrary unit for me as I don't know how much it means money-wise"* (A).

After ten days of deployment of the system, we sent participants a message containing information about the previous two weeks of power consumption. We asked the participants to set a goal for the power consumption for the forthcoming week, triggering their involvement in the process. Most participants enjoyed setting up goals for themselves as it motivated competition: *"This is interesting as it made me active in the process. It forced me to reflect upon my own consumption"* (I). While discussing this goal setting, a few participants mentioned that it was very important to keep reminding them about their own goals. This ensured that they were kept aware of their goals, in order to achieve them. Some participants suggested that this goal-setting function should be an integral part of the system. Involving consumers in goal setting seems to raise awareness of consumption, as suggested by Helen et al. (2010).

5.3. Power Consumption in the Community

The majority of the participants felt the information messages about how the other participants in the community were doing was very useful. Some participants expressed that they used the information about the others in the community, to compare with their own consumption. One participant noted: *"Absolutely, measuring up against other people gives me some feeling about my own consumption as I need to identify whether I'm doing something wrong or right."* (A)

It felt natural for some of the participants to be compared with others, and as one participant said: *"We are gregarious animals in a way. We measure ourselves and consider ourselves in relation to each other all the time"* (I). The majority of the participants said it was important to have information about what others were doing in order to be persuaded to conserve less electricity themselves.

In reality, there were mixed feelings about the community messages. Two out of the ten households spoke against the benefit of having information about others. The appliances in the participants' households could be different and other parameters such as income, household-size and occupation, could mean that it was hard to compare against them. One participant said: *"I really do not care how others are, it does not change anything for my consumption. So therefore it has no value to me to be compared with others".*

While discussing the visualization of the community messages, an important issue was raised. When showing information about what and how well the community is doing compared to the participant, it was important for the information being presented to the participant to be persuasive. For example, participants felt that if the community average was a bit better than the participant's average, the percentage or number should not be displayed but instead a smiley and a coloured message. This would prevent negative reactions to marginal changes. If the community average was much better than the participant then percentage and power consumption units showing the difference would be more helpful. Even so, a common discussion point was the importance of being compared to one's own consumption all the time as opposed to being able to persuade through community messages.

Some participants (e.g. C and H) argued that comparison with others did not affect their attitudes or consumption *"I don't really care about the other consumers, it does not affect my consumption. For me, there is no added value in being compared and measured against others"* (H). However, one participant from household J said that it would be more interesting to have consumption information about neighbours or friends as you could compete at a different level with people you know in advance.

Finally, involving the entire household was considered vital for successful power consumption management. Several participants (A, B, and I) stated that to reduce consumption in the household, all household members had to be involved actively, which was difficult during this study.

5.4. Reinforcement and Injunctive Information

We designed Power Advisor to include descriptive-plus-injunction-messages where usage data was associated with either approval or disapproval. Most participants indicated that smileys were easy to understand and interpret. They were then asked about when it would be good to make sue of smileys. One participant who received a happy smiley when being only 3% percent better than the average responded: *"I would here perhaps have a tendency to rest on its laurels"* (D). Another participant expressed the options faced when receiving a "bad" smiley: *"Either you think that this performance was bad and you try to do something different to avoid receiving a red smiley next time, or then you are indifferent and are opposed to the message next time"* (H).

Our findings align with the conclusions of Schultz et al. (2007) who claim that you should consider when to use injunctive messages to promote more pro-environmental behaviour. When discussing the use of smileys, several participants raised issues on the use of positive and negative comments to promote behavioural change. Participant households A, B and D said that positive comments on behaviour acted as a motivator while participants from households E and J stated that the positive messages did not make them do anything different and they suggested a more effective positive message might be: *"You are doing excellent, but you are 10% behind the best people in the group"* (J).

There were mixed attitudes towards negative messages, because they did motivate some of the participants: *"Then you become more motivated for improving your consumption and setup realistic goals"* (F), while others argued that the negative messages were discouraging: *"If there are too many negative messages, I might be thinking, this does not interest me any more – these stupid messages"* (A), and, *"I perceive a negative comment as a raised finger on your behaviour and it is not likely that I would read messages in the future"* (H).

There are mixed potentials in using positive and negative messages. While they support the persuasive principles on using praise with words (Fogg 2003) and the use of negative reinforcement to promote behaviour change (Kirman et al. 2010), they can cause frustration, which is potentially a pitfall when using negative verbal comments in persuasive feedback. However, Household E did suggest that negative messages could be used as a notification (alarm) if your power consumption suddenly increases.

5.5. Motivating Behaviour Change and Barriers for Change

While behaviour change is fundamentally longitudinal by nature, we did identify aspects of the system that motivated behaviour change. All participants stated that Power Advisor helped raise awareness of their own power consumption. One participant argued that continuous information and feedback made him more aware of consumption and made it easier to adjust his behaviour: *"Before the study, I was already tracking my consumption through the Modstrøm website, but every now and then I would forget to check usage for several weeks … and furthermore, you really need to compare your usage with others"* (F).

Several participants gave the impression that while they were happy to received information and feedback on their own consumption, they seemed less inclined to change their behaviour. Households C and E indicated that they perceived their current consumption as reasonable. Others saw barriers to changing behaviour, as they did not know exactly why they consumed more power one day compared to another day: *"I was surprised to observe a difference in power consumption even when we talked about the same weekday, same people at home, etc."* (G), and, "I had no idea on how to reduce power consumption (kWh) besides turning off the lights or watching less television" (G). Such lack of understanding potentially undermines the effects of introducing such a system into residential homes. For people to change behaviour, they need to understand and be aware of their different options.

Some participants reported that changing behaviour towards power consumption had to do with cultural change and thus, they required continuous support and feedback on their actions – as exemplified by this participant: *"You have to keep reminding people to change behaviour. I remember when I was a child; our parents kept telling us not to let the water running while brushing teeth. We don't tell this to our children today as it is not necessary"* (G).

In Household G, feeling that the advice given was tailored to their interests motivated a change in behaviour. In the past they had tried to control power consumption by adjusting their fridge temperature. On getting the advice from the system, they again changed the temperature in the fridge and placed a glass of water with a thermometer in it to check the actual temperature. However, it was by chance that this piece of advice resonated so strongly with this household.

Changing behaviour or attitude is extremely individual and motivated by several different factors. Most participants would reduce their power consumption to gain economical benefits (Households A, D, F, and J), but other factors were also given, including reduction of pollution in the environment: *"Instead of using kilowatt-hour as energy unit, one could also apply environmental units, e.g. how much your consumption affects the environment with pollution"* (E).

6. DISCUSSION

Our aim was to study electricity consumption in residential households to achieve insights into how people view different types of feedback on their household electricity consumption and how they could use such knowledge to reduce electricity consumption. We found that people in our study gained a significant understanding of their own power consumption by interacting with our mobile solution. They especially found the different views of consumption and the prompting in the message service (with SMS) very useful as these provided multiple ways of usage visualization and triggered use of the application. Whilst we achieved insight on power consumption and people's need for feedback as one contribution for the paper, we identified a number of themes that constitute a second contribution of the paper. These are elaborated below.

6.1. Comparative Electricity Consumption

Electricity consumption is still very difficult to understand and assess for ordinary people. To understand power usage, households have to track consumption systematically and regularly to achieve awareness and it requires knowledge of "reasonable" usage, e.g. measured in recommendations or average usage in similar households. As Froehlich argues, energy consumption is abstract, invisible, and untouchable and without tangible manifestation, energy usage often goes unnoticed [8]. Back in the late 1970's, Winett et al. wrote that people are unaware of when and where electricity usage occurs in the home (Winett et al. 1979). In 2011, we still experienced that problem. Usage varies a lot from day-to-day or week-to-week, as conditions for consumption changes over time, e.g. seasons, guests, extra laundry. Some households articulated this and they stated that they had no idea on why their usage would be very different for the same weekday having the same people at home. Additionally, the unit of electricity use (kWh) was rather poorly understood by people making it impossible for them to achieve awareness and then even more impossible to change behavior. While people increase their consumption knowledge over time, it is questionable if all people will achieve a basic level of knowledge that enables them to fully understand electricity consumption (as the study of Winett et al. (1979) shows).

We found that multiple views (visualizations) of usage data could assist consumers' awareness and understanding of their own consumption. We need ways of communicating electricity consumption where comparable visualizations compliment absolute measures (e.g. last weeks usage in kWh) with other measures (e.g. previous week or other households). This potentially enables the consumer to judge own consumption as a comparable condition where the user does not have to understand if e.g. 15 kWh is a highly daily electricity consumption rate. However, they can rather see their consumption against similar households. We found that people appreciated more abstract representations of electricity consumption, e.g. an assessment of usage as low, medium, or high as compared to other households. The comparative usage visualization was found useful not only by households with limited awareness of electricity usage but also those households with high awareness.

6.2. Social Power in Consumption Communities

Sustainability and energy resource conservation literature has mainly focused on doing research where target users are seen as individual consumers rather than groups or societies (DiSalvo et al. 2010, Foth et al. 2009). DiSalvo et al. (2010) state that studies tend to perceive user behaviour as causing environmental problems and therefore we need to change the individual. Petersen et al. (2007) found that residents in dormitories could reduce their electricity consumption when exposed to real-time visual feedback. Interestingly, their study showed that reduction effects were achieved at not only for individuals, but also at the collective level (the entire dormitory community) where residents started to educate each other on usage to achieve lower collective consumption. Thus, targeting users as part of communities may produce even stronger results.

Including data on community members introduces roles of social power between community members. Whilst the dormitory study displayed strong social power (Petersen et al. 2007), social relations between our participants were significant lower as they did now know each other. The lack of influence from social power in our study was exemplified as households argued that comparison with others did not affect their attitudes or usage, although some appreciated the included community usage data. As a participant argued, it does not make you any better that other people perform really well or really poor. Thus, while usage data of other households can induce increased user awareness on consumption, user attitudes seem unchanged. Integrating information about members with a stronger social relation, like the dormitory, could potentially lead to action. One household member illustrated this by saying that it would be more interesting to include consumption data about neighbours, families, or friends, as this would provide other opportunities for action, e.g. discussing consumption in person.

6.3. Motivation, Reward, and Charity

People's motivation for reducing usage of electricity (and perhaps other energy resources) differs quite substantially. Monetary reasons were often raised as the single primary reason for becoming more aware of your consumption and hence, therefore being able to take action and reduce usage. Our study seems to confirm this observation, as several households requested new and different measures (not only kWh kilowatt-hour as used in our study), but also units like consumption as an absolute cost (local currency) or as a relative cost where you can see how much you have earned (or lost) during the previous period. Froehlich (2009) also categorizes different units measures, e.g. kWh, cost, or environmental impact. This triggers questions of reward. How can we reward people when they try to make an effort to reduce their own consumption? While saving money on the monthly bill, different reward schemes seem suitable for motivating people. E.g. people could collect member points that could be used for purchase of goods or services (like airlines' mileage reward programs). It potentially provides another tangible manifestation of usage reduction attempts.

Furthermore, a second type of motivation was uncovered in our study. Some people found it appealing to reduce their usage if their reduction could be used for donation, e.g. money to an official charity organisation or a local football club. We have recently seen similar arrangements in several Danish supermarkets where you can donate money to charity from returned deposit bottles and cans. This has been quite successful (measured in revenue). Brandon and Lewis (1999) found that people with positive environmental attitudes were more likely to change their consumption during a 9-month study. Environmental attitudes played also a major role for some our households, as they would like to know how their actions could reduce pollution. Thus, some mentioned that a motivational factor could, for example, be to express power consumption as CO^2 units to illustrate the environmental impact. Summarized, people are highly individual when it comes to motivation and our work support that systems should be tailored to individual households to capture the more fine details of reward and motivation.

CONCLUSION

This paper has explored power consumption in residential households introducing the mobile application called Power Advisor that enables feedback on electricity usage. We integrated different kinds of information in the application (i) data on own consumption, and (ii) data on usage of other consumers. We studied electricity consumption and power consumption awareness in a case study over seven weeks where we conducted interviews, power usage monitoring, and deployment of Power Advisor.

Our findings suggest that households in our study gained a deeper understanding of their own power consumption by interacting with our mobile solution. They especially found the different views of consumption and the prompting in the message service (with SMS) very useful as these provided multiple ways of usage visualization and triggered use of the application. We further identified three themes. First, households found comparative usage visualizations useful as they enabled them to compare their consumption with other community members. This helped raise awareness. Secondly, the social power between community members influenced motivation of the households in terms of behavior change. Finally, households were rather different on aspects of motivation for reducing consumption, and how reward should be implemented.

While we conducted our case study over seven weeks, we profoundly need more and longer longitudinal studies to uncover motivation for change and sustained change. From our study, we see at least two avenues for future research. First, we need to investigate how such technology can support people over longer periods of time, and how people will adopt and alter such technologies. Are they primary educational tools, where people stop using them after some period of time? Or will they serve as tools that continuously support and persuade them to change and maintain actions. Secondly, we need to explore communities and their roles in power (and energy) consumption. Making data and information accessible to community members could have an effect on consumers' attitudes towards themselves and towards others.

ACKNOWLEDGMENTS

The authors would like to Modstrøm A/S and members of the 10 households participating in the study.

REFERENCES

Brandon, G. and Lewis, A. (1999) Reducing Household Energy Consumption: A Qualitative and Quantitative Field Study. Journal of Environment Psychology, Vol. 19, pp. 75-85

Darby, S. (2006) The effectiveness of feedback on energy consumption. University of Oxford, UK

DiSalvo, C., Sengers, P., and Brynjarsdöttir (2010) Mapping the Landscape of Sustainable HCI. In Proceedings of Human Factors in Computing Systems (CHI'10), ACM Press, pp. 1975-1984

Fischer, C. (2008) Feedback on household electricity consumption: A tool for saving energy? Energy Efficiency, Vol 1, No. 1, pp. 79-104

Fogg, B.J. (2003) Persuasive Technology, Morgan Kaufmann, San Francisco, USA

Foth, M., Paulos, E., Satchell, C., and Dourish, P. (2009) Pervasive Computing and Environmental Sustainability: Two Conference Workshops. IEEE Pervasive Computing, Vol. 8(1), pp. 78-81

Froehlich, J., Findlater, L., and Landay, J. (2010) The design of eco-feedback technology. In Proceedings of Human Factors in Computing Systems (CHI'10), ACM Press pp. 1999-2008

Froehlich, J. (2009) Promoting energy efficient behaviours in the home through feedback: The role of human-computer interaction. In Proceedings of HCIC 2009 Winter Workshop, p. 10

Helen, H.A., Greenberg, S. and Huang, E.M. (2010) One size does not fit all: Applying the Transtheoretical Model to Energy Feedback Technology Design. In Proceedings of Human Factors in Computing Systems (CHI'10), pp. 927-936

Kirman, B., Linehan, C., Lawson, S., Foster, D., and Doughty, M. (2010) There's a monster in my kitchen: using aversive feedback to motivate behaviour change. In Proceedings of Human Factors in Computing Systems (CHI'10), ACM Press, pp. 2685-2694

Modstrøm A/S, (2011) http://www.modstroem.dk/

Petersen, J. E., Shunturov, V., Janda, K., Platt, G., and Weinberger, K. (2007) Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives. International Journal of Sustainability in Higher Education. Vol. 8(1), pp. 16-33

Petersen, D., Steele, J. and Wilkerson, J. (2009) WattBot: a residential electricity monitoring and feedback system. In Proceedings of extended abstracts in Human Factors in Computing Systems (CHI'09), pp. 2847-2852

Pierce, J., Schiano, D. J., and Paulos, E. (2010) Home, Habits, and Energy: Examining Domestic Interactions and Energy Consumption. In Proceedings of Human Factors in Computing Systems (CHI'10), ACM Press pp. 1985-1994

Prochaska, J. O. and Velicer, W. F. (1997) The Transtheoretical Model of Health Behavior Change. American Journal of Health Promotion. Vol. 12, pp. 38-48

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein N. J., Griskevicius, V. (2007) The constructive, destructive, and reconstructive power of social norms. Psychological Science, Vol. 18(5), pp. 429-434

Shiraishi, M., Washio, Y., Takayama, C., Lehdonvirta, V., Kimura, H., and Nakajima, T. (2009) Tracking behaviour in persuasive apps: Is sensor-based detection always better than user self-reporting. In Proceedings of extended abstracts in Human Factors in Computing Systems (CHI'09), pp. 4045-4050

Strauss, A. and Corbin, J. M. (1990) Basics of Qualitative Research, SAGE

The Danish Energy Saving Trust (2011) http://www.goenergi.dk

Weiss, M., Graml, T., Staake, T., Mattern, F., Fleisch, E. (2009) Handy feedback: Connecting smart meters with mobile phones. Proceedings of MUM 2009, Cambridge, UK, ACM, pp. 1-4

Winett, R. A., Neale, M. S., and Grier, H. C. (1979) Effects of self monitoring and feedback on residential electricity consumption. Applied Behavior Analysis, Vol. 12(2), pp.173-184

Yann, R., Dodge, J., and Metoyer, R. A. (2010) Studying always-on electricity feedback in the home. In Proceedings of Human Factors in Computing Systems (CHI'10), ACM Press, pp. 1995-1998

Part V

Understanding

UNDERSTANDING

Part V addresses the question *how can we abstractly describe and understand the relationships between interactive mobile systems, users and context?* In order to facilitate research progress beyond small incremental steps from one design to the next, we need to develop a cumulative body of knowledge that can help explain, theoretically, the relationship between people, technology, and their context. Four of my own contributions to this are included in chapters 20-23. These chapters each take an individual approach to the issue of sense making in context on different levels of abstraction – the way we perceive the world by identifying meaningful patterns and wholes, the way we interpret the world by assigning meaning to signs, the way we use our joint embodied presence in the world to create shared meanings, and the way we organise and orchestrate the world around us.

Principles of perceptual organisation
Chapter 20 presents and discusses five principles that can be applied for explaining how people identify meaningful patterns and wholes from ensembles of mobile systems and their context. The discussion is informed by a field study of mobile user experience at Federation Square in Melbourne, Australia, and takes its theoretical inspiration from the discipline of Gestalt psychology. Based on a theoretical analysis of our empirical findings, we argue that the user experience of location-based mobile interaction designs in context can be described and understood through Gestalt theory's five principles of perceptual organisation: *proximity, closure, symmetry, continuity,* and *similarity*. Specifically, we argue that these principles assist us in explaining how people create meaningful wholes from incomplete and fragmented information on mobile devices. They do so by "drawing from a larger canvas" to which both mobile devices and their context are contributing. Consequently, as mobile interaction designers we need to design for this "larger canvas" rather than merely for the "smaller canvas" of the mobile device in isolation.

Indexical interaction design
Chapter 21 discusses how people interpret information representations on mobile devices in context by assigning meaning to indexical signs. The discussion is informed by three field studies of mobile user experience in Denmark and Australia, and is grounded theoretically in the discipline of semiotics. Based on our empirical and analytical work, we argue that the relationship between users, user interface representations, and context can be described and understood through the semiotic concept of *indexicality.* In relation to mobile interaction design, we argue that information in an interactive mobile system can be understood as a special type of indexical sign, where meaning is created through interpretation of the ensemble of system and context. We also argue that increasing the level of indexicality in an information representation by locating it in time and space results in a reduction of symbolic and iconic representations required to communicate a specific piece of information. Of particular importance for the design of mobile interactions, this allows a reduction of explicit information presented to the user.

Proxemics and interactional spaces

Chapter 22 discusses how people create shared meanings through embodiment in shared physical spaces, and how an understanding of embodiment and proxemics can be used to guide the design of interactional spaces or digital artefact ecologies. The discussion is informed by a theoretical analysis of human interaction in shared physical spaces drawing on the philosophical foundations of Husserl, Heidegger and Merleau-Ponty. The key point is that coordinated action, meaning-making and intersubjective understanding are shaped, in part, from our embodied actions in space and the availability to others of these actions, for example, the way we move, point, touch and gesture in relation to objects and other people. In respect to understanding mobile device user experiences we argue that this is profoundly influenced by spatial factors such as proxemics and the physical design of the interactional spaces in which they are used. We exemplify this through the "blended interaction space" prototype including an ecosystem of interactive surfaces and devices and facilitating various forms of "proxemic interactions".

Orchestrating mobile devices

Chapter 23 discusses how people create and orchestrate meaningful digital ecosystems of interactive mobile systems and devices. The discussion is informed by a cultural probe study of early adopters of mobile devices in Melbourne, Australia, and takes its offset in the debate about convergence versus divergence as principles for mobile interaction design. The chapter presents three seemingly irreconcilable perspectives on the relationship between functionality and user experience drawn from the literature, and argues that these are, in fact, complementary views, when observed in a broader perspective. The key point in this argument is the observation that convergence and divergence are not just principles of design, but also principles of orchestration in use.

Chapter 20

Principles of perceptual organisation

Jeni Paay and Jesper Kjeldskov

Abstract. Within recent years, the development of location based services have received increasing attention from the software industry as well as from researchers within a wide range of computing disciplines as a particular interesting class of context-aware mobile systems. However, while a lot of research has been done into sensing, adapting to, and philosophising over the complex concept of "context", little theoretically based knowledge exists about why, from a user experience perspective, some system designs work well and why others don't. Contributing to this discussion, this paper suggests the perspective of "Gestalt theory" as a theoretical framework for understanding the use of this class of computer systems. Based on findings from an empirical study, we argue that the user experience of location based services can be understood through Gestalt theory's five principles of perceptual organisation: proximity, closure, symmetry, continuity, and similarity. Specifically, we argue, that these principles assists us in explaining the interplay between context and technology in the user experience of location based services, and how people make sense of small and fragmented pieces of information on mobile devices in context.

1. INTRODUCTION

Location based services represent an emerging class of computer systems providing mobile device users with information and functionality related to their geographical location. Within recent years, this class of *context-aware* mobile computer systems has received increasing attention from researchers within a range of computer science disciplines as well as from industry. Location based services open a new market for network operators and service providers to develop and set up value-adding new services for users on the move, such as helping find nearby shops or friends, advertising traffic conditions, supplying routing information, and augmenting the built environment

---

Originally published as Paay, J. and Kjeldskov, J. (2008) Understanding the user experience of location based services: five principles of perceptual organization applied. *Journal of Location-Based Services,* 2(4), 267-286.

of cities with an ubiquitous layer of information about, for example, people, places, and activities. Recent advances in technology have fuelled the development and uptake of a wide range of location based services. PDAs and 3G mobile phones with GPS and other positioning capabilities have become increasingly affordable and popular, and more and more service providers have begun to develop and offer innovative information services that integrate wide-area broadband Internet access, web resources and geographical information. Also, generally available systems such as Google Earth and Google Maps have rapidly become popular media for people to access location-related information and even publish it themselves (i.e. relating picture blog content to geographical places and publishing GPS coordinate trails for others to follow).

The development of location based services for mobile devices faces many challenges ranging from issues of determining people's location and orientation in physical space, how to combine satellite imaging, 3D models and cartography, through to issues of what information to provide in response to a particular location, and how to facilitate suitable user interaction with this content. Within the Mobile HCI community it has been widely argued that researchers, designers and software developers need to look more broadly at the context of use of mobile devices and systems in order to understand mobile use better and to be able to produce good and relevant solutions (Johnson 1998). In response to this, much effort has been put into both ethnographic-style studies of mobile work activities, and field studies of technology in use. However, while a lot of research has been done into sensing, modelling and adapting to context, as well as philosophising over the complex concept of "context" (e.g. Dourish 2004), very little work has provided theoretically informed foundations for interface and interaction design of context-aware and location based systems, or explained, from a user experience perspective, why some solutions work well and why others don't. Hence, generally applicable rules and guidelines for interaction design, as we know them for desktop and web applications, do not exist for context-aware and location based systems, and more research is needed into the user experience of this emerging class of applications.

Contributing to this research, this paper presents a user experience study of a prototype location based service and looks at how people perceive the ensemble of information on mobile devices and use context from the perspective of five principles of perceptual organisation from Gestalt theory: *proximity, closure, symmetry, continuity,* and *similarity* as a potential theoretical framework for understanding the user experience of location based services. In suggesting this perspective on human perception and thinking in relation to technology use, we are inspired by and align ourselves closely with the acclaimed work of Erik Frøkjær and Kasper Hornbæk (2002, 2008) on metaphors of human thinking for describing aspects of human-computer interaction. Reporting on this work in the Journal of Location Based Services, we contribute to the "urgently needed" area of research of *interaction design for LBS* outlined in Raper et al.'s (2007) editorial lead paper of the first issue of the journal. We also respond to the lead paper's recommendation for more user-centred system conception of location based services (Raper et al. 2007, p. 22).

We have been studying the user experience of location based services through a 2-year project investigating the deployment of mobile and pervasive computing technologies in urban environments. The e-Spective project took its inspiration from the newly opened civic structure of Federation Square in Melbourne, Australia. It involved a series of field studies of urban socialising behaviour within the built environment of inner cities as well as the development and evaluation of a prototype location based service providing an informational overlay to Federation Square (Kjeldskov and Paay 2006, Paay and Kjeldskov 2006). In addition to learning about interaction design for location based services on mobile devices, one of our most interesting (and somewhat surprising) findings from studying people's use of the prototype location based service was that people were extremely good at making sense from small and fragmented pieces of information. When faced with incomplete or ambiguous information, people wanted to put the pieces together, and they did so with high success rates. This finding prompted two questions: 1) How can we explain this phenomenon, and 2) How can knowledge about this phenomenon inform the design of similar location based services? Motivated by these questions we analysed our video and interview data from field evaluations of the prototype system from multiple theoretical angles. As a result of this, we found that the perspective of Gestalt theory's principle of perceptual organisation provides a very useful, and yet relatively simple, lens for describing and explaining how people make sense of the content of mobile information systems situated in context.

First we introduce a background of related work. We then introduce Gestalt theory and how this theoretical approach to human perception and thinking has been applied within HCI to explain and inform qualities of graphical screen design. Following this, we present the Federation Square case study, our prototype system, and a field study of the prototype in use. We then apply five principles of perceptual organisation from Gestalt theory to our empirical data, describing and explaining the ensemble of mobile device and use context in the user experience of our location based service. From our findings, we distil a framework for the design and evaluation of location based service user experiences. Finally, we conclude on our work and outline avenues for further research.

2. THE MOBILE INTERNET AND LOCATION BASED SERVICES

Within the last decade there has been a huge focus on the development of mobile information and communication technologies bringing the potentials of the Internet to the mobile user within a wide range of use domains for work as well as for leisure. Following the widespread uptake and commercial success of the Short Message Service (SMS) on mobile phones significant attention and resources have been devoted to the development of the next generations of mobile network services, protocols and infrastructure, known as 2½G and 3G mobile telephony (Sacher and Loudon 2002). MMS was developed to allow exchange of rich media content, and WAP allowed mobile access to a downscaled version of the web. On the network level, the development of UMTS means that the speed of mobile data connections now matches many of their hard-wired counterparts, thus allowing realistic mobile use of, for example, the World Wide Web. However, while commercially available technologies have made the Internet mobile, fast, accessible and relatively cheap, uptake by the general population has not met the IT industry's expectations (e.g. Costolo 2006, BBC 2002). While people are generally increasingly interested in the mobile Internet, very few are using it (Ericsson 2007).

From a user experience perspective there are several reasons for this. Firstly, while unquestionably containing information and functionality relevant to mobile users, most Internet services are not well designed for mobile use (e.g. Forrester 2006). They are designed for desktop use, and require a lot of user input and visual and cognitive attention. In contrast, the mobile Internet is typically accessed through devices with small screens and limited means of input used in very dynamic settings. Thirdly, services for the mobile Internet are currently designed to facilitate doing, while mobile, the things we do at our desktop. They do not relate to mobile use context but look the same at home, on the bus, in a café, or walking down the street – situations with very different requirements (e.g. Lee et al. 2005). If we want to bridge the gap between interest and actual use of the mobile Internet, we must do better. We must support a user experience that takes into account the wholeness of technology and context, and *we must enable people to do relevant things that they couldn't do before.* The area of location based services has the potential to fill this void.

Advances in technology have made it possible for mobile computers to sense or access information about users' context such as location, social setting, activity, computational resources, etc. (e.g. Hinckley et al. 2005). Recent research has demonstrated that using such information to make mobile computer systems "context-aware" can increase usability within highly specialised domains such as healthcare and industrial process control. One of the most promising aspects of user context for a mobile computer system to respond to is *location* (Jones et al. 2004, Fithian et al. 2003, Kaasinen 2003). The potential benefits of location based services for the mobile Internet are several. By making a mobile Internet service aware of the user's location, developers can streamline it to present information and functionality that is particularly relevant at a specific place or within a specific distance. As a fictive example, a location based mobile Internet service for a train station could respond to the user's location, time, activity etc., by presenting only information about departures from this or nearby stations, within a short period of time, to destinations matching upcoming appointments.

Figure 1. Example location based service (http://www.viewranger.com)

The number of commercially available location based services has increased rapidly over the last 2 years. Fuelled by developments of new technology, new services are emerging that integrate wide-area broadband wireless Internet access, web resources and geographical information for increasingly affordable and popular PDAs and 3G mobile phones with GPS and other positioning capabilities such as the new iPhone and the BlackBerry. As an early example, ViewRanger (Figure 1) provides 3D models of the user's surroundings with superimposed information links on GPS enabled 3G phones in parts of the UK. Other systems and services, such as TrackStick and Phone2Gearth, allow people to track their geographical movements, annotate it with media content such as text, images and video, and then publish it through systems such as Google Earth (Figure 2). In a similar fashion, some of Sony's newest cameras and camcorders are able to record GPS position data and allow people to publish their media on an online map.

Figure 2. Location based information posted in Google Earth

Yet creating quality user experiences for location based services for the Mobile Internet is still not trivial. Given the novelty of location based services, little is known about the user experience of such of services. It is unclear how users perceive and use information provided through location based services, what content is considered relevant (and what is not) and how people will adopt and appropriate information services that react to their location and combine, for example, web content, satellite imaging, 3D graphics, and cartography. With the work presented in this paper, we contribute to an increased understanding of some of these important factors of location based service design and development. We do this through theoretically informed explanations of the interplay between context and technology in the user experience of location based services.

3. GESTALT THEORY

In this section, we turn our attention towards Gestalt theory and how it has previously been applied to the field of human-computer interaction.

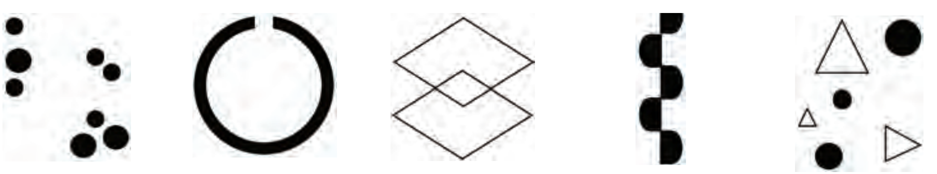
Gestalt theory evolved from explorations of human perception in the discipline of Psychology in the early twentieth century aiming to explain how people organize different information from their environment. The founders of Gestalt psychology are acknowledged as Max Wertheimer, Wolfgang Kohler and Kurt Koffka. Wertheimer applied Gestalt psychology to problem solving, Koffka to applied psychology and child psychology, and Kohler to learning strategies. Gestalt theory has over a hundred different laws that pertain to human perception, including visual and auditory. These laws of Gestalt psychology are fundamental in understanding the way people see and understand their surroundings (Borchers et al. 1996).

Gestalt theory explains how we perceive objects in our environment. The Gestalt viewpoint says that “things are affected by where they are and by what surrounds them” (Behrens 1984, p. 49), acknowledging the importance of context in how we perceived things. From the Gestalt perspective, new information is seen as organized and bridged to prior knowledge to form an organized whole, and it is the combination of the context that something sits in as well as our prior knowledge that allows us to interpret what we are looking at or listening to (Preece et al. 1994). Hence, Gestaltists believe that we intuitively perceive things as a coherent unit or object and that this is an innate human ability (Lauesen 2005). When presented with something in our physical environment that is ambiguous, we use our prior knowledge of the world to make sense of it by filling in the blanks in the current information (Smith-Gratto and Fisher 1998-1999).

Although the Gestalt laws are most often applied to visual perception, they also apply to other senses and cognitive processes. As expressed by Köhler (1947, p. 178), “the concept ‘Gestalt’ may be applied far beyond the limits of sensory experience. According to the most general functional definition of the term, the processes of learning, of recall, of striving, of emotional attitude, of thinking, acting, and so forth, may have to be included”. As a good example of this extension of the concept of Gestalt, Max Wertheimer, in his address before the Kant Society in Berlin in 1924 (Ellis 1938) raised the question of whether the listener’s experience of a melody is simply the sum of individual notes. He concluded that what we experience is rather determined by the character of the whole and that what takes place in each point in a musical piece depends upon the whole. Hence, we perceive patterns and form in music as wholes, and when a song is transposed to another key, we can still recognize it although all of the notes have changed.

3.1. Gestalt theory in human-computer interaction

Gestalt theory is included in several prominent human-computer interaction (HCI) primers (e.g. Benyon et al. 2005, Lauesen, 2005, Dix et al. 1998, Preece et al. 1994), and is introduced for its general application to the design of information screens and as providing interface designers with a theoretically informed understanding of how information screens are likely to be perceived by users. Gestalt theory, as it has been explored within human-computer interaction, consists only of a small subset of the original Gestalt laws in the form of a set of principles of perceptual organisation that can



Proximity
Things that are located near each other, in space or time, are perceived as belonging together

Closure
Things are perceived as complete or whole, even when part of the information is missing

Symmetry
When things have symmetrical borders or parts, they are perceived as a coherent whole

Continuity
We perceive things as continuous patterns rather than disjointed ones

Similarity
Things that look or sound alike are perceived as belonging together

Figure 3. Principles of perceptual organisation for interaction design derived from Gestalt theory

be applied to interface design to improve communication between user and system. For the purpose of this study, we have identified five such key principles, which are generally acknowledged within HCI. These are illustrated in figure 3.

The principles of perceptual organisation and other concepts derived from Gestalt laws have influenced many research areas related to HCI such as map reading, graph drawing, image retrieval, computer vision, pattern recognition, design of auditory displays and musical studies. Of particular interest to this paper is the research that has applied Gestalt principles to HCI analysis, design and evaluation.

As the graphical user interface became the predominant interface of computer systems during the 1990s, it was important for HCI researchers to provide interface designers with an understanding of the human perception of operating these new information-rich types of interfaces and with guidance on how to design them better. In response, researchers applied Gestalt principles of perceptual organisation to general screen design in order to create sets of design principles and guidelines for optimizing their graphical layouts. Important concepts of interface design, which are now considered elementary, such as consistency, visual hierarchy, grouping, legibility and contrast can be seen as derived from Gestalt theory (Roth 1995). This application of the Gestalt laws to interface design is also demonstrated by, for example, Mullet and Sano (1995), Roth (1995), Borchers et al. (1996) and Lauesen (2005). In more recent work within HCI, Gestalt theory's principles of perceptual organisation have been used as a basis for developing design guidelines for new paradigms of interaction design with multi-sensory displays that combine visual, auditory and haptic elements (Chang and Nesbitt 2005). Gestalt theory principles have also recently been linked with a pattern methodology for creation of a theoretical framework facilitating the incorporation of knowledge about human perception into the early stages of user interface design (Flieder and Mödritscher 2006). Beyond the creation of principles and guidelines for screen design, recent research has also explicitly used Gestalt theory in the evaluation of existing interface designs. In acknowledging the importance of human perception in measuring the quality of web page design, Hsiao and Chou (2006) used a combination of Gestalt grouping principles and fuzzy set theory from mathematics to develop a method to measure the Gestalt-like perceptual degrees of a web page design to evaluate the "wholeness" of that page.

Of significance to HCI is also the development of the concept of affordances and related cognitive design guidelines. Based on the seminal work of James Gibson (1979), who situates the origin of his Affordance Theory in work by Gestalt psychologists, the concept of affordances was introduced to the general HCI community by Donald Norman in "The Psychology of Everyday Things" (1988) and has had huge impact on interface and interaction design. According to Gibson's theory, perceiving one's environment leads to behaviour guided by clues indicating possible actions. Buttons are for pushing, handles are for pulling, knobs are for turning, etc. Extending Gibson's original affordances concept of "actual" action possibilities available in the environment independent of any individuals, Norman included objects' *perceived* properties highlighting the importance of the human observer/user and aligning it closely with the design philosophy and process of user-centred design (Norman 1999).

While much of the Gestalt related research within HCI is about graphical interface design, Oviatt et al. (2003) take a broader application of Gestalt principles by using Gestalt theory to analyse not just the computer screen but also the *interaction situation.* Aimed at the design of adaptive multimodal interfaces and providing a framework for understanding user interaction with multimodal information systems, they studied technology use in context, and observed how speakers tailor their language to accommodate the listener's perceptual capabilities. Explaining this and other phenomena, they applied Gestalt principles as a theoretical lens to look at both users perception of the interface and their production of communication patterns during its use.

Our approach, as reported in this paper, is similar to that of Oviatt et al. (2003). Rather than applying Gestalt principles only to the design of the interface, we view the computer screen as merely a small area within the larger context of the physical environment in which it is situated. In line with Gestalt theory's concept of "wholeness", the environment and the computer screen are seen as creating a unique perceptual whole, rather than as the simple sum of the individual parts. From this perspective, our study looks at the cognitively perceived ensemble of technology and surroundings as experienced when people are using location based services in context.

In terms of related work into providing better understanding of the user experience of location based services, our work is particularly related to research systematically and theoretically describing user experiences in relation to important contextual factors beyond peoples' location, such as activity, preferences, and information needs. As examples, Timpf (2002) uses ontologies of wayfinding derived from travelers' perspectives to reflect human models of the world and understand the different needs of a traveller at different stages of a trip. Focussing on the importance of individual preferences and time constraints, Raubal et al. (2002) propose a user-centred spatio-temporal theory of location based services combining time geography with an extended theory of affordances.

Another interesting piece of related work is Raubal and Winter (2002) presenting a method to automatically extract landmarks from geo-coded spatial datasets based on analysis of "landmark saliency". This information is then used to improve navigation services with wayfinding descriptions making use of concepts closer to the human user by referring to prominent features in the user's physical environment. What is particularly interesting in the context of the Gestalt approach proposed in this paper for understanding the user experience of location based services is the proposed measures for formally specifying landmark saliency based on visual, semantic, and structural attraction. Related to the Gestalt principle of closure, these measures define more specifically what combined properties of a feature in the physical environment make people perceive it as a prominent whole that stands out as a landmark.

4. CASE STUDY: THE "JUST-FOR-US" LOCATION BASED SERVICE

Inquiring into the user experience of location based services we have designed, implemented and evaluated a prototype location based service, *Just-for-Us,* providing an informational overlay to the civic space of Federation Square in Melbourne, Australia

(Figure 4). Federation Square was chosen because it was a relatively new civic structure, opened to the public in October 2002. It covers an entire city block and provides the people of Melbourne with a creative mix of attractions and public spaces for socializing including restaurants, cafes, bars, a museum, galleries, cinemas, retail shops and several public forums. In just a few years, Federation Square has become a highly popular place to socialize for Melbournians. It is open from early until late, every day of the week, and it hosts a rich range of planned and ad hoc activities. Located in the centre of the city, on major tram routes, and adjacent to a major train station, Federation Square is easily accessible, is considered a landmark in itself, and is a convenient place for people to arrange to meet up at the beginning of a night out on the town.

Figure 4. Federation Square, Melbourne, Australia, with surrounding skyline, train station and river

The Just-for-Us system (Figure 5) keeps track of the location of the user and friends within close proximity. It also keeps a history of visits to places around the city (for details see Kjeldskov and Paay 2006). On the basis of this, the service allows the user to explore his or her immediate surroundings through a series of annotated panoramic photographs. It also provides an overview of the level and nature of social activity taking place within proximity, and can make suggestions for places to go based on convenience, history, and social setting.

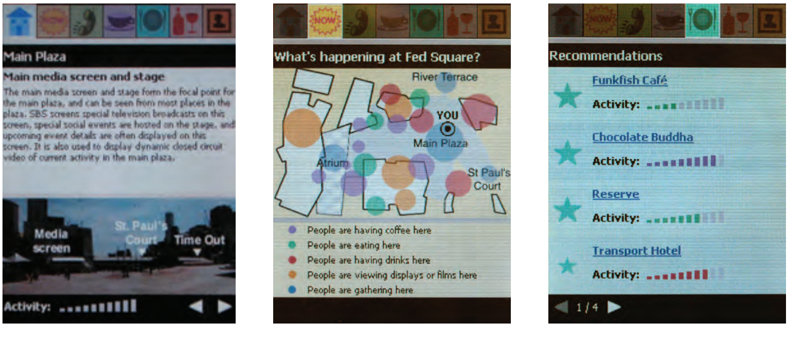


Figure 5. Screens from the Just-for-Us location based service

The design of the prototype system was informed by a field study at Federation Square exploring the interplay between people, technology and interactions in place guided by the categories of McCullough's (2001) typology of "on the town" everyday situations.

Three different established social groups participated in the study. Each group consisted of three young urban people, mixed gender, between the ages of 20 and 35, who had a shared history of socialising at Federation Square. Prior to the field visits each group received a 10-minute introduction to the study followed by a 20-minute interview about their socialising experiences and preferences. Each field visit lasted approx. 3 hours.

The use evaluation of the prototype location based service involved 20 established social pairs of mixed gender familiar with Federation Square (Figure 6). Inspired by rapid ethnography (Millen 2000) we gave the participants five overall tasks and scenarios for socialising, which prompted them to explore different parts of the system. Inspired by the constructive interaction approach to thinking-aloud studies with more than one user, the groups were asked to talk among themselves about their perception of and interaction with the system interrupted only with questions for clarification. The evaluations were video recorded by means of a miniature wireless camera attached to the mobile device mixed with a third-person view of the users. Participants wore directional wireless microphones, ensuring high-quality sound. Before taking part in the study, each participant pair jointly completed a history survey of their previous visits to Federation Square to simulate history data that the real system would have collected automatically. Each evaluation session lasted approximately 1.5 hours.

For testing purposes, the user's position, people and friends in vicinity etc. were entered manually using the "Wizard of Oz" technique. Inspired by the 1939 movie by the same name, Wizard of Oz is a technique for simulating system components commonly used for evaluation within the field of human-computer interaction (Dahlbäck et al. 1993, Buxton 2007). Using this technique, parts of a system's functionality that are not yet fully implemented are instead simulated "behind the scenes" without the knowledge of the test subjects. This is done in order to get rapid feedback on user interface design and envisioned functionality early in a design process.

Figure 6. Studying the location based service in use at Federation Square with wireless micro camera attached to the mobile device

Due to the fact that this was not a grounded theory building exercise but an exploration of the user experience of an example location based service, we used the rapid ethnography method of collaborative data analysis (Millen 2000) to provide the level of analysis needed. The collaborative data analysis approach was combined with the analytical technique of identifying critical incidents from the video data to produce a list of observations (Sharp et al. 2007) each associated with one of the five overall tasks.

The video data from the evaluations was then analysed by two researchers using content analysis, and observations were subsequently affinity diagrammed into higher-level issues (Beyer and Holtzblatt 1998). The outcome of this analysis was a list of 74 issues. In a second round of affinity diagramming, the two researchers first independently grouped these 74 issues in relation to Gestalt theory's five concepts of proximity, closure, symmetry, continuity, and similarity. Following this, the two independent groupings were then merged into one final set of groups in a structured collaborative effort. In case of disagreement between the individual groupings, the grouping of an issue was discussed until consensus was reached. In this process, 11 issues were not associated with any of the principles of proximity, closure, symmetry, continuity, and similarity, while the remaining 63 were associated with one or more of the five principles.

5. FIVE PRINCIPLES OF PERCEPTUAL ORGANISATION APPLIED

In this section we discuss qualitative findings from our field study of location based service use from the perspective of Gestalt theory's five principles of perceptual organisation as presented earlier as a lens for describing and explaining how people perceive the relationship between the mobile location based service and their environment. Like the illustrations depicted in figure 3, figures 7-11 below are designed simply to illustrate the Gestalt principles for location based services and are not meant as specific designs suggestions.

5.1. Proximity

The principle of Proximity defines that spatial or temporal proximity of elements may induce the mind to perceive a collective or totality. Things that are located near each other, in space or time, are perceived as belonging together.

Figure 7. Proximity: perceiving information as an annotation of a place

Proximity played an important role in the way that people interpreted the information presented by the Just-for-Us prototype. This happened on two levels: 1) in relation to the mobile device screen as an element of the environment, and 2) in relation to the mobile device screen on its own. Below, we discuss these two individually.

Firstly, and particularly interestingly for the design of location based services, information on the mobile device itself was seen as an annotation of the place people were situated in. People were grouping the system with objects in the environment, that is, the physical space acted as a "larger canvas" to draw from, on which the location based service was just another piece of information to be perceived and integrated into the whole experience. Relating digital information to physical locations is the essence of a location based service. The Gestalt principle of proximity explains, from a theoretical point of view, why this relationship makes sense to users of such services.

In our evaluation we found that people easily understand when information presented by the system is specific to their current physical location, and they like it when they are automatically given information relevant to where they are. In fact, the close proximate relationship between the system and the world made people perceive the information content of the service as true. For example, this happened when given the menu while they were at that particular café, or when presented with an annotated panoramic photograph of their location.

Secondly, the principle of proximity played an important role within the screen, in line with general screen design principles. The onscreen annotations on the panoramic photographs (figure 5 left) were perceived as belonging to the object or location that they were directly placed on top of, and also grouped annotations were perceived as belonging together. In addition, circles on the map (figure 5 middle), which represented the number of people at particular places, were perceived as applying to the places they were located near and groups of circles on the map were perceived as representing "busy" areas.

5.2. Closure

The principle of Closure defines that the mind may experience elements it does not perceive through sensation, in order to complete a whole. Things are perceived as complete or whole, even when part of the information is missing.

In our evaluation we found several examples of this. Although the maps used in the prototype were extremely simplistic line drawings with only a few annotations, people naturally perceived this as representing the much more complex real world around them. Annotations on the panoramic images supported people in "completing the picture" of what was behind the surrounding facades even though a large part of that picture was not visible to them. People also used their knowledge of familiar places referred to by the system as anchor points to resolve the layout of unfamiliar areas. As another example, closure played a major role in the manner in which people used the wayfinding information provided by the system. The visual perception principle of closure describes how people mentally complete incomplete graphical figures, such as a partial circle. In relation to wayfinding we found that this principle also applied to visualising a series of transition points as a complete path from A to B. As opposed to some guidance systems that give highly detailed step-by-step instructions our findings confirm other research showing that people only needed fragmented detail to find their way around urban spaces. Useful types of transition points were found to be references to familiar places, major entrances, landmarks (i.e., the river), or distinct architectural elements (i.e., the green glass wall). Another important finding in relation to closure was that in reducing the information presented to the user of a location based service, the significance of remaining information increases. This means that even though people are highly capable of connecting the dots, they still require carefully chosen "dots" to do so.

Figure 8. Closure: filling in the blanks of ambiguous information

Closure is the Gestalt principle that best describes the phenomenon that people are capable of making sense from small bits of fragmented and ambiguous information. Pieces of information on the mobile device are combined with pieces of information from the physical environment to create a "whole", and missing parts of this combined picture are filled in on the basis of peoples' prior knowledge and sense-making abilities. As described by this Gestalt principle, people supply the missing information themselves, drawing from a larger canvas in "connecting the dots" to make it easier to understand their environment.

5.3. Symmetry

The principle of Symmetry defines that symmetrical images are perceived collectively, even in spite of distance. When things have symmetrical parts or borders, they are perceived as a coherent whole.

Because people have a preference for symmetry, they made an effort to eliminate any asymmetry between the system and the real world. This was not used as much to piece together information to be able to understand it, as it was for the comfort of creating a coherent base on which to build understanding. In our evaluation we observed that people strived for symmetry between the system and the world. Visually this was evident as they worked to align the panoramic images on the mobile device screen with the buildings around them even though this was not actually necessary to operate the system. Some people even expressed that they would like the panoramic images to automatically correspond to the direction they were physically facing. Offsets between the viewpoint of the panoramic images on the screen and the user's location affected the symmetry between the two. People found this disconcerting, even though it was only a few degrees difference in view and they could still easily make sense of the representation. The same phenomenon was observed when using maps in the system. In this situation many people changed the orientation of the mobile device so that it aligned with their surroundings in striving for symmetry between the system and the physical environment and providing an egocentric frame of reference.

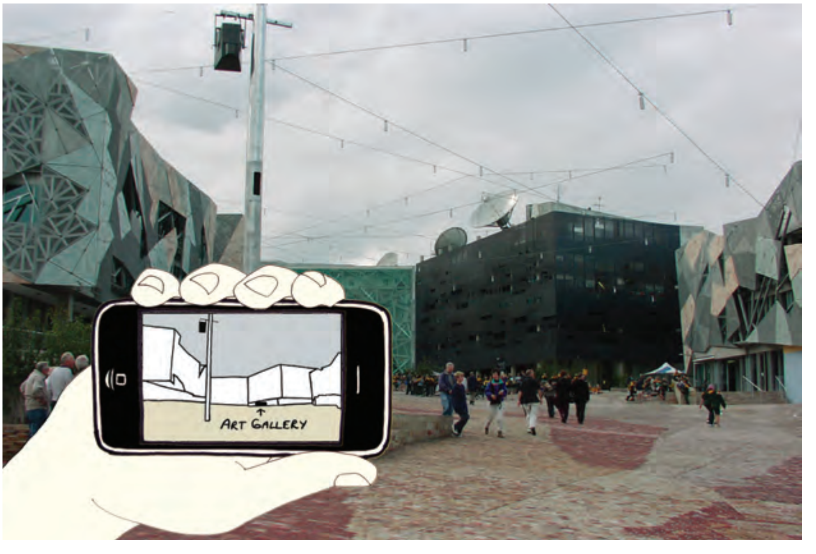


Figure 9. Symmetry: making a symmetrical alignment between system and surroundings

5.4. Continuity

The principle of Continuity defines that the mind continues a pattern, even after it stops. We perceive things as continuous patterns rather than disjointed ones. The principle of Continuity applies not only to visual sequences, but also to sequences perceived over time, such as a series of tones being perceived as music.

In looking at the use of location based services in context, the principle of continuity applies strongly to interaction over time. The fact that people have preferences for familiar places and paths indicates that interactions in a place do not happen as isolated events but are often an extrapolation of past experiences there. People have a trail of past interactions that they like to share with others, as much as they like to incorporate the trails of others into their own current experience. Rather than a random set of disjointed events, people tend to perceive their past experiences as interwoven in a continuous pattern. Events experienced close together in time are perceived as a continuous whole.

Figure 10. Continuity: experiencing a place over time

In this sense the larger canvas, which people draw from, consists not only of their mobile device and physical surroundings but also of their memories.

In our evaluation we found that although interested in exploring new places, people were primarily interested in information about current events at their familiar places. In this way, they continue to weave a story of interactions over time. When exploring new places, people preferred places that had been recommended to them by friends, and other trusted sources (i.e., reputable food guides), drawing on the experiences of others rather than starting from scratch. In this way they are adding to the continuity of other peoples' stories as well as enriching their own. Continuity also played a role in relation to the interpretation of descriptors used in the system. Here there was a clear preference for persistent descriptors, for example, "the black building", which refers to a constant quality of that building, rather than "the sitting steps", which refers to a transient activity.

The importance of continuity also came to our attention in relation to a part of our prototype system where people misunderstood or were surprised and disconcerted that the location based service adapted not only to location but also to their history of visits. Specifically, this happened when the system made suggestions for places to go based on where they had been in the past, but without indicating the rationale behind these recommendations. From the perspective of continuity, we had failed to represent to the user the trajectory of experiences from which these suggestions were drawn, thus making it impossible for people to see the recommendations as a part of their continuing experience with a place and an extrapolation of their past experiences.

5.5. Similarity

The principle of Similarity defines that the mind groups similar elements into collective entities or totalities. This similarity might depend on relationships of form, colour, size, or brightness. Things that look, sound or feel alike are perceived as belonging together.

Similarity played an enormous part in people's ability to make sense of the location based service. Things in the physical environment were continuously aligned with images and other representations on the screen that matched or looked alike. Through this, information in the system was perceived as belonging to the corresponding location or object in the world. This was not necessarily always a visual matching process. People were also able to draw on similarities between images and annotations on the screen and their knowledge about the physical environment, within and beyond visual range.

In our evaluation, similarity was primarily evident in matching physical objects and structures, such as media screens, a satellite dish, etc., to images on the screen. People looked for similarities in the outlines in their immediate, as well as distant, surroundings, such as the shape of buildings and the general skyline. They also used distinct features in their environment as anchor points for matching up the system and the world, for example, landmarks, unique patterns and colours on buildings. Finally, they used similarities between the visual style of the places surrounding them and the logos and other graphical elements in the system. Making sense on the basis of similarity happened not only on an iconic level, but also on a symbolic level. People often matched annotations on the screen, such as "the river", to the corresponding places in the world, and also matched up names in the system with signage in the physical environment. In fact, people found it perplexing if dominant signage in the world was not matched on the screen. Additionally, names in the system that hinted at the activity of a place, for example "Chinotto Café", were easily matched to a place if that activity was visually evident, in this case by the presence café tables and umbrellas. Again it was evident that in the use of the prototype location based service people were drawing conclusions from the larger canvas – not just from the system or from the context.

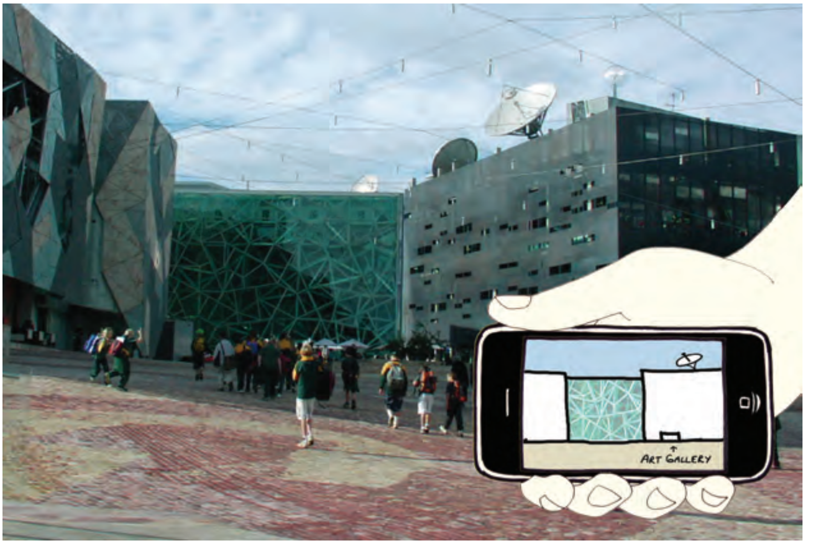


Figure 11. Similarity: matching things in system and surroundings that are alike

6. INFORMING LOCATION BASED SERVICE DESIGN AND EVALUATION

The implications of this work for the development of successful location based services are that the Gestalt principles can assist us in guiding the design and evaluation of quality user experiences of such services. By acknowledging and understanding the close interplay between technology and context in the use of location based services, and the specific implications for location based services highlighted by the five applied principles of perceptual organisation, we can distil a series of design-oriented questions. These questions can be applied either as a part of the design process of a location based service as prompts for functionality, or they can be applied as a part of an evaluation of an existing location based service as a partial set of heuristics used in combination with general interface design heuristics (e.g. Nielsen and Molich 1990).

Table 1. Five Gestalt principles, their implications for LBS, and questions for design and evaluation

Gestalt principle	Implications for LBS	Key questions for design and evaluation
Proximity	People perceive information content as closely related to their immediate location.	Does the system act as an annotation of its physical location? Can content in the user interface be grouped with objects in the environment?
Closure	People are able to fill in the blanks of ambiguous and fragmented information on the basis of their prior knowledge and sense-making abilities.	Can pieces of information in the user interface and in the surroundings be combined to create a meaningful larger whole? Does the system include clues about the relation between information in the user interface and its context? Does the system leave out redundant information already apparent through the context?
Symmetry	People have a preference for symmetry between system and surroundings in order to create a coherent base from which to interpret the relation between the two.	Does the system align information representations in the user interface with the user's surroundings? Does the system allow the user to align information representations in the user interface with their surroundings?
Continuity	People perceive their own and others' past experiences as an interwoven continuous story that develops over time and not as a series of disjoint events.	Does the system support user experiences that evolve over longer periods of time rather than a set of disjoint interactions? Does the system let the users extrapolate on their own and others' previous experiences in a place? Does the content of the system accumulate over time on the basis of peoples' use of it?
Similarity	People group content and representations in the system with elements in their physical surroundings based on iconic and symbolic similarity.	Does the user interface make use of representations that have similarities with corresponding objects in the physical surroundings? Does the user interface match elements in the surroundings, such as prominent signage, landmarks, and visual style of a place or area?

In Table 1 we summarise the observed implications for the design and use of location based services, and give examples of key questions to be asked in the design and evaluation of such services. It is important to note that the presented list of questions is open-ended and not complete. 7.

7. CONCLUSIONS

This paper has addressed the issue of explaining how people perceive and make sense of mobile location based services situated in context. Prompted by the finding from our research that people are extremely good at making sense from small and fragmented pieces of information when using location based services, we have analysed empirical data from a user study of such a system in pursuit of explanations of this phenomenon. In response, we have suggested the application of Gestalt theory as an analytical perspective for describing and explaining the interplay between people, mobile devices, and context of use through five principles of perceptual organisation. Informed by qualitative findings from our use evaluation, we have shown how Gestalt principles can be applied to the user experience of location based services as a way of explaining peoples' use of well functioning as well as problematic system design. In their use of location based services, people are not just drawing conclusions from their mobile device or their surroundings alone; they are drawing from "a larger canvas" to which both are contributing. As system designers of location based services, we need to focus on this larger canvas when designing rather than merely focussing on the "smaller canvas" of the mobile device.

Proximity explains how information on the mobile device screen was seen as belonging to peoples' current physical location. Closure explains the phenomenon of people relating and making sense of fragmented information and adding the missing bits themselves. Symmetry describes the desire to align representations in the location based service with the real world in order to obtain a coherent image from which to act. Continuity adds a temporal dimension and describes how information in a location based service does not exist in isolation from peoples' history of interactions with it. Similarity describes the mechanism by which people are able to group graphical elements in the system with corresponding elements in the surroundings. Acknowledging the importance of context, the Gestalt viewpoint is that things are affected by where they are and by what surrounds them. Hence, applying a Gestalt theoretic perspective to the user experience of context-aware mobile computer systems captures, in essence, the cognitively perceived ensemble of technology and context, and provides a foundation for rules about how this relationship can be exploited in interaction design.

While Gestalt theory's principles of perceptual organization might provide a useful umbrella theory for understanding, or a lens for characterizing, aspects of the user experience of location based services in a structured and systematic way, it is important to note that other concepts in the literature are dealing with the description and understanding of the observed characteristics of location based services in use. Proximity is closely related to the concept of *location* itself. Closure is related to *salience* or *abstraction.* Symmetry is related to *alignment* and *egocentric frames of reference.* Continuity is related to *patterns of movement.* Similarity is related to *matching*. However, although each is useful for describing one particular aspect of the user experience of a location based service, these concepts do not jointly make up a coherent whole within a common theoretical foundation, as is the case of the Gestalt principles of perceptual organisation. In contrast to the individual concepts outlined above, the Gestalt approach suggested in this paper provides a broader framework and a way of thinking about the user perception of location based services that explicitly promotes a holistic view on the ensemble of elements perceived by the users, that is, mobile devices as part of their context of use. It is our hope that applying a Gestalt theory perspective can add successfully to the repertoire of concepts and theoretical foundations for understanding the user experience of location based services.

This research is still ongoing and evolving. Motivated by the promising outcomes from the analysis of our empirical data presented in this paper, we are in the process of collecting further empirical data to extend our analysis. From this research we aim to expand on the descriptions of Gestalt theory principles as experienced by people in relation to their use of location based services. As a part of this, we are continuously elaborating on the framework presented in Table 1 towards the refinement of an established set of Gestalt-based design heuristics for the user experience of location based services on mobile devices.

8. ACKNOWEDGEMENTS

This research is supported by the Danish Technical Research Council through the Indexical Interaction Design project (project ref. 26-04-0026). The authors thank everyone participating in the field studies. We also thank Steve Howard and Bharat Dave for valuable input throughout the project.

REFERENCES

BBC (2002) Mobile users unexcited by 3G [online]. BBC News. Available from: http://news.bbc.co.uk/1/hi/business/2050994.stm [Accessed 22 Oct 2008].

Behrens, R., 1984. Design in the Visual Arts. Englewood Cliffs: Prentice-Hall.

Benyon, D., Turner, P. and Turner, S., 2005. Designing Interactive Systems. Harlow: Addison-Wesley.

Beyer, H. and Holtzblatt, K., 1998. Contextual design: Defining customer-centred systems. San Francisco: Morgan Kaufmann.

Borchers, J., Deussen, O., Klingert, A. and Knorzer, C., 1996. Layout Rules for Graphical Web Documents. Computer Graphics, 20 (3), 415-426.

Buxton, B., 2007. Sketching User Experiences: Getting the Design Right and the Right Design. San Francisco: Morgan Kaufmann.

Chang, D., Dooley, L. and Tuovinen, J., 2001. Gestalt Theory in Visual Screen Design – A New Look at an Old Subject. In: A. McDougall, J. Murnane and D. Chambers, eds. Proceedings of the Seventh World Conference on Computers in Education, WCCE'01, 29 July-3August 2001 Copenhagen. Darlinghurst: ACS, 5-12.

Chang, D. and Nesbitt, K., 2005. Developing Gestalt-based Design Guidelines for Multi-sensory Displays. In: F. Chen and J. Epps, eds. Proceedings of the 2005 NICTA-HCSNet Multimodal User Interaction Workshop, MMUI'05, 13-14 September 2005 Sydney. Sydney: ACS, 9-16.

Costolo, D., 2005. Mobile Operators Seek Incremental Revenue from the Mobile Internet [online]. SYS-CON Media: http://wbt.sys-con.com/read/40977.htm [Accessed 22 Oct 2008].

Dix, A., Finlay, J., Abowd, G. and Beale, R., 1998. Human-Computer Interaction. London: Prentice Hall Europe.

Dahlbäck, N., Jönsson, N. and Ahrenberg, L., 1993. Wizard of Oz studies: why and how. In: W.D. Gray, W.E. Hefley and D. Murray, eds. Proceedings of the 1st international conference on intelligent user interfaces, 4-7 January 1993 Orlando. New York: ACM, 193-200.

Dourish, P., 2004. What do we talk about when we talk about context? Personal and Ubiquitous Computing, 2004(8), 19-30.

Ellis, W.D., 1938. Source Book of Gestalt Psychology. New York: Hartcourt, Brace and Co.

Ericsson, 2007. The 2006 Australian Survey. Ericsson Consumer and Enterprise Lab [online]. Ericsson. Available from: http://www.ericsson.com/au/ericsson/press/2006/Consumer_Lab_2007_public.pdf [Accessed 22 Oct 2008].

Forrester, 2006. Breaking The Mobile Internet's Low Adoption Spell - Best Practices That Will Boost The Mobile Internet's Appeal For Content Firms And Consumers. Forrester Research. Available from: http://www.forrester.com/Research/Document/Excerpt/0,7211,39880,00.html [accessed 2 March 2007].

Fithian, R., Iachello, G., Moghazy, J., Pousman, Z. and Stasko J., 2003. The Design and Evaluation of a Mobile Location-Aware Handheld Event Planner. In: L. Chittaro, ed. Proceedings of the 5th International Symposium on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI 2003. 8-11 September 2003 Udine. Berlin: Springer, 145-160.

Flieder, K. and Mödritscher, F., 2006. Foundations of a Pattern Language Based on Gestalt Principles. In: R. Jeffries, ed. CHI '06 extended abstracts on Human factors in computing systems, 22-27 April 2006 Montreal. New York: ACM, 773-778.

Frøkjær E. and Hornbæk K., 2008. Metaphors of Human Thinking for Usability Inspection and Design. Transactions on Computer-Human Interactions, 24(4), 1-33.

Frøkjær E. and Hornbæk K., 2002. Metaphors of Human Thinking in HCI: Habit, Stream of Thought, Awareness, Utterance, and Knowing. In: F. Vetere, L. Johnston and R. Kushinsky, eds. Proceedings of the HF2002/OzCHI 2002 Human Factors Conference, 25-27 November 2002 Melbourne Australia. Melbourne: HFESA and CHISIG, 964-971.

Gibson, J., 1979. The Ecological Approach to Visual Perception. New Jersey: Lawrence Erlbaum Associates.

Hinckley, K., Pierce, J., Horvitz, E. and Sinclair, M., 2005. Foreground and Background Interaction with Sensor-Enhanced Mobile Devices. Transactions on Computer-Human Interaction, 12(1), 31-52.

Hsiao, S. and Chou, J., 2006. A Gestalt-like perceptual measure for home page design using fuzzy entropy approach. International Journal of Human-Computer Studies, 64, 137-156.

Johnson, P., 1998. Usability and Mobility; Interactions on the move. In: C. Johnson, ed. Proceedings of the first Workshop on Human Computer Interaction with Mobile Devices, Mobile HCI 1998, 22 May 1998 Glasgow. Glasgow: University of Glasgow, Department of Computing Science.

Jones, Q., Grandhi, S.A., Terveen, L. and Whittaker, S., 2004. People-to-People-to-Geographical-Places: The P3 Framework for Location-Based Community Systems. Computer Supported Cooperative Work, 2004(13), 249-282.

Kaasinen, E., 2003. User Needs for Location-Aware Mobile Services. Personal and Ubiquitous Computing, 2003(7), 70-79.

Kjeldskov, J. and Paay, J., 2006. Public Pervasive Computing: making the Invisible Visible. IEEE Computer, 39(9), 60-65.

Köhler, W., 1947. Gestalt Psychology: An Introduction to New Concepts in Modern Psychology. New York: Liveright Publishing Corporation.

Lauesen, S., 2005. User Interface Design: A software engineering perspective. Harlow: Addison-Wesley.

Lee, I., Kim, J. and Kim, J., 2005. Use Contexts for the Mobile Internet: A Longitudinal Study Monitoring Actual Use of Mobile Internet Services. International Journal of Human Computer Studies, 18(3), 269-292.

Millen, D. R., 2000. Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In: D. Boyarski and W.A. Kellogg, eds. Proceedings of the 3rd conference on Designing Interactive Systems, DIS 2000, 17-19 August 2000 New York City. New York: ACM, 280-286.

Mullet, K. and Sano, D., 1995. Designing Visual Interfaces: Communication oriented techniques. Englewood Cliffs, NJ: Prentice Hall.

Nielsen, J. and Molich, R., 1990. Heuristic evaluation of user interfaces. In: J.C. Chew and J. Whiteside, eds. Proceedings of the SIGCHI conference on Human factors in computing systems, CHI 1990, 1-5 April 1990 Seattle. New York: ACM, 249-256.

Norman, D., 1988. The psychology of everyday things. New York: Basic Books.

Norman, D., 1999. Affordance, conventions, and design. Interactions, 6(3), 38-43.

Oviatt, S., Coulston, R., Tomko, S., Xiao, B., Lunsford, R., Wesson, M. and Carmichael, L., 2003. Toward a Theory of Organized Multimodal Integration Patterns during Human-Computer Interaction. In: T. Darrell, M. Maybury and W. Wahlster, eds. Proceedings of the 5th international Conference on Multimodal interfaces, ICMI '03, 5-7 November 2003 Vancouver. New York: ACM, 44-51.

Paay, J. and Kjeldskov, J., 2008. Situated Social Interactions: A Case Study of Public Places in the City. Computer-Supported Cooperative Work, 17(2-3): 275-290.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey, T., 1994. Human-Computer Interaction. Workingham: Addison-Wesley.

Raper, J., Gartner, G., Karimi, H. and Rizos, C., 2007. A critical evaluation of location based services and their potential. Journal of Location Based Services, 1(1), 5-45.

Raubal, M., Miller, H.J. and Bridwell, S., 2004. User-Centered Time Geography for Location-Based Services. Geografiska Annaler: Series B, Human Geography, 86(4), 245 – 265.

Raubal, M. and Winter, S., 2002. Enriching Wayfinding Instructions with Local Landmarks. In: G. Goos, J. Hartmanis, and J. van Leeuwen, eds. Proceedings of Second International Geographic Information Science Conference, GIScience 2002, 25-28 September 2002 Boulder. Berlin: Springer, 243-259.

Roth, S., 1995. Visual Literacy and the Design of Digital Media. Computer Graphics, November 1995, 45-47.

Sacher, H. and Loudon, G., 2002. Uncovering the New Wireless Interaction Paradigm. Interactions, 9(1), 17-23.

Sharp, H., Rogers, Y. and Preece, J., 2007. Interaction Design: beyond human-computer interaction. 2nd Ed. Chichester: John Wiley and Sons.

Smith-Gratto, K. and Fisher, M., 1998-1999. Gestalt Theory: A Foundation for Instructional Screen Design. Journal of Educational Technology Systems, 27 (4), 361-371.

Timpf, S., 2002. Ontologies of Wayfinding: A Traveler's Perspective. Networks and Spatial Economics, 2 (2002), 9-33.

Chapter 21

Indexical interaction design

Jesper Kjeldskov and Jeni Paay

Abstract. A lot of research has been done within the area of mobile computing and context-awareness over the last 15 years, and the idea of systems adapting to their context has produced promising results for overcoming some of the challenges of user interaction with mobile devices within various specialised domains. However, today it is still the case that only a limited body of theoretically grounded knowledge exists that can explain the relationship between users, mobile system user interfaces, and their context. Lack of such knowledge limits our ability to elevate learning from the mobile systems we develop and study from a concrete to an abstract level. Consequently, the research field is impeded in its ability to leap forward and is limited to incremental steps from one design to the next. Addressing the problem of this void, this article contributes to the body of knowledge about mobile interaction design by promoting a theoretical approach for describing and understanding the relationship between user interface representations and user context. Specifically, we promote the concept of indexicality derived from semiotics as an analytical concept that can be used to describe and understand a design. We illustrate the value of the indexicality concept through an analysis of empirical data from evaluations of three prototype systems in use. Based on our analytical and empirical work we promote the view that users interpret information in a mobile computer user interface through creation of meaningful indexical signs based on the ensemble of context and system.

1. INTRODUCTION

Emerging technologies have made it possible for mobile computers to sense or access information about their user's contextual setting such as their physical environment, their location, their social setting, and their current activity (Bardram 2009, Hinckley et al. 2005, Jones et al. 2004, Dey 2001, Dix et al. 2000, Gaver et al. 1999, Crabtree and Rhodes 1998). Enabled by this, research in mobile human-computer interaction

Originally published as Kjeldskov, J. and Paay, J. (2010) Indexicality: understanding mobile human-computer interaction in context. *ACM Transactions on Computer-Human Interaction,* 17(4): article 14.

has demonstrated that the usability of mobile computer systems can benefit from making them "context-aware" in the sense that contextual information is used to tailor information and functionality to the given situation (Bardram 2009, Barkhuus and Dey 2003, Kaasinen 2003, Cheverst et al. 2000). The potential benefits of context-awareness are several. By making mobile computer systems aware of their user's contextual setting, designers can present information and functionality relevant only in specific situations (Barkhuus and Dey 2003, Cheverst et al. 2001). In this way, the user interface can be simplified and the demand for user interaction can be reduced (Crabtree and Rhodes 1998). Tailoring the interface to its context may facilitate partial automation of repetitive and trivial tasks (Gaver et al. 1999) and making the system react to contextual changes can also be used to increase security of data and users (Rantanen et al. 2002) and to improve safety-critical applications (Bardram and Nørskov 2008). Cataloguing information by automatically sensed contextual meta-data can be used to supplement human memory through intelligent mobile information retrieval systems (Lamming and Newman 1992). An example of combining these potentials is systems for the contextually complex domain of healthcare (see for example Bardram and Bossen 2005). In such systems, work, mobility, and collaboration can be supported through "Activity-Based Computing" and awareness about co-workers (Bardram 2009, Bardram et al. 2006), and information complexity in mobile patient record systems can be reduced by tailoring the interface to the nurse's location, current work activity, patients within proximity etc. (Skov and Høegh 2006). Information access could also be supported by making relevant related data and documents from previous similar work activities immediately available (Lamming and Newman 1992).

However, although a lot of research has been done within the area of context-awareness over the last 15 years since the term was first introduced by Schilit and Theimer (1994) (Schmidt et al. 2004), the promise of context-awareness for mobile human-computer interaction has not yet been fully realized in practice. So far, the impact on commercial products has been small, and mostly focussed on location. Exceptions include the iPhone's use of context sensors to work out the orientation of the device and adjust the interface accordingly. There are many reasons for the relatively slow transition from the fundamentally sound idea of context-awareness to useful and usable real world mobile systems. Below we outline a few.

Firstly, the seemingly simple and attractive idea of making technology context-aware hides a large degree of complexity (Brown and Randell 2004). In practice, even context aware applications that appear to be very simple, like the mobile phone that only rings where, or more importantly when, appropriate, are highly complicated to realise in practice because of the of the deeply complicated nature of context interpretation – for humans as well as for machines.

Secondly, although a lot of effort has been dedicated to sensing, adapting to, and examining the very complex concept of 'context' (Dourish 2004, Chalmers 2004), and although many definitions exist, mobile computer use context is still not well understood in a way that translates well into mobile interaction design. For example, it is still unclear how the different elements of a user's context influence their interpretation and use of mobile systems. It is also unclear how to utilize knowledge about context, decide what information and functionality to present, what to leave out, and how to make use of information already implicitly present in the user's surroundings (Paay et al. 2009).

Thirdly, only a limited body of knowledge exists that can help explain, theoretically, the relationship between users, mobile system user interfaces, and their context. This lack of knowledge limits our ability to elevate our learning from the mobile systems we develop, and study in use, from a concrete level focussing on the specific characteristics of specific systems, to an abstract level where knowledge can be generalized and transferred to other design cases, other technologies, domains, users, purposes, etc. Consequently, the research field is impeded in its ability to move forward in a pace beyond the incremental steps from one design to the next.

It is our belief that addressing and progressing the third issue would also help progress the first two. We believe that expanding the body of theoretical knowledge about the relationship between users, systems, and context holds a key to understanding the concept of context in a way that could inform interaction design better. Jointly, these will reduce the complexity of creating context-aware mobile computer systems and support realising real world applications in practice.

Contributing to the body of knowledge about mobile interaction design, this article promotes and discusses a theoretical approach for describing and understanding the relationship between user-interface representations and user context. Our purpose has been to create a theoretical foundation for future research and design by developing the concept of indexicality as an analytical lens. This lens applies to mobile user interfaces that carry a major part of their meaning implicitly through the context in which they are used. Achieving this, we have conducted theoretical, technical and empirical research. Our theoretical work has explored the concept of indexicality as a lens for describing and understanding the interpretation of information on mobile computers in context. Our technical work has explored design and implementation of prototype systems making use of indexical interface representations. Finally, our empirical work has used these mobile prototypes as vehicles for studying user interaction in context.

This paper advances from our previous work on the topic (Kjeldskov 2002, Kjeldskov and Paay 2006) by presenting the indexicality-approach as a detailed and coherent argument, and by presenting further empirically grounded analyses of the interplay between users, mobile systems and their context. We also discuss how the concept of indexicality could be used to inform a design process. We do not, however, aspire to present a complete coverage of the topics of context and context-awareness. Nor are we going to provide a step-by-step recipe for how to use indexicality in the design of such systems, but on the basis of the analytical approach presented, we hope that others will be inspired to make such a contribution.

In section 2 we discuss the concept of context and present a number of definitions and views from related literature. In section 3 we turn our attention towards the concept of indexicality and how this can be used to explain the relationship between information representations and context. Section 4 presents three mobile prototype systems used for gathering empirical data about use in context, and section 5 presents evaluations of those prototypes. In section 6 we present and discuss findings across the evaluations, using the concept of indexicality as a theoretical lens for analysing and understanding the relationships between users, mobile system user interface, and their context. Section 7 discusses how indexicality could be used to inform design. Finally, we summarize, conclude and point towards future research and plans for extending this work.

2. CONTEXT

Understanding context in as important part of informing design (Alexander 1964). There are many different definitions of context, and the debate on what constitutes context for mobile computing is ongoing. Early works on context-aware computing referred to context as primarily the location o people and objects (Schilit and Theimer 1994). In more recent works, context has been extended to include a broader collection of factors such as physical and social aspects of the environment (McCullough 2004, Dourish 2004, Bradley and Dunlop 2002, Agre 2001, Dey 2001, Abowd and Mynatt 2000, Schmidt et al. 1999, Crabtree and Rhodes 1998), as well as the activities of users (Bardram 2009).

Dey (2001) characterizes context in the following way: "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and the application themselves." Although this definition is quite complete, it is not very specific about what type of information could in fact be used to characterize such a situation. In contrast to this, Schmidt et al. (1999) present a model of context with two distinct categories: human factors and physical environment. Human factors consist of the three categories: information about the user (profile, emotional state, etc), the user's social environment (presence of other people, group dynamics, etc), and the user's tasks (current activity, goals etc.). Physical environment consist of the three categories: location (absolute and relative position, etc.), infrastructure (computational resources, etc.), and physical conditions (noise, light, etc.). This model provides a good catalogue of specific contextual factors to complement broader definitions like the one by Dey (2001).

Other works are not as comprehensive in their coverage of different contextual factors but go into detail about one or a few. In the works of Agre (2001) and McCullough (2004) particular importance is given to physical context consisting of architectural structures and elements of the built environment, for example, landmarks and pathways. In the works of Dourish (2001, 2004) particular importance is given to social context including interaction with, and the behaviour of, people in an environment. Dourish (2004) also states that context cannot be defined as a stable description of a setting, but instead arises from, and is sustained by, the activities of people. Hence it is continually being renegotiated and redefined in the course of action. These works provide us with additional contextual factors of particular relevance to mobile computing in context, and Dourish teaches us that what defines context is in itself contextually dependant.

The purpose of our work has not been to define context or challenge the existing definitions proposed in the literature. Instead, we subscribe to the definition by Dey (2001) and to the fact that several dimensions of context exist, and that the relevance of each of these for a particular system or use situation is itself dependent on context. From this starting point we are interested in explaining and describing relationships between particular dimensions of context and information representations on mobile devices. Our work does not address all dimensions of context mentioned here or in the literature. We have focussed on spatial context (absolute and relative location), physical context and social context. The reason for choosing these aspects of context is pragmatic. These are aspects of context that are often used in context-aware systems and often discussed in the literature. Hence we found this to be a suitable starting point. Other aspects of context are of course relevant as well. As an example, our three prototype systems also index to activity and temporal aspects of context, albeit not as strongly.

3. INDEXICALITY

An interesting theoretical concept for describing and understanding the user interface on a context-aware mobile computer system is that of indexicality. Indexicality is a concept drawn from semiotics describing the relation between representations and the context in which an interpreter perceives them. Taking an indexical/semiotic approach to the analysis of user interface design can contribute to a theoretical understanding of peoples' interpretation of information representations on context-aware mobile devices. Semiotics is "the study of the social production of meaning through signs" (Scollon and Scollon, 2003, p. 215). A sign is any material object that refers to something other than itself and in semiotic theory includes language, discourse, books, conventional signage (e.g., street signs), the built environment (e.g. roadways and paths indicating places to transit), and people (e.g. through physical presence, movements and gestures) (Scollon & Scollon). Peirce (1931) developed a triadic model of the sign, commonly known as the semiotic triangle, which considers the representamen (the form a sign takes), the interpretant (the sense made of that sign), and an object to which the sign refers (Chandler 2002). Simply speaking, signs are viewed as representations of something else (their object), and faced with a human interpreter these representations cause a reaction or interpretation (figure 1).

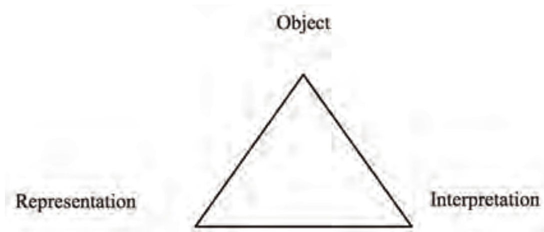


Figure 1. The semiotic relation between object, representation and interpretation

Pierce also developed three main categories of signs: symbolic (conventional), iconic (similarity) and indexical (material or causal). Symbols and icons are ways of representing information independent of context. A symbol is a sign that is a completely arbitrary representation of something in the world; the sign does not resemble what it is signifying. Examples include alphabetic letters, numbers, Morse code, or national flags. An icon is a picture of something in the world, which is perceived as resembling what it signifies. An obvious example is the use of icons in graphical user interfaces. Other examples include portraits, cartoons, or sound (Chandler 2002, Scollon and Scollon 2003).

Indexes on the other hand, are ways of representing information with a strong relation to their context (for example, spatial and/or temporal) exploiting information present in the interpreter's surroundings. An index is a sign that means something because of where and when it is located in the world. It is not arbitrary, and is directly connected, either explicitly or implicitly, to the thing it signifies. Indexical representations are, for example, used on signposts and information boards. Other examples include indexical words such as 'here', 'there', etc. (Chandler 2002, Scollon and Scollon 2003).

There are many contrasting views in semiotic theory, and taking a purist view of indexicality, Peirce (1931, Vol. 2, p. 306) states that "it would be difficult if not impossible, to instance an absolutely pure index, or to find a sign absolutely devoid of the indexical quality". A true indexical reference does not require the object of reference to be explicitly indicated, so that in order for it to be successfully interpreted, the interpreter needs to understand the detailed context in which it is given (Chandler 2002). For the purpose of the work presented in this article we take a more pragmatic view, where indexicality is based on association by contiguity (Martinovski 1995). An indexical reference is one that relies on a direct connection to an object in the world, through an implicit or explicit representation that "points to" that object, and where the interpretation is reliant on the context of that communication for understanding. Hence, we define indexicality as a property of an information representation that has context-specific meaning. This means that it is dependent on a referent with which it has a relation for its meaning. For example, if a digital display in a train carriage in Denmark reads "Aalborg" when approaching Aalborg train station, it is indexical because of the train's (and therefore the sign's) proximity-based relationship to that station. The full meaning of the digital display is "Aalborg is the next station", but some of this information can be left out as it is given implicitly in the context that the sign is indexing to.

3.1. Reducing information representations by increasing indexicality

Elaborating on this line of thinking, it is clear that symbolic and iconic representations can be converted into temporal and spatial indexical representations by locating them in time and space. As shown in Kjeldskov (2002), increasing the level indexicality in an information representation by locating it in time and space results in a reduction of symbolic and iconic representations required to communicate a particular piece of information. This inverse relationship is exemplified below and illustrated in figures 2, 3 and 4, which show three different types of information representations related to train departures: a timetable book, a timetable poster in a foyer, and an electronic timetable display on the platform of a train station.

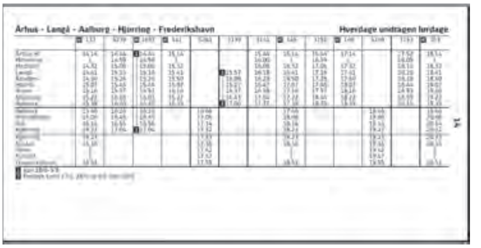


Figure 2. Page from paper based timetable book: symbolic representations with no indexicality

The page from a timetable book shown in figure 2 exemplifies symbolic (and potentially also iconic) information representation with no indexicality. It contains information about train departures at all times and at all places (within the coverage and valid lifetime of the book). Hence this representation is valid, and useful, independent of the user's location in space and time. Consequently, the amount of information contained in a book like this is quite extensive.

Figure 3 shows a paper based timetable poster commonly put up in a central location of a train station. This representation contains a selection of information from the timetable book namely information about departures at all times from here (the train station where the poster is on display). Hence, this representation of information is only valid at a particular location, and would be wrong if put on display at a different train station. In relation to the timetable book depicted in figure 2, the timetable poster in figure 3 is spatially indexical. As a result, the amount of information is greatly reduced.

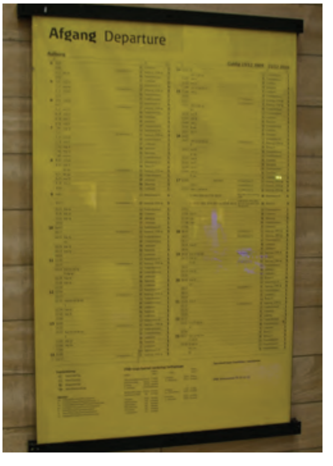


Figure 3. Timetable poster at a train station: symbolic representation with spatial indexicality

The electronic timetable display in figure 4 exemplifies a further increase of indexicality leading to a further reduction of information. This display contains information about all train departures from here, within a short time, and is only valid (and relevant) at a specific location and at a specific time. It is a symbolic information representation with spatial and temporal indexicality.

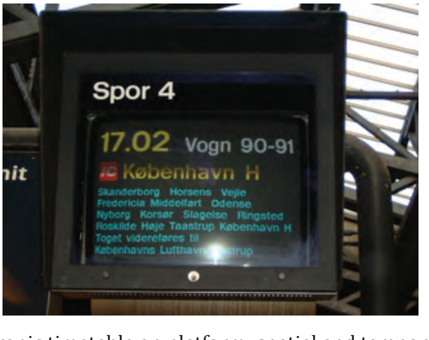


Figure 4. Electronic timetable on platform: spatial and temporal indexicality

As can be seen from this example, increasing the indexicality from the paper based book to the situated electronic display results in a huge reduction of information to be displayed and of the amount of user interaction required. Instead of having to look up departures from a specific location at a specific time (figure 2), the user is presented with a reduced selection of information tailored to his or her location (figure 3) or to location and current time (figure 4).

From the examples above it is also clear that an information representation that has the property of indexicality can only be understood correctly in a particular context. If removed from its context, information will, at best, lose its indexical properties and make little or no sense. At worst, the information representation may take on false meaning. If the digital sign in the train saying "Aalborg" were displayed when leaving Aalborg station, it would just be confusing or redundant. However, if displayed on a train leaving from a different station and not going to Aalborg it would communicate false information about its destination.

3.2 Indexical Interface Design

Andersen (2002) extended the concept of indexical representations into the digital domain by stating that "pervasive and mobile computing tend towards producing indexical signs, and since the sign can be adapted to its physical context, parts of its meaning can be located in the surrounding place". He also emphasized that this connection between space, time and information is especially important in situations where signs move with the user, such as in the case of mobile computers. Hence, if indexical-type representations can mediate context and exploit knowledge-in-the-world to increase the communicative power of situated information representations, then the concept of indexicality should also be of value for interface design of context-aware mobile computers: to explain, for existing systems, the relationship between information on the screen and its context; and to guide, for new systems, designs that actively explore this indexical relationship.

The potential of applying indexicality as an analytical lens when looking at existing mobile systems in use is that it provides a theoretical foundation on which the relationship between system and context can be analysed and understood. Looking at the ensemble of context and mobile computer system as a joint indexical sign can help explain why some design solutions work well while others don't. The potential of using the concept of indexicality to inform design is to explore this theoretical understanding by explicitly drawing on the fact that if information and functionality on a mobile computer can be indexed to the user's situation, then information already provided by the context becomes implicit and does not need to be displayed explicitly. In this paper, we focus on the potential of using indexicality as an analytical lens.

It is clear that indexicality and context-awareness are closely related. The difference between indexicality and context-awareness is that indexicality is a theoretical concept while context-awareness is a technical property of a system. Context-aware systems adapt information content to its user's context. Indexicality describes how this contextually adapted information is interpreted. Hence, in short, indexicality can be used to describe, from a theoretical point of view, how and why context-aware systems make sense.

When interacting with a mobile context-aware system the world outside the computer system becomes a part of the interface (Crabtree and Rhodes 1998) and the system's output is interpreted in light of a rich backdrop of implicit information in the context. As in the examples above, increasing the level of indexicality means that the amount of information explicitly presented to the user can be reduced. This is of great value when designing for the limited screen real estate of a mobile device. As an illustration, a context-aware mobile information service for patrons entering a cinema complex could reduce information in the interface by means of indexical references to time, location and social context. It could, for example, provide only information about upcoming movies playing within a limited frame of time (temporal indexicality) in that specific cinema (spatial indexicality) of interest to a group of users (social context) (Kjeldskov 2002).

4. THREE PROTOTYPE SYSTEMS

The examples above illustrate that the concept of indexicality can be used as a lens for analyzing and describing information systems in context. However, as the value of an indexical user interface relies strongly on the user's interpretation of and knowledge about their own context, for instance where and when they are situated, it is important to complement this type of descriptive theoretical analysis with analysis based on empirical data about actual use of such systems. From a theoretically informed analysis grounded in empirical data it is possible to gain a deeper understanding of how indexicality between user-interfaces and user-contexts functions in practice, and what potentials and limitations this way of thinking has for interface design.

Following on from our theoretical work we have developed a series of functional context-aware mobile prototype systems that have served as vehicles for studies of use in real world contexts. Three of these systems are particularly relevant for the argument presented in this article and are described briefly in the following sections. The three systems are similar in that they all run on handheld mobile computers and all represent some level of context-awareness. However, they are very different in terms of their target use domain and purpose. The first prototype, TramMate II, is a mobile route-planning service. The second one, MobileWARD, is a mobile electronic patient record terminal. The third one, Just-for-Us, is a mobile urban guide system.

4.1. TramMate II

In 2003 we explored ways of supporting use of the tram-based public transport system in Melbourne, Australia, by means of mobile information systems. This was done through field studies on the use of transportation by business employees attending appointments at different physical locations in the city during a typical workday (Kjeldskov et al. 2003). As a part of the project, a functional mobile guide prototype was developed by researchers at the University of Melbourne's Department of Geomatics (Smith et al. 2004). The prototype, here referred to as TramMate II, provided route-planning facilities for the Melbourne tram system based on the user's current location (figure 5). This was done through a combination of textual instructions and annotated maps.

The TramMate II prototype had three basic functions supporting the use of public transport. The first function was a "Timetable Lookup". This provided timetable information based on stop and route numbers entered by the user (origin and destination), and was aimed at regular tram users who are very familiar with their route. The second function, “Plan Trip”, provided information about the whole route containing route descriptions and maps of the individual segments of the journey. This was based on user entry of origin and destination suburbs and street corners, and also allowed entry of desired arrival or departure time. The third function, “Determine Route”, provided a simplified “Plan Trip” function where the user’s origin was resolved via GPS and the system automatically computed a travel plan to a manually entered destination.

Figure 5. TramMate II prototype system: route-planning information for the public transport system in Melbourne, Australia indexed to location, time and physical objects.

The TramMate II prototype was implemented for a Compaq iPAQ handheld computer equipped with a WAP browser. The device was connected to the Internet via GPRS.

4.2. MobileWARD

As a part of a larger research activity studying the use of information systems in the healthcare domain, a prototype mobile context-aware electronic patient record (EPR) terminal, refereed to as MobileWARD, was developed at Aalborg University, Denmark (Skov and Høegh 2006). MobileWARD supports the work activities of nurses during their morning round by keeping track of contextual factors such as the nurse’s location, patients and staff in proximity, upcoming tasks etc. and automatically presents relevant data from the electronic patient record database to the nurse based on this (figure 6).

In our previous studies of stationary EPR system use at a large regional hospital, we had found that the usefulness of such systems suffered from issues related to mobility, complexity, and lack of relation to work activities (Kjeldskov and Skov 2007). Firstly, most nurses were concerned about the EPR system not being mobile while many of their work tasks required them to move between different locations. Due to the complexity of information in the EPR system, nurses also had difficulties finding the information necessary for doing their work. Finally, they experienced problems with the use of the EPR system because the data and structure of information in the system did not relate clearly to work activities, locations, and people (nurses, doctors and patients).

Figure 6. MobileWARD prototype system: Indexing patient information at a large regional hospital in Denmark to patients in proximity, location and upcoming work activities

MobileWARD responds to these observations by providing patient data filtered by and indexed to context. When the nurse is in the corridor, the system lists all patients admitted to the ward, highlighting the ones assigned to her. For each patient, MobileWARD provides information about previous tasks, upcoming tasks and upcoming operations. If the nurse wants to view data about a specific patient, she can click on one of the patients on the list. When the nurse enters a ward, the system automatically reduces the list of patients to the ones in that room hence indexing to that location. By clicking on a patient's name, a detailed view appears with information about previous and upcoming tasks (figure 8). In order to enter new data into the system, the nurse has to scan a barcode on the patient's wristband. The subsequent information screen indexes to that patient.

The MobileWARD prototype was implemented for a Compaq iPAQ handheld computer connected to an IEEE 802.11b wireless network.

4.3. Just-for-Us

The third prototype system is a context-aware urban social guide, referred to as Just-for-Us, developed as part of a collaborative project between The University of Melbourne, Australia and Aalborg University, Denmark. Just-for-Us facilitates social interactions in the city of Melbourne by providing the user with a simplified digital layer of information about people, places and activities within proximity; adapted to users' physical and social context, and their history of social interactions in the city (Paay et al. 2009). Based on field studies of groups of friends socializing "out on the town", we identified key properties of the physical and social context which people used as reference points in their situated social interactions – the way they communicated and the way they made sense of the world around them. Informed by this, we designed and implemented a functional prototype, which pushed the use of indexical references to further extremes than in the previous two designs in order to gain deeper insight into the use of mobile user interfaces with this particular characteristic.

Figure 7. Just-for-Us: Indexing to the user's physical surroundings and history of visits

In the Just-for-Us prototype, indexical links were created between the information in the system and the world surrounding the user through augmented panoramic photographs pushed to the user on the basis of their location. In this way, information in the system is indexed to the user's physical context mediated by an interactive photographic representation. Interacting with this "augmented reality" type of representation the user can align information in the system with the physical world using information cues in the environment such as the shape and colour of buildings and major structures. Secondly, information content, such as recommendations of places to go for a certain activity, was reduced by tailoring it to the users current social group (who they are with at that time) and this group's shared history of socializing out on the town. Thirdly, indexical references were used to generate way-finding descriptions referring to the user's familiar paths and places, rather than coordinates, directions and distances, and to visually prominent objects and structures in the user's surroundings (figure 7).

Just-for-Us was implemented as a web service accessible through a mobile browser. For the prototype we used an HP iPAQ h5550 connected to the Internet through GPRS.

5. THREE EVALUATIONS

Because context plays a central role in the interpretation of interaction with a context-aware mobile device from the perspective of indexicality, all three prototypes described above have been studied during use in the field, and not only in laboratory settings. Below we describe the evaluations of the three prototypes involving a total of 62 users.

5.1. Evaluating on public transport: TramMate II

The TramMate II prototype system was evaluated in Melbourne, Australia in 2003. This evaluation involved 10 people using the system for 40-90 minutes. All users were familiar with mobile devices and frequent users of the public transport system. Half of the evaluations were carried out in a usability laboratory with the user seated at a desk. The other half was carried out in the field while the user was commuting around the inner city on trams (figure 8). The evaluations were structured by a series of tasks identical in the lab and in the field. During the evaluations, the users were asked to think aloud and respond to questions from an interviewer. The evaluations were recorded on digital video. In the lab, this included close-up views of the mobile device screen as well as overviews of the user and the interviewer. In the field, the cameraman shifted focus between close-up of the mobile device screen, the user and the interviewer, and overviews of the overall use situation. The TramMate II evaluation is described in detail in (Kjeldskov et al. 2005).

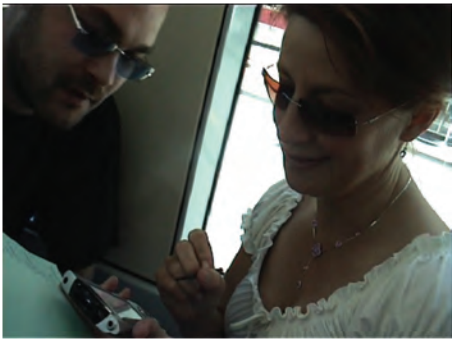


Figure 8. Field evaluation of TramMate II on board a tram in Melbourne, Australia

5.2. Evaluating at the hospital: MobileWARD

The MobileWARD prototype system was evaluated in Northern Jutland, Denmark in 2004. This evaluation involved 12 people using the system for 15-40 minutes. All users were trained nurses and familiar with the use of electronic patient records. Half of the evaluations were carried out in a usability laboratory at Aalborg University consisting of several rooms and a hallway furnished to resemble a section of a hospital ward with actors impersonating hospitalized patients (figure 9). The other half was carried out in situ at a large regional hospital in Fredrikshavn involving real work tasks and real patients. The evaluations in the laboratory were structured by a series of tasks derived from an earlier field study of work activities at the hospital. In the field, we did not enforce researcher control on the evaluations but let real world work tasks prompt use of the system. During the laboratory evaluations, the users were asked to think aloud. For ethical reasons this was not possible at all times at the hospital. Hence, interview questions were asked during times where the nurses were in the hallway and after the evaluation. In the laboratory, ceiling-mounted motorized cameras captured overviews of the nurses and "patients". Close-up views of the mobile device and user interaction were captured by a small wireless camera attached to the device. In the field, obvious ethical concerns restricted us from filming the nurses' interactions with patients. Therefore only the close-up view of the device was captured while nurses were working in the wards. The MobileWARD evaluation is described in detail in (Skov and Høegh 2006).

Figure 9. Laboratory evaluation of MobileWARD in a usability lab emulating a hospital ward

5.3. Evaluating in the city: Just-for-Us

The Just-for-Us prototype system was evaluated in Melbourne, Australia in 2005. This evaluation involved 40 people (grouped in pairs of two) interacting with the system for 45-70 minutes. Again, half of the evaluation was carried out in a usability laboratory and half of them in the field (Figure 10). All pairs of users were familiar with Federation Square and frequently socialized there together. Being primarily interested in peoples' use of the system and their response to its indexical information content, neither laboratory nor field evaluations were structured by tasks in the traditional usability evaluation sense. Instead, the evaluations were structured by a set of overall prompts for use of different parts of the system and a list of corresponding interview questions.

Figure 10. Field evaluation of Just-for-Us at Federation Square, Melbourne, Australia

Data was collected through note taking and by means of mobile audio/video equipment carried by a cameraman. A wireless camera was attached to the mobile device capturing a close-up image of the screen. This was mixed on the fly with a third-person view of the users allowing high-quality data collection as well as unobstructed user interaction (figure 11). Users and interviewer were wearing wireless directional microphones.

Figure 11. Wireless camera on PDA and video of participants, interviewer, surroundings and screen

6. DISCUSSION OF FINDINGS FROM EVALUATIONS

Below, we present and discuss some of our qualitative findings about the use of mobile computer systems with indexical interface and interaction design from the evaluations of the three prototype systems described above.

Most importantly we found that, even with a minimum of clues, people are extremely capable of making sense from small pieces of information and information implicitly present in their surroundings. They typically have no problem aligning information in the system with corresponding elements in the physical world surrounding them, including attributing names on the screen to physical places and correlating representations of activities with real work tasks and real people. The indexical references that we found to be most easily understood were those that related very directly to the users' perceived setting such as their location and the current time. Other well-functioning references were those that align with visually match-able elements in the user's surroundings such as physical structures and objects nearby.

It was also evident that people used redundant indexical interface references to corroborate their interpretation of the information provided by the systems. They used a redundancy of indexes (i.e., labels, images, signs, structures and activities of others) to make sense of the information presented to them on their mobile computer display, and as alternating strategies for matching information in the system to the world around them.

However, we also found that the use of indexical references in digital interaction is different, and represents a greater risk when getting it wrong, compared to the use of indexical references in face-to-face communication. It is much more difficult for a computer system than for a person to gauge a persons' reaction to an instruction or a piece of information, and adapt to that reaction through additional information or meta-communication if we realize that more information is needed for clarification or that we have been misunderstood. In reducing the amount of information presented in the user interface of a mobile computer user interface it becomes increasingly important that the information that remains still provides the right clues for the user to interpret it correctly on the basis of their context.

Below we present and discuss findings related specifically to indexing to physical context physical context (architectural structures and elements of the built environment), spatial context (location), and social context (presence and behaviour of other people).

6.1. Indexing to physical context

The three prototype systems all indexed to the user's physical context. They all had information content directly related to specific physical entities and structures in the users' physical surroundings, such as trams, tram stops, wards, venues, and landmarks. The indexical relationship between the information in the system (the sign) and the entity they were referring to (their object) was supported through textual-type references such as descriptors like "the next tram" or " the black building", and through iconic-type references like pictures of noticeable structures, logos, drawings, and maps.

From our evaluations we found that indexing mobile computer systems to physical context is not difficult for people to interpret and understand, and that people readily use a variety of indexes to physical context to create meaningful indexical signs out of the ensemble of user interface and use context.

When looking at the usefulness of particular types of indexical references to physical context, we observed that people were particularly good at using visually prominent outlines of their immediate surroundings, for example, the layout of a room, the shape of nearby buildings and structures, the shape of distant structures, or parts of the skyline, to align the system with their surroundings. From there they would create a meaningful indexical sign out of the information presented on the screen in their particular physical context. People also used the presence of distinct physical objects to create indexical signs. This included distinct physical objects in their immediate surrounds, for example, coloured walls, media screens, satellite dishes, tram stops, trams, beds, as well as in their distant surrounds, for example, a river, a church, a train station, and a distant tram.

Related to this, we also found that people frequently used labels and headings in the system to match up with labels and signposts in their physical surroundings in order to make meaning out of their mobile computer system. They expressed a clear expectation to be able to find such matches, and conversely they also expected clear labels and signs in the physical world to appear in the system too. The matching up of labels and signs happened not only textually but also iconicly in terms of the visual style of labels and signposts across system and context, notably that of logos.

Other types of useful indexical references to physical context were not based on support through visual similarity. Using a distinct physical quality of an object or a place as descriptor, for example, "the black building", "the glass wall" or "the old tram", was also found to be very useful reference type for the creation of meaningful indexical signs. It was also observed that using these types of indexical references allowed information content to be indexed to physical context beyond the users immediately visible physical surroundings to, for example, familiar places nearby such as landmarks, specific high-rise buildings, railway stations, tram stops, wards, offices, etc. with distinct physical features.

Indexing to physical context through descriptive references to distinct features that the user know through their familiarity with a place was found to be a valuable way of reducing information for expert or repeat users of a system, such as nurses using MobileWARD or frequent travellers in the city using TramMate. Indexing to distinct features in the physical context was also found to be of value as key anchor points in way-finding instructions for people who are new to a place, such as tourists using the Just-for-Us system. In terms of way finding, we found that such indexical references work well because they replicate the way we often give directions to others: pointing out distinct feature in the physical surroundings along a path to the given destination, allowing people to operate with greater flexibility in between those anchor points.

Redundancy and the sufficiency of approximation

On a general level, we found that when indexing content in a mobile computer system to physical context there are two things to take particular notice of: redundancy and the sufficiency of approximation.

When faced with an indexical reference there is always a risk of misinterpreting what is being referred to. From our evaluations we observed that people appear quite used to this risk and make use of a redundancy of references to support and confirm their interpretation (not dissimilar to triangulation in data analysis). As an example of redundant indexical references in our prototype systems, textual labels often complemented images and maps, which allowed people to double-check their interpretation based on one of those by testing it through the other. For example, having interpreted from a layout representation of a ward that the patient in the bed to the right is named “Marie Frandsen”, this can be confirmed by comparing this name to the label on the bed or patient and vice versa.

In the use of text-based and icon-based indexical references described above, we found that there was not a strong need for representations that matched exactly with the user’s physical context. Peoples’ responses to the use of skylines, outlines of buildings, layout of rooms and areas, and overall shape and appearance of distinct physical objects and landmarks indicates that these approximated representations are sufficient enough for meaningful indexical signs to be formed. This sufficiency of approximation was also observed for textual references to features in physical context where it was found that people apply quite broad interpretations of descriptors like “black”, “tall”, “old” etc. in order for the indexical reference to make sense. Sufficiency of approximation was also observed in relation to the text and visual appearance of labels, signage and logos.

6.2. Indexing to spatial context

The three prototype systems all indexed to the user’s spatial context. They all presented information related to the users’ absolute as well as relative location, such as what tram stop they were at, how far they were from a tram stop, which room they were in at the hospital (ward, hallway, office, etc) and which venue at Federation Square they were at. The indexical relationship between information in the system and the location it related to was usually not supported by means other than the users’ presence at a particular location and a simple label describing the system’s interpretation of this location such as “Stop 7”, “Ward 254” or “Main Plaza”.

From our evaluations we found that people generally understand when a mobile system adapts to their current location and that this type of adaptation is usually found to be useful. Our observations showed that when indexing information on the screen of a mobile computer system to the user’s spatial context, this relationship is easily understood, and the ensemble of system and context is interpreted as a joint indexical sign. We believe that this observation reflects the fact that we are very familiar with spatial indexicality through our life long experience of language and signage that relates specifically to its location, and that we are very experienced with the interpretation of such signs. Hence, as we observed, it is easily accepted that an electronic “sign” (the mobile computer screen) should be understood in a similar way through interpretation of its implicit references to the location in which it is situated.

As with indexing to physical context we found that approximation of spatial context was also sufficient when it was possible to identify notable spatial areas (e.g. trams, wards, plazas, and bars). Hence, for the specific prototype systems evaluated, using such places or areas as an approximation of current location, rather than using precise Cartesian coordinates, was found to be entirely adequate for the correct interpretation of information. In fact, approximating spatial context to specific places or areas corresponds well to the way we are used to experiencing our spatial context as human beings situated in the world. Hence making use of such approximations explores our life long experience of interpreting our own spatial context. In some situations, however, it might not be possible to make such approximations, in which case more precise location information is needed, or indexical references will have to be made open for broader interpretation.

Although people understood the indexical signs produced jointly by the mobile computer system and their spatial context, and appreciated this kind of system behaviour, we also observed that the reduction of information resulting from indexing to spatial context sometimes had the unintended effect of overly limiting the amount of information available to the user. In response to the prototype system automatically converting information into spatially indexical signs, we observed that people were sometimes not fully satisfied with this reduced subset of information and expressed a need or desire for information beyond their current spatial context: information about other trams, other wards, other venues, etc. This finding led us to conclude that although spatial indexicality can be a powerful means of reducing information on mobile computer systems through the creation of indexical signs, the fact that this particular “sign” is also interactive and networked comes with a inherent set of user expectations about information being available any time and anywhere – that is, independent of context.

Trust and control

Taking a step back, we found two things to be particularly aware of when indexing content on a mobile computer system to spatial context: trust and control.

In response to the system knowing the users’ location, we observed the unexpected side effect that people perceived the system’s information content as true. For the purpose of the evaluations of MobileWARD and Just-for-Us this was actually not always the case because of ethical and copyright-related issues. Nevertheless, for example, some people actually pursued ordering from the made up menu of a café presented to them by the system at that place, trusting that the content of the spatially indexical sign was indeed true. The reverse effect was, however, just as strong when the system got the user’s location wrong and therefore appeared “unpredictable”, as discussed by Cheverst et al. (2000). Such loss of trust was observed when, for example, TramMate II displayed a wrong stop number at a critical point of the journey, and when MobileWARD displayed information related to a ward when the user was actually in the corridor. These and similar behaviours were caused by technical bugs, but the effect on the user experience brings to attention the importance of spatially indexical systems being robust in their ability to sense and adapt correctly to their spatial context, as also discussed by Schmidt-Belz (2003).

Somewhat related to the issue of trust is the issue of control. While people understood that the systems adapted to location, they were sometimes uncertain about how to then control the system. The observed issues of control related to situations where the users wanted to stop the system from automatically pushing new content due to a change of location. This happened either because the users still needed the information automatically presented to them at their previous location, or because they had manually navigated to a piece of information that they wanted to keep ready at hand. In both cases, the systems sometimes took this level of control away from the user. Giving control back to the user could be done through ways of stopping or pausing automatic updates, making newly pushed information appear without completely taking away what was already there, or, at the very least, by giving the user an option for browsing backwards at any point in time (or place).

6.3. Indexing to social context

MobileWARD and Just-for-Us both contained indexical references to social context. They presented information related to the presence and activity or state of other people in the user's surroundings such as patients, friends, and groups in vicinity. The indexical relationship between information in the system and the user's social context was supported only by simple textual labels like "Marie Frandsen" in MobileWARD, or through "social activity meters" depicting the amount of people at different places in Just-for-Us, revealing the system's simplified interpretation of the user's social context.

From our evaluations we found that indexing mobile computer systems to social context can be more difficult for people to understand than when indexing to physical or spatial context. People often had difficulties interpreting the presented information as a meaningful indexical sign when it was indexed to their social context. In order to interpret this type of indexical reference, they often needed a more detailed explanation about what the system was doing and what the system knew about its context. Once having learned how a system adapted to the user's social context, this interpretation improved.

Relating back to the use of descriptors in indexing to physical context, we found that for social context using descriptors based on social activities, such as "the sitting steps" was not found to be useful. This is because they relied on ephemeral conditions surrounding those places and objects, and people were concerned they were too unclear and open for wrong interpretations when, for example, nobody was actually sitting on those steps. This was surprising, because most people knew which steps were meant based on their experience of the place and on the physical affordances of those steps.

One of the ways where indexing to social context did work well was in relation to way-finding descriptions provided to a social group rather than to an individual. Here we found that the use of references to shared familiar places and shared past social visits were useful. Instructions were related to the groups' joint memory and knowledge of an area, and also anchored naturally in to their unique shared history and patterns of social interactions. Similarly to the way rhythms of work activities over time were observed to facilitate information seeking by Reddy and Dourish (2002), rhythms of social activities over time also seemed to facilitate information interpretation, thus linking social context closely to activity and temporal context. We believe that this observation reflects the fact that this is, again, how we are used to making use of social context to reduce complexity in face-to-face interaction: describing the location of places with reference to other places that we know that person is familiar with or that we both know through prior shared experiences. For example, meeting at "the place we met for dinner last time".

Rhythms of activities are integrated parts of our everyday life and manifest themselves in many ways (Zerubavel 1979, 1985). This pervasiveness of rhythms makes them a compelling focus for the development of information tools (Reddy and Dourish 2002, Begole and Tang 2007, Bellotti et al. 2008) because people have a strong and shared sense of temporal patterns of activities and use these to coordinate, form expectations etc. From the perspective of context-aware mobile systems, rhythms of social activities over time could in themselves very well constitute an important, derivative, dimension of context.

On a side note to this, once knowing how a system made use of peoples' history and rhythm of social interactions, many people expressed concerns and uncertainty about how to control this system behaviour in relation to issues of privacy.

Subtle context and making the implicit explicit

In looking at peoples' use of indexes to social context we found the following two things particularly important to consider: subtle context and making the implicit explicit.

As described above, we found that socially contextual factors indexed to by the system were much less obvious to people than their physical and spatial context. Thus peoples' interpretation of information on the screen often failed to take those subtle social context factors into consideration. People sometimes simply didn't expect, or understand very well, that the systems knew about their current social setting and was capable of adapting and indexing its content to this context. Missing the subtle clues of social context was mostly obvious in the evaluation of Just-for-Us, which was designed specifically to facilitate social interactions. This system had access to socially contextual information like whom you were with at the time, and your friends' and your individual as well as shared history and rhythm of social interactions. It then generated ranked suggestions for where to go based on patterns in the current social group's shared history and rhythm. In the use of this particular, and quite advanced, functionality we observed that people completely failed to interpret the indexical reference to their social context. Consequently, the information held no meaning for them or was misinterpreted in different ways (i.e. vendors paying for rankings). We believe that this observation reflects a fundamental difference between social and physical/spatial context. As social context is not only about whom you are with, but also very much about your history and rhythm of social interactions with this group of people, social context is not only often implicit but also largely invisible and something that is peripheral to us. This makes social context harder to index to in a computer system, and it makes it harder to interpret a socially indexical sign correctly.

One thing that we found did work very well in terms of indexing to social context in our evaluations of both MobileWARD and Just-for-Us was representing social context information. Specifically, we found that people like to get an overview of their social context such as the presence and activities of other people in the surrounding environment. This information was presented in different ways in the MobileWARD and Just-for-Us system, but in common for both they provided not only new and valuable information in themselves, but also objectified social context, which could then be indexed to more successfully. In terms of the limitations of subtle factors of social context in the creation of meaningful indexical signs, representing social context in this way increases the potential for making interpretable indexical references to social context by taking something implicit and invisible and making it explicit and visible.

7. USING INDEXICALITY IN DESIGN

How can we use indexicality in the design of mobile and context-aware systems? In response to this challenge it is important to note that indexicality is not a design tool or method. It is purely a concept that can be used to describe and understand an aspect of a design. This is not unique to the concept of indexicality though. Exactly the same can be said for established principles within human-computer interaction, such as mapping, affordances, the Gestalt principles, and so on. These are theoretically grounded principles that can be used to describe features of a design. By understanding such principles they can be used to inform the design process. Doing the latter is perhaps the hard part though. How do we transcend from the retrospective activity of analysis through a certain theoretical lens to the proactive activity of designing through it? In line with Alexander's (1964, 1977) views on the activity of design, we believe that good design requires a solid understanding of its context and of the principles that previous solutions have shown can be successfully applied. Interaction design for mobile devices involves several such principles. Some are related to optimising limited screen real estate and some are related to the use input devices on mobiles. The principle of indexicality would relate to the interplay between user, system and context.

In Alexander's own work, such principles take the form of design patterns (Alexander et al. 1977) each exemplifying design challenge, theoretical understanding, and possible solutions. This makes them particularly accessible and useful for designers. They are grounded in massive empirical evidence and solid understanding, but provide guidance for design that is specific enough to inspire solutions while general enough not to prescribe them. One of the things that make established design principles within human-computer interaction useful in design is that, similar to Alexander's design patterns in architecture, a body of empirically grounded examples have evolved in their support. This makes underlying theories and concepts (e.g. cognition and perception) much more practically accessible, and hence those theories are in effect being used more to inform design. Such patterns and examples are yet to evolve for indexicality as a principle for mobile interaction design and would support the process of designing systems on the basis of this concept greatly. Developing such design patterns would involve analysing and describing indexical properties of other successful existing context-aware systems apart from the few ones discussed here.

Apart from understanding the indexical interplay between users, systems, and context, and having access to patterns of indexical design solutions in other systems, using indexicality in design requires knowledge about what specific elements in the users context they can be indexed to for the system being developed. This requires the identification of indexable attributes of the context during the projects' analysis phase. Our own work includes structured mappings of physical, spatial and social context using a multidisciplinary socio-physical approach (Paay et al. 2009), and illustrates one possible process to follow. Other processes may involve more stringent techniques for identifying the contextual information that a mobile system might index to.

8. SUMMARY AND CONCLUSIONS

This article has promoted the concept of indexicality as a theoretical concept for describing and understanding the relationship between user interface representations and user context for mobile human-computer interaction. We have argued that the lack of a theoretically grounded body of knowledge that explains this relationship is limiting our ability to elevate learning from the mobile systems we develop and study in use from a concrete to an abstract level. Consequently, the research field is impeded in its ability to leap forward beyond the pace of, at best, incremental steps from one design to another. In response to this lack of theoretically grounded knowledge, we have explored the semiotic concept of indexicality as an analytical concept that can be used to explain the user experience of a specific design in context. We have illustrated the analytical power of this concept through the analysis of a mobile interaction design concept, and through the analysis of empirical data from three studies of context-aware mobile computer systems in use; TramMate II, MobileWARD, and Just-for-Us.

Our findings show that by applying the lens of indexicality, new and theoretically grounded knowledge can be generated from empirical data about mobile human-computer interaction in context. We have found that even with a minimum of clues, people are extremely capable of making sense from small pieces of information in a user interface if they can be meaningfully indexed to their surroundings. People interpret mobile computer systems in context as joint indexical "signs" carrying their meaning through the ensemble of implicit context and explicit interface representations. In the design of such interfaces, this indexical interpretation allows the amount of information explicitly presented to the user to be reduced. This is particularly valuable when designing for systems with small graphical user interfaces, such as handheld computers, and for situations where users have limited or divided attention towards the system, such as most mobile use contexts.

The indexical references that we found to be most easily understood were those that related the users' objectively perceivable settings such as their location and the current time. Other well-functioning indexical references were those that related to visually match-able elements in the user's surrounding, such as prominent physical structures and objects nearby. Indexes to social context were found to be more difficult for users to interpret correctly, and we speculate that this is caused by the intangible and peripheral nature of this type of context compared to location, surroundings, activity and time.

In terms of indexing to physical context, we conclude that this is not difficult for people to interpret and understand, and that people use redundancy of indexes to physical context to create meaningful indexical signs and that they double check their interpretation of one against another. We also conclude that there is a sufficiency of approximation associated with representations that index to physical context through iconic and symbolic references.

In terms of indexing to spatial context, we conclude that this is easily understood, and that the ensemble of user interface and user context is interpreted as one joint indexical sign. In relation to this, we have highlighted the potential impact of spatial indexicality on the users' experience of control over what the system is doing, and their experience of trust in the content that is provided when a system adapts correctly to its location.

In terms of indexing to social context, we conclude that this is more difficult for people to understand, than when indexing to physical or spatial context. Social context is subtle and often invisible and implicit, and in order to understand socially indexical references, people need to be made aware about what the system knows about their social context, and what aspects of it are being indexed to. We describe this as making the implicit explicit.

Inspiring further research, the findings from the three studies of use in context discussed above also revealed a series of challenges for indexical interaction design for mobile computer systems. In relation to the issues of control and trust, people rightfully raise issues about their privacy when faced with a system that indexes to their current and history of spatial and social context. In order to make systems spatially and socially indexical, it is important that the users trust them enough to allow collection and reference to this information. One of the central components in the creation of such trust is the availability of transparent means of user control.

It also appears that different people and different situations require different levels of indexicality, and that there is no such thing as universally appropriate indexical references when it comes to complex digital signs such as interactive mobile computer systems in context. Using redundant indexical references allows some level of flexibility in interpretation, but as we are dealing with interactive signs here, it would be interesting to explore the possibility of developing a mechanism allowing the user to manually adjust the level of indexicality in the interface: reducing or increasing the "strength" of implicitness and consequently increasing or reducing the amount of explicitness.

9. FURTHER WORK

In terms of realising indexical interface design in practice, there are two particular things that we find need additional work. Firstly, for system developers and interaction designers to be able to index to elements in the users context, a solid understanding of the indexable attributes of a specific environment needs to be gathered during the projects' analysis phase. Our work within this area includes making structured mappings of physical, spatial and social context in a particular place using a interdisciplinary socio-physical approach (Paay et al. 2009). However, this work is not complete and needs to be extended further. Secondly, designing explicitly with the concept of indexicality in mind is, like any other concept is likely to benefit from additional support in the form of design heuristics, guidelines or patterns outlining challenges and generally well functioning design solutions. However, the creation of such heuristics, guidelines or patterns rely on the cumulative formation of a body of knowledge about design challenges and corresponding indexical design solutions. Here, we have described what we have learned from three specific systems through the theoretical lens of indexicality. More studies of mobile human-computer interaction in context are needed through the same theoretical lens in order to create general guidelines for indexical interaction design.

Further research also needs to extend the range of contextual factors indexed to, for example, the aspects of context related to activity, time and other information. This could also include a systematic decomposition of the different aspects of context and related sources of information that a system might provide an index to.

Finally, the generalizeability of the analytical power of the concept of indexicality for describing and explaining the user experience of mobile systems in context should be investigated beyond the three prototype systems discussed here. As a starting point, it would be interesting to look at other successful context-aware systems through the lens of indexicality.

ACKNOWLEDGEMENTS

This research is supported by the Danish Technical Research Council (26-04-0026) and The Smart Internet Technology CRC, Australia. The authors thank everyone participating in the design and evaluations of the prototype systems – especially Mikael B. Skov, Rune T. Høegh, Connor Graham, Sonja Pedell, Jessica Davies, Steve Howard, Frank Vetere and Bharat Dave. We would also like to thank Peter Bøgh Andersen who inspired our thinking about indexical interface design for mobile devices. Finally, we thank the anonymous reviewers, and associate editor Susanne Bødker, for constructive comments and suggestions that helped us revise and improve the paper.

REFERENCES

Abowd, G.D. and Mynatt E.D. 2000. Charting Past, Present and Future Research in Ubiquitous Computing. ACM TOCHI, 7(1), 29-58.

Agre, P. 2001. Changing Places: Contexts of Awareness in Computing. Human-Computer Interaction 16, 177-192.

Alexander, C. 1964. Notes on the Synthesis of Form. Harvard University Press, London.

Alexander, C., Ishikawa, S., Silverstein, M., with Jacobson, M., Fiksdahl-King, I., and Angel, S. 1977. A Pattern Language: Towns, Buildings, Construction. Oxford University Press, New York.

Andersen, P. B. 2002. Pervasive Computing and Space. In: L. Kecheng, R. J. Clarke, P. B. Andersen, R. Stamper and A. El-Sayed, Eds. Organisational Semiotics: evolving a science of information systems, IFIP TC8/WG8.1, Montreal, Quebec: Kluwer,133-152.

Bardram, J. E. 2009. Activity-based computing for medical work in hospitals. ACM Transactions of Computer-Human Interaction 16(2): 1-36.

Bardram, J. E. and Nørskov, N. 2008. A context-aware patient safety system for the operating room. In Proceedings of UbiComp'08. ACM, 272-281.

Bardram, J. E., Hansen, T. R., and Soegaard, M. 2006. AwareMedia: a shared interactive display supporting social, temporal, and spatial awareness in surgery. In Proceedings of CSCW'06. ACM, 109-118.

Bardram, J. E. and Bossen, C. 2005. A web of coordinative artifacts: collaborative work at a hospital ward. In Proceedings of GROUP'05. ACM, 168-176.

Barkhuus, L. and Dey, A. K. 2003. Is Context-Awareness Computing Taking Control away from the User? Three Levels of Interactivity Examined. In Proceedings of UbiComp 2003, Springer-Verlag, 149-156.

Begole, J. B. and Tang, J. C. 2007. Incorporating human and machine interpretation of unavailability and rhythm awareness into the design of collaborative applications. Human-Computer Interaction 22, 1, Lawrence Erlbaum Associates, 7-45.

Bellotti, V., Begole, B., Chi, E. H., Ducheneaut, N., Fang, J., Isaacs, E., King, T., Newman, M. W., Partridge, K., Price, B., Rasmussen, P., Roberts, M., Schiano, D. J., and Walendowski, A. 2008. Activity-based serendipitous recommendations with the Magitti mobile leisure guide. In Proceeding of CHI '08. ACM, 1157-1166.

Bradley, N. A. and Dunlop, M. D. 2002. Understanding Contextual Interactions to Design Navigational Context-Aware Applications. In Proceedings of Mobile HCI 2002, Springer Verlag, 349-353.

Brown, B. and Randell, R. 2004. Building a context sensitive telephone: some hopes and pitfalls for context sensitive computing. Computer Supported Cooperative Work 13, 329-345.

Chalmers, M. 2004. A historical view of context. Computer Supported Cooperative Work 13, 223-247.

Chandler, D. 2002. Semiotics: The Basics. London: Routeledge.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C. 2000. Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. In Proceedings of CHI 2000, ACM, 17-24.

Cheverst, K., Davies, N., Mitchell, K. and Efstratiou, C. 2001. Using Context as a Crystal Ball: Rewards and Pitfalls. Personal and Ubiquitous Computing 5, 1, 8-11.

Crabtree, B. and Rhodes, B. 1998 Wearable Computing and the Remembrance Agent. BT Technology Journal, 16, 3, 118-124.

Dey, A. K. and Abowd, G. D. 2000. Towards a Better Understanding of Context and Context-Awareness. In Proceedings of CHI 2000 Workshop on The What, Who, Where, When, and How of Context-Awareness, ACM.

Dey, A.K. 2001. Understanding and Using Context. Personal and Ubiquitous Computing 5, 1, 4-7.

Dix, A., Rodden, T., Davies, N., Trevor, J., Friday, A. and Palfreyman, K. 2000. Exploiting Space and Location as a Design Framework for Interactive Mobile Systems. ACM Transactions on Computer-Human Interaction 7, 3, 285-321.

Dourish, P. 2004. What we talk about when we talk about context. Personal and Ubiquitous Computing 8, 1, 19-30.

Dourish, P. 2001. Seeking a Foundation for Context-Aware Computing. Human-Computer Interaction 16, 229-241.

Gaver, W., Dunne, A. and Pacenti, E. 1999. Projected Realities: Conceptual Design for Cultural Effect. In Proceedings of CHI'99, ACM, 600-607.

Hinckley, K., Pierce, J., Horvitz, E, and Sinclair, M. 2005. Foreground and background interaction with sensor-enhanced mobile devices. ACM Transactions of Computer-Human Interaction 12, 1, 31-52.

Jones, Q., Grandhi, S.A., Terveen, L. and Whittaker, S. 2004. People-to-People-to-Geographical-Places: The P3 Framework for Location-Based Community Systems. Computer Supported Cooperative Work 13, 249-282.

Kaasinen E. 2003. User Needs for Location-Aware Mobile Services. Personal and Ubiquitous Computing 7, 1, 70-79.

Kjeldskov, J. and Skov, M.B. 2007. Exploring Context-Awareness for Ubiquitous Computing in the Healthcare Domain. Personal and Ubiquitous Computing 11, 7, 549-562.

Kjeldskov, J. and Paay, J. 2006. Indexical Interaction Design for Context-Aware Mobile Computer Systems. In Proceedings of OzCHI 2006, Sydney, Australia, ACM and CHISIG, 71-79.

Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S. and Davies, J. 2005. Evaluating the Usability of a Mobile Guide: The influence of Location, Participants and Resources. Behaviour and Information Technology 24, 1, 51-65.

Kjeldskov J., Howard S., Murphy J., Carroll J., Vetere F. and Graham C. 2003. Designing TramMate - a context aware mobile system supporting use of public transportation. In Proceedings of DUX 2003. San Francisco, CA, USA. AIGA & ACM,1-4.

Kjeldskov, J. 2002 Just-In-Place Information for Mobile Device Interfaces. In Proceedings of Mobile HCI'02, Pisa, Italy, Springer-Verlag, 271-275.

Lamming, M. G. and Newman, W. M. 1992. Activity-based Information Retrieval: Technology in Support of Personal Memory. In Proceedings of the IFIP 12th World Computer Congress on Personal Computers and intelligent Systems - Information Processing '92. North-Holland Publishing Co., 68-81.

Martinovski, B. 1995. Shifting worlds or deitic signs in www. http://www.ling.gu.se/~biljana/web.html. (accessed October 2009).

McCullough, M. 2004. Digital Ground - Architecture, Pervasive Computing, and Environmental Knowing. The MIT Press, Cambridge, MA.

Paay, J., Kjeldskov, J., Howard, S, and Dave, B. 2009. Out on the town: a socio-physical approach to the design of a context aware urban guide. Transactions on Computer-Human Interaction 16, 2, 7-34.

Peirce, C. S. 1931-58. Collected Writings (8 Vols.). In C. Hartshorne, P. Weiss, and A. Burks, Eds. Collected Papers of Charles Sanders Peirce. Cambridge MA: Harvard University Press.

Rantanen, J., Impio, J., Karinsalo, T., Reho, A., Tasanen, M. and Vanhala, J. 2002 Smart Clothing Prototype for the Artic Environment. Personal and Ubiquitous Computing 6, 1, 3-16.

Skov, M. B. and Høegh, R. T. 2006. Supporting Information Access in a Hospital Ward by a Context-Aware Mobile Electronic Patient Record. Personal and Ubiquitous Computing 10, 4, 205-214.

Schmidt, A., Gross, T., and Billinghurst, M. 2004. Introduction to special issue on context-aware computing in CSCW. Computer Supported Cooperative Work 13, 221-222.

Schmidt, A., Beigl, M., and Gellersen, H. 1999. There is more to Context than Location. Computers and Graphics Journal 23, 6, 893-902.

Schmidt-Belz, B. 2003. Aspects of user trust in mobile guides. Proceedings of Workshop on HCI in mobile guides, Mobile HCI 2003, Udine, Italy. Available from http://www.mguides.info/ (accessed May 2010)

Schilit, B, and Theimer, M. 1994. Disseminating active map information to mobile hosts. IEEE Network 88, 22-32.

Scollon, R., and Scollon, S. 2003. Discourses in Place - Language in the Material World. London: Routeledge.

Smith, J., Mackaness, W., Kealy, A. and Williamson, I.P. 2004. Spatial data infrastructure requirements for mobile location based journey planning. Transactions in GIS 8, 1.

Zerubavel, E. 1979. Patterns of time in hospital life: a sociological perspective. Chicago: University of Chicago Press.

Zerubavel, E. 1985. Hidden rhythms: schedules and calendars in social life. Berkeley: University of California Press.

Chapter 22

Proxemics and interactional spaces

Kenton O'Hara, Jesper Kjeldskov and Jeni Paay

Abstract. In recent years there has been an introduction of sophisticated new video conferencing technologies (e.g. HP Halo, Cisco Telepresence) that have led to enhancements in the collaborative user experience over traditional video conferencing technologies. Traditional video conferencing set-ups often distort the shared spatial properties of action and communication due to screen and camera orientation disparities and other asymmetries. These distortions affect access to the common resources used to mutually organize action and communication. By contrast new systems, such as Halo, are physically configured to reduce these asymmetries and orientation disparities, thereby minimizing these spatial distortions. By creating appropriate shared spatial geometries, the distributed spaces become "blended" where the spatial geometries of the local space continue coherently across the distributed boundary into the remote site providing the illusion of a single unified space. Drawing on theories of embodied action and workplace design we discuss the importance of this geometric "blending" of space for distributed collaboration and how this is achieved in systems such as Halo. We then extend these arguments to explore the concept of Blended Interaction Spaces - blended spaces in which interactive groupware is incorporated in ways spatially consistent with the physical geometries of the video mediated setup. We illustrate this discussion through a system called BISi that introduces interactive horizontal and vertical multipoint surfaces into a blended video mediated collaboration space. In presenting this system, we highlight some of the particular challenges of creating these systems arising from the spatial consequences of different interaction mechanisms (e.g. direct touch or remote control) and how they affect movement and spatial configuration of people in these spaces.

Originally published as O'Hara, K., Kjeldskov, J. and Paay, J. (2011) Blended Interaction Spaces for Distributed Team Collaboration. *ACM Transactions on Computer-Human Interaction,* 18(1), article 3.

1. INTRODUCTION

Over the last couple of decades, we have seen continual shifts in the way organizations are structured and the way that work gets done. As organizations operate more within a global environment this has created a greater imperative for work practices to operate more and more over distance. Within this climate, teams of workers are no longer assembled to just work in collocated settings. The adoption of a myriad of computer-mediated communication technologies have enabled teams to be assembled according to the appropriate expertise required wherever it is located. The issue here is not simply about enabling a distributed set of individuals to work together, which arguably are supported by a range of conventional desktop collaboration products such as video chat, NetMeeting, Google Docs and so forth. Rather the concern is also about how to enable multiple collocated teams situated at different locations to work more effectively together when collaborating across a distance.

In spite of the progress made in collaborative technologies, the experience of distributed teamwork remains a difficult and frustrating one. Travel remains an important component in any effective operation of such teams in order to enable face-to-face interaction. Aside from significant environmental impacts, this need to travel is also costly and time consuming, creating intermittent collaboration rather than the more fluid, regular and serendipitous interaction that characterizes collocated teamwork.

Traditional video conferencing technologies sought to overcome some of these collaborative difficulties but it is well documented that the rhetoric behind such systems never quite matched the reality of collaborative experience with such systems. There has been much research over the last 20 years that sought to understand and explain why this was the case (e.g. Buxton 2009, Dourish et al. 1996, Finn et al. 1997, Gaver et al. 1993, Harrison 2009, Heath and Luff 1991, Heath and Luff 1992, Hirsh et al. 2005, Mantei et al. 1991, Nguyen and Canny 2005, Noll 1992, O'Connaill et al. 1993, Olson et al. 1997, Olson and Olson 2000, Sellen 1992, Sellen et al. 1992, Sellen 1995, Sellen and Harper 1997, Short et al. 1976, Tang and Isaacs 1993). As such, while these systems support some limited communication activities within distributed teams, they often remain under utilized for collaborative activities of any realistic complexity. A recent study conducted in one particular organization has indicated that traditional video conferencing systems are used on average for only 12 hours per month (Weinstein 2005). This level of usage matches well with the research findings in the literature relating to user experience and organizational factors that often hinder use.

In recent years though, we have seen the introduction of more sophisticated video conferencing technologies that have led to a stepwise increase in the collaborative experience between remote team members. Systems such as HP Halo, Cisco Telepresence, Tandberg T3, and Polycom TPX (see Figure 1) provide an enhanced user experience with research showing usage rates of such systems as much as ten times higher than traditional video conference systems (Weinstein 2005). Our intention here is not to overplay the significance of these comparative statistics, as there are a number of factors underpinning them. For example, the fact there are more of the traditional video conferencing units relative to the higher end systems is a contributing factor. But in part they are also attributable to the enhanced user experience associated with these higher end systems with users saying the technology "disappears" enabling them to focus on the collaboration (Gorzynski et al. 2009; Weinstein 2005). The experience and usage rates with these high-end systems, then, appear to challenge some of the well-established difficulties with traditional video conferencing systems.

Figure 1. HP Halo and Tandberg Telepresence T3

Articulating the reasons for why these systems challenge some of our assumptions about the value of video mediated communication is one of the concerns of this paper. While there are a number of possibilities for how we may approach this, the way we aim to do so in this paper is through a closer look at the design characteristics of these systems using the Halo environment as an example, drawing on and extending the work of Gorzynski et al. (2009). Our concerns here are more than simply a characterization and explanation of the Halo system design. Rather, we want to use this characterization to exemplify the broader design philosophy behind such systems. In particular, we want to articulate further what Gorzynski et al. have come to call "Blended Spaces"[2]. That is, distributed set-ups in which the design of the whole environment produces a geometrically coherent representation of the remote site, faithfully representing its spatial geometries with respect to the local site. This provides the perception of unified spatial frame of reference for all parties.

Taking this as a starting point, we want to argue how the basic blended space philosophy can be extended to think about an ecology of other distributed work place environments that support a broader variety of distributed collaborative work practices. In particular, one of the key aims of the paper is to extend the Blended Spaces work to see how collaborative practices can further be supported by the introduction of interactive information workspace elements into such environments. That is we will move towards what we have come to call "Blended Interaction Spaces". Our concerns here are to illustrate some of the design challenges of making such a move and the ways that different interaction mechanisms and approaches can affect the way that we achieve geometrically coherent representation of the remote site with respect to the local site. As with the characterization of the Halo system we will adopt a similar approach of articulating the characteristics of Blended Interaction Spaces through a system we have developed for data intensive collaboration for distributed small groups.

2) While the work of Gorzynski et al. 2009 articulate characteristics of blended spaces, they do not actually make explicit reference to the term in this article. The term, though, is attributable to Gorzynski and his Halo colleagues but one that has been shared with the current authors through personal communication while working with the Halo team at HP.

Before we delve into a deeper articulation of Blended Spaces and Blended Interaction Spaces, we want to first take a step back and set some context within which to ground some of the subsequent discussion. We discuss first of all, notions of *embodied action* and the importance of physical space as the basis for coordinated action, meaning making and intersubjective understanding. We see how the notion of embodied action is used to make sense of the communication and collaboration behaviors arising in media spaces and traditional video mediated communication and through this set the scene for understanding some of the design directions of Blended Spaces. Building on the ideas of how action is enacted embodied in space, we move on to a discussion of the ways that our social action and behaviour are influenced by our spatial configuration with respect to others and objects. This discussion draws in particular on Hall's (1966) notions of *proxemics* and Kendon's theory of *F-formations* of interactional space. We go on then to discuss workplace design where again, notions of physical space are fundamental. Like CSCW, workplace design has as one of its central concerns the way effective workplace collaboration is achieved. In contrast to much of CSCW, workplace design approaches this from the point of view of designing physical space, from architectural elements through to specifics of furniture design. We then articulate where CSCW and media space research have drawn on notions of spatiality in the design of distributed environments. This leads on to our discussion of Blended Spaces and Blended Interaction Spaces.

2. EMBODIED ACTION AND MEDIA SPACES

One of the key ways that CSCW has concerned itself with notions of space is through the notions of embodied action (e.g. Dourish 2001; Robertson 1997). Drawing on the philosophical foundations of phenomenologist thinkers such as Husserl, Heidegger and Merleau-Ponty, these authors take as their starting point that consciousness and perception are active, interpretive and embodied, arising from our presence and action in the world. It is through our actions that we are able to create shared meanings with other people. Robertson (1997) argues that of particular importance is Merleau-Ponty's notion of *reversibility* (Merleau-Ponty 1962, 1968). Quoting Robertson (1997):

> "Reversibility is the complex, reciprocal insertion and intertwining of the sensed and the sensing, that is the essential condition of our interaction with the world and with others. In a shared physical space a lived body can simultaneously see and be seen, touch and be touched, make sounds and be heard, move and reorient its perspective and cross over these sensory modes; that is, see both itself and others being touched or touching, moving, making sounds etc. The fact that we are able to perceive our own bodily surfaces at the same time as we live our acting bodies enables us to organise our actions. The public availability of these actions to the perceptions of others enables them to organise their own actions in relation to ours."

The key issue here is that coordinated action, meaning making and intersubjective understanding are shaped, in part, from our embodied actions in space and the availability of these actions to others and the availability of others' actions to ourselves, for example, the way we move, point, touch and gesture in relation to objects and other people in that physical space. Thinking about coordinated action and meaning in this way was an important foundation for much of the early analytic enquiries into media spaces and video-mediated communication. In particular, of interest were the ways video mediated communication distorted some of the essential shared spatial properties of action and communication in the form of various asymmetries (Dourish 2001, Gaver et al. 1993, Heath and Luff 1992, 2000). These asymmetries have been argued to stem from orientation disparities arising from particular camera and screen configurations and affect access to the common resources used to mutually organize action and communication. As Heath and Luff (2000) articulate:

> "… gesture and other forms of bodily activity are systematically designed with respect to the local environment and the emergent conduct within the interaction. In video co-presence, mediated through audio-visual technology, the camera and the monitor inevitably delimit and distort your access to the co-participants. Your view of the other is from a particular angle and severely circumscribes access not only to the other and their bodily conduct but to the local environment in which it is produced. In consequence your ability to design a bodily movement such as a gesture which is sensitive to orientation to the other and their relationship to their own environment is problematic. Moreover, the limited access to the other also means that you are relatively unaware of changes within their local environment with which your visual conduct may well be competing; for example, changes to the content of their computer screen or their workstation, or even people entering the room…The technology therefore provides physically distributed individuals with incongruent environments for interaction. What I see is not what you see, and I am unable to see how you see me and the actions in which I engage. Despite this incongruity, individuals presuppose the effectiveness of their conduct and assume that their frame of reference is 'parallel' with their co-participant's. They presuppose, for the practicalities at hand and their mutually coordinated activity, that their image of the other is congruent with the image of them… This presupposition of a common frame of reference, a reciprocity of perspectives, is a foundation of socially organized conduct." (Heath and Luff 2000, p.198)

Within traditional video mediated communication the lack of common spatial reference points with which coordinated action and meaning can be facilitated resulted in well-documented difficulties with certain tasks in these environments or the need to adapt behavioral practices within these new spatial contexts. Pointing and looking at objects, orienting towards other people within these video mediated spaces were all things that needed to be consciously reinterpreted rather than "natural" forms of interaction.

3. SPATIALITY AND HUMAN INTERACTION

As well as the argument that action is embodied and enacted in physical space, it is important too to understand the ways that our social action and behaviour are influenced by our spatial configuration with respect to others and objects. These understandings provide important foundations for the blended space arguments, but are informative too when thinking in more detail about the specifics of their design. A key contribution to this issue can be found in the work of Ed Hall and his notion of *proxemics* (Hall 1966) - man's use of space as a specialized elaboration of culture. One of the key themes within Hall's work concerns the notion of physical distances that people maintain between each other according to their relationship and type of interaction. He characterises four main spatial distances that exist around a person, *intimate* distance, *personal* distance, *social* distance and *public* distance. Intimate distance is somewhere in the range of 0 to 45cm and is reserved to interactions with lovers, family and close friends. Personal distance is in the range of 45 cm to 1.2m and is the distance that we naturally would maintain from strangers in everyday life. Social distance is between 1.2m and 3.6m, getting gradually more formal towards the further end of the scale. Hall argues that it is in the social range that work and business meetings typically occur. The public distance is anything beyond 1.2m where one loses any sense of personal involvement with the other actor, for example, a person giving a speech in a conference hall. These zones then are of significance for how we conduct interaction and can lead to psychological discomfort if they are broken. This is not simply arbitrary but in part relates to the opportunities for certain types of actions that are possible in these different spaces. For example, what can be seen and taken in with a visual sweep of a certain angle is dependent upon the distance from the subject. This, in turn, affects the patterns of gaze and eye movements that are made and interpreted as acceptable at particular distances. Hall also discusses how height differences in these zones can subtly alter the nature of the interaction and interpersonal relationships. Consequently, to stand and look down at someone at in the social zone can confer a particular advantage over the person sitting. While the interpersonal distances are the most familiar aspects of Halls' work, his arguments are much broader and concern more general ways in which human interaction is spatially organised. Indeed, his discussion highlights the significance of different spatial configurations of interacting parties in relation to communication characteristics. For example, sitting next to someone at a table vs. sitting across the table from them can influence the communication dynamics between dyads.

An understanding of these different spatial configurations of interacting parties is developed further in the seminal work of Kendon (1990), in particular in his articulation of the F-formation, which is essentially the "spatial and orientational organisation of participants". The transactional segments (the area in front of each body) overlap to form the o-space, the shared interactional space to which the participants have mutual access. Participants continually arrange themselves throughout an interaction in a way that maintains the mutual access to the space. There are a number of points of significance that relate to the spatial organisation of interaction. The first is that different f-formations alter the nature of conversation and interaction. Even different positions within particular f-formations can also have an impact on the conversational relationship between different parties (e.g. Sommer and Ross 1958; Sommer 1969; Steinzor 1950). Second, studies show that gestures (including body orientation and gaze) are very much impacted by and dependent upon the spatial arrangement of bodies in the formation (e.g. Ozyurek 2000, 2002). Healey and Battersby (2009), for example, argue that the mutually-known arrangement of participants, gestures and orientation are used to create what they call interactional maps of conversational contributions. These interactional geometries support a number of things such as references to locations as representative of prior turns, as representative of the turns' producers and as ways of keeping sub dialogues visually distinct from one another. In this respect, physical reference space is used not just to gesture to actual objects and artefacts in the shared space but is also used to organise gestural reference to more abstract concepts – for example "earlier" or "later" might be represented in spatial terms such as left and right.

Spatial features of the world and the way things are organised are also laden with social meaning and protocols that people make use of in their everyday life to interpret what is going on and how to behave (Buxton 2009). For example, as Buxton suggests, sitting at the head of the table confers an understood social status that can be understood; or two people walking in the park can be interpreted as friends or lovers depending on their observed interpersonal distances mentioned above. These space-function-distance relationships are important to reflect upon in design.

4. PHYSICAL WORKPLACE DESIGN

While CSCW has as its core focus the impact of technology design on collaborative work practices, other fields of research and design have oriented more closely to the relationship between spatiality and human interaction and the impact of space-function-distance design on effective collaborative practices (e.g. Becker and Steele 1995; Duffy 1997; Laing et al. 1998; Steelcase360 2007). Authors such as these and others in the workplace design and facilities management literatures essentially argue that the design of different physical spaces affects the social and informational aspects of an organization and its work practices. Different physical spaces are designed to support a range of different individual and collaborative work practices and the fluid movement between these (Laing et al. 1998). These spaces are characterized by different architectural dimensions, wall sizes, lighting, furniture configurations and information artefacts and technologies. Room dimensions, for example, affect the size of groups that can be accommodated in collaborative settings. Wall and boundary design can affect the openness, visibility and audibility of a space with respect to the larger office environment (Laing et al. 1998). The design of table shapes can be used to affect proxemic arrangements such as interpersonal distance between collaborating parties and also the f-formations that people construct in relation to each other and information artefacts in the room. Table heights can be used to cause people to sit or stand. Different seating arrangements can be used to create different postures and orientations to people and information artefacts in a space. The point here is that different spaces are composed of particular configurations of architecture, furniture and technology, the interacting dimensions of which profoundly affect the social, informational and collaborative practices that can take place within them along the lines of the arguments made by Hall (1966) and Kendon (1990). The point too is that these spatial dimensions can and are deliberately designed with particular social effects in mind.

What is also important in this argument is that people's work practices are not static and homogenous over time. During any given work day people will move between different types of work practices, from quiet individual work, to serendipitous dyadic interactions, through to informal small group collaborations and onto larger more formal collaborations (Becker and Steele 1995, Duffy 1997, Laing et al. 1998, Steelcase360 2007). To reflect and enable this, any modern office environment is essentially an ecology of different types of spaces with different social, spatial and informational properties. For example, there may be individual offices or workstations. Larger individual offices may also incorporate tables for small group collaboration. There may be informal meeting spaces, bookable enclosed meeting rooms and large formal conference rooms. Within such an ecology of different physical space environments, people can make choices about workspace use according to the particular type of collaboration they are engaged in.

The point of these works is that "space matters" and that architectural arrangements and furniture design impact as significantly on collaborative work practice as technology. Many CSCW researchers over the years have of course oriented to these matters in various ways in their discussions of work practice. We see this in our earlier discussions of embodied action for example as well as key works discussing space and place and their relationship in the configuration of collaborative work practices (Dourish 2001, 2006, Fitzpatrick 2003, Harrison and Dourish 1996). Others such as Olson and Olson have made important references to spatial characteristics' of the workplace and how these pertain to the social organization of work, most notably in their "Distance Matters" paper and in the discussions of radical collocation practices (Covi et al. 1998; Olson and Olson 2000). More specifically within the context of video mediated communication design, there has been some important foundational knowledge that relates to the concerns in this paper. Of note here is Buxton (2009) in his retrospective of the media space work in which he raises the issue of space-function-distance relationships. This discussion points to his early media space experiments with screen positioning and spatial audio with a view to maintaining certain spatial configurations of the virtual in relation to the physical space. The idea is that by appropriate configuration of the remote participant display in physical space, the local participants can use the same mechanisms to address and understand the remote participant as are used to address and understand the other local participants, e.g. using gaze, head turning or gestures. One of the most notable of the systems discussed by Buxton is the Hydra system (Buxton et al. 1997, Sellen et al. 1992). In this system, each participant is represented by small video surrogates that are configured in a round table arrangement. For each local participant, their particular arrangement of video surrogates corresponded to the respective view arrangement of the other participants comprising a meeting around a round table. While Hydra did not address all the issues of spatiality it was important in laying down the foundations of the concerns with geometrical configurations in video mediated environments that relate to the arguments in this paper. Some of the issues with Hydra, for example, were the small sizes of the remote participants images. This issue was highlighted by Okada et al. (1994) who argued that the small image sizes make the remote participants feel further away – perceptually placing them, in Hall's (1966) terms, at a different interpersonal distance than desired. In response to this, Okada et al. demonstrated the benefits of life sized images in their Majic system, presenting life sized images of remote participants at a virtual distance of about 4 feet – in Hall's "social" distance as appropriate for these kinds of meetings.

Some more recent work addresses different concerns with respect to spatial organisation in video mediated environments (Yamashita et al. 2008). In particular, this work orients to Hall's (1966) and Kendon's (1990) arguments about the influence of different spatial configurations of people or different f-formations on conversational dynamics and psychological states (see also Sommer 1969, Steinzor 1950). In their t-room system, Yamashita and colleagues were able to create different arrangements of local and remote participants, comparing configurations in which remote participants were arranged to be spatially adjacent to local participants versus arranged to be opposite the remote participants. The findings indicated differences in turn taking behaviour and perceptions of unity in the different settings, highlighting the importance of spatial considerations in video mediated environments.

While acknowledging the importance of spatial configurations and settings, when it comes to designing interventions, CSCW's primary concern is typically and understandably focused on the technological. Our argument in raising the workplace design literature is that there are times when effective intervention in particular circumstances needs to articulate the design details of the technology in conjunction with those of architecture and furniture. As the work practices research literature would acknowledge, each of these components shape collaborative practice in a mutually dependent manner. But when it comes to designing new collaborative systems, it is often the case (though not without exception) that less concern is given over to the architectural and furniture elements within a design. It is the relationship between these components that is of interest in the current paper and our discussion of Blended Spaces. It is precisely this attention to the details of architecture, furniture and technology design together that is important in the way that they work.

The arguments presented here, then, begin to articulate the value of achieving geometrically coherent representations of the remote site with respect to the local site in distributed collaboration settings. Doing this, we can better enable the interpretation of embodied action across the sites and move towards improved intersubjective understanding, meaning making through space and coordination of action. The Halo system and the blended spaces argument builds on the foundational concepts of these earlier explorations but significantly extends it in terms of detailed attention to a much broader range of architectural, furniture and technological elements and the interaction between these elements. It is these interactions that we now turn our attention to.

5. BLENDED SPACES: THE EXAMPLE OF HALO

The considerations for physical space design and the impact on collaborative practices are the starting point for understanding the design choices in Halo and their contribution to the user experience. In designing the system, the Halo team focused on four aspects of the user experience: *work* (what kind of work will take place in that space and what tools are needed), *communication* (what dynamic verbal and non-verbal elements are required), *interaction*[3] (control of the set-up and work tools) and *service* (support services for maintenance, optimization and troubleshooting) (Gorzynski et al. 2009). With this in mind, the team started first with a look at what typical traditional video conferencing set-ups would be if considered in terms of equivalent physical spaces. What the team concluded was that with traditional videoconference set-ups screens and cameras are positioned in ways that are primarily appropriate for local participants or placed simply due to installation pragmatics. Often, then, they are placed out of the way where televisions or media screens would be placed. Considering this in terms of a meeting happening in physical space, the Halo team argued this was equivalent to seating some meeting participants in the corner of the room away from the main table (cf. Buxton 2009). This creates lots of well-documented problems and presence disparities in terms of participants not being included in the meeting or being forced to watch the meeting from a corner rather than be actively involved. With this in mind, the starting point for Halo design was to think of what would be the appropriate configuration of collocated physical space design if it were to effectively support collaboration of a particular type and size. The next step would then be how to transform the physical dimensions and characteristics of the ideal physical room and recreate these properties perceptually in a distributed setting through a carefully crafted configuration of architecture, furniture and technology elements. This would include manipulating things such as camera positioning, display arrangement, wall design, table design, lighting and audio design to create geometrically correct representations of the remote site – to provide, in essence the visual sensation of remote parties sitting in the same ideally designed physical environment. This is what Gorzynski et al. (2009) refers to as the spaces becoming "blended" – that is where spatial geometries are preserved across the distributed settings providing a unified perceptual frame of reference that facilitates interpretation of embodied actions.

Within this basic philosophy, it would have been possible to develop a range of different Blended Space environments that broadly map onto the ecology of different collaboration environments we see in the physical workspace as outlined above. As the first development, though, the initial Halo studio chose a base physical work setting where the design was oriented to a small boardroom type environment that would accommodate meeting sizes of up to 12 people. The type of work being supported here were relatively small meetings for critical business discussion where the focus of the meeting is conversation and where the group non-verbal communication aspects are too important for a voice only conference call (Gorzynski et al. 2009).

The base physical space was chiefly based around a table designed to position six people either side of an elliptical table. This base design would position people at a physical distance and orientation from each other appropriate for a meeting of this type.

---

3) As with the Gorzynski et al. framework, we share some concerns with the interactive tools. The focus of our arguments in this paper extend the discussion of these beyond just thinking about what tools are needed for a particular type of work. Critical to our arguments here is that the tools and particular interaction mechanisms are an integral part of the communication experience, changing assumptions we can make with respect to proxemic configurations of people, the collaboration dynamics and the fundamental nature of their expressive possibilities.

In Hall's (1966) terms, this kind of work takes place within the range of *social* distance – somewhere between 1.2m and 3.6m. Starting from this base physical space concept, the design of the Halo system then sought to recreate these physical dimensions of this space across the distributed setting (Gorzynski et al. 2009). In doing this, the design was guided by a number of key principles: *magnification constancy* of images from multiple sites; correct *eye heights*; consistent *foreground and table height* across sites; perspective *distortion reduction;* alignment of AV components for *correct eye contact and gesture awareness*; and representation of *spatial audio.* This involved management of the interacting factors of room geometry, lighting, furniture, screen size and positioning and camera arrangement.

In each Halo room, there is a panoramic array of three displays on the front wall. The width of the three-display arrangement matches the width of the table. The screens are placed at the appropriate physical height with respect to seated participants, so that the eye levels of the local and remote participants are at the same height as if the remote participants were sitting on the opposite side of the table. With cameras directly above the screens at this height and physical distance, vertical eye gaze contact error is also mimimised (Gorzynski et al. 2009). With regards to the data screen provided in the Halo rooms, positioning of the display was also a key consideration for Gorzynski et al. Their argument was that placing the display above the participants kept it close to the participants, making it easy to see both information and participants at the same time without distracting focus form the conversational and nonverbal communication elements.

The table placement is such that it positions the participants at the appropriate physical distance from the screens. This combined with the correct camera zoom creates the necessary size representation of the remote participants that creates the perception of the desired physical distance between participants set out in the base physical space boardroom design. The three cameras are calibrated to provide a consistent zoom level such that they can be combined into a single consistent panoramic view across the three screens. Likewise, color and balance are fixed rather than automatically adjusting, and consistent across the cameras again to facilitate the perception of a single panoramic image. The representations of the remote site show upper body and head views at the correct size. This is consistent with the recent findings by Nguyen and Canny (2009) that demonstrate how such framing can facilitate the development of trust across the distributed settings relative to, for example, head only views.

Other aspects of the table design are also crucial. For example, the curvature of the table helps align people with the three camera views so that they are facing in a direction that helps the perception of eye gaze[4]. Table legs are strategically positioned so that pairs of participants sit between them. These positions correspond again to the camera views and in doing this people naturally sit so that they appear on a single screen. Likewise there are subtle joins on the table that also provide cues to seating position.

4) Eye gaze in Halo is never going to be perfect as the same image is presented to participants regardless of where they sit. But there are emerging solutions to this problem in which different images can be presented to participants according to where they sit, greatly enhancing the possibilities for accurate eye gaze throughout the environment. For example see Nguyen and Canny's (2005) MultiView technology.

What is important about the design here is that the constraints and cues are embodied in the environment. As such, position can be implicitly oriented to without the need for conscious consideration. We can draw some theoretical parallels here with early work in cognitive psychology and HCI (e.g. Larkin 1989, Zhang and Norman 1994) and more recently with advocates of embodied interaction (e.g. Dourish 2001, Hornecker and Buur 2006) that discuss the important ways that physical constraints on action can be embodied in the environment. Because these cues help position people appropriately, this contributes to Halo's design working without the need to provide visual feedback to the participants about how they appear on camera. This is a significant design consideration in the sense that it promotes greater transparency of the technology and the illusion of a unified space since no picture-in-picture arrangement is present to break the illusion.

The table size and positioning with respect to the cameras plays a further role in providing continuity across the distributed space. The set-up of these elements is such that a small proportion of the table is made visible on the displayed representation of the remote site. The table edge then provides a continuous line across the panorama of screens and gives the impression of completing the table in the local space. As such, the perception is of the participants being sat on the opposite side of the table. Achieving this effect is actually a more involved configuration than it might first appear. Because of the camera placement above the screens and its physical distance from the seated participants, there are inevitably perspective distortions introduced into the image. If too much of the table edge is shown in the displayed image, then these camera distortions are more salient in the image in a way that is damaging to the illusion. There is a necessary balance between showing just enough of the table to represent the continuity across the spaces but not enough to reveal inherent camera distortions. This again is achieved through the careful design and arrangement of the technology and furniture elements.

These potential distortions in the image arising from the camera placement have a relationship with the design of the architectural elements of the rooms, namely the walls. The walls are designed to be as neutral as possible. No decorative detail is included. The camera angles and focal length are such that horizontal and vertical lines are not present in the represented image on the displays. For example the joins between the walls and the ceiling, or the walls and the floor are deliberately not visible in the image. Traditional video conferencing shows no real concern for such details. But the point here is that such horizontal and vertical lines again make the camera distortions much more salient, appearing curved. By carefully combining architectural detail appropriately with camera views, these effects can be minimized.

As with film and photographic studios, lighting was also an important consideration with side and overhead lighting carefully chosen to illuminate participants naturally without introducing shadows and too many depth cues. These lighting choices interacted with other architecture, furniture and technology elements in the space. For example, with the front wall housing the monitors, the lighting of this wall had to be dimmer to avoid glare on the monitors. Within this arrangement, the front walls appeared darker than the rear walls and consequently did not match the color of the image appearing on the screens. This made the screens feel discontinuous from the front wall. In order to compensate for this and achieve a consistent color across the front wall and wall color as represented on the screen, slightly lighter colored wall panels from the rest of the room were chosen. Again we see here a solution based on a combination of the architectural with the technical.

Other aspects of the Halo design are also kept constant in an attempt to maintain the transparency of the room. A good example here is the full duplex audio that is kept at a constant volume. There is no interface control presented to the user to allow adjustment of such features of the environment. From one perspective, this may be seen as inflexible and a limitation in terms of the level of control afforded the user. But actually, such design decisions represent an important part of the blended space design philosophy, namely that the space be as immediate to use as walking into a normal physical meeting room. Removing layers of control is part of creating this feeling. So for example, if in a normal physical space there are difficulties hearing, one simply requests that a person speak up. So too are such protocols an inherent part of what is trying to be achieved in Blended Spaces such as Halo. That is, these things are dealt with through social mediation rather than technological manipulation.

Finally, of importance to the Halo design is the whole notion of symmetry. Each and every Halo room is identical in its configuration of architectural furniture and technology elements. In the distributed set-up, each Halo room acts as the completing half of the other room, to maintain the perception of the single continuous physical space. Using symmetry in this way obviously brings its limitations in terms of flexibility of the system, for example, in terms of the openness and number of sites that can be connected to relative to technologies such as Access Grid. The intention here though is not to cover all value points in the arena of video conferencing since these lead to particular compromises and trade-offs in the design. The intention rather is to take a certain type and size of collaboration and create an environment that works well for this, as opposed to designing a more flexible system that compromises the experience. There is no right and wrong answer here in design terms, but simply one of focusing on different value propositions. Symmetry is also helpful in maintaining a geometrical coherence necessary for supporting the correct spatial representations of gestures and attention direction. As Heath and Luff (1991, 1992, 2000), Gaver et al. (1993) and others have argued, asymmetries in media space arrangements can lead to difficulties with the lack of common frame of reference necessary for effective communication. Within the conceptual framework of embodied interaction (Dourish 2001; Robertson 1997) these symmetrical arrangements better facilitate achievement of intersubjective understanding of action. It is this intersubjective understanding of action that allows Halo to dispense with the picture-in-picture model of video conferencing showing an image of the local participants - symmetry and its accurate spatial representation facilitate an embodied understanding of how others appear and how one appears to others.

6. BLENDED INTERACTION SPACES

As has been discussed, the current Halo system and other similar commercial Blended Space offerings are designed specifically to support a certain type of collaborative work, in particular, critical business discussion with a focus on conversation and group non-verbal cues. The document and information sharing tools needed for this kind of collaboration are fairly minimal which is reflected in the design choices within the current Blended Spaces. For example, in Halo, there is a VGA out solution from a connecting laptop to the 4th display above the three panoramic displays. This realistically only allows single documents to be viewed at any one time. Any interaction with the information on this display is restricted to one person in front of the laptop. This use of a single document where control is maintained with the presenter is appropriate for the kind of collaborative work being supported by Halo.

What happens, though, when we consider collaborative work in which there is more complex, data intensive collaboration around documents (e.g. scientific analysis, emergency response, military planning). Research has shown how shared document and information spaces provide an important resource in scaffolding collaborative talk and work, playing a number of different roles. For example, they are something that can be pointed at and worked on during conversation allowing the use of more efficient deictic language (e.g., Bly 1988, Bly and Minneman 1990, Heath and Luff 1991, Luff et al. 1992, Sellen and Harper 2000, Spinnelli and O'Hara 2001, Whittaker 2003). They can provide an ongoing history of a meeting as they are manipulated. These representations embody the knowledge and decisions created through the collaborative work. The process of seeing things being added to the shared surface creates a sense of commitment and ownership to the ideas and information that is important during evaluative phases and assignment of action (e.g., Moran et al. 1999, Spinelli 2003, Whittaker and Schwartz 1995). Providing persistent representations of information is also important for ongoing referral throughout meetings. People need to draw those visual resources into conversations that are immediately to hand (e.g. Sellen and Harper 2000). If there is effort necessary to bring in a shared visual reference and interact with it then it is much less likely to take place (Perry et al. 2001, Spinelli et al. 2005, Spinelli and O'Hara, 2001). In these more document intensive collaboration tasks, large amounts of information and data from different sources need to be viewed and interacted with concurrently in order to facilitate the cognitive demands and social dynamics of the collaboration. In physically collocated situations, the ability to assemble a collage of information artefacts is easier: information resources such as documents can be arranged adjacent to each other and navigation back and forth between the different information sources is simpler and more tangible; information can also be spatially arranged to facilitate the cognitive processes of information synthesis (e.g., Kirsh 1995, Hutchins 1995, Rogers and Bellotti 1997, Spinelli et al. 2005).

But the issues are not only about supporting the information processing in these tasks. As a collaborative activity, the information artefact-centric interactions, gestures and manipulations are a fundamental component in the conduct of communication within Kendon's *o-space* (1990). The embodied nature of actions in relation to a commonly understood spatial frame of reference is particularly significant for these kinds of data intensive tasks and the coordination of talk and activity that occurs around the artefact space. Again, in physically collocated situations, the artefact space provides the shared spatial references that provide the foundations of intersubjectively produced and interpreted actions and talk. With distributed collaboration there are no simple solutions for providing this in any rich way and certainly not within the existing high-end commercial Blended Space systems such as Halo. Providing such support in distributed settings is key for this kind of collaborative work. Doing so in a spatially faithful way across distributed settings is even more important for this kind of work activity than in the social business communication work supported by Halo, because of the richer coupling of talk to embodied action.

Another key component of this these data and information intensive collaborative activities, is the use of multiple *entry points* in to the task, whereby different team members have equal opportunities to simultaneously access and interact with the information as the needs of the collaboration dictate (Rogers et al. 2009). Collocated ecologies of physical information artefacts provide multiple entry points, offering this scope for multiple people to interact with information from different sources at the same time. This is not simply significant because it allows people to organise this collaborative work in different ways (e.g. working on part of a task together or on different parts of the task in parallel); it is also significant for how people can choreograph their talk with reference to information artefacts. Imagine if you have to put your hand up every time you want to talk in a group setting. This would rapidly become burdensome and ultimately affect the fluidity with which group discussion could take place – the same is true for single *entry point* interaction mechanisms. The dynamics and impact of multi-point interaction mechanisms on more data intensive collaborative activity are now beginning to be demonstrated and explored in collocated settings (e.g. Hornecker et al. 2008, Jiang et al. 2008, Nacenta et al. 2007, Rogers and Lindley 2004, Rogers et al. 2009, Wigor et al. 2009). This ability for all team members to concurrently interact with multiple pieces of information from disparate sources, is not well supported in distributed settings beyond bespoke application level implementations, and certainly not supported in the current commercial high-end "Blended Space" systems.

In thinking about how to extend "Blended Spaces" to become "Blended Interaction Spaces", there have been a number of significant advances in recent years that open up opportunities for a more effective type of this data intensive collaboration between distributed teams. Large, high definition data panels, multi-touch sensitive horizontal and vertical surfaces, multi-person input capabilities and multi-display environments and techniques for viewing data from multiple sources offer intriguing new ways for collaborative teams to simultaneously view and concurrently interact with large amounts of data. A large number of research projects have investigated many aspects of these new possibilities for collocated settings, some notable examples being i-land (Streitz et al. 1997, 1999) Stanford iRoom (Johanson et al. 2002) and more recently WeSpace (Jiang et al. 2008, Wigor et al. 2009) to name but a few. Such systems though focus primarily only on collocated collaborative settings.

Research on these kinds of systems and technologies has shown how different configurations of interactive surfaces have particular properties with regards to how information can come to be shared and used within collaborative contexts (e.g. Huang et al. 2006, Russell and Sue 2003, Trimble et al. 2003). Particular form factors and size of displays, horizontal and vertical orientations, and the ways they are configured with respect to other artefacts and architectural elements within a collaborative work setting, as well as a range of different interaction mechanisms (touch vs. remote interaction) all have a significant impact on the way that people spatially arrange themselves with respect to the information and with respect to other members of the group. This in turn shapes the social dynamics of collaborative activity impacting on, for example, things such as control structures within the group. For a group to simultaneously interact around a large vertical touch screen for example requires people to arrange themselves in a line adjacent to the surface. With horizontal surfaces people sit or stand around a tabletop and interact with the information from a variety of orientations according to the size, shape and height of the table. Rotation mechanisms for objects on these horizontal surfaces and the ability for multiple people to interact at the same time mean that the people gathered round the table could interact with the information from wherever they are sitting or standing (e.g. Hornecker et al. 2008, Nacenta et al. 2007, Rogers and Lindley 2004, Rogers et al. 2009).

The argument we want to make here is that there are proxemic consequences of these interactive surfaces and mechanisms – what we call *interaction proxemics*. These interaction proxemics create particular considerations when trying to introduce them into a Blended Space environment. As we discussed in the previous section, Halo and other Blended Spaces are carefully crafted in terms of architectural, furniture and technology dimensions and arrangement to introduce necessary spatial constraints and cues that lead to geometrically coherent representations of distributed spaces. This allows assumptions to be made about how people will arrange themselves that enables appropriate camera and display configuration. Given the proxemic consequences of different interactive surfaces and mechanisms, the introduction of such systems into a Blended Space needs to be done in the same carefully crafted manner. As a simple illustrative example, if one were to introduce a large vertical touch sensitive surface into a Halo type environment the interactional properties of the surface will encourage people to stand up and interact with it leading people away from the carefully positioned camera and display arrangement.

The arguments here in relation to the *interaction proxemics* are twofold. On the one hand, it is important to understand how the physical configuration of the environment and information artefacts can be designed to accommodate the particular spatial dynamics introduced by these different interaction mechanisms. On another hand, it is important to consider how other elements in any particular Blended Space configuration (e.g. video and data display arrangement, camera positioning, microphone and speaker positioning), impact on the ways particular interaction mechanisms are subsequently used because of potentially competing spatial requirements. For example, if we take the same simple interactive whiteboard example as illustration, would the need to be positioned in the line of sight of the camera and in the vicinity of microphone setup in a Blended Space actually prevent people getting up to use the whiteboard? The design of a Blended Interaction Space, then, needs to consider these interacting factors.

In order to unpack this further, we follow the approach adopted in the Blended Space argument in which Halo was used as an illustrative exemplar to highlight key issues of Blended Spaces with reference to particular design features. In our extension of this discussion to Blended Interaction Spaces, we use as an illustrative example, a system called BISi (see Broughton et al. 2009), a blended interaction space developed for small group data intensive collaboration. It is important to recognise this system as just one example in a much larger design space of blended interaction spaces for different types of collaborative work. The purpose of our discussion using this example is to highlight issues, challenges and considerations arising from attending to *interaction proxemics* in the creation of geometrically coherent distributed collaboration tools.

7. A BLENDED INTERACTION SPACE FOR SMALL GROUPS

Just as Halo had taken as its starting point a boardroom-type physical space for supporting meetings of up to 12 people, the base physical starting point for the BISi systems was aimed at supporting smaller group collaborations between of two or more people and no more than eight. In addition, BISi was designed to support more data-intensive collaboration around multiple data sources and documents. In assembling a Blended Interaction Space of this type, the concerns needed to be with the provision of a shared digital workspace across the sites as well as with the geometric configuration of the physical set-up. In presenting this discussion, then we begin with a brief overview of the shared digital workspace application, TAPESTRY. The intention here is not to engage in a detailed technical specification of the system (which will be the subject of other research papers) but rather a sufficient overview of the functional characteristics that are necessary for a discussion of the interaction proxemics associated with the BISi system. That is, particular features of TAPESTRY have implications for the spatial consequences of interacting with the system that relate to how blending may be achieved. And it is these that we wish to discuss in relation to the design of BISi. Following the description of TAPESTRY, we will discuss the set-up of BISi and the ways in which the configuration of architecture, furniture and technology elements create geometrically correct representations of the remote site in light of the new interactive properties.

7.1. TAPESTRY

TAPESTRY is a distributed application framework shared across remote sites designed to provide a common interactive workspace. It can be considered as a conceptual "surface" shared by all sites onto which local and remote participants can share "windows" and "applications" or complete "desktops". The philosophy of TAPESTRY is to connect the everyday computing resources used by participants in their everyday work (e.g. laptops, desktop machines, specialist data clusters) to the shared workspace allowing users to work with their existing applications and move fluidly from their individual work (i-work) to collaborative work scenarios (we-work) and back[5]. As Huang et al. (2006) have demonstrated with NASA's MERBoard, people will reject the technology in spite of its interactional affordances if it is cumbersome to get information and applications to and from the system.

The TAPESTRY system combines a light-weight and extensible distributed system infrastructure for synchronous collaboration featuring wide-area federation and

5) One of the continued themes throughout this work is that work practices take place within an ecology of artefacts and spaces that people move through. In thinking about the design of blended interaction spaces, it was our explicit intention not simply to design a circumscribed system but one that would sit well within the broader ecology of the workspaces and enable fluid movement to and from that space. The ability to transfer digital as well as physical artefacts between these spaces formed a strong part of the design focus.

uniform service modeling (Livespaces (Phillips 2008)), OpenGL-based desktop and advanced application sharing capable of efficient rendering of data and graphics intensive applications (Virtual Terminal (Jiang et al. 2007)). Applications and files on the TAPESTRY then essentially run from their source machines. TAPESTRY also incorporates the Multi-Pointer X server (MPX[6]), a multi-input-device-capable windowing system that enables multi-point interaction for new and legacy applications (within the particular architectural constraints of the legacy applications). As such, TAPESTRY enables multiple applications, windows and desktops from different source machines to be viewed and interacted with concurrently on the same interactive surface.

7.2. BISi set-up

The physical set-up for BISi can be seen in Figure 2 and 3. A 2x2 array of LCD displays occupies the front wall. Each of these displays is 102.7cm x 63.5 cm operating at a resolution of 1920x1080. As with the Halo system, the lower displays are used for presenting video conferencing streams from the remote sites, again positioned at the appropriate height to enable across-the-table eye-gaze. Above the cameras is the second row of displays that are used to display the TAPESTRY collaborative workspace surface continuously across both displays. The top edges of these displays are tilted 18 degrees away from the front wall. The LCD panels are surrounded by wall-panels mounted flush with the front edge of the displays – the colour of the walls is a continuation of the colour of the TAPESTRY surface and the background of the video displays - creating the effect of the displays being a continuation of the wall surface. In-between the upper and lower displays are two high definition video cameras providing parallel video streams to the two corresponding displays of the far end set-up. These are positioned at 51 cm to the right and left of the vertical centre line between the two displays respectively. The system also incorporates a multi-touch interactive horizontal surface for interacting with information in the shared TAPESTRY environment. The table is set at a standard conference table height of 72.5cm. Measuring from the centre line of the vertical display grid, the rear end of the table is positioned 16 cm from the vertical displays. Continuing along this central axis of the table, the front most point of the curved edge is positioned 140cm away from the vertical displays. The multi-touch surface within the table frame is a 1920x1080 high definition display measuring 105.2 cm by 60.5 (46" diagonal) and is capable of tracking multiple fingers and hands from multiple users.

From the set-up depicted in Figures 2 and 3, we can see how the system draws on the blended spaces design philosophy in terms of display arrangements, front wall design, table positioning and geometry, etc. But there are particular design features, challenges, issues and compromises arising from the attempts to position the space for the support of smaller groups and more intense interactive collaboration tasks. These issues begin to form our agenda and understanding around Blended Interaction Spaces in a way that extends beyond the initial blended spaces philosophy highlighted through our discussion of Halo above.

---

6) http://wearables.unisa.edu.au/mpx/?q=mpx

Figure 2. The BISi setup

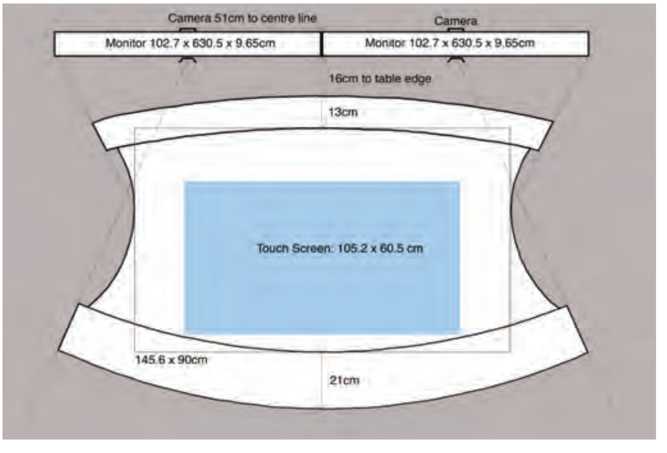


Figure 3. Schematic overview of the BISi setup

The first challenge with this part of the design space concerned the physical dimensions of the space. For this type of collaboration, we were dealing with smaller group sizes as well as more intense discussion around shared data artefacts. In accommodating this group size and type of discussion, the geometries of the space needed to be smaller than Halo with less interpersonal distance between the collaborating participants than was seen in Halo. While still in the Hall's (1966) range of interpersonal *social distance* (1.2m-3.6m), appropriate for work related meetings, the BISi system aimed to work at the more *intimate* end (1.2m) of this interpersonal social distance range, in a way that would enable the shared visual and physical access to the tabletop based *o-space* (Kendon, 1990) between the collaborating parties. Table shape and size are key issues here that present some initial challenges in the development of such spaces. While early paper-based sketching allowed exploration of initial shape concepts for the BISi system to meet these criteria, it was quickly apparent that such sketches were limited in terms of understanding of scale and proxemic arrangements of collaborators in relation to work artefacts and a sense of what these might feel like. In developing our understanding of size, shape and positioning, then, it was important to conduct design iterations with full-scale mockups (see Figure 4). Cardboard and foam board were ideal materials here for creating different table shapes and sizes, and different workspace display configurations. These materials could be quickly reshaped and were light enough to be reconfigured in many arrangements that allowed many iterations to be experienced. Experiencing these early iterations was an important component in understanding relationships and tradeoffs between factors such as physical proximity from other people, proximity from information and displays, sense of reach and sense of object manipulation. For example, in one iteration where a mockup semi-circular table design was created, this seemed to provide appropriate interpersonal distances and shared reach access to tabletop documents for local collaborators. However, it was also apparent that this design positioned certain participants uncomfortably close to mockup video displays – this lead to exploration of other shapes and configurations to avoid this consequence.

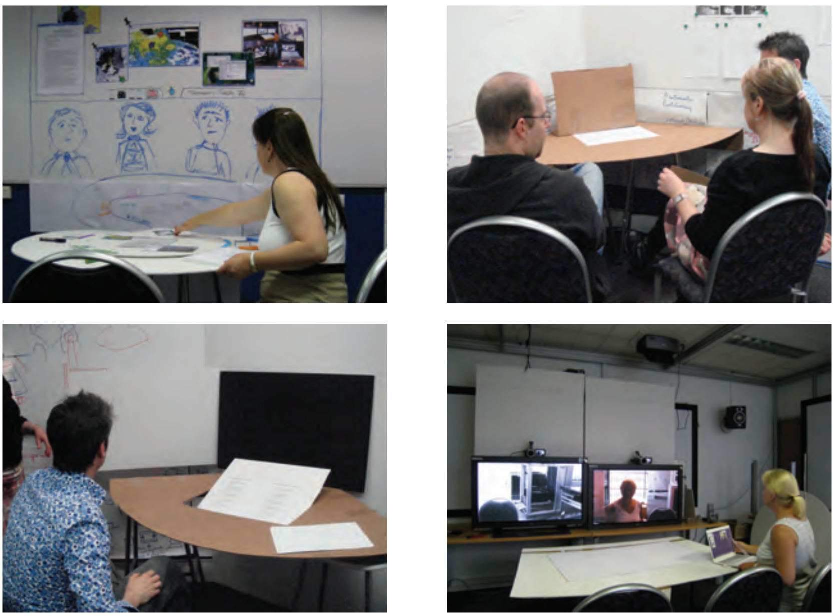


Figure 4. Iterations of full-scale mockups of Blended Interaction Spaces

Again the low fidelity of the foam board table prototypes is important at this stage, enabling fine-grained adjustment of camera parameters in combination with adjustments in table sizes, shapes and distances. This refinement of the setup is necessarily a painstaking process that involves an *informed* experimentation with different parameter changes until the desired combination is reached[7]. At the smaller distances (relative to Halo) appropriate for these group sizes and type of collaborative work, this balance of interdependencies is somewhat brittle. But there are further challenges with this balance arising from the larger range of focal planes to be accommodated due to the proxemic consequences of the interaction mechanisms in the space. With Halo, one could make the assumption that people would be positioned in the plane set by the table edge – the interaction mechanisms present in Halo do not cause people to shift from this plane. Indeed, this is one of the key features of the design in which the table creates this fixity of position. With BISi, the interactive capabilities and their proxemic consequences introduce a new set of behavioral assumptions that need to be considered. For instance, when people need to point to objects on the upper part of the table they tend to lean over, thereby moving closer to the camera. This results in certain magnification anomalies in the image as well as potential breaks in the panorama continuity. Such magnification anomalies are affected by the combined factors of camera zoom levels, focal length and depth of field – for example being increased at high zoom levels. As such this issue needs to be compensated for in our choice of zoom and focal length as well as table dimensions and distances from the camera. So table dimensions and distances from the displays are designed in part with these factors in mind to reduce where possible the amount of leaning forward and to minimize magnification distortions through choice of lower zoom levels[8]. Alternative design considerations here would be to introduce tabletop interaction mechanisms that do not encourage such physical leaning and which would allow an assumption of narrower focal planes. An example of such a technique here would be the Pantograph that provides a greater virtual reach than direct touch mechanisms employed in BISi (e.g. Nacenta et al. 2007)[9].

The geometrical arrangement of information on the interactive surfaces is also a particular challenge in the context of a blended environment arranged as a distributed environment. The data screens being positioned above the video screens need to be tilted forward by an angle of 18 degrees for more comfortable viewing at these small interpersonal distances. But the particular challenge arises out of the desire to maintain the spatial coherence of pointing and eye gaze as represented on the video screens with respect to different information sources.

---

7) The important point about these systems is that their design involves a complex series of trade-offs. Exploring these trade-offs is neither "black art" nor "exact science" but, rather, is a principled exploration within the bounds of understood socio-technical issues. One of the key aims of this paper is to provide an awareness of the proxemic issues and principles within which blended interaction spaces can be designed, explored and understood.

8) There are other factors to be considered in the design of the table dimensions that will be discussed later. As such there are a complex set of trade-offs that are being balanced in the design that go beyond optimizing table design to reduce these magnification anomalies.

9) The work of Nacenta et al. (2007) is an important example of a study that systematically articulates the spatial consequences of a range of different Tabletop interaction techniques. This kind of study of interaction proxemics in relation to other interaction technologies will form a key part of the understanding necessary to effectively introduce interaction possibilities into new blended interaction spaces.

There are well established Differential Rotation Effects in 2D representations of 3D scenes for pointing and gaze, whereby orientations of gaze and pointing are increasingly misjudged with larger viewing angles away from the perpendicular (e.g. Goldstein 1987). As distances from the screen are reduced (e.g. with the smaller interpersonal distances in BISi compared to Halo), these effects can be increased. These effects need to be compensated through use of multiple cameras and through constraints on the width of the overall system (thereby reducing viewing angle differentials). In the BISi system, a dual camera set-up was used to mitigate these effects. Attaining satisfactory levels of gaze and pointing accuracy also impacted judgments of table distance and width. Reducing table width and increasing distance from the screen can both help reducing differential viewing angles and differential rotations effects. Attempts to achieve satisfactory levels of pointing and gaze accuracy were also factors in our iterative adjustments of table dimensions and positioning distance from the screens. With these kinds of display technologies, it is not possible to remove these effects completely. In our experiences with the system, determining who is looking at who can be achieved to a satisfactory level with the BISi setup, as with Halo. Such discrepancies in accuracy do affect the level of spatial granularity with which deictic reference can operate. For higher level pointing tasks such as pointing to a large information object such as an onscreen window, this accuracy in our experience with the system is acceptable. For finer grained deictic referencing requirements, the approximations are not always sufficient and better achieved by pointing and gesturing with shared cursors. However, some of these issues can be overcome with more sophisticated multi-view display and camera setups such as in the work of Nguyen and Canny (2005).

But there are deeper levels of complexity associated with this issue. Typically in CSCW systems for remote collaboration, a key design principle for shared workspaces is based on the notion of "what you see is what I see" (WYSIWIS). Such a principle, of course, has much evidential support as a deictic conversational resource. Within a Blended Interaction Space though, if we simply present the same representation on both ends of the remote collaboration, both will see the same thing but this will not match what their perspective would be if the space's geometrical arrangement were considered. For example, if there is an information window on the left most edge of Room A's and Room B's display, when physically pointing on the video conference, both parties would appear to be pointing and looking in opposite directions - even though they would actually be looking at a common piece of information. With Halo, this was not so much of a problem in the chosen design. Having a single data screen placed in the centre of the front wall and at a further distance from the participants all contributed to the achievement of correct gaze directionality with respect to the data display. That is, when people look at any data on the data screen in Halo they will appear to stare broadly into the middle of the room - so issues of directionality are not perceived or made visible. In BISi, the problem is exacerbated by the need for a larger information space and the use of multiple data screens at a shorter distance from the participants. This means that the angle subtended by the data displays is much wider than in the Halo room, making the directional anomalies more visible when looking at a specific place on the data screen.

There are a number of solutions to be considered here. Ideally these should adopt as much as possible a *spatial reference domain* where there is a match between information organization and the arrangement of the displays – as has been shown to be effective in other multi-display environments (see Nacenta et al. (2009) for an overview). The use of spatial reference domains works well for collocated multi-display environments but new challenges are introduced when trying to do this in distributed environments - in particular, those distributed environments where the representation of user actions is both through the video channel and digital representations such as pointer movement (i.e. Blended Interaction Spaces). As such, the solutions available to us are not without some compromise along this dimension. For example, it is possible to adopt various mirror arrangements (cf. Ishii and Koboyashi 1992) whereby the screens in each room are a mirror image of those in the other room. While this solves the spatial issues it obviously leads to difficulties with the presentation of information, which, if mirrored is not readable. With this in mind one can adopt partial mirror solutions in which there is higher level mirroring of window frames on the TAPESTRY while the information within the windows is presented in the correct orientation. But this introduces problems of cursor leaps as one moves a pointer across the TAPESTRY surface and over a window. We have tried the much simpler solution of swapping things at the screen level, so that data presented on the left hand screen of the local room would be presented on the right hand screen of the remote room and vice versa. This solution achieves high-level gaze and pointing directional accuracy. It also comes with some level of cursor leap when moving from one screen to another for one of the rooms participating (since the centre line for one location maps to the far right or left in the other, depending on the direction the cursor is moving in). This was our chosen compromise, in the first instance, being broadly commensurate with some of the other achieved pointing accuracy levels arising from differential rotations effects. But in many ways this remains an open issue requiring more systematic research and innovation. What is important here though is the demonstration of the new interacting factors that arise when introducing new interactive workspace capabilities. We can see again the ways that the architecture, furniture and technology arrangements (both hardware and software) are intimately intertwined with a set of interdependencies where changes in any of these components lead to knock on consequences that need to be compensated for in the configuration of the other components.

Similar issues arise when we consider the horizontal interactive surface in the BISi set-up where there remain open questions as to how information on the horizontal surface should be presented. The arguments here are potentially at odds with standard CSCW convention of WYSIWIS. While the WYSIWIS approach is most appropriate in situations where only virtual pointing is available, the paradigm conflicts with the desire for geometric consistency with respect to pointing and gaze portrayed through the video representation. In this sense we potentially reduce the sensation of sitting on the opposite side of the table.

Again there are a number of approaches to take here. One is to adopt some of the methods outlined above for information presented on the vertical displays. A second approach is to present information on the remote table from the perspective of someone sitting at the opposite side of the table. This of course may lead to difficulties with certain pieces of information being upside down. But the arguments here are that the control and presentation of information would be socially mediated through rotation, orientation and other forms of micromobility (Heath and Luff 2000; Kruger et al. 2003) in the same way that paper and other artefacts are oriented in collocated physical settings. This approach is designed to facilitate the sensation of remote parties being on opposite sides of the physical table. It also aims to enable the portrayal of social meaning and coordination control through the actions on the digital artefacts of the surface. For example, using orientation of a document to control privacy or to signal an invite to look at the document (e.g. by rotating the document towards the remote participants). A third approach is to treat the local and remote tabletop surfaces as a continuous space extending through the plane of the video wall. In this respect, the local and remote parties will see a different part of the information surface but can "push" information from their side of the tabletop "through" the plane of the video wall to appear on the remote surface. This respects some of the physical geometries of the physical and information space, but potentially conflicts with others (for example, the information space may extend a little behind where the remote participants are perceived to be sitting).

The TAPESTRY system can be configured in different ways according to these different approaches. Again the intent here is not to propose the single perfect solution but rather to highlight the key issues, challenges and potential trade-offs arising from these different approaches to introducing interactive surfaces and artefacts in geometrically appropriate Blended Interaction Spaces. Configuring the system in different ways provides us with the basis for more systematic research enquiry into the properties of these different approaches.

An area where the spatial consequences of our interactive design choices are further apparent is in the use of the multitouch table and MPX system as a way of providing multipoint capabilities in the system. From a spatial point of view, one of the things this brings to the system is multiple *access points* (Hornecker and Buur 2006, Rogers and Lindley 2004) whereby all the participants can concurrently interact with the system[10]. These authors have demonstrated the important properties of multiple access points in terms of collaboration dynamics. The particular concerns of importance to us in the design of the BISi system is that it allows interaction by all present in the meeting from wherever they are sitting. As we discussed earlier with respect to the Halo system, one of the key things is for there to be a strong coupling of participant positions with respect to the camera and display configuration in order to achieve the desired blended effects – this being done through particular characteristics of the furniture such as table size, shape, positioning and the table legs. But within Halo, there is only single point access to the data presented on the screen being controlled through the laptop of the person presenting the information. If another participant wanted to interact with the data on the display they would need to get up and move to the location of the laptop.

10) There are obvious inherent constraints of legacy applications here that are designed for single person interaction. Multiple cursers can enter into these applications at the same time and point but there is a need for "floor control" mechanisms to manage input from these multiple pointers in a way compatible with the input capabilities of the legacy applications.

In doing this, they would break the carefully crafted coupling of their position with the camera and display set-up. In the BISi system, we can see then that the provision of multiple access points available from the users seating position, enables the close coupling between participant position and camera positions to be maintained during the course of any interaction.

A related point here concerns the use of laptops within the BISi system. As we have discussed, BISi was designed to allow participants to connect their laptops to the system. This again serves a purpose of providing multiple access points to the system from wherever people are sitting. Support for this activity goes beyond just the connectivity of these devices. There are additional physical requirements necessary to support this activity. The table rim was designed explicitly with a visual and textural delineation of bench space to enable the comfortable placement of these laptops (among other artefacts such as paper documents and input devices) on the surface in the personal space of the seated participants. It was designed as a compromise between the need to have space for these items while not interfering with the ability to reach the horizontal touch screen. A distance of 21cm, the width of small laptops, and A4 documents in landscape orientation was found to be suitable, as illustrated in Figure 5.

Figure 5. Table rim in BISi designed to accommodate keyboards, pointing devices as well as laptops that people might bring into the collaboration without interfering with the touch surface.

8. DISCUSSION

Our concerns in this paper have been to articulate the key features and theoretical foundations of what we call *Blended Interaction Spaces.* These are distributed collaboration environments that attempt to faithfully incorporate geometrical properties and configuration of space, which have been shown to be significant in the organisation of communication behaviour in physical space. In developing these arguments, we have drawn extensively on a long tradition of literature within media space research and work examining the relationship between spatiality and human interaction. Of particular significance here have been the theoretical framework of embodied interaction and theories of the spatial organisation of behaviour such as Hall's (1966) proxemics and Kendon's (1990) interactional spaces. These understandings of the spatial organisation of behaviour provide important analytical lenses through which developments in media space research and design have and can continue to be understood. More specifically, they provide the resources with which we can think about new kinds of media space environments, of concern in this paper, that explicitly address these spatial concerns in their design. In order to more concretely articulate the design ideas, principles and issues involved, the paper takes as a launch point, a discussion of Gorzynski's (2009) original concept of Blended Spaces. Our approach here has been to further elucidate the ideas concerned through an examination of a canonical example of a Blended Space, namely HP's Halo collaboration environment. This has allowed us to explore some of the key design features that contribute to the perception of a shared spatial geometry between users working across distributed environments. Our concerns here are not simply with the specifics of the Halo environment, but more broadly how they help communicate some of the more general design considerations and ways of thinking that need to be applied in developing different types of Blended Spaces.

It is our thinking about other types of Blended Spaces that requires us to extend the original notion to encompass Blended *Interaction* Spaces – distributed environments that encompass possibilities for richer collaboration over distributed interactive workspaces in which the communication is an integral part of a collaborative analysis and interpretation of presented information; where the group needs to gesture, point, interact with and manipulate presented information in the context of their ongoing discussion. While such environments in some senses form part of the larger set of Blended Spaces, there are important reasons why such a conceptual extension is necessary and valuable. By introducing interactive workspace capabilities within such environments, we fundamentally change some of the behavioural and spatial assumptions that can be relied upon in the creation of a perception of shared spatial geometries. Through making the extension to Blended Interaction Spaces, we specifically draw attention to the spatial consequences of the interactive workspaces. Again, as with the Halo discussion, our approach in articulating these concerns is through an examination of a canonical example of a Blended Interaction Space called BISi. And again, our aims in articulating these design details lie in more than just the specifics of the BISi system itself. Our intent, through this illustration is to highlight some of the key design considerations, issues and more general way of thinking that is pertinent to the design of other systems of this genre. With other systems, some of the specifics will change, but the overarching design philosophy will remain the same.

In presenting this discussion, there are a number of key issues that the work has highlighted. First is the important notion of *interaction proxemics*. Drawing on Hall's (1966) original concept of proxemics, this notion makes particular reference to the spatial consequences of different interaction mechanisms in terms of the way they position people with respect to information resources and in terms of how they dynamically configure people with respect to each other during the ongoing course of collaboration. What we are arguing for, then, is for greater understanding and reflection on the spatial consequences of particular interaction mechanisms so that they can be combined appropriately with architectural and furniture elements with particular effect. Understanding the interaction proxemics of particular interaction mechanisms forms an essential part of the spatial assumptions that need to be worked with in the creation of a blended environment. So in the same way that table design allows embodiment of assumptions about spatial positioning of people in the environment, we need to do something similar with the interaction mechanisms. We are beginning to see some informative work to this end, for example in the domain of Tabletop interaction mechanisms and territoriality (e.g. Nacenta et al. 2007). While this work has focused its attention on collocated collaboration, it is nevertheless informative to our more specific spatial concerns with distributed collaboration environments. What we would argue is that this kind of work needs to be extended to a broader range of interaction mechanisms and configurations that might usefully be incorporated within other Blended Interaction Spaces (e.g. touchless gesture based interaction).

A second set of issues arising out of the move from Blended Spaces to Blended Interaction Spaces comes from the spatial qualities of the data space and in particular how to map this onto the geometric properties of the space envisioned. In our discussion of BISi we have seen ways this can be informed from other work in multi-display environments but also the additional complexities of trying to do this across distributed environments. What we have also seen is that the attempts to pursue spatial continuity across the distributed boundary can potentially lead to design conflicts with well-established design philosophies in CSCW such as WYSIWIS. In developing the Blended Interaction Space paradigm, such potential conflicts need to be the subject of a systematic research effort to fully understand their implications for distributed collaboration.

What we also hope to have highlighted in our arguments about Blended Interaction Space systems, is an attention not simply to the details of technology design but also to the details of architectural and furniture elements. In presenting our discussion of Halo and BISi, we have shown how these components are tightly coupled together in mutually dependent relationships and that adjustment in one component intimately affects our experience with other components. By paying attention to all these different components together, it is possible to exploit these interdependencies to good effect. For example, we have seen the use of furniture design to embody assumptions in the environment about the positioning of people. This information is then something that can be used in configuration and placement of cameras and in configuration and placement of displays. We have also seen examples where architectural design has been used to compensate for certain aspects of the technology such as image distortions. An example here is the lack of features on the walls avoiding horizontal and vertical lines that make salient any distortions. As well as thinking about this issue in relation to Blended Interaction Spaces, we believe that these considerations of architectural, technological and furniture relationships can be applied more generally to CSCW design in collocated situations. That is, applying attention not just to the technological but also giving consideration to the technological within the context of architectural and furniture elements. By attending to these other elements, it is possible to change people's spatial relationships with the technological. Hall's (1966) proxemics are again informative here in his articulation of how architecture and furniture can affect the nature of interpersonal interactions and social action.

Following on from this in blended interaction space philosophy is the argument, drawn from traditional work place design literature, for a varied set of spaces each designed specifically to afford different aspects of individual and collaborative work practice. That is, in every day work practices, people work in an ecology of interlinked spaces moving between them and exploiting their respective affordances for the unfolding demands of work. While we have illustrated the concept of Blended Interaction Spaces through the discussion of the BISi system, this system is really a starting point in thinking about the development of a larger ecology of these spaces. This is an important reason why our concerns go beyond the specifics of the Halo and BISi systems to encourage thinking about this broader potential ecology of Blended Interaction Spaces. Each of these points in the design space can draw inspiration from the varied physical workspaces that make up the spatial ecology of the modern office place. This approach attempts to deal with the varied nature of work practices through a combined system of purpose designed spaces as opposed to designing single general purpose spaces or technologies that accommodate varied practices and circumstances through configurability and tailorability. In taking this system-of-spaces approach, Blended Interaction Spaces trade off the values of purpose-built design with the values of flexibility and extendibility (e.g. extending to more and more nodes than one gets, for example, with Access Grid type technology). In this respect, we can draw some parallels between the appliance (purpose built) and convergence (general purpose) debate discussed by authors such as Norman (1998) with Blended Interaction Spaces being at the appliance end of the design approach – well designed spaces with very particular purpose. As with the appliance/convergence model these devices can co-exist (cf. Dourish 2001). In advocating the importance of Blended Spaces, our intent is not that these should be at the expense of other more general-purpose video mediated communication tools. Rather, it is envisioned that these different kinds of systems will coexist according to the particular values they have acquired.

In presenting the ideas here, and in particular the emphasis on providing shared spatial geometries across distributed environments, there is some concern that the Blended Interaction Space approach be viewed simply as an uncritical and blind pursuit of the "being there" experience. However, just because we are highlighting spatial characteristics and the importance of physical space design does not mean that we are advocating what Fitzpatrick (2003) and others such as Harrison and Dourish (1996) might characterize as a "Space-based" approach to the design of these systems. Rather our concerns are much more in line with the "Place-based" approach (Fitzpatrick 2003, Harrison and Dourish 1996) to the design of collaboration systems – that is the design and use of space for place-making where meaning making takes place through the actions in the space (Fitzpatrick 2003). With this in mind, our reasons for trying to maintain spatial geometries in the design of distributed environments can be grounded in the ideas of embodied action. From these it is argued that the representation of geometrically accurate spatial continuity across distributed spaces can facilitate visible access to and understanding of the gestures, actions and orientations of others in relation to the spatial environment and understanding of how our own actions, gestures and orientations appear to others in relation to the environment. Having a shared spatial geometry can help overcome negative consequences arising from orientation and perspective disparities that can affect in-the-moment coordination of conversational mechanics. Through this, it is possible to overcome some of the difficulties associated with the fractured and disembodied experiences with some early video mediated communication and have a space that affords "place-making" through the more embodied experience.

When making these arguments, it is important to treat them with an appropriate perspective. As we discussed earlier, the claims being made are not about the *necessity* or *sufficiency* of shared spatial geometries as a basis for successful collaboration. The numerous examples of distributed collaboration technologies that are usefully incorporated into people's lives, making no attempt to represent spatial geometries, are testament to this (telephone, instant messaging, etc). So while people can and do communicate without coherently represented spatial geometries, our arguments are that the provision of shared spatial geometries can facilitate and make easier aspects of communication, coordination and collaboration. As Healey and Battersby (2009) argue, people make use of them when they are available, offloading the mental demands of conversational and gestural reference to the spatial geometries of physical and interactional space. The facilitation effects of these cues are particularly relevant where visual communication channels are used to provide spatial reference points and which are "assumed" to be congruous but where orientation disparities lead to a confusing incongruity – such as in video mediated communication.

9. CONCLUSION

In this paper, we have presented an analysis of the concept of Blended Interaction Spaces. Our chosen approach here has been to combine the theoretical underpinnings of the concept with important design issues illustrated through analysis of concrete system examples. In adopting this approach, our aims go beyond the specifics of the systems explored in the paper. Rather, our intention has been to provide some of the important analytic foundations through which such systems can be conceived and studied in a more principled way.

Out of this analytic foundation there arises a broad set of interesting empirical issues and questions that can form part of a larger research agenda. These relate to a number of different areas. The first concerns the impact of these environments and proxemic concerns on the mechanical aspects of communication and collaboration. As with other media space research, these concerns apply not simply to performance *outcome* measures but also to characterisation of *process* of communication and collaboration. That is, these systems can be evaluated in terms of particular spatial configurations that result in more effective decision-making or more effective collaborative interpretations. They can also be measured in terms of the ways that they change communication and collaboration style – for example, do they result in different utterance or turn taking patterns, or different patterns of gesture. As well as these quantitative analyses, it is important to complement this with more detailed qualitative interaction analysis of these kinds of systems such as those applied by Heath and Luff (1991, 1992, 2000) in earlier media space research. In the same way that they highlighted some of the problems of media space asymmetries, these kinds of analyses can help relate particular interaction behaviours and utterances to the geometrically faithful representations of Blended Interaction Spaces. This kind of research into the mechanical aspects of communication and collaboration applies not just to the specific Halo and BISi systems but also to a broader ecology of Blended Interaction Spaces with different proxemic configurations.

A second area of empirical interest arising from the discussion here relates to the specific question of interaction proxemics. The systems discussed here highlighted issues related to the particular interaction mechanisms and configurations used in their design. This points to a larger agenda of understanding the spatial consequences of existing input technologies and interaction techniques. Experimental studies around this issue can of course utilise spatial tracking technologies of bodies and body parts as well as input logging technologies to create important data components in the visualization and systematic analyses of these spatial consequences. However, the research agenda here is more complicated than simply taking a particular interaction technique and mapping out the particular spatial behaviours around it. As we have argued earlier, these spatial consequences will also be shaped by the presence of other competing influences on spatial behaviour within blended space environments. For example, the need to be positioned on camera, or the presence of other interaction mechanisms are likely to interact with the basic interaction proxemics of a particular input device. As such, in conducting experimental analyses of interaction proxemics, it will be important not just to examine interaction mechanisms in isolation but within the context of other artefacts and issues with distributed collaboration and Blended Interaction Spaces.

A final area of empirical work, drawing on arguments made by Schmidt (2009), concerns the less mechanical aspects of communication and collaboration. The current presentation of Blended Interaction Spaces put forward in this paper has focused on the mechanical aspects of collaboration and conversation and how this is facilitated by particular material properties and design characteristics of these spaces. This focus has enabled a level of clarity in the presentation of the arguments but it is important to recognize this as just the initial part of the understanding of these kinds of Blended Interaction Spaces. Missing from this more mechanical emphasis are any details of actual work practice and behaviors within particular organizational contexts. There is a need then for an empirical assessment of Blended Space systems, such as Halo, BISi and others, within real world organizational contexts. Rather than focusing on the mechanical properties of communication, such an empirical agenda will aim to understand how and why particular details and characteristics of work practice within particular organisational settings relate to the geometrical properties of these distributed spaces. Through more ethnographic enquiry into these work practices, it should be possible to articulate further reasons why, in particular organisational contexts, people do and don't orient to common spatial references in the shaping of their work practices and the ways embodied interaction within a range of organisationally situated Blended Interaction Spaces creates meaning and the production of "place".

ACKNOWLEDGEMENTS

The ideas developed here have been informed and influenced by a number of key people. We would like to thank the HP Halo team, in particular, Mike Derocher, Mark Gorzynski and April Slayden-Mitchell, who formulated the original Blended Space concept and, through many helpful discussions, informed some of the ideas articulated in this paper. We would like to thank too those at Steelcase Workplace Futures, in particular Terry West and Mark Bologa, for their insights into the relationship between architecture, furniture and technology that have been an important inspiration for this work. Thanks also, to all the members of the HxI project in Australia who contributed to the development of the BISi system. Finally, many thanks to Richard Harper and all the reviewers for their critical insights and comments on earlier drafts of this paper.

REFERENCES

Becker, F. and Steele, F. 1995. *Workplace by Design: Mapping the High Performance Workscape*, Jossey-Bass, San Francisco.

Bly, S. 1988. A use of drawing surfaces in different collaborative settings. In *Proceedings of CSCW 88 Conference on Computer Supported Cooperative Work*. 250-256.

Bly, S. and Minneman, S. 1990. Commune: A Shared Drawing Surface. In *Proceedings of the 5th Conference on Office Information Systems*. 184-192.

Broughton, B., Paay, J., Kjeldskov, J., O'Hara, K., Li, J., Phillips, M., and Rittenbruch, M. 2009. Being here: designing for distributed hands-on collaboration in blended interaction spaces. In *Proceedings of OzCHI 2009*. 73-80.

Buxton, W., Sellen, A., and Sheasby, M. 1997. Interfaces for Multiparty Videoconferencing. In Finn, K., Sellen, A., Wilber, S. (Eds.). *Video Mediated Communication,* Hillsdale, N.J., Erlbaum, 385-400.

Buxton, W. 2009. Mediaspace - Meaningspace - Meetingspace. In Harrison, S. (Ed) *Media Space: 20+ Years of Mediated Life,* Springer, London, 217-231.

Covi, L. M., Olson, J. S., and Rocco, E. 1998. A room of your own: What do we learn about support of teamwork from assessing teams in dedicated project rooms? In Streitz, N., Konomi, S., Burkhardt, H. J. (Eds) *Cooperative Buildings*, Springer, Amsterdam, 53-65.

Dourish, P. 2001. *Where the Action Is: The Foundations of Embodied Interaction*, MIT Press, Cambridge.

Dourish, P. 2006. Re-Space-ing Place: Place and Space Ten Years On. In *Proceedings of ACM Conf. Computer-Supported Cooperative Work CSCW 2006*. 299-308.

Dourish, P., Adler, A., Bellotti, V., and Henderson, A. 1996. Your Place or Mine? Learning from Long-Term Use of Audio-Video Communication. *Computer-Supported Cooperative Work*, 5, 1, 33-62.

Duffy, F. 1997. *The New Office*, Conran Octopus, London.

Finn, K., Sellen, A., and Wilbur, S. 1997. *Video-mediated Communication*, Erlbaum, New Jersey.

Fitzpatrick, G. 2003. *The Locales Framework: Understanding and Designing for Wicked Problems*, Kluwer Academic Publishers.

Gaver, W., Sellen, A., Heath, C., and Luff, P. 1993. One is not enough: Multiple views in a media space. In *Proceedings of INTERCHI '93*, 335-341.

Goldstein, E.B. 1987. Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance,* 13, 2, 256–266.

Gorzynski, M., Derocher, M., and Slayden Mitchell, A. 2009. The Halo B2B Studio. In Harrison, S. (Ed) *Media Space 20+ Years of Mediated Life*, Springer, London.

Hall, E. T. 1966. *The Hidden Dimension,* Doubleday, New York.

Harper, R.H.R. 1998. *Inside the IMF: An ethnography of documents, technology and organizational action*, Academic Press, London.

Harrison, S. 2009. *Media Space: 20+ Years of Mediated Life*. Springer, London:.

Harrison, S. and Dourish, P. 1996. Re-Place-ing Space: The Roles of Place and Space in Collaborative Systems. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work CSCW'96*, ACM, 67-76.

Healey, P.H. and Battersby, S.A. 2009. The Interactional Geometry of a Three-way Conversation. In *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society*, 785 - 790.

Heath, C. and Luff, P. 1991. Disembodied conduct: Communication through video in a multi-media environment. In *Proceedings of CHI '91*, 99-103.

Heath, C. and Luff, P. 1992. Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction. *Human-Computer Interaction*, 7, 315-346.

Heath, C. and Luff, P. 2000. *Technology in Action*, Cambridge University Press, Cambridge.

Hirsh, S., Sellen, A., and Brokopp, N. 2005. *Why HP People Do and Don't Use Videoconferencing Systems*. HP Technical Report, HPL-2004-140(R.1).

Hornecker, E. and Buur, J. 2006. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of CHI 2006*, ACM, 437–446.

Hornecker, E., Marshall, P., Dalton, N.S., and Rogers, Y. 2008. Collaboration and interference: Awareness with Mice or Touch Input. In *Proceedings of CSCW '08*, ACM, 167-176.

Huang, E. M., Mynatt, E. D., and Trimble, J. P. 2006. Displays in the wild: Understanding the dynamics and evolution of a display ecology. In *Proceedings of Pervasive 2006*.

Hutchins, E. 1995. *Cognition in the Wild*. The MIT Press, Cambridge, MA.

Ishii, Hz. and Kobayashi, M. 1992. ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact. In *Proceedings of CHI'92*, ACM, 525-532.

Jiang, H., Wigdor, D., Forlines, C., and Shen, C. 2008. System Design for the WeSpace: Linking Personal Devices to a Table-Centered Multi-User, Multi-Surface Environment. In *Proceedings of Tabletop 2008*.

Jiang, L., Krumm-Heller, A., and Mueller-Tomfelde, C. 2007. The Virtual Terminal A Cross Platform and High Performance Remote Desktop Sharing Program. Presented at *Fourth CSIRO ICT Centre Conference 2007*, Sydney, Australia.

Johanson, B., Fox, A., and Winograd, T. 2002. The Stanford iRoom and Interactive Workspaces Project. In *Proceedings of IEEE Pervasive Computing*, 67-75.

Kendon, A. 1990. Spatial Organization in Social Encounters: the F-formation System. A. Kendon, A. (Ed) *Conducting Interaction: Patterns of Behavior in Focused Encounters*, Cambridge University Press, 209-237.

Kirsh, D. 1995. The intelligent use of space. *Artificial Intelligence*. 73, 1-2, 31-68.

Kruger, R., Carpendale, S., Scott, S.D., and Greenberg, S. 2003. How people use orientation on tables: comprehension, coordination and communication. In *Proceedings of GROUP 2003*. ACM, 369-378.

Laing, A., Duffy, F., Jaunzens, D., and Willis, S. 1998. *New Environments for Working: The Re-Design of Offices and Environmental Systems for New Ways of Working*, Construction Research Communications.

Larkin, J.H. 1989. Display-Based Problem Solving. In Klahr, D. and Kotovsky, K. (Eds) *Complex Information Processing: The impact of Herbert A. Simon*, Erlbaum, Hillsdale, NJ.

Mantei, M., Baecker, R., Sellen, A., Buxton, W., Milligan, T., and Wellman, B. 1991. Experiences in the use of a media space. In *Proceedings of CHI '91, ACM Conference on Human Factors in Software*, 203-208. Reprinted in Marca, D. and Bock, G. (Eds). 1992. *Groupware: software for computer-supported collaborative work*, IEEE Computer Society Press, Los Alamitos, CA, 372 - 377.

Merleau-Ponty, M 1962. *Phenomenology of Perception.* Routledge, UK.

Merleau-Ponty, M. 1968. The Intertwining - the Chiasm. In *The Visible and the Invisible.* Northwestern University Press, Ilhnios, USA.

Moran, T., Saund, E., Van Melle, W., Gujar A., Fishkin K., and Harrison, B. 1999. Design and technology for Collaborage: collaborative collages of information on physical walls. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*, 197-206.

Nacenta, M. A., Gutwin, C., Aliakseyeu, D., and Subramanian, S. 2009. There and Back Again: Cross-Display Object Movement in Multi-Display Environments. *Human-Computer Interaction*, 24, 1,170-229.

Nacenta, M.A., Pinelle, D., Stuckel, D., and Gutwin, C. 2007. The Effects of Interaction Technique on Coordination in Tabletop Groupware. In *Proceedings of GI 2007*, ACM, 191-198.

Nguyen, D. and Canny, J. 2005. MultiView: spatially faithful group video conferencing. In *Proceedings of CHI '05*, ACM, 799-808.

Nguyen, D. and Canny, J. 2009. More than face-to-face: empathy effects of video framing. In *Proceedings of CHI '09*, ACM, 423-432.

Noll, A. 1992. Anatomy of a Failure: Picturephone revisited. *Telecommunications Policy*, 16, 307-316.

Norman, D. 1998. *The Invisible Computer: Why Good Products Can Fail, the Personal Computer Is So Complex, and Information Appliances Are the Solution*. MIT Press, Cambridge, MA.

O'Connaill, B., Whittaker, S., and Wilbur, S. 1993. Conversations over video conferences: An evaluation of video mediated interaction. *Human-Computer Interaction*, 8, 389-428.

Olson, G. M. and Olson, J. S. 2000. Distance Matters. *Human-Computer Interaction*, 15, 2/3, 139-178.

Olson, J. S., Olson, G. M., and Meader, D. 1997. Face-to-face group work compared to remote group work with and without video. In Finn, K., Sellen, A., Wilbur, S. (Eds) *Video Mediated Communication*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Okada, K., Maeda, F., Ichikawaa, Y., and Matsushita, Y. 1994. Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design. In *Proceedings of CSCW 94*, 385-393.

Ozyurek, A. 2000. The influence of addressee location on spatial language and representational gestures of direction. In McNeil, D. (Ed), *Language and gesture*, Cambridge University Press, 64-83.

Ozyurek, A. 2002. Do speakers design their co-speech gestures for their addresees? The effects of addressee location on representational gestures*. Journal of Memory and Language*, 46, 688-704.

Phillips, M. 2008. *Livespaces technical overview*. Defence Science and Technology Organisation, Technical Report DSTO-TR-2188.

Robertson, T. 1997. Cooperative Work and Lived Cognition: A Taxonomy of Embodied Actions. In *Proceedings of ECSCW '97*, 205-220.

Rogers, Y. and Bellotti, V. 1997. From Web press to Web pressure: multimedia representations and multimedia publishing. In *Proceedings of CHI '97*, ACM, 279-286.

Rogers, Y. and Lindley, S. 2004. Collaborating around vertical and horizontal displays: which way is best? *Interacting With Computers*, 16, 1133-1152.

Rogers, Y., Lim, Y.-K., Hazlewood, W., and Marshall, P. 2009. Equal Opportunities: Do Shareable Interfaces Promote More Group Participation Than Single User Displays? *Human-Computer Interaction,* 24, 2, 79-116.

Russell, D. and Sue, A. 2003. Large interactive public displays: use patterns, support patterns, community patterns. In O'Hara, K., Perry, M., Churchill, E., Russell, D. (Eds) *Public and Situated Displays: Social and Interactional aspects of Shared Display Technologies*. Kluwer Academic Publishers, London.

Schmidt, K. 2009. Divided by a common acronym: On the fragmentation of CSCW. In *Proceedings of ECSCW'*09.

Sellen, A. J. 1992. Speech patterns in video mediated conversations. In *Proceedings of CHI'92*, ACM, 49-59.

Sellen, A.J. 1995. Remote conversations: The effects of mediating talk with technology. *Human-Computer Interaction*, 10, 4, 401-444.

Sellen, A. J., Buxton, W., and Arnott, J. 1992. Using spatial cues to improve videoconferencing. In *Proceedings of CHI '92*, 651-652. (Videotape in CHI '92 Video Proceedings)

Sellen, A.J. and Harper, R.H.R. 1997. Video in support of organisational talk. In Finn, K., Sellen, A., Wilbur, S. (Eds), *Video-mediated Communication*, Erlbaum, NJ, 225-243.

Sellen A.J. and Harper R.H.R. 2000. *The myth of the paperless office*, MIT Press, Cambridge, MA.

Short, J., Williams, E., and Christie, B. 1976. *The social psychology of telecommunications*, Wiley, London.

Sommer, R. 1969. *Personal Space*, Prentice-Hall, Englewood Cliffs, N.J.

Sommer, R. and Ross, H. 1958. Social Interaction on a Geriatric Ward. *International Journal of Social Psychology*, 4, 123-133.

Spinelli, G. 2003. *Socially Distributed Cognition: Computational Space and Collaborative Artefacts For The Workspace*, PhD Thesis, Brunel University.

Spinelli, G. and O'Hara, K. 2001. *Observing High Performance Team Work: How Space and Information Artefacts Structure Team Problem Solving, Communication and Fluid Movement From Individual to Collaborative Work*. Appliance Studio, UK.

Spinelli, G., Perry, M., and O'Hara, K. 2005. Understanding complex cognitive systems: the role of space in the organisation of collaborative work. *Cognition Technology and Work*, 7, 2, 111-118.

Steelcase360 2007. Working in Four-Part Harmony. *Steelcase360*, 1, 44, 2007. http://content.ll-0.com/360steelcase/June2007-ReadingLayout.pdf?i=060607172333

Steinzor, R., B. 1950. The spatial factor in face to face discussion groups. *Journal of Abnormal and Social Psychology*, 45, 552-555.

Streitz, N.A., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. 1999. i-LAND: An interactive Landscape for Creativity and Innovation. In *Proceedings of CHI '99*, ACM, 120- 127.

Streitz, N.A., Rexroth, P., and Holmer, T. 1997. Does roomware matter? Investigating the role of personal and public information devices and their combination in meeting room collaboration. In *Proceedings of ECSCW '97 European Conference on Computer-Supported Cooperative Work*, 297-312.

Tang, J. and Isaacs, E. 1993. Why Do Users Like Video? Studies of Multimedia-Supported Collaboration. *Computer-Supported Cooperative Work*, 1, 3, 163-196.

Trimble, J., Wales, R., and Gossweiler, R. 2003. NASA's MERBoard. In O'Hara, K., Perry, M., Churchill, E., Russell, D. (Eds) *Public and Situated Displays: Social and Interactional aspects of Shared Display Technologies*, Kluwer Academic Publishers, London.

Whittaker, S. 2003. Things to talk about when talking about things. *Human-Computer Interaction*, 18, 149-170.

Whittaker, S. and Schwartz, H. 1995. Back to the future: Pen and paper technology supports complex group coordination. In *Proceedings of CHI'95*, ACM, 495-502.

Weinstein I. M. 2005. *HP Halo – Performance Assessment: Analysis and opinions on capabilities, features, usability, performance, and business model,* Wainhouse Research Report.

Wigor, D., Jiang, H., Forlines, C., Borkin, M., and Shen, C. 2009. WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. In Proceedings of CHI '09, ACM, 1237-1246.

Yamashita, N., Hirata, K., Aoyagi, S., Kuzuoka, H., and Yasunori Harada, Y. 2008. Impact of Seating Positions on Group Video Communication. In *Proceedings of CSCW '08*, 177-186.

Zhang, J., and Norman, D. A. 1994. Representations in Distributed Cognitive tasks. *Cognitive Science*,18, 87-122.

Chapter 23

Orchestrating mobile devices

John Murphy, Jesper Kjeldskov, Steve Howard,
Graeme Shanks and Elizabeth Hartnell-Young

Abstract. The last decade has seen convergence marketed as one response to the challenge of users having to juggle an increasingly wide array of digital services, technologies and media. Key to this view is the assumption that by converging computer devices, and digital media, the value of technology for end users can be maximised whilst the overheads involved in purchasing, maintaining and orchestrating a variety of different technology solutions can be minimised. In contrast however, some authors have argued that convergence creates weak-general solutions, and rather we should be aiming for strong-specific technology by means of the deliberate design of multiple diverged devices. This paper contributes to the ongoing discussion of convergence and divergence. We discuss three apparently irreconcilable perspectives on the relationship between functionality and usability, and show that they are in fact complementary views of convergence. To ground this discussion we draw on the results of a recent cultural probes study of a cohort of early adopters of converged devices.

1. INTRODUCTION

Within the last decade of ICT development, convergence has often been suggested as the ultimate answer to the challenge of users having to juggle an increasing number of different digital technologies and media. The driving force behind this line of thought is that by converging computer devices, and digital media, the value of technology for end users can be maximised while the overhead required for maintaining and combining the different technologies is minimised. Many examples of this approach to technology design exist. Today's digital video cameras often offer high-quality still-photography (see e.g. Sony 2005) and digital still cameras ape some of the functionality more typical of video recorders (see e.g. Canon 2005). USB memory sticks come with the capability

Originally published as Murphy, J., Kjeldskov, J., Howard, S., Shanks, G. and Hartnell-Young, E. (2005) The Converged Appliance: "I Love it… But I Hate it". In *Proceedings of OzCHI 2005,* Canberra, Australia (pp. 1-10). New York: ACM.

of playing their music files directly through headphones (see e.g. Packard Bell 2005) and MP3-players moonlight as large external high-speed hard disk drives (see e.g. Apple 2005a). Other music players show an ability to store digital photos and can be plugged directly into a TV (Apple 2005b). Personal digital assistants (PDAs) are increasingly being combined with mobile phones and multimedia technologies, extending their traditional calendar and contact functionality to include capabilities for voice communication, Internet browsing, digital still photography and video recording as well as personal audio/video playback (see e.g. HP 2005, PalmOne 2005, Archos 2005). Approaching the same functionality, but from a very different origin, the same trend is now clear within the mobile phone market (see e.g. Nokia 2005).

While many of the examples above are 'strong-specific' (Buxton, 2001) solutions that have inherited one or more relatively weak implementations of other technologies (such as the personal video players shipping with light-weight PDA functionality or mobile phones with poor quality FM radios), this boundary between primary and secondary functionality is now beginning to blur. As an example of this, digital cameras built into mobile phones are now achieving picture quality approaching their stand-alone counterparts, making one ask if the converged device is a phone with a camera or a camera with a phone (figure 1). In relation to the challenge of interaction design, should such technology mutations inherit the interaction paradigm of the telephone or the camera (Sacher and Louden 2002) Or are either of these paradigms really appropriate?

Figure 1. Camera phone or phone-camera?

In contrast to the approach to technology design depicted above, others have argued that convergence creates 'weak-general' solutions, with usability comparable to the Swiss army knife: clumsy technology with a wide range of functions, none of which are ideal in isolation (see e.g. Norman 1998, Bergman 2000, Buxton 2001). Responding to the 'one size fits all' view, they suggest a single-function/many-device or 'information appliance' approach where each device is "designed to perform a specific activity, such as music, photography, or writing" (Bergman 2000). The driving force behind this line of thought is that having a wide range of good specialized tools to choose from is better than having a general one that does not perform any individual task particularly well. Specialized tools facilitate optimisation of functionality over time and the refinement of well-known paradigms of use. In contrast, weak-general tools share much in common with the proverbial camel ('a horse designed by a committee'); they are hostage to compromise, limited by the accumulated complexity of too many ways of possible use.

As pointed out by Pemberton (2001), although convergence has been a popular topic of discussion within the field of HCI for more than a decade, very little empirical data has been reported on the use of converged solutions. This paper contributes to this ongoing discussion, grounding the contribution in empirical data drawn from studies of technology in use. We present and discuss three contrasting perspectives on convergence, and report early reflections on data from a cultural probes study of technology use.

Section 2 of the paper presents and discusses three perspectives on convergences, all drawn from previous work. Section 3 outlines four concepts useful in understanding convergence and divergence in both the design and use of technology. Section 4 briefly presents an empirical study of technology use. Section 5 presents some of the key findings drawn from the analysis of the empirical data. Finally, section 6 discusses the findings in the light of the three perspectives on convergence discussed earlier; we argue that the three apparently irreconcilable perspectives are in point of fact quite complementary.

2. THREE PERSPECTIVES ON CONVERGENCE

The HCI literature on technology design and usability includes, on first reading, three quite different views of the relationship between functional scope and user experience. Let us call them:

- Utopian perspective: User experience is proportionally related to functional scope, "more means more".
- Dystopian perspective: User experience is inversely proportional to functional scope, "less is more"
- Hybrid perspective: User experience is positively related to functional scope but only up to a threshold value, the 'tipping point', after which the Dystopian view prevails.

2.1. Utopian

Drawing parallels with Moore's Law (i.e. the number of transistors fitting on a single chip doubles every 18 months) a Utopian view of the relationship between functionality and usability seems to be that the more functionality that can be accommodated by one single technology, the higher its usability will be. This view reflects the basic assumptions underlying the push to convergence discussed above. If one device can function as a music player, wireless headset and data storage device, the associated user experience will be higher than with a 'toolbox' of individual devices each performing only one function (see figure 2, Blueant Bluetooth stereo headphones and audio player).

From a Utopian perspective, increasing convergence is the path to enhanced user experience. However, defining the factors relevant to determining a good functional grouping is a non-trivial problem. Further, the Utopian view, held by many technology vendors, does not explain the to date limited success of several converged technologies such as the PDA-phone or the memory stick-MP3 player, and the failure of earlier converged solutions such as video telephony.

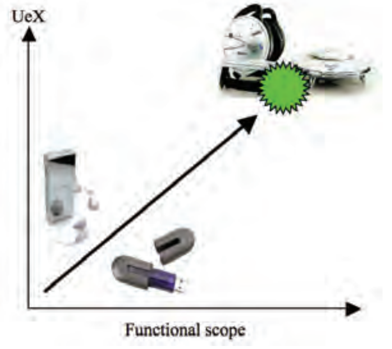


Figure 2. Utopian view of the relationship between functionality and usability

2.2. Dystopian

Motivated by the observations that the functional complexity of many ICT's has already, in a real sense, exceeded the threshold of human problem solving capacity, and that this capacity limitation is stable (referred to as "God's Law") Buxton (2001) stands in stark contrast to the Utopians in proposing an inverse relationship between usability and functionality. As the number of functions supported by a device climbs, its usability approximates zero. Unless convergence design is conducted with genuine sensitivity to users, their characteristics and their practices, convergence merely adds complexity to the technology, encouraging workarounds, increasing frustration and introducing inefficiencies. Exemplifying the Dystopian view, dedicated mobile phones and torches each have higher usability than a mobile phone with built-in flashlight (as illustrated in figure 3).

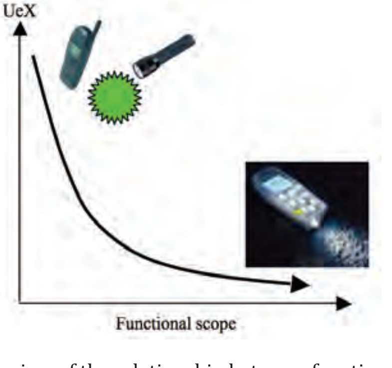


Figure 3. Dystopian view of the relationship between functionality and usability

The Dystopians sit in the 'information appliance' camp introduced above. Introducing a "threshold level of frustration" (Buxton 2001) where the level of functionality intersects with the threshold level of human capacity, Buxton's view may help explain why some highly advanced novel technologies such as virtual reality, interactive television, MMS (Multimedia Messaging Service) and video telephony have never really achieved a commercial foothold whilst other strong-specific solutions, such as mobile telephony and SMS (Short Message Service), have had huge success.

While it is relatively simple to implement Buxton's philosophy (simply limit functional scope), this approach does not help in understanding why some technologies have indeed been successfully converged, offering both a high level of functional capability and a positive user experience, such as the camera-phone.

2.3. Hybrid

In contrast to both the Utopian and Dystopian views, researchers at Nokia have argued that user experience is not simply proportional (inversely or otherwise) to functional scope (Kiljander and Järnström 2003). Rather they argue that the usability of a specific technology can remain relatively stable (high, low or otherwise) whilst functionality is increased up to a certain threshold value, beyond which usability will drop rapidly (as illustrated in figure 4).

Figure 4. Hybrid view of the relationship between functionality and usability

They term this tipping point in the relationship between functionality and usability the 'usability knee'. The line of thought behind the usability knee is that each technology platform (such as a specific series of mobile phones) has a certain threshold or critical mass of functionality or functional complexity determined by, for example, it's input and output capabilities. Thus, the position of the usability knee on the horizontal (functional scope) axis can be moved further to the right by improving the capabilities of the platform, for example, increasing screen size, screen resolution or creating better input capabilities.

Using Buxton's (2001) terminology, the art of technology design according to Nokia lies in striking the right balance between weak-general (i.e. converged) and strong-specific (i.e. diverged) solutions, the right balance between functionality and usability. The trick is to maximise functional power without going beyond the knee of usability.

3. CONVERGENCE AND DIVERGENCE BY DESIGN AND IN USE

All three perspectives outlined above are rather technology-centred. Clearly though, user experience is not only influenced by decisions made during design and development, but also through the act of technology use itself, and the contexts within which that use occurs. Howard et al. (2004) proposed that the interplay between technology, designers and users could be conceptualised in terms of:

- Convergence-by-design
- Divergence-by-design
- Convergence-in-use
- Divergence-in-use

These are illustrated in figure 5 below. Whereas convergence-by-design and divergence-by-design describe the two opposing approaches to technology design, and are unpacked above in terms of the three perspectives, convergence-in-use and divergence-in-use describe opposing approaches to convergence and divergence as practiced by users. Convergence-in-use and divergence-in-use describe how users, not designers, converge and diverge technologies by, for example, integrating them, combining them, taking them apart or using them consciously to supplement each other.

Convergence-by-design

Divergence-by-design

Convergence-in-use

Divergence-in-use

Figure 5. Convergence- and divergence-by-design and convergence- and divergence-in-use

Where convergence-by-design creates solutions which can be seen as general (and sometimes criticised for being weak), and divergence-by-design creates solutions which can be described as strong (and criticised for being too specific), the concepts of convergence- and divergence-*in-use* capture the mechanisms by which users compensate for the frustrations of a designed technology by creating solutions themselves, which fit their specific needs and contexts of use.

Convergence-in-use describes the phenomenon of users putting together and manually configuring different technologies for certain purposes beyond their originally intended use, exploiting the joint potentials of a range of technologies or overcoming the limitations of one technology by means of another. Convergence-in-use does not apply only to technologies that have been diverged-by-design (i.e. putting them back together), but also to the further 'use-side' convergence of technologies that has already been converged 'design-side'. A basic convergence-in-use solution to controlling a home-entertainment system consists of multiple remote control units physically combined (see figure 5). Stretching this example a little, the increasingly popular home-entertainment system is no longer limited to traditional audio and video components but now includes personal computers functioning as media servers, wireless local networks for remote control and streaming of media files between distributed devices.

Divergence-in-use conversely describes the phenomenon of users deliberately deploying multiple separate technologies to support a single activity, exploiting their individual strengths and compensating for any weakness by supplementing with other technologies. As with convergence-in-use, divergence-in-use does not only apply to the divergence of previously converged-by-design solutions (i.e. ripping them apart again), but also to the situation where a range of different technologies, converged- or diverged-by-design, are being used comfortably side-by-side. An example of divergence-in-use is the parallel use of electronic and paper-based calendars (see figure 5).

4. EMPIRICAL STUDY

We seek a better understanding of convergence and divergence as they are experienced in use, and therefore we aim to ground that understanding in the experience of end users of converged solutions. Here we focus on the use of converged devices by young urban professionals. We employed 'cultural probes' (Gaver et al. 1999, Vetere et al, 2005) in order to empower our early adopters in the telling of 'their stories' of convergence.

Studying people's use of technology is common within HCI. However, such user studies are often limited to relatively brief snapshots of use in limited or artificial settings. Whilst this approach can generate valuable knowledge about interface design, usability and the usefulness of technology solutions for specific activities or work tasks, it provides only limited insight into *real* technology use over extended periods of time. Studying technology use in-situ is a particular challenge as technology pervades our everyday life. Technology's influence extends beyond specific physical, temporal and social situations; its very ubiquity renders it 'unremarked upon'. Hence, traditional empirical methods within HCI, and more broadly within the social sciences, fall short in capturing a comprehensive picture. 'Cultural probes' are designed to give researchers access to these secluded and unfamiliar territories, in part by shifting the task of data collection onto the participants themselves. Supporting this, researchers equip their participants with a series of tools and materials to assist focus, data collection and data reporting, such as digital cameras, voice recorders, notebooks, diaries, scrapbooks, pens, scissors, glue and technology prototypes, all designed specifically for the purposes of the particular study or phenomenon of interest. Whilst being a relatively new approach to ethnographic user studies within the field of HCI, the use of cultural probes has grown in popularity within the last couple of years, and the method has been refined through several studies (see e.g. Gaver et al. 1999, Cheverst et al. 2003, Vetere et al. 2005).

The probe pack used in this study consisted of a diary, a scrapbook, a Polaroid camera and a set of catchphrases prompting the participants to reflect on their use of technology. The participants were asked to use the probes for a period of four weeks. Supplementing the probes, two interviews were carried out with each participant. The first interview occurred one week into the probe period and inquired into a series of general technology questions, and established the degree to which the participants were comfortable with the use of the probes. The second interview was carried out a few weeks after the probes had been returned to the researchers, and aimed to delve deeper into recorded incidents.

All participants were recruited through a professional recruitment company and each was paid approximately Aus$300 for their involvement. The study originally consisted of six participants, although half way through the study one participant decided not to continue with the probe pack and interviews. Thus, in total, the collected data amounted to 5 illustrated scrapbooks, 20 weeks of written diaries and 10 hours of interview recordings (each interview lasted roughly 1 hour).

The collected data were analysed in two rounds. Firstly, the probe data were provisionally reviewed after participants returned their diaries and scrapbooks, generating individual interview guides for the second round of interviews. Later, probe data and audio recordings from the 10 final interviews were analysed in order to identify themes related to convergence and divergence-in-use.

5. FINDINGS

The primary finding from this study is empirical evidence confirming Utopian, Dystopian and Hybrid experiences of convergence. We may have hoped, naively, to find support for one perspective only, and challenges for the remaining two. Instead, people are evidently using technology converged-by-design, actively engaging in convergence-in-use, frequently experiencing the frustrations of the usability knee, and responding to the usability knee by selecting a suite of alternative strong-specific solutions, thereby demonstrating divergence-in-use.

On the one hand, participants often reported the added value of converged devices with increased functionality compared to what they used to have. Examples of this include the highly popular Blackberry devices, mobile phones with cameras and email capabilities etc. Adopting the Utopian perspective that usability will increase with more functionality, this line of thought was often extended further through expressions of desire for one device that would be good at everything (primarily in the diaries and interviews notes) and through evidence of being engaged in the search for such a 'perfect device' (primarily through the scrapbooks, as illustrated in figure 6).

Figure 6. Extract from scrapbook: "in search of the perfect device"

On the other hand, participants also often reported frustrations related to the use of converged devices not providing a positive user experience, designs pushed beyond their usability knee (clear in diaries, scrapbooks and interviews). For example, several participants were using Blackberry devices for diary, contacts and mobile email access with high levels of satisfaction, though all complained about the usability of the built-in mobile phone when compared to a dedicated mobile phone. The points of criticism of the Blackberry's phone related to both technical issues such as poor user interface, limited contact list functionality, and sound quality, as well as form factor issues such as the device being too big for 'out of office' usage (slipping in small handbags or the back pocket of a pair of jeans).

In response to a breach of the usability knee, participants responded in line with Buxton's view, demonstrating either divergence-in-use by deploying multiple often redundant devices, or divergence-by-design by adapting highly specialized devices. For example, none of the participants actually used the Blackberry device's mobile phone. Instead, they all chose to carry an additional device, a strong-specific mobile phone.

In summary, the participants optimised their user experience by selecting functionally powerful devices, whilst also carefully spreading activity across multiple devices, in deliberated and intentional ways.

5.1. Increasing usability by means of convergence

The data contain multiple examples of convergence improving user experience by-design and in-use. As an example, most participants made extensive use of mobile access to email and diaries on their mobile devices such as mobile phones, iPAQs or Blackberry. Mobile devices with integrated cameras were also seen as highly useful for people working in the field who had a need for visual documentation – especially when integrated with functionality for sending images directly via email. One of the main motivations reported for adopting converged devices was integration and portability by means of keeping down the number of devices having to be carried around. Increased functionality by means of convergence was not only identified on the form factor level but also on the data level, such as receiving incoming faxes as image files attached to emails. Examples of convergence-in-use included use of the same technologies for work purposes, private purposes and leisure (such as PC's used for work email, personal net banking and online music downloads).

5.2. Beyond the tipping point

The location of the usability knee, the point at which user experience is harmed by increasing functional complexity, depends not only on the technical capabilities of the platform, but also on the users' context: work/private, office/field etc. Movement between the worlds of work and play can change the position of the usability knee (due in part to changing requirements), e.g. some participants were happy to carry a highly converged PDA with foldout keyboard, camera extension, etc. during work hours, but switched to a strong-specific mobile during the weekend.

Managing the dynamic position of the usability knee across the work/play boundary involves exploiting convergence during design (by providing a PDA with fold-out keyboard etc.) whilst still facilitating divergence-in-use (by keeping the phone separate, or at least detachable).

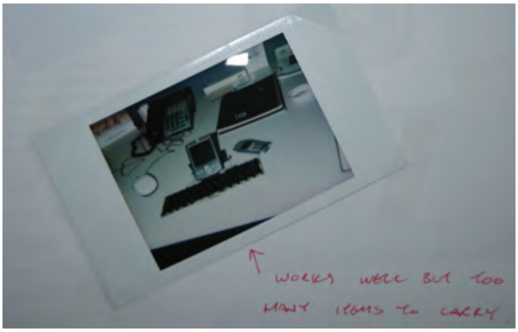


Figure 6. Extract from scrapbook: “I wish these technologies worked better together…”

5.3. Increasing usability by means of divergence

Evidence of divergence-by-design improving the end user experience was found in relation to both work and leisure. Some participants were frequent users of dedicated personal music players, such as iPods, even though their PDA or mobile phone was quite capable of functioning as a music player as well. However, because listening to music was typically done outside the work domain, while jogging, going to the gym or during other activities for leisure, the participants were clear that in those situations they did not always require the functionality of their PDA or mobile phone. On the contrary, some of them suggested that this device choice helped them draw a line between work- and leisure-time. Exemplifying divergence-in-use, some participants reported the use of multiple mobile and landline phones in order to separate work and private life. More outspokenly, divergence-in-use was also observed as a reaction to reaching a specific technology’s threshold of usability. Examples of this include integration of Blackberry devices with laptop computers for better viewing of attached documents and with mobile phones for better voice communication capabilities. Another example reported in the data concerned the combination of camera phones with high-quality stand-alone digital cameras according to the situation of use (e.g. low quality for leisure and high quality for work). Another observed trigger for divergence-in-use was the separation of work and private activity, due to e.g. fear of Trojan infestation from private use (music download) to business use (net banking).

6. DISCUSSION AND CONCLUSIONS

Design and use overlap, coexist and interplay in compelling ways; the design-side practice that results in interactive products, is continued as a use-side dialogue between the user and the technology. The social and technical influences active as convergence and divergence-*by-design* transmute into divergence and convergence-*in-use* are the key to matching converged solutions with converged practice.

A number of such social and technical influences are evident in the current data. For illustrative purposes we have highlighted one key influence, that of the interplay between work and leisure. The circumstantial differences between work and play often preferenced diverged solutions. The penchant for converged solutions capable of supporting the complex nomadic work habits of many of our early adopters, inverted during periods of leisure. ‘Week-end technology’ was characterised by stand-alone music players and strong specific, physically small mobile phones. Further our participants maintained the distinction between work and play by using different sim-cards for weekdays and weekends, and separating work related data from personal data by using multiple address books and calendars. The shift between general-work and specific-play technologies was of psychological significance to our participants. Just as playtime preferred strong specific solutions, so the use of strong specific devices marked that playtime, thus users denoted through device choice the difference between the activities.

The work/play influence is one example of many, and no one perspective on convergence captures the complexity we see in our data. Figure 7 shows that the three apparently irreconcilable views on convergence introduced earlier in the paper may be quite complementary, and reflect subtle shifts in the needs and circumstances of use. The conceptual graph again plots functional scope against user experience (UeX).

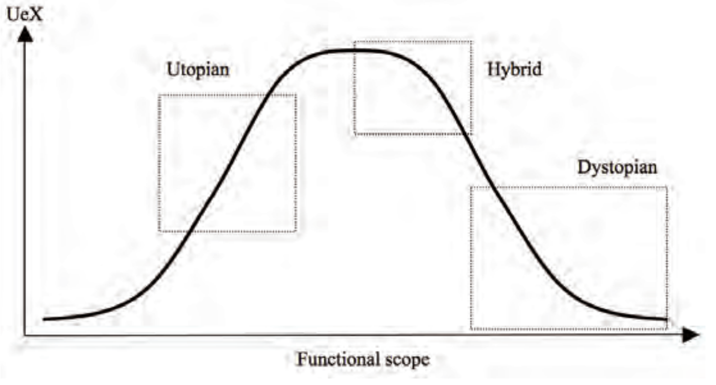


Figure 7. The Convergence Curve - Three complementary perspectives

Each perspective on convergence is a useful spur to design innovation. The Utopian perspective challenges us to seek out convergence opportunities, to explore synergistic collections of capabilities that marry seamlessly to users and the circumstances of use. In relation to work, and their particularly nomadic and multifarious work practices, all of our participants were Utopians. Counter to Norman and Buxton’s information appliance proselytising, our participants were eager to adopt cutting edge solutions that supported their converged work practices, keen to seek out an idealised ‘strong but general’ solution.

In other situations a distinctly Dystopian view emerges. At times of play and leisure, participants preferred niche appliances, and the advantages of strong specific solutions, perhaps tightly integrated through synchronisation support, over their converged counterparts were clearly evident.

Our cautionary tale in this paper, captured in the Hybrid view of Figure 7, is that if designers push convergence too far and in ignorance of user practice, users will push back. From a device vendor’s perspective, at best that pushback involves user frustration and divergence-in-use, at worst it results in the user rejecting the technology.

ACKNOWLEDGEMENTS

Thanks to all the participants in the empirical study. Data collection with cultural probes as reported here occurred as part of a broader research programme, using a variety of techniques and foci, and involving chief investigators S.Howard (University of Melbourne), J.Carroll (University of Melbourne), G.Shanks (Monash University), and supported on behalf of Novell P/L by J.Murphy. The research is funded by the Australian Research Council and Novell P/L through Linkage Grant LP0347459 and is contributing to Novell's 'Customers of the Future' programme through complementary research themes.

REFERENCES

Apple (2005a) Apple iPod. http://www.apple.com/ipod/ accessed 8/30/05.

Apple (2005b) Apple iPod Photo. http://www.apple.com/ipodphoto/ accessed 8/30/05.

Archos (2005) Archos Pocket Media Assistant PMA400. http://www.archos.com/products/overview/pma_400.html accessed 8/30/05.

Bergman, E. (2000) Information Appliances & Beyond. New York: Morgan Kaufman.

Buxton, W. (2001) Less is More (More or Less): Uncommon Sense and the Design of Computers. In P. Denning (Ed.), The Invisible Future: The seamless integration of technology in everyday life. (pp. 145-179). New York: McGraw Hill.

Canon (2005) Digital IXUS 50. http://www.canon.com.au/products/cameras/digital_compact_cameras.html accessed 8/30/05.

Cheverst, K., Dix, A., Fitton, D. and Rouncefield, M. (2003) "Out To Lunch": Exploring the Sharing of Personal Context through Office Door Displays. In Proceedings of OzCHI 2003 (pp. 74-83), Brisbane, Australia.

Gaver, B., Dunne, T. and Pacenti, E. (1999) Design: Cultural Probes. Interactions, 6(1), 21-29.

Howard, S., Hartnell-Young, E., Shanks, G., Murphy, J. and Carroll, Jennie (2004) When the Whole is Less than the Sum of the Parts: Humanising Convergence in Interactive Systems Design. In Proceedings of OzCHI 2004, University of Wollongong, Australia.

HP (2005) HP iPAQ h6315 Pocket PC – Phone Edition. http://www.hp.com/ accessed 8/30/05.

Kiljander, H., and Järnström, J. (2003) User Interface Styles. In C. Lindholm, T. Keinonen & H. Kiljander (Eds.), Mobile Usability: How Nokia Changed the Face of the Mobile Phone. (pp. 15-44). New York: McGraw-Hill.

Lindholm, C., Keinonen, T. and Kiljander, H. (2003) Mobile Usability: How Nokia Changed the Face of the Mobile Phone. New York: McGraw-Hill.

Nokia (2005) Nokia 9300. http://www.nokia.dk/campaign/9300/ accessed 8/30/05.

Norman, D.A. (1998) The Invisible Computer - Why Good Products Can Fail, the Personal Computer Is So Complex and Information Appliances Are the Solution. Cambridge, MA: MIT Press.

Packard Bell (2005) Packard Bell AudioKey. http://www.packardbell.co.uk/products/node1650.asp accessed 8/30/05.

PalmOne (2005) PalmOne Treo 650 Smartphone. http://www.palmone.com/us/products/smartphones/ accessed 8/30/05.

Pemberton, S. (2001) Did convergence kill the clock? Interactions of the ACM, 8(6), Nov-Dec, 2001.

Sony (2005) DCR-HC42 MiniDV Camcorder. http://www.sonystyle.com/ accessed 8/30/05.

Sacher, H. & Loudon, G. (2002) Uncovering the new wireless interaction paradigm. ACM Interactions, 9,1, 17-23. Jan-Feb 2002.

Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Pedell, S., Mecoles, K., and Mueller, F. (2005) Mediating Intimacy: Designing Technologies to Support Strong-Tie Relationships. In Proceedings of CHI'05. Portland, Oregon. ACM.

List of contributions

Chapter 2: Kjeldskov, J. and Graham, C. (2003) A Review of Mobile HCI Research Methods. In *Proceedings of Mobile HCI 2003*, Udine, Italy, LNCS (pp. 317-335). Berlin: Springer-Verlag.

Chapter 3: Paay, J. and Kjeldskov, J. (2005) Understanding and Modelling the Built Environment for Mobile Guide Interface Design. *Behaviour and Information Technology, 24*(1), 21-35.

Chapter 4: Paay, J. and Kjeldskov, J. (2008) Situated Social Interactions: a Case Study of Public Places in the City. *Computer-Supported Cooperative Work, 17*(2-3), 275-290.

Chapter 5: Kjeldskov J., Gibbs M., Vetere F., Howard S., Pedell S., Mecoles K. and Bunyan M. (2004) Using Cultural Probes to Explore Mediated Intimacy. *Australasian Journal of Information Systems, 11*(2), 102-115.

Chapter 6: Kjeldskov, J. and Stage, J. (2006) Exploring "Canned Communication" for Coordinating Distributed Mobile Work Activities. *Interacting with Computers, 2006*(18), 1310-1335.

Chapter 7: Kjeldskov, J. and Howard, S. (2004) Envisioning Mobile Information Services: Combining User- and Technology-Centered Design. In *Proceedings of APCHI 2004*, Rotorua, New Zealand, LNCS (180-190). Berlin: Springer-Verlag. With epilogue added in October 2010.

Chapter 8: Paay, J., Kjeldskov, J., Howard S. and Dave, B. (2009) Out on the town: a socio-physical approach to the design of a context aware urban guide. *Transactions on Computer-Human Interaction, 16*(2), 7-34.

Chapter 9: Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoles, K. and Bunyan, M. (2005) Mediating Intimacy: Designing Technologies to Support Strong-Tie Relationships. In *Proceedings of CHI 2005*, Portland, Oregon, USA (pp. 471-480). New York: ACM.

Chapter 10: Kjeldskov, J. and Stage, J. (2012) Combining ethnography and object-orientation: contextual richness and abstract models for mobile interaction design. *International Journal of Human-Computer Studies, 70*(3), 197–217.

Chapter 11: Kjeldskov, J. and Stage, J. (2004) New Techniques for Usability Evaluation of Mobile Systems. *International Journal of Human-Computer Studies, 60*(5-6), 599-620.

Chapter 12: Kjeldskov, J. and Skov, M. B. (2007) Studying Usability in Sitro: Simulating Real World Phenomena in Controlled Environments. *International Journal of Human-Computer Interaction, 22*(1), 7-37.

Chapter 13: Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S. and Davies, J. (2005) Evaluating the Usability of a Mobile Guide: The influence of Location, Participants and Resources. *Behaviour and Information Technology, 24*(1), 51-65.

Chapter 14: Høegh, R. T., Kjeldskov, J., Skov, M. B. and Stage J. (2008) A Field Laboratory for Evaluating In Situ. In J. Lumsden (Ed.), *Handbook of Research on User Interface Design and Evaluation for Mobile Technology* (pp. 982-996). PA, USA: Idea Group Inc (IGI).

Chapter 15: Kjeldskov, J. and Skov, M. B. (2007) Exploring Context-Awareness for Ubiquitous Computing in the Healthcare Domain. *Personal and Ubiquitous Computing, 11*(7), 549-562.

Chapter 16: Kjeldskov, J. and Paay, J. (2005) Just-for-Us: A Context-Aware Mobile Information System Facilitating Sociality. In *Proceedings of Mobile HCI 2005*, Salzburg, Austria (pp. 23-30). New York: ACM. (A condensed version appeared in IEEE Computer: Kjeldskov J. and Paay J. (2006) Public Pervasive Computing in the City: Making the Invisible Visible. *IEEE Computer, 39*(9), 60-65.)

Chapter 17: Kjeldskov, J., Christensen, C. M. and Rasmussen, K. K. (2010) GeoHealth: a location-based service for home healthcare workers. *Journal of Location-Based Services, 4*(1), 3-27.

Chapter 18: Skov, M. B, Kjeldskov, J., Paay, J., Husted, N., Nørskov, J. and Pedersen, K. (2012) Designing on-site: Facilitating Participatory Contextual Architecture with Mobile Phones. *Journal of Pervasive and Mobile Computing.* (in press)

Chapter 19: Kjeldskov, J., Skov, M. B., Paay, J. and Pathmanathan, R. (2012) Using Mobile Phones to Support Sustainability: A Field Study of Residential Electricity Consumption. In *Proceedings of CHI 2012*, Austin, Texas, USA. New York: ACM.

Chapter 20: Paay, J. and Kjeldskov, J. (2008) Understanding the user experience of location based services: five principles of perceptual organization applied. *Journal of Location-Based Services, 2*(4), 267-286.

Chapter 21: Kjeldskov, J. and Paay, J. (2010) Indexicality: understanding mobile human-computer interaction in context. *ACM Transactions on Computer-Human Interaction, 17*(4): article 14.

Chapter 22: O'Hara, K., Kjeldskov, J. and Paay, J. (2011) Blended Interaction Spaces for Distributed Team Collaboration. *ACM Transactions on Computer-Human Interaction, 18*(1), article 3.

Chapter 23: Murphy, J., Kjeldskov, J., Howard, S., Shanks, G. and Hartnell-Young, E. (2005) The Converged Appliance: "I Love it… But I Hate it". In *Proceedings of OzCHI 2005*, Canberra, Australia (pp. 1-10). New York: ACM.

List of co-authors

Bharat Dave	The University of Melbourne, Australia
Claus M. Christensen	Clearhaus A/S, Aarhus, Denmark
Connor Graham	National University of Singapore, Singapore
Elizabeth Hartnell-Young	DEECD, Melbourne, Australia
Florian Mueller	RMIT University, Melbourne, Australia
Frank Vetere	The University of Melbourne, Australia
Graeme Shanks	The University of Melbourne, Australia
Jacob Nørskov	Trifork, Aarhus, Denmark
Jan Stage	Aalborg University, Denmark
Jessica Davies	Geomatic Technologies, Melbourne, Australia
Jeni Paay	Aalborg University, Denmark
John Murphy	Design4Use, Melbourne, Australia
Karen Mecoles	The University of Melbourne, Australia
Kenneth Pedersen	Trifork, Aarhus, Denmark
Kenton O'Hara	Microsoft Research, Cambridge, UK
Klaus K. Rasmussen	eBay Classifieds Group, Aarhus
Marcus Bunyan	The University of Melbourne, Australia
Martin R. Gibbs	The University of Melbourne, Australia
Mikael B. Skov	Aalborg University, Denmark
Niels Husted	Trifork, Aarhus, Denmark
Rahuvaran Pathmanathan	Danske Bank, Copenhagen, Denmark
Rune T. Høegh	Nykredit Data, Aalborg, Denmark
Sandrine Balbo	Nuance Communications, Melbourne, Australia
Sonja Pedell	The University of Melbourne, Australia
Steve Howard	The University of Melbourne, Australia

Design af mobile interaktioner

- den kontinuerlige konvergens mellem form og kontekst

Jesper Kjeldskov

Ved udgangen af 2010 blev der, for første gang, solgt flere smartphones på verdensplan end personlige computere, hvilket indikerer, at vi for alvor er på vej ind i den epoke, der er blevet døbt “post-PC æraen” med allestedsnærværende computerkraft og “IT i alting”. Den massive udbredelse af mobile computere har haft enorm indflydelse på den måde, vi opfatter og bruger computerteknologi i vores arbejds- og privatliv. Interaktive mobile systemer og enheder er blevet funktionelle designobjekter, som vi interesserer os dybt for udseendet, udformningen, og oplevelsen af, og som vi i vores hverdag organiserer, dirigerer, og iscenesætter i samspil med et væld af andre computerteknologier. For at være succesfulde skal sådanne interaktive mobile systemer og enheder designes til at passe ind i den større kontekstuelle helhed som de indgår i. Dette opnås ikke godt gennem traditionelle metoder til bruger-centreret design og usability-engineering. I stedet kræver det design-orienterede tilgange til interaktionsdesign, der kan hjælpe os med at designe til helheder, skabe ny praksis, og håndtere de ofte vagt definerede, og skiftende, mål der opstår i løbet af en designproces. Denne afhandling præsenterer en sådan kontekstuel tilgang til design af interaktive mobile computersystemer. Den præsenterede tilgang hæver sig ud over eksisterende bruger- eller teknologi-orienterede tilgange ved i stedet at fokusere på helheden og samspillet mellem form og kontekst, og anvende dette som den centrale analytiske enhed. Herigennem fremmes en design-orienteret måde at opnå konvergens mellem form og kontekst gennem en kontekstuelt forankret, helhedsorienteret, og kontinuerligt udfoldende designproces.

“Fantasi er vigtigere end viden. For viden er begrænset, mens fantasien favner hele verden” sagde Albert Einstein i 1931. Inden for datalogi og interaktionsdesign i dag er fantasi lige så vigtig for at fremme vores viden og praksis som for videnskaben i 1930’erne. Uden fantasi og kreativitet er vi ikke i stand til at sætte os ud over hvordan vi tænker og agerer i dag. Dette er den tidløse måde at designe på, og mit udgangspunkt for beskæftige mig med mobil interaktionsdesign. Vores nuværende landskab af interaktive teknologier er i sig selv vokset ud af paradigmatiske skift i den måde, vi har tænkt om, og udviklet, computersystemer. Disse skift har bragt computerteknologi ind på arbejdspladsen, hjemmekontoret, og i privatsfæren, og har fået anvendelsen af computere til at handle om menneskers arbejde, samarbejde, kommunikation, brug af massemedier, og sociale netværk. Men hvad er næste skridt herfra? Hvordan kan vi igen sætte os ud over vores etablerede måder at tænke og agere, og aktivt fremme designet af morgendagens interaktive computersystemer? Dette er det overordnede spørgsmål, jeg i denne afhandling vil behandle i lyset af mine forskningsbidrag inden for området interaktionsdesign til mobile computersystemer.

Afhandlingen bygger på et fundament af empirisk, teknisk, kreativ og teoretisk forskning udført af mig selv, og i samarbejde med kolleger, på Aalborg Universitet i Danmark og ved The University of Melbourne i Australien mellem 2001 og 2012. Denne forskning har involveret flere projekter om mobilt interaktionsdesign med forskellige brugergrupper og inden for brugsdomæner – fra officerer på store containerskibe og sygeplejersker på sygehuse og i hjemmeplejen, til unge mennesker i storbyen, og intime familiemedlemmer i hjemmet. Fælles for disse projekter er, at de har været kontekstuelle af natur, fulgt en design-orienteret tilgang, indeholdt elementer af undersøgelse, analyse, design og konstruktion, og forsøgt at kombinere og integrere metoder og viden fra forskellige discipliner.

Kapitel 1: Design af mobile interaktioner

Afhandlingen indledes med en sammenfatning af min forskningsposition i kapitel 1. I dette kapitel præsenterer jeg baggrunden for min forskning i interaktionsdesign for mobile computersystemer og opsummerer forskningen og resultaterne præsenteret i de følgende 22 kapitler i afhandlingen. Jeg tager udgangspunkt i en historisk gennemgang af udviklingen inden for mobile computersystemer opdelt i syv forskellige faser. Herefter introducerer jeg til disciplinen interaktionsdesign, og beskriver de mest centrale tilgange til design inden for denne. I afsnit 3 beskriver jeg udgangspunktet for min egen forskning i mobil interaktionsdesign, og skitserer fem forskningsemner jeg har arbejdet med. Dette leder til frem en beskrivelse af to nye udfordringer indenfor området; behovet for 1) at hæve sig ud over todelingen mellem menneske- eller teknolog-orienteret forskning og design, og 2) at udvide fokus fra den enkelte mobile enhed og den enkelte brugers interaktion med denne. I respons til disse udfordringer drøfter jeg, i afsnit 4, mulighederne for at udføre og tænke om interaktionsdesign på en design-orienteret måde snarere end på en traditionel videnskabelig måde. Dette gør jeg ved at inddrage designtænkning og forskning i design fra litteratur der i vid udstrækning ligger udenfor området interaktionsdesign, for eksempel Donald Schöns begreber om problemformulering og refleksion-i-praksis, Stephen Pepper's begreb om kontextualisme, Charles Saunders Pierce's begreb om abduktiv tænkning og ræsonnement, og Christopher Alexander's begreber om designudfoldelse og helhed, og ved at vise hvordan disse tanker kan hjælpe os til at berige den måde, vi tænker om interaktionsdesign. Dette efterfølges af en kritisk diskussion af den etablerede bruger-centrerede design model og en række en ændringsforslag hertil. I afsnit 5 præsenterer og diskuterer jeg, som en alternativ tilgang, mit perspektiv på interaktionsdesign som en kontinuerlig konvergens mellem form og kontekst. I afsnit 6 præsenterer jeg en oversigt over mine individuelle forskningsbidrag, og relaterer disse til de centrale aktiviteter, processer, og begreber præsenteret i afsnit 6. Slutteligt sammenfatter og konkluderer jeg i afsnit 7 på mit arbejde.

Kapitel 1 efterfølges af et kapitel, der beskriver den undersøgelse, der i 2002 dannede udgangspunkt for min egen forskning inden for interaktionsdesign til mobile enheder og systemer. Afhandlingens efterfølgende 21 kapitler er inddelt i fem dele, der hver indledes med en oversigt over dets overordnede fokus og de indeholdte kapitler.

Kapitel 2: Udfordringer og muligheder
Kapitel 2 giver et øjebliksbillede af forskningspraksis inden for mobil menneske-maskine interaktion og interaktionsdesign i starten af 2000-tallet. Kapitlet beskriver et litteraturstudie af forskningsmetode og fokus inden for feltet baseret på 102 publicerede artikler, og udleder heraf en række tendenser og antagelser der prægede forskningsretningen på det pågældende tidspunkt. Disse diskuteres kritisk, og der angives en række udfordringer og muligheder for videre forskning. Kapitlet understøtter argumentationen i afsnit 3 af Kapitel 1.

Del I: Undersøgelse og analyse
Del I behandler spørgsmålet hvordan vi kan studere, analysere og forstå aspekter af kontekst, der er relevante for mobil interaktionsdesign. For at gøre interaktionsdesign kontekstuel er vi nødt til at forstå, hvad kontekst er for et bestemt designopgave. Hvilke dimensioner af kontekst er vigtigt, hvilken rolle spiller de, hvordan interagerer de med og påvirker hinanden osv. Fire af mine egne forskningsbidrag til dette findes i Kapitel 3-6. Disse kapitler adresser og udforsker hver især en specifik dimension af kontekst for mobil interaktionsdesign.

Kapitel 3 adresserer fysisk kontekst. Kapitlet præsenterer en feltundersøgelse og arkitektonisk analyse af en konkret bymiljø, Federation Square i Melbourne, Australien. Undersøgelsen blev gennemført ved at kombinere empiriske og analytiske metoder fra arkitektur og byplanlægning til at undersøge, modellere og repræsentere fremtrædende aspekter af en fysisk omgivelse som distrikter, bygninger, strukturer og skiltning. Kapitel 4 adresserer social kontekst, og præsenterer en feltundersøgelse og sociologisk analyse af små gruppers sociale samvær, ligeledes på Federation Square, og leder til udarbejdelsen af et begrebsmæssigt rammeværk om samspillet mellem mennesker, aktivitet og steder. Kapitel 5 adresserer personlig kontekst og præsenterer et længerevarende feltstudie af teknologi-medieret intimitet og tætte personlige forhold udført med seks familier i Melbourne, Australien. Dette leder frem til en tematisk forståelse af intimitet og hvordan interaktive systemer bruges og tilpasses i denne kontekst. Kapitel 6 adresserer arbejdskontekst og præsenterer et etnografisk feltstudie af kommunikation om bord på store containerskibe samt et simulatorstudie af prototypeteknologi i brug. Resultatet af undersøgelsen er en forståelse af talehandling i kommunikation og koordinering.

Del II: Design og konstruktion
Del II behandler spørgsmålet hvordan vi kan designe og konstruere interaktive mobile systemer funderet i kontekst. For at kunne gøre brug af vores empiriske og teoretiske forståelse af kontekst i afsøgningen og udforskningen af designmuligheder, må vi udvikle måder til at overføre denne viden til designaktiviteten, måder til at fundere ideudvikling og udforskning i sådan viden, og måder til at sammenfatte mangfoldighed af idéer til konkrete interaktive systemer. Fire af mine egne forskningsbidrag til dette findes i Kapitel 7-10. Disse kapitler adresserer og drøfter mulige måder til at støtte den kreative proces med at bevæge sig fra abstrakt forståelse til konkrete artefakter gennem kombination af inspirationalistisk og strukturalistisk kreativitet.

Kapitel 7 diskuterer bruger- og teknologi-centrering og kommenterer på de relative styrker og svagheder ved inspirationalistisk og strukturalistisk kreativitet i forhold til to tilgange til det samme designoplæg. Kapitel 8 udforsker en socio-fysisk tilgang til interaktionsdesign der kombinerer inspirationalistisk og strukturalistisk kreativitet ved, på den ene side, at supportere fri ideudvikling, og på den anden side, tilbyde en trinvis og struktureret proces. Kapitel 9 illustrerer brugen af skitser og mock-ups i en semi-struktureret designproces i samarbejde med potentielle brugere, hvorved der blev tilføjet en grad af struktur og metode til designteamets primært inspirationalistiske kreativitet. Kapitel 10 adresserer kombinationen af etnografi og objekt-orientering og kombinerer inspirationalistisk og strukturalistisk kreativitet ved at integrere en meget åben metode med en meget struktureret metode.

Del III: Forbedring af evaluering
Del III behandler spørgsmålet hvordan vi kan forbedre vores teknikker til at studere brugerens oplevelse af mobil interaktion design i kontekst? For at realistisk vurdere kvaliteten af et mobil interaktion design er vi nødt til at studere det under forhold, der er repræsentative for den fremtidige brugssituation. Fire af mine egne forskningsbidrag til dette findes i Kapitel 11-14. Disse kapitler viser forskellige måder at evaluere mobil interaktion design med både realisme og kontrol. De viser, hvordan elementer af kontekst kan simuleres i kontrollerede omgivelser, og hvordan data af høj kvalitet kan indsamles i naturlige omgivelser.

Kapitel 11 undersøger hvordan mobilitet kan simuleres i en kontrolleret omgivelse, og præsenterer fem laboratorieteknikker der involverer forskellige aspekter af fysisk bevægelse. Kapitel 12 undersøger hvordan brugsdomænet kan simuleres i en kontrolleret omgivelse og præsenterer to sammenlignende empiriske studier. I begge kapitler præsenteres en evaluering i felten som sammenligningsgrundlag. Kapitel 13 undersøger hvordan en struktureret evaluering kan udføres i felten og præsenterer et empirisk studie hvor resultaterne fra en felt og en laboratorieevaluering sammenlignes. Kapitel 14 undersøger hvordan dataindsamling kan forbedres i feltet og præsenterer udviklingen og brugen af et "felt-laboratorie" til evaluering af mobil interaktionsdesign gennem anvendelse af små trådløse kameraer og mobilt video og lydudstyr.

Del IV: Artefakter
Del IV behandler spørgsmålet hvordan vi kan gøre brug af kontekst i forbindelse med konstruktion af konkrete interaktive mobile systemer. For at skabe konkrete interaktionsdesign-artefakter har vi brug for at vide, hvad der er teknisk muligt nu og i den nærmeste fremtid, og vi er nødt til at vide, hvordan nye teknologier kan bruges til at flytte denne grænse yderligere. Fem af mine egne forskningsbidrag til dette findes i Kapitel 15-19. Disse kapitler præsenterer erfaringerne med konstruktionen af fem interaktive mobile prototypesystemer, der på forskellig vis tilpasser sig deres kontekst.

Kapitel 15 udforsker konstruktionen af et mobilt prototypesystem til et sygehus, kaldet MobileWARD. Det præsenterer detaljer om systemets konstruktion og kontekstuelle interaktionsdesign, og en empirisk felt- og laboratorieundersøgelse af dets anvendelse. Kapitel 16 udforsker konstruktionen af et mobilt prototypesystem tiltænkt at fremme socialt samvær i byen, kaldet Just-for-Us. Systemet var et tidligt forsøg på at gøre mobile kontekst-opmærksomme systemer web-baserede, og kapitlet beskriver hvordan dette blev opnået og hvilket interaktionsdesign det muliggjorde. Kapitel 17 udforsker konstruktionen af et Web 2.0 baseret mobilt prototypesystem, GeoHealth, til hjemmesygeplejersker fordelt over et større geografisk område. Systemet bygger videre på de tekniske erfaringer fra de to tidligere systemer, og kapitlet præsenterer det resulterende interaktionsdesign i detaljer og en empirisk feltundersøgelse af det i brug. Kapitel 18 udforsker konstruktionen af et mobilt "augmented reality" prototypesystem til visualisering af arkitektur i kontekst kaldet ArchiLens. Systemet gør brug af grafikpotentialet og indbyggede kontekstsensorer i moderne smartphones til at skabe en interaktiv 3D repræsentation. Kapitlet præsenterer systemet i detaljer og en empirisk feltundersøgelse af dets anvendelse. Kapitel 19 udforsker konstruktionen af et mobilt prototypesystem til at fremme miljømæssig bæredygtighed ved at gøre det muligt for folk at overvåge og justere deres elforbrug på en Android eller iOS smartphone. Systemet, kaldet Power Advisor, præsenteres i detaljer efterfulgt af resultaterne fra en 7-ugers empirisk undersøgelse af det i brug.

Del V: Forståelse

Del V behandler spørgsmålet hvordan vi kan beskrive og forstå forholdet mellem interaktive mobile systemer, brugere og kontekst? For at fremme forskningen er vi nødt til at skabe viden der på et teoretisk plan kan forklare forholdet mellem mennesker, teknologi, og deres kontekst. Fire af mine egne forskningsbidrag til dette findes i Kapitel 20-23. Disse kapitler præsenterer, på forskellige abstraktionsniveau, hver især et perspektiv på meningsskabelse i kontekst – den måde vi opfatter verden gennem identifikation af meningsfulde mønstre og helheder, den måde vi fortolker verden gennem tildeling af mening til tegn, den måde vi bruger vores fælles fysiske tilstedeværelse i verden til at skabe fælles betydninger, og den måde vi organiserer og orkestrerer verden omkring os.

Kapitel 20 præsenterer og diskuterer på baggrund Gestalt Teori fem principper, der kan anvendes til at forklare, hvordan mennesker identificerer meningsfulde mønstre og helheder ud fra samspil mellem mobile systemer og deres kontekst. Kapitel 21 diskuterer på baggrund af Semiotik hvordan mennesker fortolker information på mobile enheder i kontekst ved at forstå dem en speciel type af indeksikalske tegn. I begge kapitler illustreres anvendelsen af det teoretiske perspektiv gennem analyse af et eller flere interaktive mobile prototypesystemer. Kapitel 22 diskuterer med udgangspunkt i Fænomenologi hvordan mennesker skaber fælles betydning gennem tilstedeværelse i fælles fysiske rum, og illustrerer anvendelsen af denne tankegang i designet af et delt interaktionsrum indeholdende, blandt andet, mobile enheder og systemer. Endeligt diskuterer Kapitel 23 på baggrund af litteraturen om divergens og konvergens, og et empirisk feltstudie, hvordan mennesker orkestrerer interaktive mobile systemer og enheder gennem skabelsen af meningsfulde digitale økosystemer og helheder.

Opsummering af sammenfatningen

Afhandlingens fem dele adresserer tilsammen de forskellige elementer, principper, og dynamikker i den kontekstuelle tilgang, og eksemplificerer de aktiviteter og resultater, den tilstræber at fremme. Det er mit forhåbning, at dette perspektiv på interaktionsdesign vil inspirere læseren, forskere såvel som praktikere, til selv at søge ud over bruger- eller teknologi-centreret design, og gribe designet af mobile interaktioner an med et bredere fokus på samspillet mellem form og kontekst, og med en design-orienteret tilgang.