Designing Mobile Interactions: The continual convergence of form and context

Volume 2

Kjeldskov, Jesper

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

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

Preface *i*

1. Designing mobile interactions 1
2. A review of mobile interaction design research 109

Part I. Studying and analysing *129*

3. Physical context 131
4. Social context 153
5. Personal context 167
6. Work context 183

Part II. Designing and building *209*

7. User- and technology-centredness 211
8. Socio-physical design 225
9. Sketches and mock-ups 259
10. Ethnography and object-orientation 277

Part III. Improving evaluation *313*

11. Simulating mobility 315
12. Simulating the domain 335
13. Bringing the system in to the field 363
14. Taking the lab with you 385

Part IV. Artefacts *403*

15. MobileWARD 405
16. Just-for-Us 427
17. GeoHealth 443
18. ArchiLens 467
19. Power Advisor 485

Part V. Understanding *505*

20. Principles of perceptual organisation 507
21. Indexical interaction design 529
22. Proxemics and interactional spaces 555
23. Orchestrating mobile devices 591

List of contributions *605*

List of co-authors *607*

Sammenfatning på dansk *609*

Part III

Improving evaluation

Chapter 11. Simulating mobility

Chapter 12. Simulating the domain

Chapter 13. Bringing the system into the field

Chapter 14. Taking the lab with you

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).

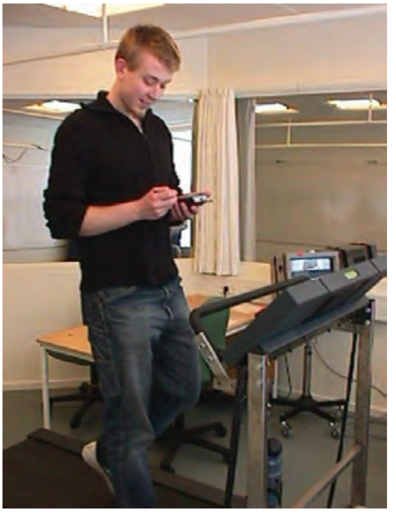

Figure 1. Test subject walking on a treadmill in the usability laboratory

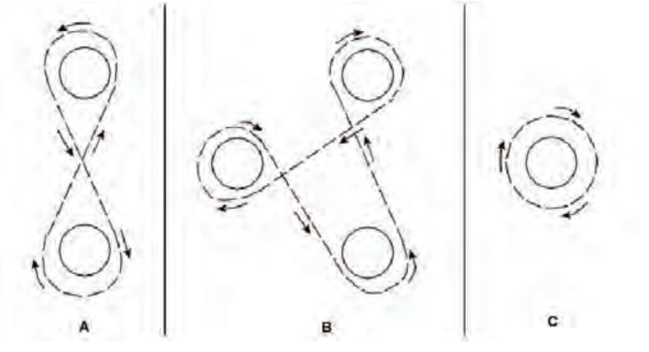

Figure 2. The sequence of the different courses to be walked in the laboratory (a, b, c)

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.

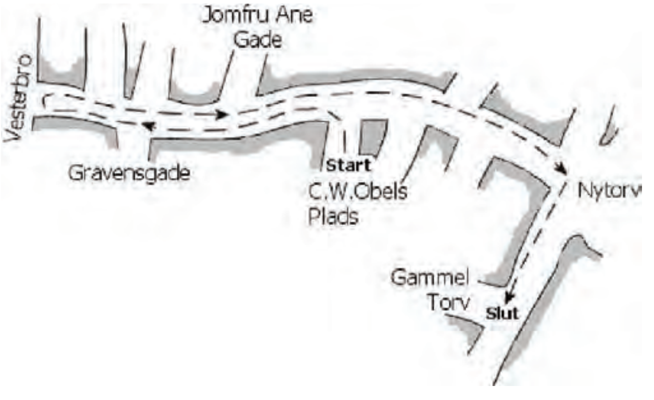

Figure 3. The route to be walked in the evaluations in the pedestrian street

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.

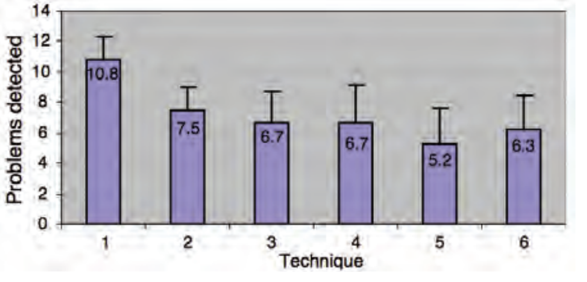

Figure 5. Number of usability problems detected with each of the six techniques

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).

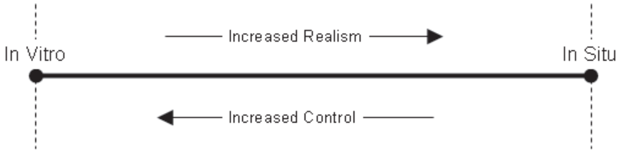

Figure 1. Trade-offs between a high level of control and a high level of realism.

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.

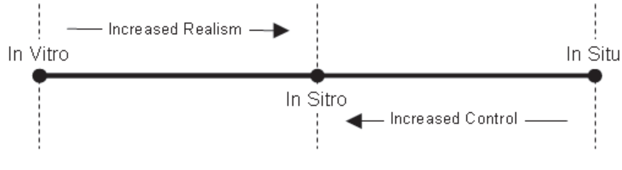

Figure 2. In sitro evaluations with increased levels of control and realism

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.

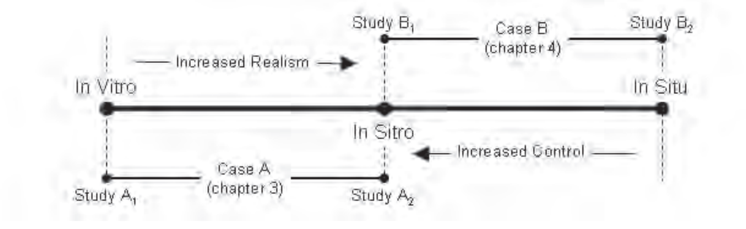

Figure 3. Illustration of our two studies investigating in sitro experiments.

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.

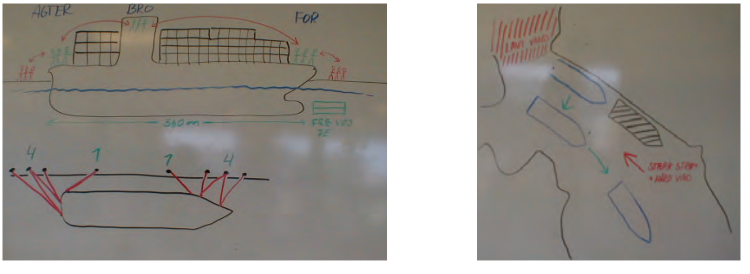

Figure 6. Introduction to use context and a possible use scenario drawn on whiteboard

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.

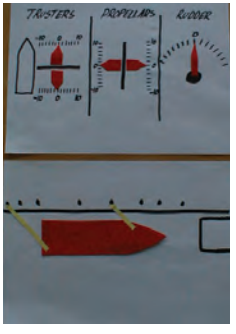

Figure 7. Cardboard mock-up of the bridge, ship and mooring

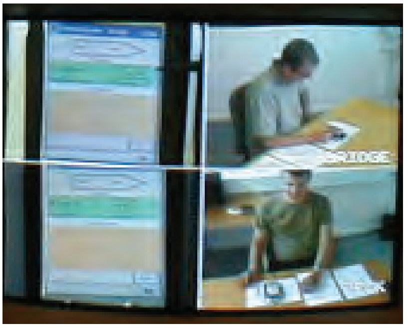

Figure 8. Video recording from evaluation in the usability laboratory

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.

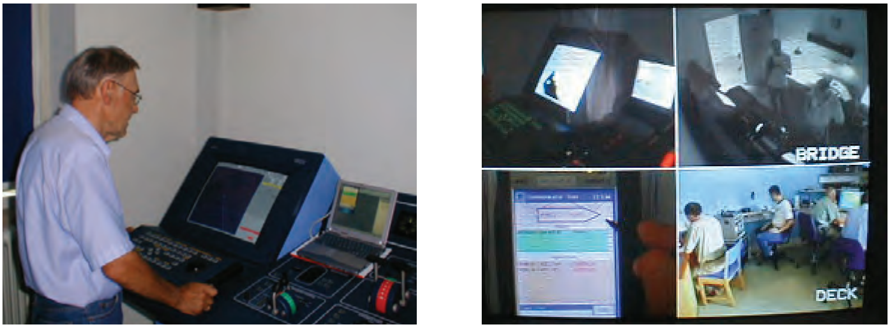

Figure 10. Studying use in the ship simulator: the simulator and the video recording

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.

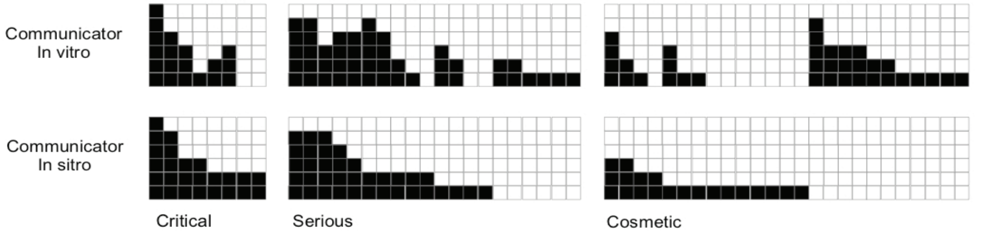

Figure 11. Distribution of identified usability problems

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.

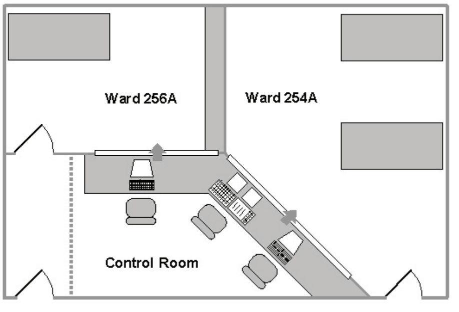

Figure 13. Physical layout of the usability laboratory simulating the hospital ward

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

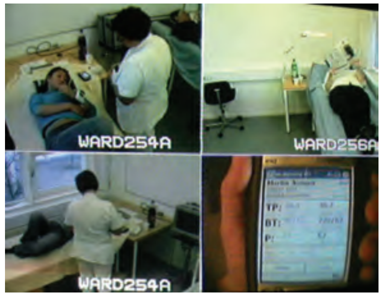

Figure 15.
Video images from simulated ward and PDA

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.

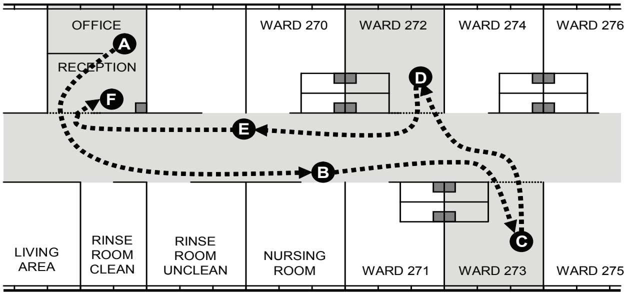

Figure 16. Physical layout of the hospital wards

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.

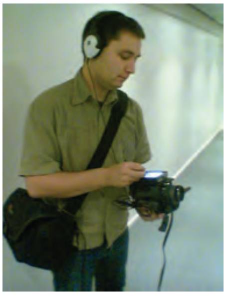

Figure 17. Observer (left) carrying and operating portable audio/video equipment (right) for capturing high-quality data in the field.

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.

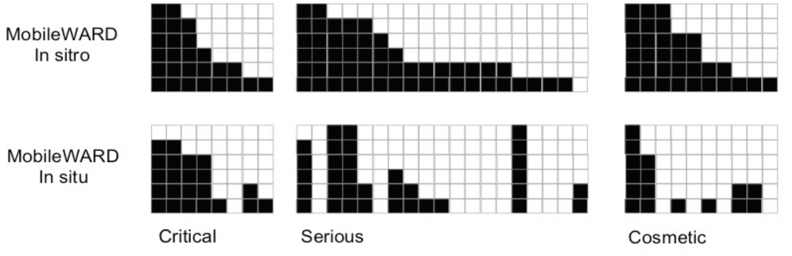

Figure 18. Distribution of identified usability problems

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.

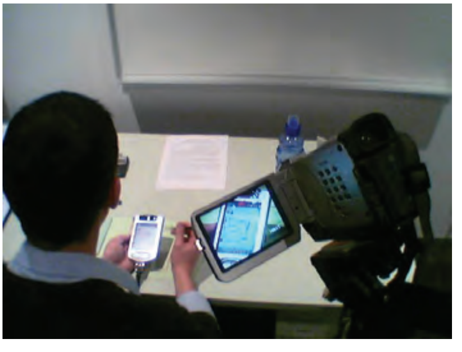

Figure 5. Close-up of interaction with the mobile guide.

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.

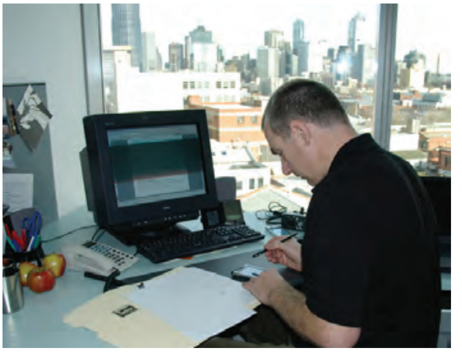

Figure 6. Heuristic walkthrough.

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.

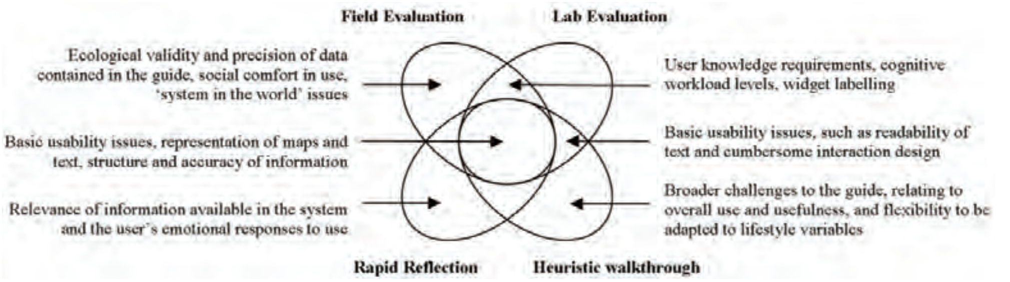

Figure 9. Overview of the types of usability issues identified in overlaps between techniques.

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.

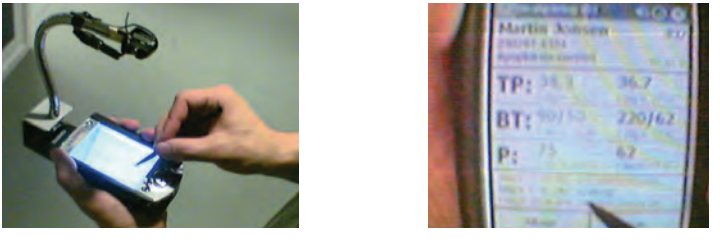

Figure 2. PDA on camera-mount allowing for close-up view of screen and user interaction.

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.

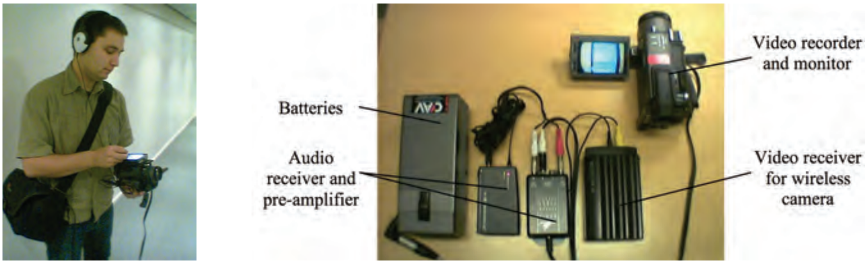

Figure 3. Observer (left) carrying and operating portable audio/video equipment (right) for capturing close-up view of screen and user interaction.

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.

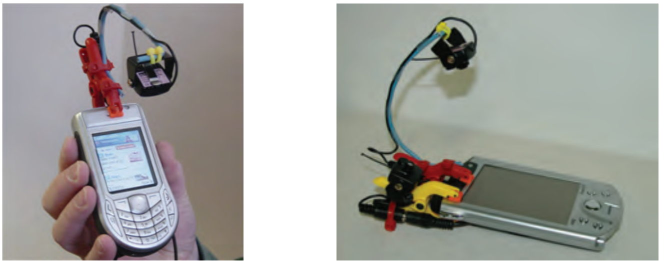

Figure 4. Lightweight camera-clamps attached to mobile devices

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).

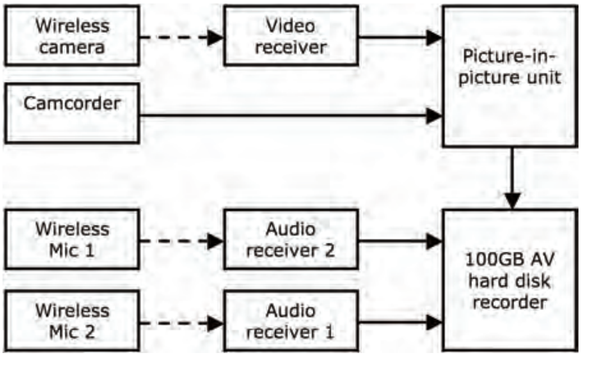

Figure 7. Schematic configuration of the current version of the field laboratory with two video sources and two audio sources recorded in one composite digital file.

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

Chapter 15. MobileWARD

Chapter 16. Just-for-Us

Chapter 17. GeoHealth

Chapter 18. ArchiLens

Chapter 19. Power Advisor

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.

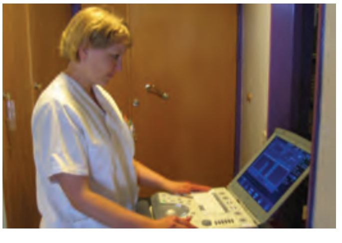

Figure 1. Ubiquitous EPR use in the healthcare domain

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.

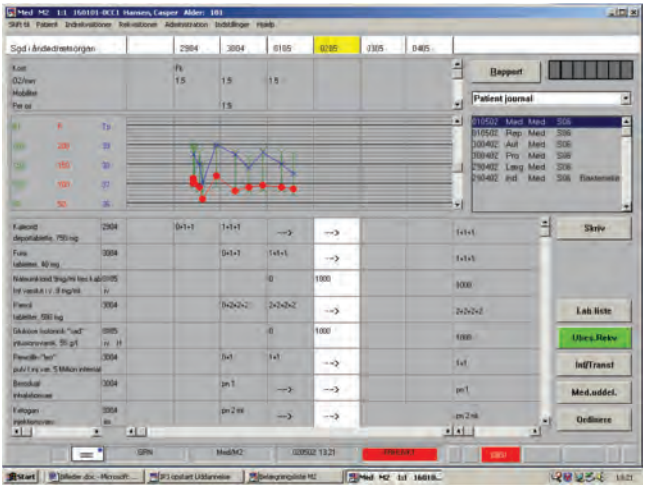

Figure 2. Screen shot from IPJ 2.3 illustrating the primary information, for example, previous measured temperatures, heart rates, on a specific patient

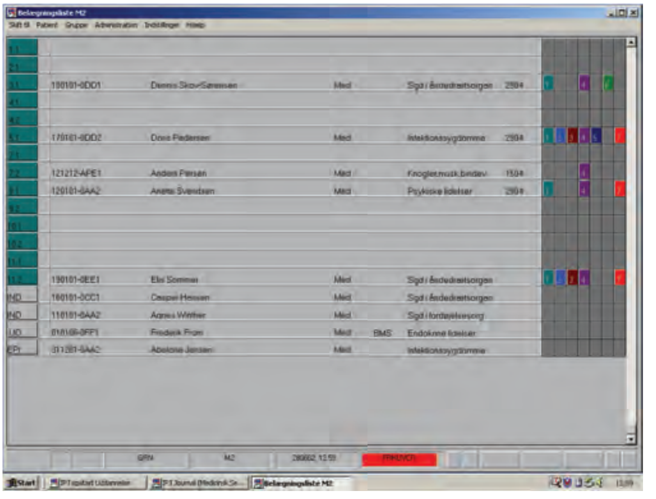

Figure 3. Screen shot from IPJ 2.3 showing the central list of patients on the ward with overall information about diagnosis and treatment.

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.

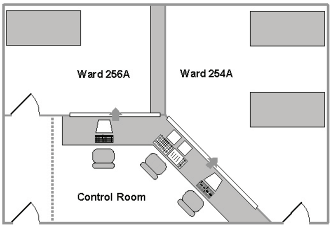

Figure 6. Physical layout of the lab set up to emulate a section of the hospital ward

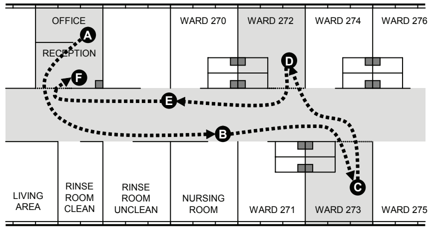

Figure 7. Physical layout of the hospital ward used for the field evaluation

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

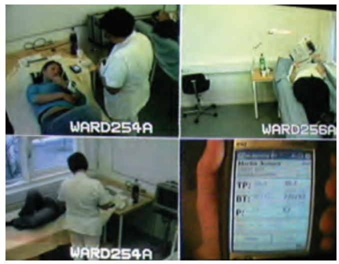

Figure 9. Video recording from usability evaluation

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).

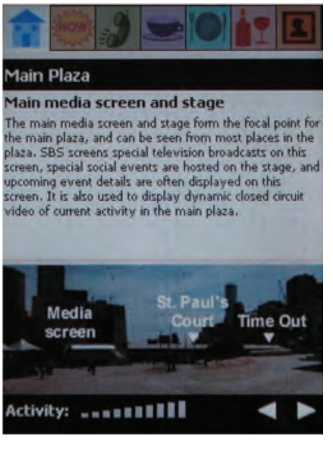

Figure 4. Home screen: augmentation of the user's built surroundings

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).

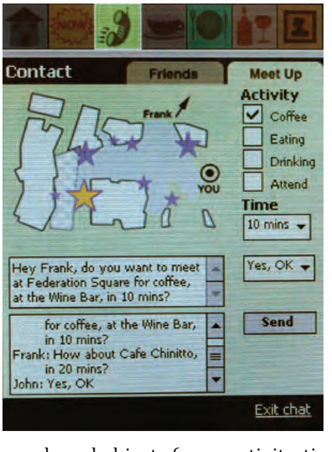

Figure 5. Contact screen: shared objects (map, activity, time) and chat window

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).

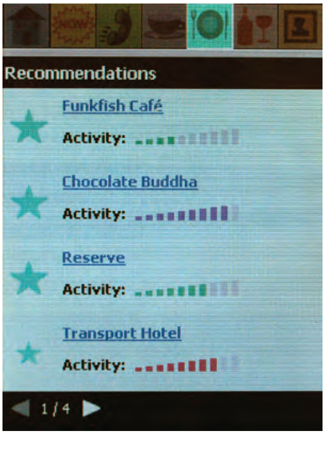

Figure 6. Recommendations screen: ranked list of places to go for food

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.

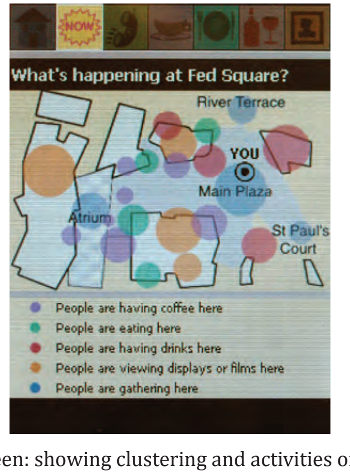

Figure 7. Now screen: showing clustering and activities of close-by people

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.

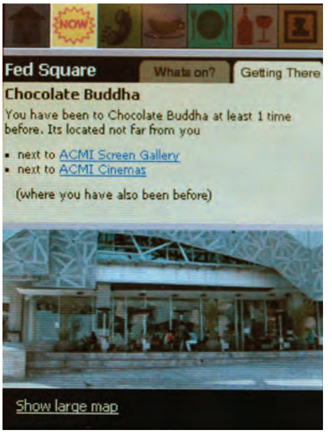

Figure 9. Now screen: way-finding instructions based on history of visits

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).

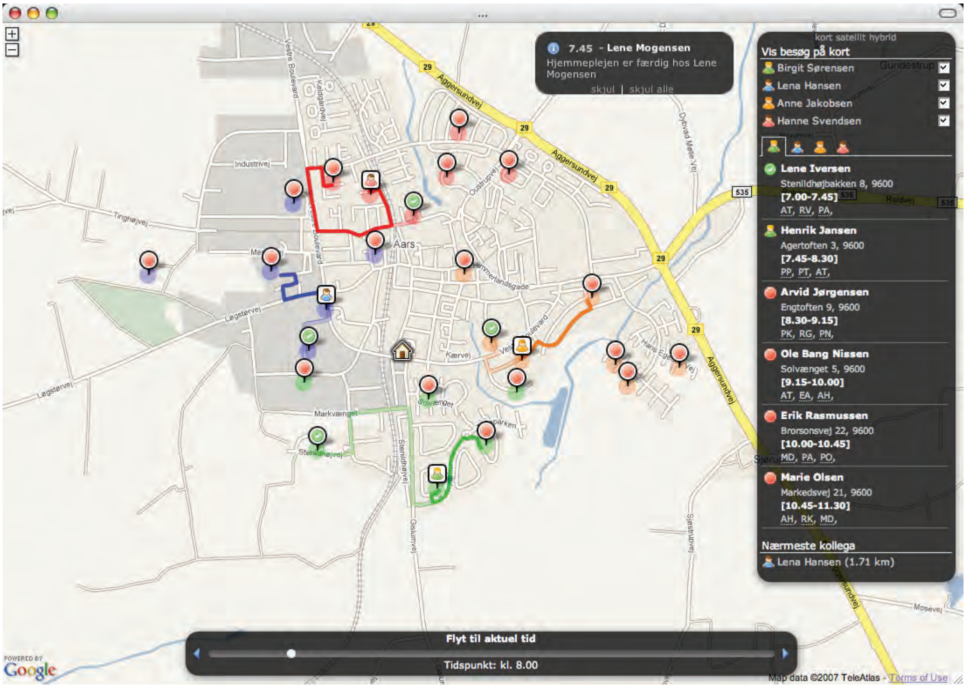

Figure 1. The main screen of the GeoHealth prototype system with the layers of four healthcare workers' patients, routes and current location enabled. In the palette on the right, the user can see her own scheduled work tasks and keep an eye on which colleague is closest to her. Using the timeline, the 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.*

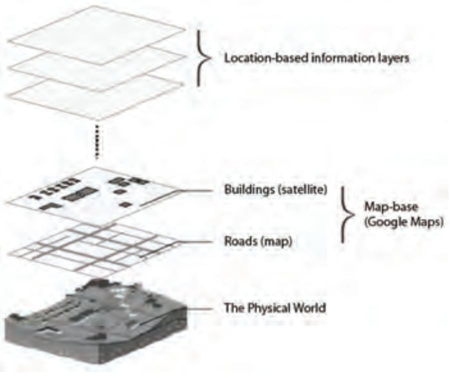

Figure 2. The geographical information-layering model of GeoHealth

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.

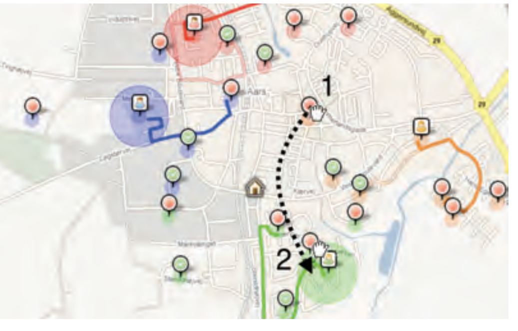

Figure 5. Location-based ad-hoc exchange of work tasks by dragging a patient icon onto the location of another healthcare worker.

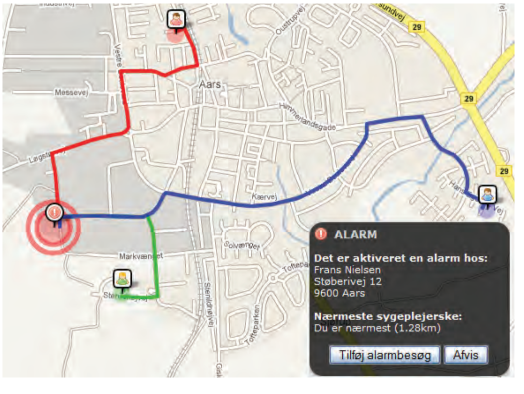

Figure 6. Patient alarm appearing as a pop-up message on the closest healthcare worker's screen along with information about location of the two most nearby colleagues.

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.

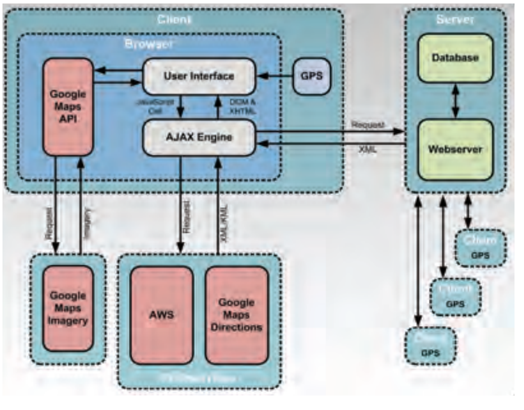

Figure 7. Technical architecture of the prototype system

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.

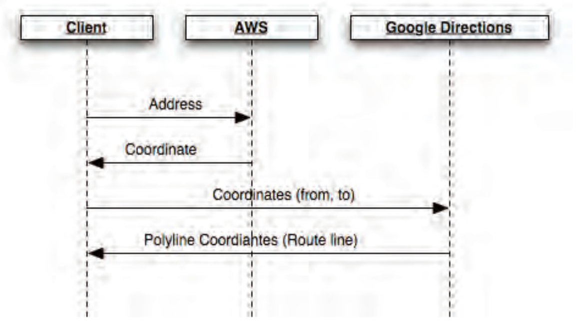

Figure 8. Information flow in the prototype system

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).

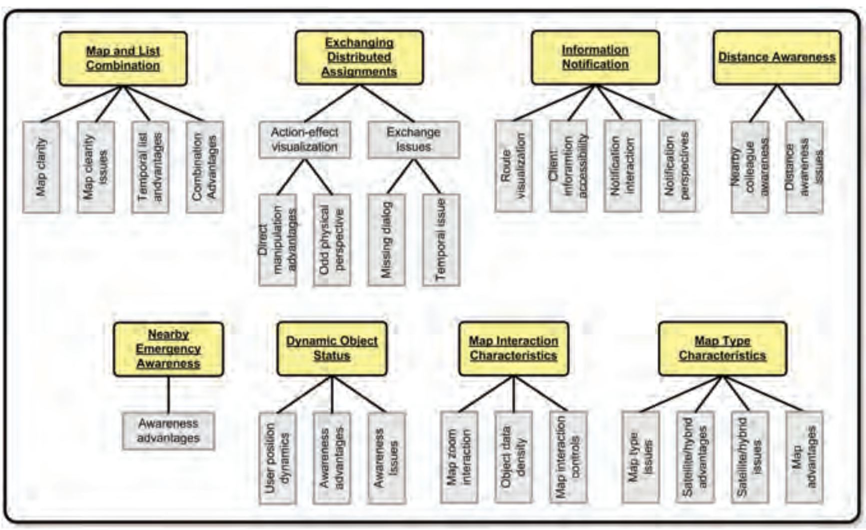

Figure 10. Thematic clustering of findings from the evaluations.

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.

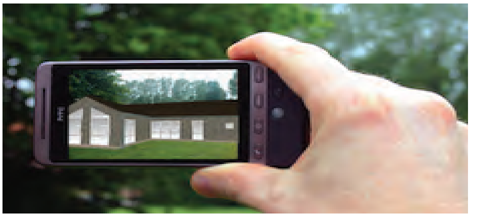

Figure 1. Visualizing architecture in context on a mobile phone while standing at the building site

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).

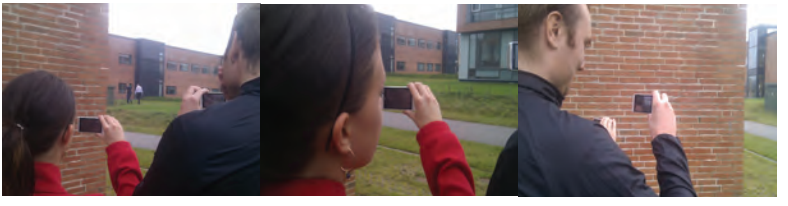

Figure 4. Participants experiencing problems with visualizing a house extension related to inaccurate 3D registration.

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.

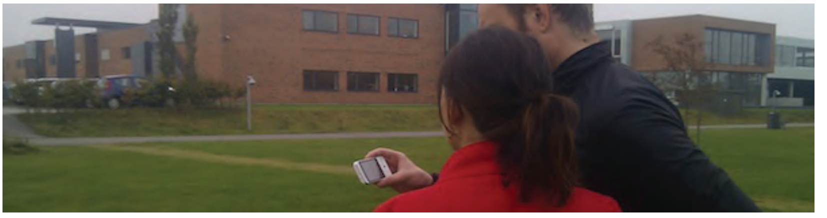

Figure 5. Participants trying to understand contextual influences of nearby houses.

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.

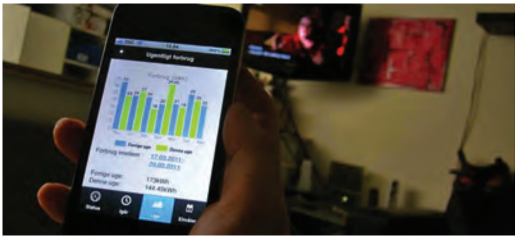

Figure 1. Power Advisor in use at home while watching TV

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.

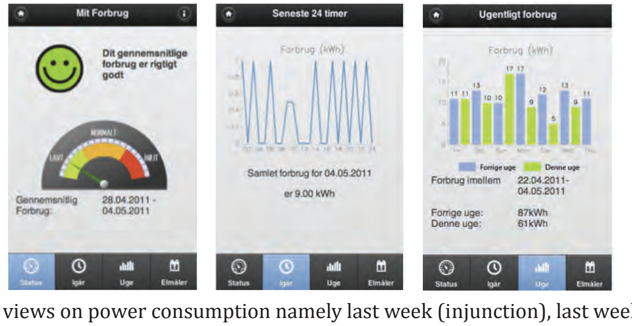

Figure 2. Three views on power consumption namely last week (injunction), last week consumption (descriptive), and last week compared to the week before (descriptive).

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

Chapter 20. Principles of perceptual organisation

Chapter 21. Indexical interaction design

Chapter 22. Proxemics and interactional spaces

Chapter 23. Orchestrating mobile devices

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.

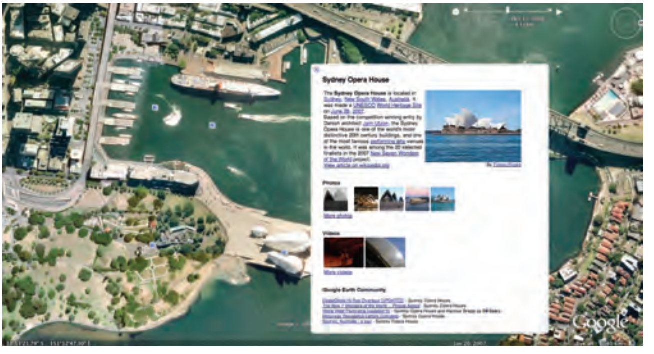

Figure 2. Location based information posted in Google Earth

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

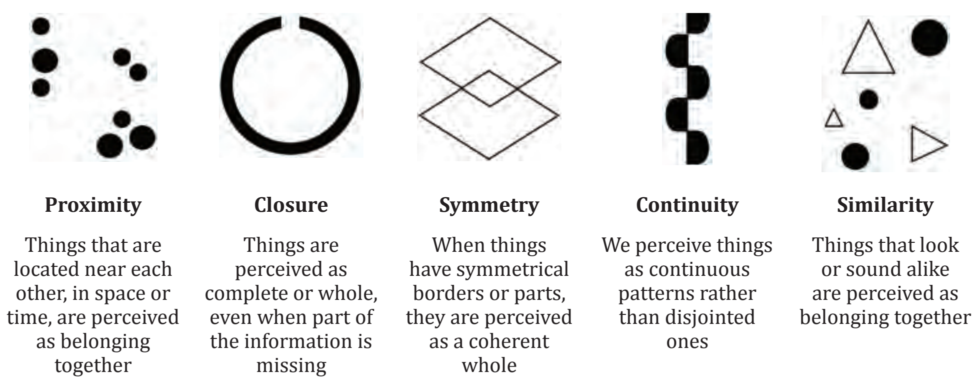

Figure 3. Principles of perceptual organisation for interaction design derived from Gestalt theory

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.

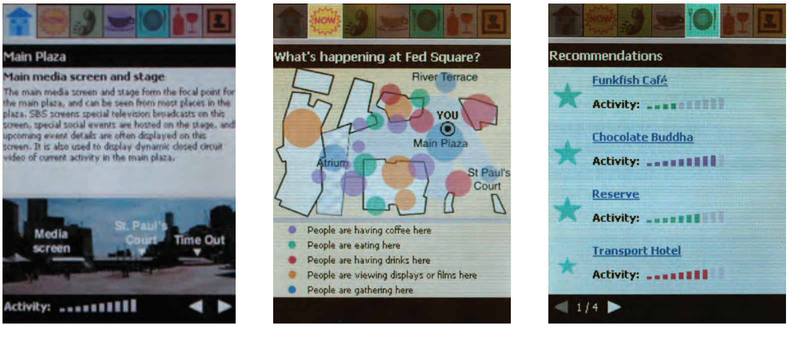

Figure 5. Screens from the Just-for-Us location based service

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.

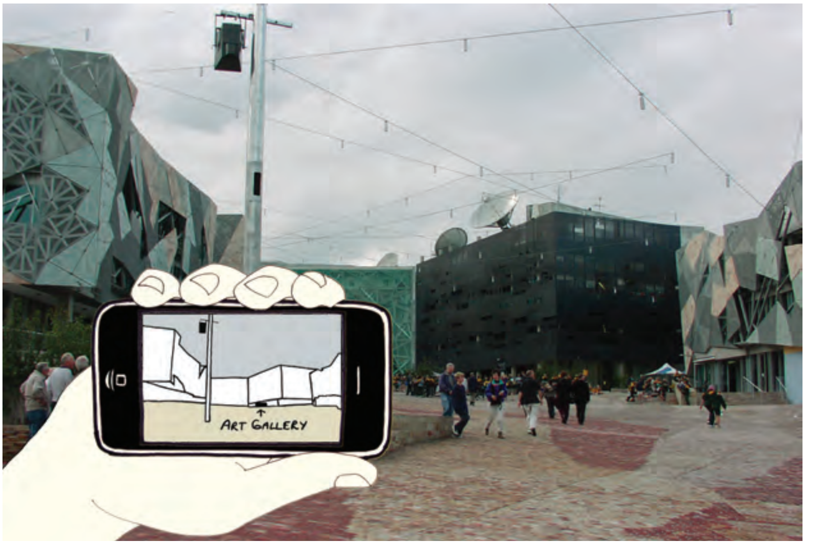

Figure 9. Symmetry: making a symmetrical alignment between system and surroundings

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.

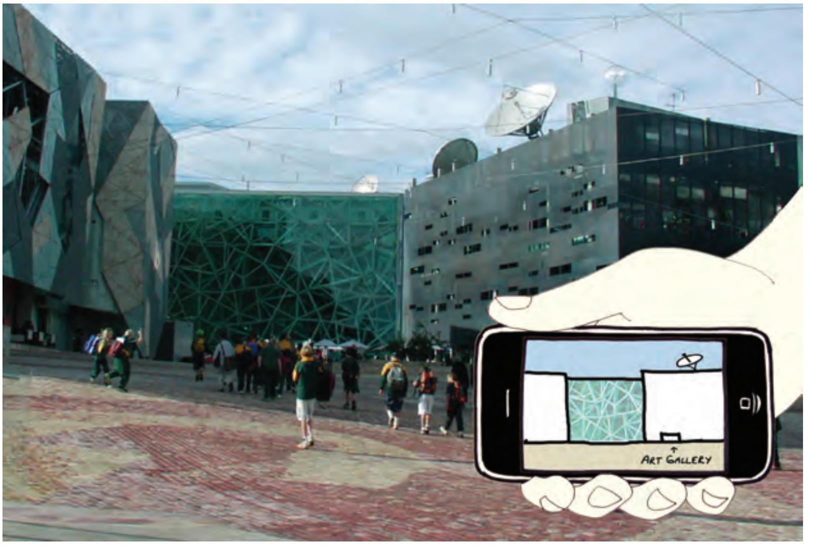

Figure 11. Similarity: matching things in system and surroundings that are alike

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).

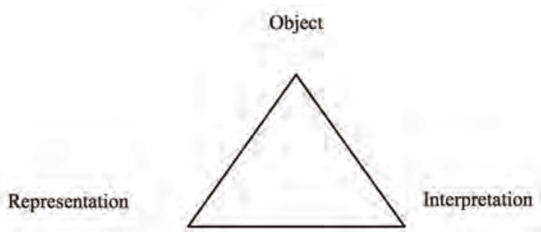

Figure 1. The semiotic relation between object, representation and interpretation

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.

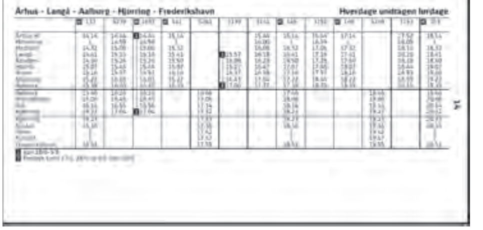
Århus - Langå - Aalborg - Hjørring - Frederikshavn
Hverdage undtagen lørdage

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.

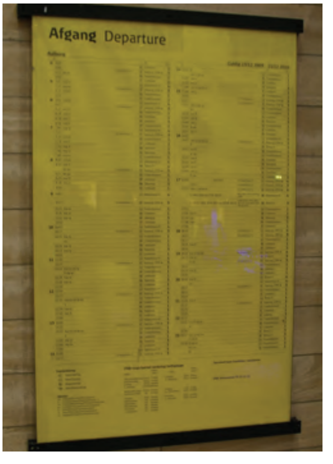

Figure 3. Timetable poster at a train station: symbolic representation with spatial indexicality

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).

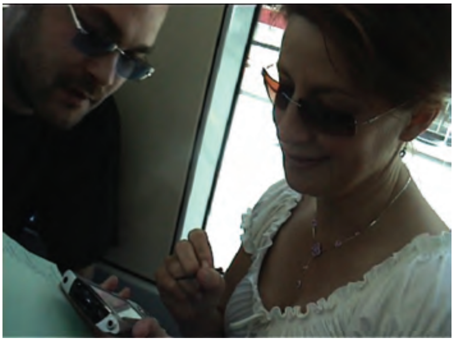

Figure 8. Field evaluation of TramMate II on board a tram in Melbourne, Australia

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

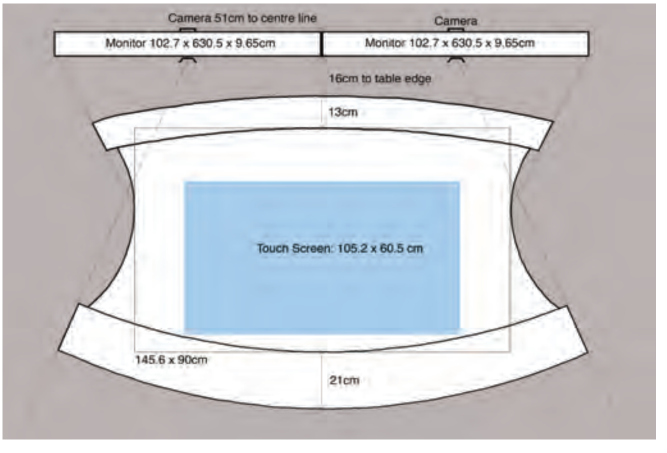

Figure 3. Schematic overview of the BISi setup

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.

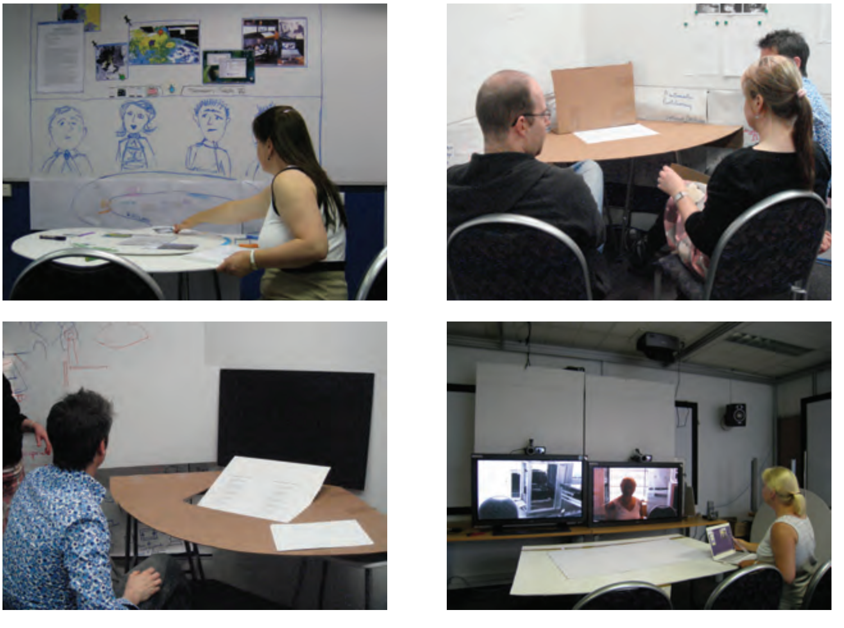

Figure 4. Iterations of full-scale mockups of Blended Interaction Spaces

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.

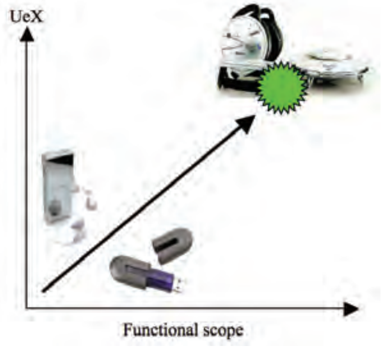

Figure 2. Utopian view of the relationship between functionality and usability

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).

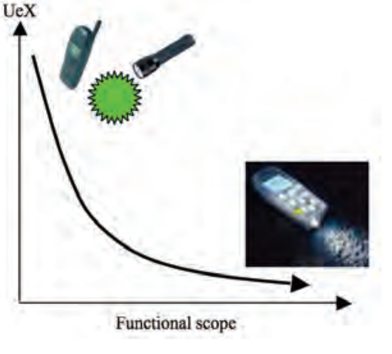

Figure 3. Dystopian view of the relationship between functionality and usability

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).

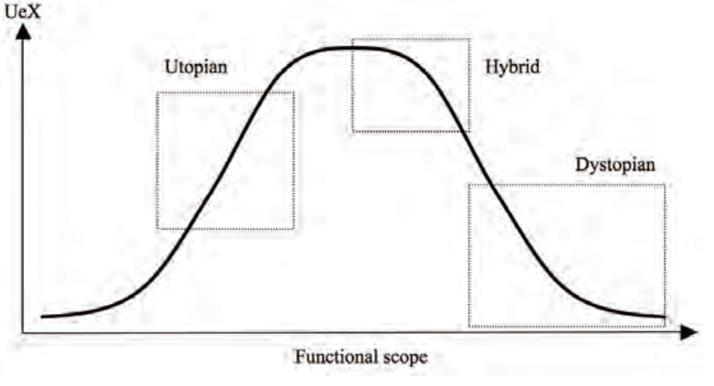

Figure 7. The Convergence Curve - Three complementary perspectives

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.