Scientific Part

On the understanding of industrial usability work in IT-systems development

Tobias Uldall-Espersen
PhD Thesis
Department of Computing
Faculty of Science
University of Copenhagen
2008

Tracing impact in a usability improvement process

Tobias Uldall-Espersen *, Erik Frøkjær, Kasper Hornbæk

Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen OE, Denmark

Abstract

Analyzing usability improvement processes as they take place in real-life organizations is necessary to understand the practice of usability work. This paper describes a case study where the usability of an information system is improved and a relationship between the improvements and the evaluation efforts is established. Results show that evaluation techniques complemented each other by suggesting different kinds of usability improvement. Among the techniques applied, a combination of questionnaires and Metaphors of Human Thinking (MOT) showed the largest mean impact and MOT produced the largest number of impacts. Logging of real-life use of the system over 6 months indicated six aspects of improved usability, where significant differences among evaluation techniques were found. Concerning five of the six aspects Think Aloud evaluations and the above-mentioned combination of questionnaire and MOT performed equally well, and better than MOT. Based on the evaluations 40 redesign proposals were developed and 30 of these were implemented. Four of the implemented redesigns where considered especially important. These evolved with inspiration from multiple evaluations and were informed by stakeholders with different kinds of expertise. Our results suggest that practitioners should not rely on isolated evaluations. Instead complementing techniques should be combined, and people with different expertise should be involved.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Usability engineering; Case study; Usability improvement process; Metaphors of human thinking; Think loud; Questionnaire

1. Introduction

To support practitioners in developing usable IT-systems we need to understand the processes leading to systems with better usability. These processes are in this paper collectively named the usability improvement process (UIP), and are defined as the activities conducted in order to ensure or improve the usability of a system. The UIP includes activities like defining the main objectives of the usability evaluation activities, selecting usability evaluation techniques, assigning personnel with the proper expertise (such as work-domain experts, usability experts, or technical experts), identifying and processing usability problems, developing and prioritizing solutions, and implementing and evaluating improvements. All of these activities are interdependent and shaped by their organizational setting. Nevertheless, the whole UIP is rarely covered in published studies, although the success of software development projects depends on this whole. Practitioners cannot rely on studies that exclusively focus on finding usability problems; this is only one part of the UIP and perhaps even a minor part. As Wixon (2003, p. 31) stated: 'problems should be fixed and not just found.' Therefore we aim at studying the whole process of improving usability.

This paper describes a case study of a usability-driven redesign process where the usability of an information system in a financial company was improved and documented. We propose a new way of analyzing impact from usability evaluation techniques that establish a relationship between conducted evaluations and achieved improvements. Through our analyses of the data from the case study we give a more coherent description of the whole UIP than previous studies.

* Corresponding author. Tel.: +45 35321451; fax: +45 35321401.
  E-mail addresses: tobiasue@diku.dk (T. Uldall-Espersen), erikf@diku.dk (E. Frøkjær), kash@diku.dk (K. Hornbæk).

2. Related work

In the last 20 years usability work has been given much attention and the literature concerning usability is comprehensive. For example, a search for the term usability in the ACM Digital Library finds more than 9200 publications after January 1986. Few of these studies discuss how the UIP is influenced by its organizational setting, for example how usability work may inform the software development process and the quality of the developed software. Even fewer studies measure actual improvement of usability based on real-life use of software, as fixing the identified usability problems are rarely within the studies' scopes. In this section we argue the importance and usefulness of studying the UIP as a whole by presenting some studies that extend the traditional scope of usability research. First we argue that the majority of studies of usability work disregard the organizational settings in which the work typically is situated, and then we present a number of studies with extended scopes.

Among the studies listed in comprehensive reviews like Gray and Salzman (1998) and Hartson et al. (2001) none account for improved end-product usability or the influence of real-life organizational settings when working with the UIP. Two of these studies were conducted in large organizations (Karat et al., 1992; Jeffries et al., 1991), but did not include redesign or implementation efforts. The study by Jeffries et al. (1991) was done in collaboration with Hewlett Packard Laboratories. A pre-release version of a software product was evaluated using four evaluation techniques, but only with limited involvement of the product developers. The aim of Jeffries et al.'s study was to evaluate the techniques and so it did not document an impact on the subsequent software development process or end-product usability. Karat et al. (1992) tested two office systems at IBM T. J. Watson Research Center and compared the effectiveness of walkthrough methods and usability testing. They found that about a third of the significant problems were identified by all of the applied methods. Karat et al. did not make recommendations on how to solve the identified problems and did not implement any solutions. According to Wixon (2003, p. 32), attempts to transfer results from such studies to an industrial context could be "... wasteful and damaging to the practice of usability work in corporations", suggesting that real-life organizational settings are important when studying practical usability work. Unfortunately only few of such studies exist; Ruthford and Ramey (2000, p. 315) noted "... it is difficult to find published studies that actually document the impact of usability test findings and recommendations on product design".

Organizational aspects of usability problem identification and reporting were investigated by for example Hornbæk and Frøkjær (2005). They reported a study analyzing a major web-application for finding jobs. They extended the traditional scope of usability studies by letting the system developers assess the value of usability problems and redesign proposals. Hornbæk and Frøkjær found that the system developers appreciated and found useful both problem descriptions and redesign proposals. Thereby they connected identification and processing of usability problems with real-life software development and suggested a way to improve communication about usability issues. In a recent laboratory study, Kelkar et al. (2005) investigated the use of process measures and performance measures as feedback to interface design. In the first of two iterations, performance measures were used to improve a system; in the second iteration Kelkar et al. used both performance and process measures. Their study showed that different techniques improved different aspects of usability. Selection of techniques thus not only influences the size of the improvements, but also what kind of improvement to expect. Følstad (2007) compared the impact from usability inspections by domain experts against inspections by usability experts. Følstad found that usability issues identified by domain experts are much more likely to be addressed by the developers than usability issues identified by usability experts. Ruthford and Ramey (2000) examined the impact of a usability test during the development of an MS Windows based synchronization product. They documented that the evaluation had a significant impact on the system design, but did not measure the impact on usability.

Few real-life studies covering the majority of the UIP exist, with Butler (1985) and Good et al. (1986) as notable exceptions. Butler investigated how usability goals were integrated and applied in the definition, design, and testing of a new financial analysis package. Good et al. described the improvements of a VMS/VAX based IT-system at Digital Equipment Corporation. They followed a process that involved defining usability requirements, analyzing the system, implementing improvement, and measuring the impact, and were able to document a 37% improvement of users' efficiency.

The last couple of years a number of real-life case studies have been published at the CHI-conferences, for example under the track titles Design Case Studies (CHI2004) and Design Expo (CHI2005). These studies offer qualitative descriptions of real-life usability work and suggest that large projects focusing on usability must be based on a genuine interest from the stakeholders, including management, users, IT representatives and usability experts. In most of these studies several usability work techniques were used, and time and resources typically posed significant challenges to the success of the projects. Thus these studies indicate that we need to do usability work involving a multitude of stakeholders without overloading the project plan and budget with time consuming and expensive activities. While these real-life case studies describe their organizational settings they do not account for the usability improvements in enough detail so that the relationship between evaluations and improvements can be established.

Bias and Mayhew (2005) took another approach to studying the effects of real-world usability work by focus-
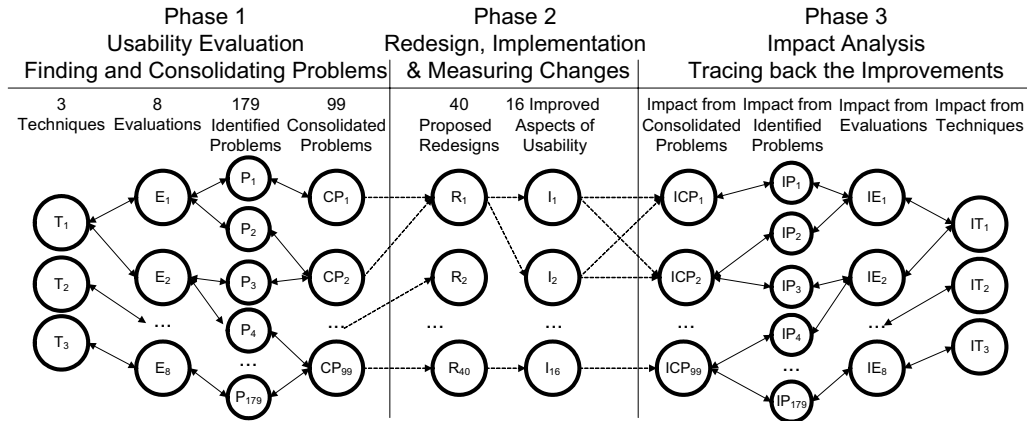
Fig. 1. The process from finding usability problems to estimating the impact from each technique. Phase 1 and Phase 3 are symmetrical and double-arrows indicate transitions that can be reversed, that is if a set of transitions leads from T_1 to CP_2 then the mirrored set leads from ICP_2 to IT_1. The transitions of Phase 2 are marked with one-directional dotted arrows to indicate an influence of judgment, which cannot be unambiguously reversed.

ing on how to cost-justify it. In their book they describe usability work in various organizational settings and in a broad real-world context. They have dedicated a chapter to examples and statistics extracted from various sources. Throughout the book a number of minor case studies are presented, but as the major focus is on costs and benefits, only few details about the usability work are given.

The approach taken in the current study holds important similarities with the user-centered design (UCD) traditions where design teams in iterative processes develop designs based on thorough studies of real users, their tasks, needs, and as responses to changes in the surrounding environment (ISO 13407, 1999). Already the seminal paper by Gould and Lewis (1985) described the central principles of what was later to become known as user-centered design (Norman and Draper, 1986); Gould and Lewis also included a case study that applied these principles in the development of a computer-based message system at IBM. Other UCD case studies include descriptions of usability improvements (e.g., Jokela et al., 2006), but none that we are aware of account for how different usability techniques contribute to specific improvements in usability measures, as is the aim in the current study.

The current study is inspired by a study done by Ebling and John (2000). They linked the expected impact from a set of validated usability problems to the evaluations identifying the problems. Ebling and John assumed that a problem was valid only if identified by more than one evaluation. The number of valid usability problems was then used as a measure of evaluation impact. Their study was done without implementing and measuring actual improvements.

3. Method

This section describes the method applied in the study, and – with the aim of clarifying the analyses – also includes an example of how two usability problems were identified

and dealt with. The study is divided into three phases as shown in Fig. 1. Each phase describes an independent set of activities and has a specific output used in the following phase. The three phases are as follows:

- Phase 1: Usability evaluations. In Phase 1, three usability evaluation techniques are used to conduct a total of eight usability evaluations. The output of Phase 1 is a set of consolidated usability problems.
- Phase 2: Redesign, implementation and measuring changes. In Phase 2, the consolidated problems from Phase 1 help develop a set of redesign proposals, which are subsequently prioritized as to their expected impact on usability and their expected implementation costs. A subset of these proposals is implemented and changes in the use of the system are measured. Output from Phase 2 is a list of improved aspects of usability.
- Phase 3: Impact analysis. In Phase 3 the improved aspects of usability from Phase 2 are linked to the consolidated problems from Phase 1 through a set of estimated contributions. A new measure based on impacts and impact values is introduced to support the analysis. Next, the symmetry of Phase 1 and Phase 3 makes it possible to calculate the contribution from each evaluation and each evaluation technique.

Phase 1 and 2 cover the UIP. They were part of the original case study presented in Uldall-Espersen (2004) and Uldall-Espersen (2005). Phase 3 was planned and conducted afterwards.

3.1. The context of the study

The case study is based on a usability-driven redesign process conducted in a real-life industrial environment so as to increase its realism compared to laboratory-style studies (McGrath, 1995). The original aim was to explore the possibilities of how software developers could use low

cost usability work to drive software redesign and whether it paid off to do so. Therefore, the focus was on keeping evaluations low cost and on balancing the time used with the expected improvement of usability. The system under study was in use in two departments a couple of years before the case study and it was developed and maintained by the first author, hereafter called the responsible systems developer. The organisation in which the study was conducted had no formal procedures regarding usability and the responsible systems developer was made responsible for usability of the end-product.

3.2. System

The system under study was an MS Windows based information system used in a Danish bank. The system was programmed in PowerBuilder and used for administrative purposes, reporting, and risk management of a specialized loan concept with a total loan amount of about 150 million euros. The system included a database and could exchange data with the main banking system. The initial system (version 1) consisted of 12 primary windows, where data could be searched, inserted, and updated. It was the primary tool for administration of the specialized loans. The improved system (version 3) had 11 primary windows. During the study, all users worked with three different versions of the system. Only one version was used at a time, and the use of each version were logged for at least 11/2 month. The changes in usability described in this paper are measured between version 1 and version 3 of the system. Version 2 was a by-product of the iterative work process with no measurement of usability.

3.3. Users

The system was studied for 7 months. During this period the system had 10 users. All users were bank employees with a financial education and worked in two different departments. At the beginning of the case study the newest employee had been in the department for about 9 months. Hence all of the users knew the internal routines and business rules well.

The users were involved in the process from the beginning. They were informed of the purpose of the study, its overall plan, and the importance of their participation. Also, they had an early opportunity to ask questions about the study. One important question came up just after the first meeting where a couple of users contacted the responsible systems developer. The users were concerned that the expected improvements of efficiency would cause a reduction of the staff. This concern was clearly dismissed. The concern was understandable, because a large number of sackings had been announced. The company's insurance activities had suffered from a great storm in Denmark in December 1999, reinsurance obligations to the destruction of the World Trade Center, and the post 9-11 financial recession.

3.4. Design of the study

3.4.1. Phase 1: Usability evaluations

In Phase 1 of the study, the objective was to identify usability problems and prepare a list of consolidated problems. To define usability the ISO 9241-11 (1998) standard was used, namely: "Usability: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." Problems were identified through a questionnaire followed by two iterations of usability evaluations, where a total of three Think Aloud (TA) evaluations and four evaluations using Metaphors of Human Thinking (MOT) were conducted. The three TA evaluations (TA1, TA2, TA3) were based on the descriptions of Contextual Inquiry (Beyer and Holtzblatt, 1998, pp. 41–66) and of TA test (Molich, 1994, pp. 100–115). The TA evaluation approach was selected because we wanted a thorough understanding of the actual use of the system. The MOT evaluation technique is an analytical technique based on five important aspects of human thinking. The evaluations were based on the description by Hornbæk and Frøkjær (2002). The MOT technique was selected as two earlier studies had suggested it a reasonable alternative to widely used inspection techniques such as Heuristic Evaluation and Cognitive Walkthrough (Hornbæk and Frøkjær, 2004a,b). Furthermore, we had an interest in testing the technique in an industrial UIP. The three techniques – questionnaire, TA, and MOT – were selected based on the assumption that they would complement each other, and thus be interesting to combine (Frøkjær and Lárusdóttir, 1999). The questionnaire gave a broad coverage of users experienced with the system; the TA covered a limited number of users more qualitatively; the MOT evaluations were based on theoretical knowledge about human thinking and HCI issues. Also, the evaluations were planned so that the questionnaire could be used to point out critical parts of the system and support the later selection of users who complemented each other for the TA evaluations. An overview of the evaluation process is shown in Fig. 2.

The same questionnaire was administered twice by the sales and marketing department using a web-based tool: first at the start of the study when the logging was enabled and second when the final, redesigned system had been used for about 6 weeks. The users were asked background information (age, gender, etc.) and information about their system usage patterns, that is information about the most important tasks they solve using the system, information about which functionality they find most important, and whether they primarily use the system to retrieve information, update/enter information, or to draw reports. Furthermore, the users were asked 19 questions about their satisfaction with the system, and 20 questions about their experience with the use of the system. These 39 questions were answered on a five-point Likert scale supplemented with freeform comments, where the users could go into details with their answers. The 39 questions are listed in the Appendix.
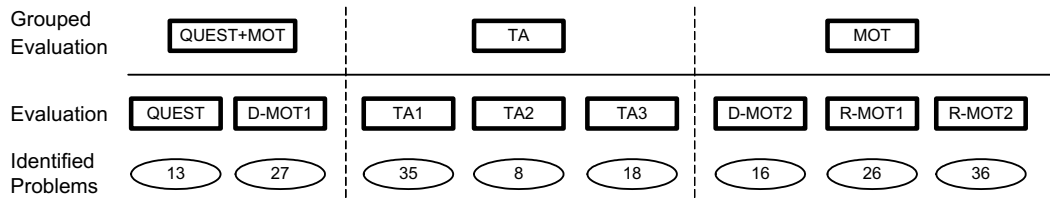
Fig. 2. An overview of the evaluation process. Eight different evaluations were conducted and the number of usability problems identified by each evaluation is shown. To support later analyses the evaluations were grouped into three groups. (QUEST = questionnaire, TA = think aloud evaluation, R-MOT = reviewer-based MOT evaluation, D-MOT = developer-based MOT evaluation).

Based on the tasks stated in the questionnaire three users were selected and asked to collect a couple of hours of typical work, which could be completed during the test sessions. Letting the users bring their own work was necessary as the TA evaluations were conducted using the real production system. By doing so we aimed for the highest possible realism of the tasks and the technical and contextual environments. The use of such very open tasks was done at the expense of the control of the performed tasks, but the tasks had considerable diversity and implied the most important functionality. During the sessions the users were asked to think aloud and the problems were registered when observed. To keep the evaluations at low cost, no sound or video recordings were done. On the one hand some problems might therefore have been missed during each evaluation, but on the other hand it enabled us to do more evaluations within the same timeframe. As our main goal was to identify enough problems to drive the redesign process effectively, we found this a reasonable solution.

The responsible systems developer conducted two developer-based MOT evaluations (D-MOT1, D-MOT2). D-MOT1 was guided by the results of the questionnaire and thus a combined evaluation approach named QUEST+MOT was applied. This approach was assumed useful because both user comments and low usability ratings could be used as a guide to point out important usability issues. D-MOT2 was conducted after the TA-evaluations and followed the procedure of Hornbæk and Frøkjær (2002). After D-MOT2 had been conducted, the two creators of the MOT technique each did one reviewer-based MOT evaluation (R-MOT1, R-MOT2) together with the responsible systems developer. R-MOT1 and R-MOT2 started with an introduction to the system and a presentation of the tasks. Then the reviewers were asked to evaluate the system. During the evaluation, problems were registered and possible solutions discussed, and the responsible systems developer served to answer clarifying questions demanding domain expertise not otherwise available to the external reviewers.

During the evaluations, problems or usability issues were considered to exist when it was assumed that the users work process could be improved by revising the design of specific parts of the system. The assumption could be based on observed user behaviour, user comments, or misfits between parts of the system and the MOT key questions and guidelines. The granularity of the problems was very

different. At the low level a problem could concern how entering a single value was problematic due to some properties of the input field. On the higher level a problem could concern how the process of entering a new loan commitment into the system was too demanding. After the evaluations problems were consolidated. The aim of the consolidation process was to reduce the number of problems to be handled. Similar problems were grouped together and problems were considered similar when they (1) were caused by interaction with the same object or type of object, and (2) the interaction caused the same inconvenience or erroneous situation.

3.4.2. Phase 2: Redesign, implementation and measuring changes

In Phase 2 of the case study, the consolidated problems were used to generate a number of high-level redesign proposals, each dealing with one or more consolidated problems. One of the main concerns in the UIP was to secure a reasonable relationship between costs and benefits. Based on this objective, all of the redesigns were prioritized before implementation, so that no change would be implemented without a reasonable expectation of increased usability. The possible effect of false consolidated problems was considered to be limited, because the consolidated problems were used as inspiration when developing each redesign. In the overall process the responsible systems developer selected to implement redesigns and not to fix single consolidated problems.

The redesign proposals were prioritized from the following four factors, aiming at considering both user interests and business interests.

- How many users are affected by the redesign?
  1. Few (1–10%)
  2. Some (11–50%)
  3. Many (51–100%)
- How often do the covered problems occur?
  1. Rarely (less than once a month)
  2. Often (less than once a week)
  3. Frequent (once a week or more)
- How serious are the covered problems for the business?
  1. Cosmetic (minimal influence on business performance)
  2. Critical (increased use of resources)

3. Catastrophe (erroneous information or loss of data)
• How serious are the covered problems for the users?
  1. Cosmetic (users live with it without problems)
  2. Critical (users are annoyed)
  3. Catastrophe (users give it up)

Each factor was scored on a scale from one to three and the four numbers were then multiplied and divided with the expected time needed to implement the solution. This cost-benefit value was used as an initial guide to select what redesigns to implement, but strategic factors were also taken into account. Section 4.2 presents an example of a strategic prioritization.

As suggested by Frøkjær et al. (2000) the general usability factors effectiveness, efficiency and satisfaction were a priori considered independent factors, and indicators of all three factors were measured in order to assess the total change in usability. The changes in efficiency were measured by logging, changes in satisfaction measured by questionnaires, and changes in effectiveness measured both by logging and by questionnaires.

Throughout the study data from use of the system were logged. The log registered 35 different events. These events covered interaction with windows (Open, Close, Pre-close, Activate, Deactivate), various interactions with the database, clicking, changing values, and change of focus on visual objects (list boxes, edit fields, control buttons). All events were time stamped and related to a user and an action, allowing us to measure a wide range of performance issues and to identify unintended interaction with the system. The initial questionnaire was administered again at the end of the study to uncover changes in satisfaction and effectiveness. The Likert scale ratings from the two questionnaires were quantitatively compared. No qualitative analyses of the comments were conducted on data from the second questionnaire, as this was not within the scope of the initial study. During both Phase 1 and Phase 2 of the study, the time used by the responsible systems developer was registered. An interview with the department manager was also conducted.

3.4.3. Phase 3: Impact analysis

The aim of Phase 3 is to establish a connection from the improved aspects of usability back to the evaluation techniques. First, a description of the impact analysis process is presented, and then the process and calculations are illustrated by an example.

The impact analysis is based on the 99 consolidated problems and the 16 improved aspects of usability (cf. Fig. 1). The responsible systems developer estimated the expected contribution from each consolidated problem in relation to each of the improved aspect of usability. To reduce possible bias caused by the problems' connection to different evaluations and evaluators, problems were treated alphabetically with no information about their relation to evaluation techniques or to individual evaluations. The following question helped estimate the contribution:

"When the problem occurs, how much will the implemented solution contribute to the improved aspects of usability?" A scale from zero (no contribution) to five (very large contribution) was used. This estimated value is called the *Impact Value* and a non-zero impact value is in the following referred to as an *Impact*. In this way an initial estimate of impact values were assigned to the consolidated problems.

These impact values were re-iterated in order to establish a more comprehensive and balanced assessment. For example, if one consolidated problem concerned maintaining investor data, the impact value from this problem was compared to impact values from other problems concerning data maintenance and from other problems concerning investor data. Our concern was that the impact assessments could change as the assessor understood the problem set better and that the assessments could depend on the problem description, and not only on the core of the problem.

This re-iteration was done in a step-by-step process where each of the 99 consolidated problems was categorized using the following three properties: (1) Which of the users' work processes are influenced by the problem? Eleven work processes were distinguished, for example typing in new data, maintaining data, and information search and retrieval. Each of the consolidated problems was related to one or more work processes. (2) What objects of the user interface are influenced by the problem? Twenty-three objects were distinguished, for example presentation of commitments, presentation of mortgages, icons, and drop-down menus. Each of the consolidated problems was related to one or more user interface objects. (3) A short generic description of the consolidated problems without references to work processes or user interface objects. Each problem had one description, for example "undetected error is made by the user" or "missing support of the user's goal". When possible discrepancies were identified the consolidated problems sharing the same categorization properties were further analyzed and adjustments of the assigned impact values were made.

From the impact value of the consolidated problems, it was possible to calculate the impact value of each of the identified problems and then assess the impact value from each of the evaluations and the used techniques. As mentioned earlier, note that Phase 3 is a mirror of Phase 1 in Fig. 1. This symmetry made it possible to reverse the flow in Phase 1, and thereby link the impact value to the evaluations.

The number of impacts indicates how many times a contribution to the improved aspects of usability was made. Very specific problems have a low number of impacts as they influence only few aspects while very general problems have a higher number of impacts because they contribute to several of the improved aspects. The *mean impact value* indicates how much each problem with an impact on average has contributed to the improved aspects of usability.

The measure based on impact values is suggested to support the analysis of the effect of the different evaluations. Also, the degree of overlap between evaluations is detectable through the measures, making it possible to uncover if different evaluations tend to suggest different kinds of usability improvements.

3.5. Example

This section presents an example of how two usability problems were identified and dealt with, explaining how the impact on the improved aspects of usability were estimated and how the impact was traced back to the techniques. The example aims to clarify the process and the calculations used in the analysis.

3.5.1. Problem 1 – Write-down of principal amounts

During one of the TA-evaluations it was observed how principal amounts were kept up to date and synchronized with the main banking system (an MVS based system accessed through a 3270 terminal emulator). The user was switching between the two systems, comparing numbers and, if the conditions were right, updating data in both systems. It was a time consuming and demanding task, which only few users were able to solve. Furthermore, a lot of time was wasted on checking each account, as there was no way to decide whether an account should be updated without comparing data manually. The solution to this problem was to make an interactive report showing data from both systems and, when the conditions were right, allowing the user to update the accounts by pressing a single key. The user still needed to update the main banking system manually, but much time was saved. The "write-down of principal amounts" problem was only identified through the TA evaluation and was not consolidated with other problems. No other identified problems were related to the problem either, so it was the single source to the implemented redesign proposal.

3.5.2. Problem 2 – Entering a commitment twice

When entering a new commitment, a number of usability problems were identified. One major problem in the original system was that it was easy to save the commitment twice without noticing. This problem was identified three times, once in the questionnaire, once during a TA-evaluation and once during a MOT-evaluation. The problem did not influence the users immediately, but caused problems later when the users were adding data and drawing reports. Then the system behaved unpredictably and it was difficult for users to identify the source of the problem. The three problems were consolidated to one unique problem, which together with a handful of other consolidated problems inspired a redesign proposal that changed the process of entering new commitments.

3.5.3. Prioritizing the redesigns

After the two redesigns were developed it was estimated how many hours of work it would take to implement the

changes, the redesigns were scored according to the four criteria, and the priority score calculated (see Table 1). It was estimated that the redesign of the write-down process could be implemented within 5 h and the redesign for entering commitment within 10 h. The redesign of the write-down procedure were considered very important for the business as the failures in the process could introduce erroneous data that would be very hard to find and could give a misleading impression of the current risk. The "entering commitment twice" could also lead to a misleading impression of the current economic risk, but since this error was easily spotted when drawing the risk reports it was considered less serious. Both redesigns would affect similar number of users who despite being annoyed by the problems were expected to be able to fix the problems.

3.5.4. Estimating the impact values of the two consolidated problems

For each of the consolidated problems the impact values were estimated in relation to the improved aspects of usability and the total set of impact values were re-iterated and cross-compared to ensure that the assessments were comprehensive and well balanced. The impact values of the problems in this example are shown in Table 2.

Solving both problems should improve efficiency as the task solving process would be improved and the time wasted on comparing numbers and detecting errors would be reduced. Likewise, satisfaction should increase because the users would get better control of the system behaviour and use less time on non-productive activities. Avoiding the "entering a commitment twice" problem should improve the effectiveness because it reduces the number of uncontrolled and erroneous situations.

During the re-iteration process the estimated impact values from the "write-down of principal amounts" problem were compared to the estimated impact values from other problems with similar properties. The "write-down of principal amounts" problem was categorized with three main

Table 1
Example of redesign prioritising

	Redesign for write-down of principal amounts	Redesign for entering a commitment
Estimated time for implementation (hours)	5	10
How many users are affected by the redesign?	2 (Some)	2 (Some)
How often do the covered problems occur?	1 (Rarely)	2 (Often)
How serious are the covered problems for the business?	3 (Catastrophe)	2 (Critical)
How serious are the covered problems for the users?	2 (Critical)	2 (Critical)
Priority score:	2.4	1.6

The table shows how two redesigns were prioritized.

Table 2
Example of estimated impact values

Aspects of usability	Consolidated problem	
	Write-down of principal amounts	Entering a commitment twice
Task solution time	4	3
Time used to orientate	0	0
Starting time	0	0
Time used to navigate	0	0
General improvements of efficiency	3	3
Reduction of interruptions	0	3
Reduction of unsaved changes	1	3
General improvements of effectiveness	3	3
Satisfaction with navigation	0	0
Satisfaction with adding and summary of properties	0	0
Satisfaction with summary of commitment	0	3
Satisfaction with summary of investor	0	0
Satisfaction with summary of guarantor	0	0
Task solving	3	4
Feedback	0	3
General satisfaction	3	3
Sum of impact values	17	28
Number of impacts	6	9
Mean impact value	2.83	3.11

The table shows the estimated impact values from the two consolidated problems in the example in relation to the improved aspects of usability. The scale goes from zero (no contribution) to five (very high contribution).

properties: (1) The user's work-process was maintaining data. (2) The main user interface object was the presentation of principal amounts data. (3) The generic problem description was "lack of support for users' work process". Five other problems concerned a lack of support for maintaining other types of data and the impact from these problems was compared to the estimated impact from the "write-down of principal amounts" problem. A few adjustments were made when the deviation or lack of deviation between the estimated impacts could not be justified. In a similar manner the impact values from the "entering a commitment twice" problem were compared to impact values from other problems categorized with the properties: (1) entering new data, (2) interacting with the commitment data presentation, and (3) lack of support for users' work process.

3.5.5. Analyzing the impact values

The impacts from each of the problems are described through three summarized measures: Sum of impact values, number of impacts with a positive impact value and mean impact value, defined as (sum of impact values)/ (number of impacts). The three rows in the bottom of Table 2 show the summarized measures for the two example problems. As the purpose of the analysis was to link the impact values to the techniques, each of the two consoli-

dated problems were split back into the original problems. This had no effect on the "write-down of principal amounts" problem as it was only identified in one evaluation. The "entering a commitment twice" problem was identified in three different evaluation sessions and was therefore contributing with one problem to each of the three source evaluations. This made it possible to summarize the impact values on evaluation and evaluation group level, which is done in Table 3. Here it is shown that two problems are identified by TA-evaluations, one by MOT and one by QUEST+MOT. This results in a higher group number of impacts and group total impact value for TA than for MOT and QUEST+MOT, but a lower group mean impact value. In our example the TA evaluation is estimated to lead to a higher number of improvements and a larger total improvement of usability, but each improvement introduced by MOT and QUEST+MOT on average is estimated to be more significant.

4. Results

4.1. Phase 1: Usability evaluations

In Phase 1 of the study 179 potential usability problems were identified. The problems were subsequently consolidated into 99 unique problems. Fig. 2 gives an overview of the evaluation process and the number of identified problems.

4.2. Phase 2: Redesign, implementation and measuring changes

In Phase 2, the 99 consolidated problems were used as source to developing 40 redesign proposals. Thirty of the proposals were implemented. Applying cost-benefit analyses at this stage in the UIP helped select which redesigns to implement, but strategic issues could not be ignored. For example, one of the major improvements in the case study was a total re-engineering of the user manual and help system. As an isolated activity this would never pay off based on expected costs and benefits, but in a larger perspective it increased the value of the system. The help system was rarely accessed and we found no reason to believe that it improved the effectiveness or efficiency of the work processes. Yet the users felt more secure when using the

Table 3
Example of impact analysis

	Number of problems	Group number of impact	Group total impact value	Group mean impact value
QUEST+MOT	1	9	28	3.11
TA	2	15	45	3
MOT	1	9	28	3.11

The table shows how the impact values from the two problems in the example are summed up in relation to the grouped evaluations.

system and the system was easier to deploy in a new department, which happened after the study was finished.

About 250 000 log entries were obtained from about 190 h of active use of the system. These entries covered 509 different sessions. Nine of the 10 users chose to answer the questionnaires. From the two questionnaires a total of 360 questions about user experience (180 in each questionnaire) and 266 about satisfaction (138 in the first and 128 in the last) were answered. Based on the log and the ratings in the questionnaires, 13 improved aspects of usability were identified. These aspects were used in the impact analysis together with the three general aspects of usability found in the ISO 9241-11 definition. The 13 aspects are shown in Table 4.

The log shows that the improvements in efficiency were primarily obtained by a better adaptation to users' work practices and by a reduction in the time used to start up, orientate, and navigate. For a number of these important processes, performance was improved by 40–76%. For example, twice a year the users had to print a summary of every commitment (about 150) and of every investor (about 600). This was done by manually selecting one commitment or investor at a time and then pressing the print icon, a rather time consuming process. In the final version of the system a "print all commitments" and a "print all investors" function were added. In a similar manner the data maintenance processes were re-engineered.

The time from starting the system to viewing relevant data was reduced from 43 to 31 s (28%). The database access time did not change significantly during the case study, and the observed effect was obtained by moving the search facilities to one window and by changing the navigation. From the measured start up and search times, the case study shows a reduction of time used to navigate in the order of 60–70%.

The time from a user opens a window until the user interacts with it (called *orientation time*) shows another improvement. For the 90% percentile of more than 7800 observations, the time has been reduced from 8.3 to 5.8 s (30%). The 90% percentile was considered to cover the normal use of the system.

Effectiveness was improved as well. A reduction in interruptions of about 88% was observed. The reduction could partly be explained by changes in feedback and error handling. The result was fewer disturbances of the users. Likewise, a 79% reduction in the number of times the users start changing data without saving was observed. From the questionnaire, two experience questions (A3, A17) and six satisfaction questions (B2, B4, B9, B10, B13, B14) showed improvements (0.5–1 point). The overall experience did only improve slightly (0.2 points), and only one user expressed a significant change (0.6 point). The total satisfaction was improved from 3.56 to 4.03 and four of nine users expressed significant improvements (0.7–0.9 point). No relationship between the users, their tasks, and their satisfaction with the system were found.

Table 4
The improved aspects of usability

Improved aspects of usability	Description of measure	Relative improvement (%)
Starting time	Time from system start up to first set of user-selected data is displayed	28
Task solution time	Time used to complete a set of different tasks	40–76
Time used to orientate [a]	Time used from a window is activated until the user interacts with it	30
Time used to navigate	Measured indirectly by the starting time subtracted the time used to search for the first set of data	60–70
Reduction of interruptions	An interruption is identified when a window lose focus without being closed	88
Reduction of unsaved changes	Unsaved changes are identified when data in a window has been edited without a following update of the database	79
Navigation (B2[b])	Satisfaction with selection of and navigation between windows	32
Adding/general view of a property (B4[b]/B10[b])[c]	Satisfaction with adding a property and with the general view of properties	23/13
General view of commitment (B9[b])	Satisfaction with general view of commitments	18
General view of investor (B13[b])	Satisfaction with general view of investors	13
General view of guarantor (B14[b])	Satisfaction with general view of guarantors	24
Task solving (A3[b])	Level of agreement with the statement: There are tasks the system should be able to solve, which cannot be solved	30
Feedback (A17[b])	Level of agreement with the statement: I miss feedback from the system	19

The table shows the 13 improved aspects of usability in the study.
[a] This measure is based on the 90% percentile. This is done to remove a number of unexplained outliers.
[b] Improved aspects of usability, which are uncovered through the stated question. Questions used in the questionnaire are shown in Appendix.
[c] As property data was added and viewed in the same window, B4 and B10 could not be separated in the impact analysis. Therefore they were treated as one aspect.

The final evaluation was an interview with the responsible department manager, who also was one of the primary users of the system. The interview revealed a general satisfaction with the process undertaken and the usability results. The manager found that the new system was better and easier to use and that it gave the users a better general view of data. Also, the users had gained more confidence with the system and they felt less insecure when using it. The manager found it difficult to estimate the actual economic benefits, but a planned replacement of the system would be postponed. He also emphasized how the improvements of work effectiveness and user satisfaction were more important and interesting to him than the efficiency gains.

4.3. Time used on evaluation and redesign in Phase 1 and Phase 2

During the study, the time used by the responsible systems developer was registered. Overall 15 h were used to study the techniques, 25 h to analyze the system, 10 h to analyze and prioritize problems and solutions, and 117 h to implement improvements. Of the 25 h used for analysis, eight were used on the questionnaire, 41/2 on the three TA sessions, 101/2 on the four MOT evaluations, and 2 h on other activities. The 101/2 h used on MOT evaluations covers the introduction, evaluations, and a discussion of the problems and possible solutions. Only few solutions were discussed during the TA sessions. When discussing the time used for the different activities, it should be noticed that both TA and MOT did benefit from the questionnaire. The tasks used in the TA and MOT evaluations and the users selected for the TA evaluation were identified from the questionnaire data.

4.4. Phase 3: Impact analysis

With the 99 consolidated problems and the 16 improved aspects of usability, the first task in Phase 3 was to estimate

the 1584 (99 consolidated problems * 16 improved aspects) possible impacts. Estimating the impact values required thorough system expertise, as fixing one usability problem could influence more than one of the improvements. An example of this is the re-engineering of the search facilities, moving the search functionality from a number of different windows to one centralized window. This influenced the usability of the search task, the usability of the windows where the users used to search, and the usability of processes where searching was a natural activity.

4.5. Analysis of impact values

Data from the impact analysis are shown in Table 5, Table 6 and Table 7. Table 5 shows the number of impacts and the mean impact values from each evaluation and each evaluation group, and the group impact per evaluation in each evaluation group. Table 6 shows the differences between the evaluations and within the evaluation groups found by an analysis of variance (ANOVA) and Bonferroni post hoc tests. Table 7 shows the differences between the evaluation groups found by an ANOVA and Bonferroni post hoc tests.

Within the evaluation groups in Table 5 we find no significant differences (Table 6). Looking at the grouped evaluations, however, we find differences between the three evaluation groups (Table 7). Among groups, MOT has the largest number of impacts, but the smallest mean impact value. QUEST+MOT has the greatest mean impact value and the greatest impact value per evaluation. TA has the lowest impact per evaluation, but a number of impacts and a mean impact value between MOT and QUEST+MOT.

Looking at the distribution of the size of the impact values between the three groups of techniques, we observe a clear trend (Fig. 3). TA and QUEST+MOT reveal more problems with a middle impact value, while MOT reveals more problems with a smaller impact values. Problems identified by MOT might be less visible to the users because

Table 5
Size of impact values grouped by evaluation

Evaluation group	Evaluation	Number of impacts	Mean impact	Group number of impacts	Group mean impact value	Group impact per evaluation
QUEST +MOT	Questionnaire	125	2.93	394	2.85	561
	D-MOT1	269	2.82			
TA	TA1	292	2.59	509	2.63	446
	TA2	66	2.58			
	TA3	161	2.73			
MOT	R-MOT1	241	2.17	703	2.22	520
	R-MOT2	326	2.29			
	D-MOT2	136	2.15			

Number of impacts and mean impact values for each of the eight evaluations and for the three evaluation groups. There are significant differences between the evaluations, $F(7, 1608) = 15.989$, $p = .000$) and between the groups, $F(2, 1613) = 53.084$, $p = .000$. The differences within the groups are not significant (cf. Table 6), but all differences are significant among the groups (cf. Table 7). Group impact per evaluation = (Group number of impacts * Group mean impact value)/(Number of evaluations in group).

Table 6
Differences of impact values between the individual evaluations

Evaluation groups	QUEST+MOT		TA			MOT		
Techniques	Quest	DMOT1	TA1	TA2	TA3	DMOT2	RMOT1	RMOT2
Quest	–							
DMOT1	$F(1,392)=1.096$ $p=.296$	–						
TA1	$F(1,415)=9.092$ $p=.003$	$F(1,559)=6.866$ $p=.009$	–					
TA2	$F(1,189)=6.158$ $p=.014$	$F(1,333)=3.423$ $p=.065$	$F(1,356)=0.009$ $p=.926$	–				
TA3	$F(1,284)=2.679$ $p=.103$	$F(1,428)=0.737$ $p=.391$	$F(1,451)=1.899$ $p=.169$	$F(1,225)=1.206$ $p=.273$	–			
DMOT2	**$F(1,259)=45.866$** **$p=.000$**	**$F(1,403)=44.777$** **$p=.000$**	**$F(1,426)=16.725$** **$p=.000$**	$F(1,200)=10.388$ $p=.001$	**$F(1,295)=26.831$** **$p=.000$**	–		
RMOT1	**$F(1,364)=46.553$** **$p=.000$**	**$F(1,508)=53.653$** **$p=.000$**	**$F(1,531)=20.912$** **$p=.000$**	$F(1,305)\ 8.686=$ $p=.003$	**$F(1,400)=29.208$** **$p=.000$**	$F(1,375)=0.012$ $p=.913$	–	
RMOT2	**$F(1,449)=30.760$** **$p=.000$**	**$F(1,593)=36.318$** **$p=.000$**	**$F(1,616)=11.353$** **$p=.001$**	$F(1,390)=3.790$ $p=.052$	**$F(1,485)=17.575$** **$p=.000$**	$F(1,460)=1.420$ $p=.234$	$F(1,565)=1.625$ $p=.203$	–

F and p values from ANOVA between evaluations and within evaluation groups. The bolded cells indicate significant differences between two evaluations confirmed by Bonferroni post hoc tests.

Table 7
Differences of impact values between the grouped evaluations

	QUEST+MOT	TA	MOT
QUEST+MOT	–		
TA	**$F(1,911)=10.672$** **$p=.001$**	–	
MOT	**$F(1,1095)=95.415$** **$p=.000$**	**$F(1,1220)=45.905$** **$p=.000$**	–

F and p values between evaluation groups computed by ANOVAAll the observed differences are confirmed significant by Bonferroni post hoc tests.
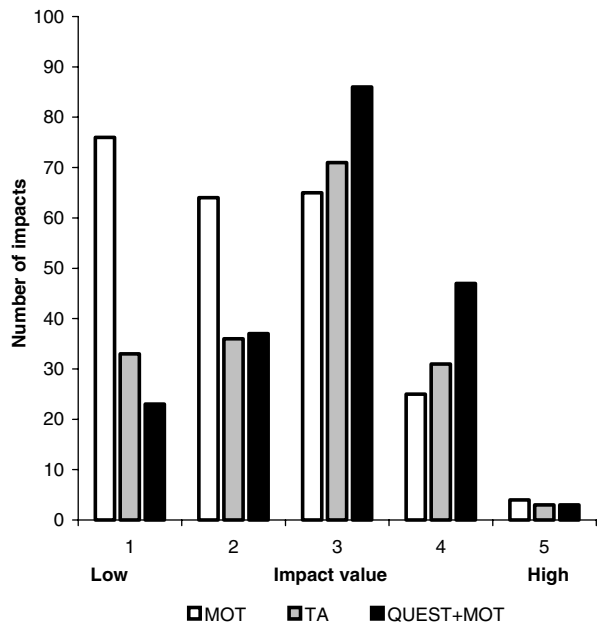

Fig. 3. The figure shows the distribution of impact values per evaluation group. MOT reveals more problems with a low impact value and fewer middle impact value problems than TA and QUEST+MOT.

evaluators lacked knowledge about the system and the users' present work practices. Instead, the MOT evaluations might be dominated by a more abstract vision and expertise about technological options including human use of technology, both through the methodology itself and through the external reviewers. Both the TA evaluations and the QUEST+MOT evaluations rely on concrete experiences and expertise about users' present work practices, which could help identifying problems with greater impact. However, the large number of small-impact problems seems to have an important influence on the total impact per evaluation where MOT performs significantly better than TA.

Comparing the impact of the eight evaluations across the 16 improved aspects of usability, significant differences are found in six cases (Table 8). The differences concern the three general usability aspects (general satisfaction, general effectiveness, and general efficiency) and three specific aspects (task solution time, task solving satisfaction, and time used to orientate). A Bonferroni post hoc analysis confirmed some significant pair-wise differences between evaluations. The most important results are that the questionnaire had a greater contribution to task-solving satisfaction and general satisfaction, with mean impact values (MIV) both on 3.73 compared to the MOT-evaluations (MIV = 2.38–2.46/2.57–2.77) ($p = .002$–.012). Other differences regarded task solving time and general efficiency, where the questionnaire (MIV = 2.91/3.09) contributed more than D-MOT2 (MIV = 1.53/1.73) ($p = .035$).

Table 8 shows the mean impact value from each group of techniques on the six improved aspects of usability. In all six cases MOT is significant different from TA; in five out of six cases MOT differs from

Table 8
Mean impact value from evaluation groups in relation to six selected aspects of usability

Number of problems	TA	MOT	QUEST +MOT	Significance test
	47	56	35	
General efficiency	2.53	2.07	2.91	$F(2, 135) = 6.298$, $p = .002$
General effectiveness	2.62	1.77	2.97	$F(2, 135) = 12.461$, $p < .001$
General satisfaction	3.13	2.70	3.54	$F(2, 135) = 7.017$, $p = .001$
Task solution time	2.47	1.86	2.77	$F(2, 135) = 6.262$, $p = .001$
Time used to orientate	1.02	1.52	1.69	$F(2, 135) = 4.104$, $p = .019$
Task solving satisfaction	2.98	2.45	3.31	$F(2, 135) = 8.664$, $p < .001$

The table shows the mean impact value of the three technique groups in relation to the three general aspect of usability and the three specific aspects of usability where the grouped evaluations perform significantly different. In five out of six cases TA and QUEST+MOT have significantly greater mean impacts values than MOT. Only in "time used to orientate" has MOT a greater mean impact value than TA.

QUEST+MOT. Only in the case of time used to orientate do QUEST+MOT and MOT perform equally well and better than TA. QUEST+MOT has the greatest mean impact value in all six cases, and in five out of six cases TA comes in second. It should be noticed that the responsible systems developer was the person identifying usability problems during the QUEST+MOT evaluations and the person who estimated the impact. This could bias the estimated impact.

4.6. Distribution of problems

To be able to estimate the effect of combining the techniques, the contribution from each technique was analyzed. This analysis is based on the original problems identified. The validity of the problems was continuously assessed during the whole process and no indication of false problems was identified. Thus every problem was included in the analysis. Table 9 shows the percentage of problems identified through combinations of the different evaluation groups. First is shown the distribution of all problems and second the distribution of problems influencing orientation time. Few problems are identified by all three techniques; the vast majority is identified by one technique only. Looking at all problems, MOT and TA seem to perform equally well and better than QUEST+MOT.[1] The same trend is observed when analyzing problems that influence task solution time and task solving satisfaction. Analyzing the time used to orientate shows a different result. Here, the major changes are that MOT performs better than both TA and

Table 9
Identified problems distributed over evaluation groups

Combination of evaluation group	All problems (%)	Problems influencing orientation time (%)
Only TA	30	15
Only MOT	37	43
Only QUEST+MOT	11	15
TA and MOT	7	11
TA and QUEST+MOT	8	6
MOT and QUEST+MOT	4	6
All Identified by TA	47	36
All Identified by MOT	50	64
All Identified by QUEST+MOT	25	31
Identified by only one technique	78	73
Identified by two techniques	19	23
Identified by all techniques	2	4

The table shows the percentage of problems identified through combination of the different evaluation groups for all problems and problems influencing orientation time. Major changes are found in the three bolded combinations.

QUEST+MOT, and that TA and QUEST+MOT performs equally well.

The distribution suggests to not rely on a single technique. Even though all of the techniques have a significant influence on the total improvement of usability, the overlap between techniques is quite small. This suggests that we shall apply different techniques to cover a broader range of usability issues.

4.7. Impact for effort

Looking on impact-for-effort could indicate how to gain the most from the invested resources. Table 10 shows the impact value per hour invested in analyzing the system. TA has the highest impact value per hour and QUEST+MOT the lowest, indicating that with at limited amount of time, TA would be the best choice of technique. It should be noted that the TA tests were conducted with a minimal usage of technological and human resources. Time was only used to observe the users, which in this case proved efficient. In contrast, the questionnaire needed to be prepared thoroughly and the qualitative analysis was time consuming, resulting in a low impact value per hour. The MOT evaluations were influenced by the fact that the external reviewers did not know the system before the evaluation started. The reviewers had to use considerable time getting acquainted with the system, which partly explains the lower impact value per hour.

From the measured impact-for-effort we cannot conclude that it would be sufficient or desirable just to use the time on TA tests because each evaluation help identify problems with different types of impact. In our case important impact would have been missed if we had relied only on TA tests. When planning the UIP it should be considered what kind of impact we aim at and this consideration should guide the resources invested in the different evalua-

---

[1] Note that the numbers for QUEST+MOT is based on two evaluations, while the numbers for MOT and TA are based on three.

Table 10
Impact-for-effort

	Total impact value	Effort in hours	I/E ratio
TA	1338	4.5	297
MOT	1560	7.5	208
QUEST+MOT	1122	11	102

The I/E ratio indicates that impact from a 1-h TA session is much higher than for a 1-h MOT or QUEST+MOT session.

tions. If our goal is to produce a balanced impact from each applied evaluation technique we might want to use relative less time on TA test and relative more time on techniques like MOT because the MOT evaluations in our case seems to be more time consuming. Such balancing supports a better overall distribution of impact and reduces the risk of sub optimization, which the use of a single evaluation technique would introduce.

5. Discussion

This paper presents a case study and a first attempt to link real-life usability improvements with the usability techniques that give rise to the improvements. As the proposed method and measures used to analyze and describe the UIP are new, we should evaluate both the validity and interpretation of the results as well as the usefulness and the validity of the method.

Conducting a case study in industrial settings allows for high realism, but limits control over the situation under study because the participants need to carry out their daily work. Phase 1 and 2 was a realistic case study of industrial usability work, where a developer improved the usability of an information system. Phase 3 attempts to link actual usability improvements and techniques. The method seems applicable, but to ensure higher validity some further steps should be taken in future projects. The main issue is that Phase 3 should be taken into careful consideration when planning and conducting Phase 1 and 2. Especially there is a need for cross-validation of the estimated impact. The impact estimation process requires a thorough understanding of the application, the implemented changes, and the improvements obtained. In the current study only the first author had this understanding. He was the only developer of the system and sharing his expertise was not realistic. Sharing the expertise with colleagues would be very time consuming, and as only limited access to the studied system could be obtained, this was not possible to do at a sufficiently detailed level. Asking users to cross-validate the estimated impact was not an option either, as they would have had to spend considerable time familiarising themselves with the task of re-estimating impacts. Their part of the project had been finalized for some time and the department had no additional interest in this research. In future studies though, this should be handled differently, for example by involving peer developers or users in estimating the impact values.

Which measures to focus on when selecting evaluation technique is open for discussion. In this study QUEST+MOT has the highest mean impact value of the three evaluation groups. This could be explained by the involvement of the large number of users adding concrete contextual information to the evaluation. A high mean impact value is desirable if communicating a limited number of usability problems is a priority, for instance when evaluations have to be quick or when a short usability report is desirable.

MOT has the highest total impact value and the highest number of impacts of the three evaluation groups. This could be explained by properties of the evaluation technique combined with the involvement of external reviewers adding a fresh set of eyes and external expertise to the evaluations. Software development in specific organizational settings relies heavily on fixed practices, for example development standards and stakeholders' habits, and external reviewers can more easily question such practices. A high total impact value or impact value per evaluation is desirable when aiming at the most substantial summarized improvement of usability. However, these measures should be considered in combination with the mean impact value, since it might inform the outcome differently if the total impact originates from high impact problems or from low impact problems. This study documents large variations in impact value per evaluation within groups and minor variations among groups, suggesting that these two measures should not be considered strong indicators for the selection of techniques. Furthermore, these measures tell us little about the distribution of impact. If only the evaluation technique with the highest impact is selected, there is a risk of sub optimizing certain usability aspects and leaving other important aspects untouched.

TA has the highest impact-for-effort value of the three evaluation groups. On explanation may be that TA in this study required low effort rather than a high impact compared to the other evaluation groups. The relationship between impact value and effort could help us select an evaluation strategy, but selecting techniques only from impact-for-effort would be a mistake. We should be aware that impact-for-effort depends on the amount of effort. For example, doubling the number of TA evaluations is unlikely to double the produced impact.

Our focus has been on suggesting and prioritizing redesigns for implementation and not on dealing with problems on an isolated level. Hornbæk and Frøkjær (2005) found that redesigns could be a useful tool to reduce the gap between usability evaluations and software development. We have applied a method for prioritizing redesign based on problems, which takes both business and user interests into account. This approach could also be useful in projects where a number of proposed redesigns, some based on usability problems and other based on different sources of inspiration, should be prioritized. This could help prior-

itize competing activities in a software development process, which could help reaching a more coherent and well-balanced view on software development and usability work.

In the usability improvement process four redesigns were considered particularly important. These redesigns were based on multiple consolidated problems and on data from multiple evaluation groups. All four redesigns evolved over time and introduced major changes for all the users, for example by changing the navigation and the search functionality and the general presentation of data. This observation suggests that it sometimes is advantageous to deal with problems as parts of a greater whole and to let solutions evolve over time. By relying on multiple sources of information we might be able to develop a deeper understanding of the problems allowing us to introduce more comprehensive and profound redesign solutions with a higher impact on usability. This coheres well with the conjecture by Hornbæk and Frøkjær (2005) who in a discussion of developers use of redesigns and usability problems wrote: 'Further, developers who for years have worked intensively with the application and its use context will not easily take up results of usability evaluations. On the contrary, changing their understanding is a process requiring time, during which new insights does not appear as something distinct and immediately clear. Rather, developers will experience nagging doubts, small changes in thinking, and challenges to their habitual understanding.'

This study has involved different persons with special kinds of expertise. The users were highly skilled domain experts who contributed their knowledge especially in answering the questionnaire, in the planning and completion of TA evaluations, and by allowing logging of their system usage. The responsible systems developer contributed concrete insight into the existing IT-system as well as more general knowledge about information systems, their development, and usage. Further he planned and managed the interplay between usability evaluations and redesigns. The experiment was only possible because of his existing understanding of the specific organizational setting. The external reviewers contributed with general expertise in HCI and systems design in their MOT evaluations. The analysis of impacts has uncovered how these different persons contributed with important expertise to the UIP. These different kinds of expertise brought into play in the evaluation and redesign cycles cohere well with the six areas of knowledge described as needed in design projects by Bødker et al. (2004, p. 62) building on Kensing and Munk-Madsen (1993). They distinguish between (a) abstract knowledge and (b) concrete experience in relation to the knowledge about (1) users' present work practices, (2) new IT usage and (3) the technological options. Also Woolrych et al. (2005) emphasize the importance of combining different knowledge resources in usability evaluations, namely (1) knowledge about users and abilities, (2) task knowledge, (3) domain knowledge, (4) interaction knowledge, (5) product knowledge, (6) design knowledge and (7) technical knowledge. All these knowledge resources influenced the evaluation and redesign activities of this study.

In summary, our study shows how the UIP gains from bringing the different areas of expertise into play. This is done partly by applying different techniques, and partly by close cooperation between people with different types of expertise.

6. Conclusion

This paper describes a real-life case study where the usability of an information system is improved. The usability improvement process is analyzed and improvements traced back to the set of evaluations, which suggested the improvements. The usability impact of the improvements is assessed and a relationship between this impact and the evaluation effort calculated.

The analysis shows that the applied techniques, namely Think Aloud (TA), Metaphors of Human Thinking (MOT), and questionnaire combined with MOT (QUEST+MOT), contribute with considerable, but different improvements of usability, suggesting not to rely on a single technique. The three techniques impact the interaction design differently and the overlap between problems identified with the three techniques is small. Thus we should apply complementing techniques to cover a broader range of usability issues. The main result is that the combination of questionnaire and MOT has the largest mean impact value and the largest impact per evaluation. Further, the MOT evaluation group has the largest number of impacts. Six improved aspects of usability show significant differences among evaluations; in five aspects TA has a mean impact value equal to QUEST+MOT, and a higher mean impact than MOT. Concerning one usability aspect, that is time used to orientate, MOT and QUEST+MOT have a significant higher mean impact than TA. Regarding problems influencing time used to orientate, MOT identifies 64% of the total problem set and 43% of the problems are identified only by MOT.

Similarly to the importance of combining complementary evaluation techniques, the involvement of people with complementing areas of expertise stands out as effective. The study shows how it was important to allow redesigns to evolve based on inspirations from multiple evaluations and over time, and four particularly important redesigns evolved throughout the work process, that is they did not emerge as sudden insights to the responsible systems developer or sudden creative ideas. Combining different kinds of expertise from the people involved in the usability improvement process was important. Think aloud tests and questionnaires as used in this experiment draw heavily on expertise about the users' present work practices. This expertise was instrumental in uncovering problems with high impact on usability. During the MOT evaluations external reviewers contributed with expertise about interac-

tion design issues, information visualization and IT-usage. The mean impact from problems identified by MOT or external reviewers is smaller than the mean impact from problems found by user-based evaluations, that is TA tests and questionnaires. Yet problems found by MOT is clearly complementing the impact from the user-based evaluations, since 37% of the problems only were identified during the MOT evaluations.

Acknowledgements

Appendix A

Questions about user experience:

How much do you agree in the following questions?

A1. There are tasks in the system, which are difficult to solve.

A2. There are tasks in the system, which are to time consuming to solve.

A3. There are tasks the system should be able to solve, which cannot be solved.

A4. There are tasks in the system I do not know how to solve.

A5. There are special tasks in the system, which I hand over to others.

A6. There are special tasks in the system, which are often handed over to me.

A7. There are tasks in the system, which I often need help from other in the department to solve.

A8. There are tasks in the system, which I have to help other solving.

A9. There are tasks in the system, which I avoid solving.

A10. There are parts of the system, which annoy me when I use it.

A11. There are parts of the system, which I use without understanding it.

A12. I am insecure about how to solve a task and try to do it my way.

A13. There are parts of the system, which I feel insecure using.

A14. There are parts of the system, which often give rise to errors.

A15. There are shortcomings in the system.

A16. It is difficult to get around in the system.

A17. I miss feedback from the system.

A18. The system does something different from what I expect.

A19. The system expects me to solve a task in another sequence, than I would have done.

A20. I often make errors in parts of the system.

Specific questions about user satisfaction:

How satisfied are you with the following parts of the system?

B1. The search function where you can select commitment, investor or guarantor from a list.

B2. Selection of and navigation between windows.

B3. Entering a new commitment.

B4. Adding a property.

B5. Adding a lease.

B6. Adding a mortgage.

B7. Entering a new investor.

B8. Entering a new guarantor.

B9. General view of commitments.

B10. General view of properties.

B11. General view of leases.

B12. General view of mortgages.

B13. General view of investors.

B14. General view of guarantors.

B15. Reports.

B16. Import of data from data warehouse.

B17. The total general view.

General questions about user satisfaction:

B18. How satisfied are you generally with the system?

B19. If a colleague in another department were thinking of starting using the system, would you recommend it?

References

Beyer, H., Holtzblatt, K., 1998. Contextual Design. Morgan Kaufmann.

Bias, R.G., Mayhew, D.J., 2005. Cost-Justifying Usability. An update for the Internet Age. Morgan Kaufmann.

Butler, K.A. 1985. Connecting theory and practice: a case study of achieving usability goals. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM Press, San Francisco, California, United States, pp. 85–88.

Bødker, K., Kensing, F., Simonsen, J., 2004. Participatory IT Design. The MIT Press, Cambridge, Massachusetts.

Ebling, M.R., John, B.E. On the contributions of different empirical data in usability testing. In: Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, ACM Press, New York City, New York, United States.

Frøkjær, E., Hertzum, M., Hornbæk, K. 2000. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM Press, The Hague, The Netherlands.

Frøkjær, E., Lárusdóttir, M., 1999. Prediction of usability: comparing method combinations. Proceedings of IRMA International Conference, 248–257.

Følstad, A., 2007. Work-domain experts as evaluators: usability inspection of domain-specific work-support systems. International Journal of Human–Computer Interaction 22 (3), 217–245.

Good, M., Spine, T.M., Whiteside, J., George, P., 1986. User-derived impact analysis as a tool for usability engineering. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM Press, Boston, Massachusetts, United States, pp. 241–246.

Gould, J.D., Lewis, C., 1985. Designing for usability: key principles and what designers think. Communications of the ACM 28 (3), 300–311.

Gray, W.D., Salzman, M.C., 1998. Damaged merchandise? a review of experiments that compare usability evaluation methods. Human–Computer Interaction 13 (3), 203–261.

Hartson, H.R., Andre, T.S., Williges, R.C., 2001. Criteria for evaluating usability evaluation methods. International Journal of Human–Computer Interaction 13 (4), 373–410.

Hornbæk, K., Frøkjær, E. 2002. Evaluating user interfaces with metaphors of human thinking. In: Proceedings of User Interfaces for All, Lecture Notes in Computer Science 2615, Springer-Verlag, pp. 486–507.

Hornbæk, K., Frøkjær, E., 2004a. Usability inspection by metaphors of human thinking compared to heuristic evaluation. International Journal of Human–Computer Interaction 17 (3), 357–374.

Hornbæk, K., Frøkjær, E. 2004b. Two psychology-based usability inspection techniques studied in a diary experiment. In: Proceedings of Nordic Conference on Human–Computer Interaction (NordiChi 2004), pp. 2–12.

Hornbæk, K., Frøkjær, E. 2005. Comparing usability problems and redesign proposals as input to practical systems development. In: Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, ACM Press, Portland, Oregon, USA, pp. 391–400.

ISO 13407, 1999. Human-Centered Design Processes for Interactive Systems, ISO/IEC.

ISO 9241-11, 1998. Ergonomic requirements for office work with visual display terminals (VDT)s – Part 11 Guidance on usability, ISO/IEC.

Jeffries, R., Millar, J.R., Wharton, C., Uyeda, K.M., 1991. User interface evaluation in the real world: a comparison of four techniques. ACM Press, New Orleans, Louisiana, USA, pp. 119–124.

Jokela, T., Koivumaa, J., Pirkola, J., Salminen, P., Kantola, N., 2006. Methods for quantitative usability requirements: a case study on the development of the user interface of a mobile phone. Personal Ubiquitous Computing 10 (6), 345–355.

Karat, C.-M., Campbell, R., Fiegel, T. 1992. Comparison of empirical testing and walkthrough methods in user interface evaluation. In: Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, ACM Press, Monterey, California, USA, pp. 397–404.

Kelkar, K., Khasawneh, M.T., Bowling, S.R., Gramopadhye, A.K., Melloy, B.J., Grimes, L., 2005. The added usefulness of process measures over performance measures in interface design. International Journal of Human–Computer Interaction 18 (1), 1–18.

Kensing, F., Munk-Madsen, A., 1993. PD: structure in the toolbox. Communications of the ACM 36 (6).

McGrath, J.E., 1995. Methodology matters: doing research in the behavioral and social sciences. In: Baecker, R.M. et al. (Eds.), Readings in Human–Computer Interaction. Morgan Kaufmann Publishers Inc.

Molich, R., 1994. Brugervenlige EDB-systemer, Second ed. Ingeniøren Bøger.

Norman, D.A., Draper, S.W. User Centered System Design New Perspectives on Human–Computer Interaction, Lawrence Erlbaum Associates, Inc.

Ruthford, M.A., Ramey, J.A. 2000. The design response to usability test findings: a case study based on artifacts and interviews. In: Proceedings of IEEE Professional Communication Society International Professional Communication Conference and Proceedings of the 18th Annual ACM International Conference on Computer Documentation: Technology & Teamwork, Cambridge, Massachusetts: IEEE Educational Activities Department, pp. 315–323.

Uldall-Espersen, T. 2004. Benefits of usability work – does it pay off?, (In Danish), Department of Computer Science, University of Copenhagen. (http://www.diku.dk/~tobiasue/thesis.exe).

Uldall-Espersen, T. 2005. Benefits of usability work – does it pay off?, Proceedings of the INTERACT 2005 International Conference on Human–Computer Interaction, Rome, Italy: CEUR-WS, pp. 7–14.

Wixon, D., 2003. Evaluating usability methods: why the current literature fails the practitioner. Interactions 10 (4), 28–34.

Woolrych, A., Cockton, G., Hindmarch, M., 2005. Knowledge resources in usability inspection. In: Proceedings of HCI 2005. Springer, Edinburgh, Scotland, pp. 15–20.

Scientific Part - Page 17

Exploring multiple usability perspectives

Tobias Uldall-Espersen
Department of computing, University of Copenhagen, Universitetsparken 1,
DK-2100 Copenhagen, Denmark
tobiasue@diku.dk

Abstract. Industrial usability work often fails to produce the expected impact on software products even though significant resources have been used on uncovering problems and suggesting improvements. So, it seems that feedback from industrial usability work lacks persuasiveness, i.e. it fails to convince the key stakeholders that actions need to be taken. This study reports from interviews with 26 stakeholders in software development projects. Our data suggests that the interviewees address usability using different perspectives and based on our observations we describe five such perspectives. Further, we discuss how applying different usability perspectives might inform the persuasiveness of usability work.

1 Introduction

One important problem when developing software is that usability work does not sufficiently inform software development even though a large number of usability issues are identified. This problem is in the literature described as lack of design-change effectiveness [4], lack of downstream utility [2], or lack of impact [1], and can partly be explained by lack of persuasive power [4] in the usability feedback. Recent studies show that a large number of usability issues are known to stakeholders prior to usability evaluations are conducted [1,3,5], and this suggests that feedback given to stakeholders are not adequate. In this paper we suggest an approach to explore and possibly increase adequacy and persuasiveness of feedback from usability work. In a resent paper [6] we argue, that usability as defined in ISO 9241-11 can be oriented towards (1) the user interface or user interests, and/or (2) the organization or other stakeholders. Here, we expand this approach by arguing that different usability perspectives are in play when developing software. Data originates from an ongoing interview study involving 26 stakeholders from six industrial software development projects in Denmark. Our observations are extracted using grounded theory (see [6]). The limited space in this paper makes it impossible to fully document our findings, but we aim at describing five frequently observed significant usability perspectives.

2 The five usability perspectives

2.1 The interaction object usability perspective
Interaction object usability concerns whether users are able to successfully perform isolated interactions with user interface objects in the product. We saw how consistency

was a concern using this perspective, and how standards and guidelines informed the visual design and interaction design of user interface objects. We also saw how developers were given considerable freedom regarding interaction object usability. Our data suggests that interaction object usability interplays with the applied technology (i.e. hardware, software and infrastructure), and thus that technology can inform the possibilities to produce usable software. For example we saw how a shift to wireless technology in a mobile sales support application significantly changed the usability of input fields. Online data validation was introduced reducing the amount of errors in data, but increased response time when entering data in the system. Furthermore we saw how development standards informed interaction object usability, e.g. by disallowing use of "mouse over" events on buttons, which in one case was requested by the designer.

Our data shows how interaction object usability was handled through use of a number of the traditional usability evaluation methods, such as user tests, expert evaluations, and use of guidelines or standards. Using this perspective our findings suggest that we need to take both the users and the context of use into consideration, and especially the users skills and familiarity with the technology seems important.

2.2 The task usability perspective
Task usability concerns whether the users are able to complete single tasks, i.e. fulfill a (sub) goal through a combination of interactions with user interface objects. We observed how some tasks received high level of attention when implementing new software in organizations and that the level of attention dedicated to individual tasks varied considerably. Simplicity and completeness of tasks received high attentions. Simplicity means that users only need limited knowledge related to the task to complete it, and lack of simplicity was to some degree counterbalanced through user education. Completeness means that tasks should embrace and successfully complete the corresponding work process. We observed how technology informed task usability since different technologies provide different possibilities and solutions. Furthermore we observed how techniques determining task flows informed task usability and the motivation to evaluate task usability. For example we saw how a strong process oriented development approach supported developing tasks that covered the entire process, but with an ineffectively implemented design.

Evaluating task usability required knowledge about the tasks, the users, the domain, and the context of use and was often done using different variations of user testing. Also, role-plays showed to be useful when testing tasks involving interaction between humans while using the system, e.g. in sales or interview situations.

2.3 The product usability perspective
Product usability concerns whether the product supports the users in reaching the coherent set of goals with efficiency, effectiveness and satisfaction. We observed how product usability was depending on whether the product provided flexibility, consistency, and completeness during usage.

Product usability seems to play a more important role in products with complex user interaction or products with an explorative nature compared with simple products. Task usability concerns having a straight way to reach a specified goal. In contrast product usability concerns interplay between different parts of the product allowing numerous

roads to reach important goals. We observed how the degrees of freedom of use made it hard to predict and evaluate product usability, since it required a thorough overview of the product and its usage. Knowing the specified tasks is not sufficient and a very open approach is needed. Furthermore our data shows that field observations were neither extensively nor widely used, but occasionally used in small scale.

2.4 The context of use usability perspective

Context of use usability concerns to what extent use of the system, possibly interplaying with other systems, in the actual context of use is effective, efficient, and satisfactory. Consistency across IT-systems and/or manual systems, systems integration, and inter-human relations during use of the system were important factors that influenced context of use usability. We saw how these factors had significant influence on business performance. Further we saw how users in complex work situations worked with and combined data from various systems, also informing context of use usability.

Context of use usability was rarely addressed systematically in the cases in our study. This could be explained by the fact that context of use experts only superficially were involved in the usability work. Also, systems interplay issues need to be addressed across projects rather than within projects, which increases the complexity of usability work using this perspective. Furthermore, we saw how important context of use usability issues were known from earlier versions of systems, and how workshops with context of use experts were used to address such issues in the early phases of the projects.

2.5 The enterprise usability perspective

Enterprise usability concerns to what extent goals of the enterprise are fulfilled effectively, efficiently, and satisfactorily through use of the system. This concern is not necessarily related to the users of the systems, but rather depends on whether use of the system informs or is informed by the enterprise. Enterprise usability seemed informed by three conditions: First, we saw how visions combined with IT-development projects supported stakeholders in working towards common goals rather than individual goals. Second, we saw how systems integration supported utilizing information across the enterprise and enabled support to related work processes in other departments. We also saw how failing to integrate systems could jeopardize the success of a project. Third, we saw how consistency and completeness in processes across the enterprise supported that individual completed tasks together made out a coherent environment.

In our study enterprise usability was addressed by involving key stakeholders in the initial phase, e.g. through workshops, and by maintaining their involvement throughout the process. Since enterprise usability concerns issues from all over the enterprise, it seems necessary to involve the widest range of stakeholders and filter the information afterward rather than limiting the number of involved stakeholders.

3 Discussion

In an ongoing interview study covering six software development projects, five perspectives on usability was observed. In relation to the conducted usability work in the projects, the observed perspectives had a significant practical importance. They rose

from different approaches to usability among the stakeholders and revealed both conflicting and coherent interests regarding identified usability issues. Our data suggests that usability issues related to different perspective have different properties, and studying these properties will be objects for further research. We will do this by addressing the following hypotheses:

- Persuasiveness of usability issues increases if different usability perspectives point to the same solution. Persuasiveness decreases if they point to conflicting solutions.
- Different usability perspectives appeal to different stakeholders and inform business value differently.
- Different usability perspectives are relevant at different stages of the software development process.

At the time of this writing we are looking for everyday examples from industry and related research to support our observations. One example comes from the eCommerce sector where business revenue and user experience are closely tied together bringing multiple usability perspectives into play. Another example comes from development of safety critical systems where usability problems can have severe consequences for the entire enterprise. Thus, multiple perspectives on usability could help ensuring that no stakeholders suffer from lack of usability.

In the cases we have studied, the multiple perspectives of usability were not treated systematically. Even in current research and in the state-of-the-art techniques, we rarely see such perspectives addressed and the literature fails to support practitioners in working with and understanding the perspectives. Consequently, the outcome of usability work could continue to be inadequate and non-persuasive constituting a significant risk of failure when developing industrial software.

4 Acknowledgement

References

1. Følstad, A. Work-Domaine Experts As Evaluators: Usability Inspection of Domain-Specific Work-Support Systems, *IJHCI*, *22*, 3 (2007), 217-245.
2. Hartson, H. R., Andre, T. S., & Williges, R. C. Criteria for Evaluation Usability Evaluation Methods, *IJHCI*, *13* (2001), 373-410.
3. Hornbæk, K. & Frøkjær, E. Comparing Usability Problems and Redesign Proposals As Input to Practical Systems Development, ACM Press (2005), 391-400.
4. John, B. E. & Marks, S. J. Tracking the Effectiveness of Usability Evaluation Methods, *BIT*, *16* (1997), 188-202.
5. Molich, R. & Dumas, J. S. Comparetive Usability Evaluation (CUE-4), *BIT*, (2007), 1 - 19.
6. Uldall-Espersen, T. & Frøkjær, E. Usability and Software Development: Roles of the Stakeholders, HCI International, Springer (2007)

Usability and Software Development: Roles of the stakeholders

Tobias Uldall-Espersen and Erik Frøkjær

Department of Computing, University of Copenhagen
Universitetsparken 1, DK-2100 Copenhagen
{tobiasue, erikf}@diku.dk

Abstract. Usability is a key issue when developing software, but how to integrate usability work and software development continues to be a problem, which the stakeholders must face. This study aims at developing a more coherent and realistic understanding of the problem based on 14 interviews in three case studies. The results indicate that usability during software development has to be considered with both a user interface focus and an organizational focus. Especially techniques to support the uncovering of organizational usability are lacking in both human computer interaction and software engineering. Further, the continued engagement of stakeholders, who carry the vision about the purpose of change, stands out as a critical factor for the realization of project goals.

1 Introduction

Integrating usability work into software development is not easy [3]. It requires thorough understanding about usability work methods and software development practices to reach a proper integration, but this understanding seems insufficient when aiming at improving end product usability. Despite heavy investments in information technology we observe deficiencies in practical usability work and significant lack of impact [4]. Even current research fails to explain why [7].

This paper reports from a study combining both an organizational and an individual approach to understanding and exploring the problem. By selecting this approach we seek an understanding of how organizational issues and stakeholders in the organization influence end product usability.

2 Method

We have conducted an interview study to explore how usability work and software development are handled under different organizational settings. We looked at cases in three companies where useful and usable software was developed. From each company four stakeholders covering the roles of users, software developers, technical project managers, and business decision makers were interviewed. To prevent influencing the respondents up front they were not informed about our special interest

in usability. In one of the companies usability was forced into the software development process by two stakeholders: a graphical user interface designer and a business representative responsible for requirement specification, test planning and user education. These two persons were interviewed as well.

The main research question was how practitioners in software development projects are working with usability and what we can learn from their practices? All interviews had the same interview guide as starting point, but there were significant differences in how they progressed. The interview guide covered four themes: (1) The software development process. (2) Software quality. (3) Developing usable software. (4) General experiences with development of usable and useful software products. During the interviews theme 1 and 3 were given most attention, and theme 1, 2 and 3 were all discussed based on one specific software development project significant to the interviewees and their organization. Each interview took 60-90 minutes.

The interviews were transcribed and analyzed using elements from grounded theory [5]. During the analyses we looked for information that directly or indirectly related to usability. This information was for instance statements about stakeholders' perception of usability, descriptions of usability related activities, and non-usability related issues that influenced end product usability.

2.1 Usability as a Concept
Our data suggests that usability is treated with different goals in mind in the various development projects and their organizational context. This leads us to look further into the relevance and practical conditions of conducting usability work in software development projects in order to examine the various stakeholders' roles and the possible risks regarding realization of the full potential of the solution.

The ISO 9241-11 standard defines usability as: "*The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*"

Using this definition, usability is depending on four variables, i.e. a product, specified users, specified goals, specified context of use. Following our organizational approach we observed how specified goals had significant influence on the handling of usability. This we found important since these goals existed more or less autonomously of the product, the users and the context of use; three variables which traditional usability work often have special focus on. Various stakeholders formulated goals and their direct or indirect roles informed each case significantly. We found it useful to distinguish between two groups of stakeholders, the users, i.e. persons who interact with the system, and the other stakeholders, i.e. persons who are directly or indirectly affected by the system or have important interests regarding it. Our data suggests that usability work is oriented towards two different dimensions, which is related to the various goals in the development project, among the stakeholders, and in the organization. The two dimensions found were: (1) Usability work oriented towards the user interface or user interests, which we refer to as *user interface usability*. (2) Usability work oriented towards the organization or other stakeholders, which we refer to as *organizational usability*. Incidents with both identical and different interests between the two usability orientations were observed, which support our assumption about the importance of analyzing these two dimensions.

3 Results

The cases had both strong similarities and differences. All projects were based on web technology and were all considered quite successful by the interviewees. In relation to their organizations the developed applications were innovative and both influencing and influenced by their organizations. All the systems had various user groups and groups of people that were influenced by the systems. The systems were all initiated centrally, and anchoring the systems locally in the organizations was a challenge. By nature the systems were very different. Two systems were custom-developed by external contractors and an in-house development team developed one system.

Case 1: Development of a new insurance sales tool. This case regards the development of a new sales tool for two groups of users, insurance agents and customer service persons. The tool was developed in-house over a period of 18 months. At the most 25 employees were working at the project. About 400-500 employees would be using the tool. The two user groups had significantly different requirements as the insurance agents were selling at the customers' locations, typically in their homes, and the customer service persons serviced customers over the phone. It was not considered possible by the project management team to make two different interfaces and considerable efforts were made to make one suitable interface. The sales tool was build as a front-end to two large insurance administration systems and it was a challenge to avoid letting administrative procedures inform the design. A customer centred approach was taken and all possible stakeholders were involved. The aim was to ensure the users the best possible tool and the main improvements were a better quality of data and an improved general view of the customers and their households. The company had a strongly centralized organization rooted at the head office, but employees at five regional offices generated the majority of the sale. A main challenge was to avoid that the tool became "*another head office's idea*" and a considerable effort was done to insure that the tool was firmly anchored locally. The project was innovative and utilized new technology, such as wireless access to the back-end systems and other relevant systems, e.g. the national civil registration number register. The new technology also caused severe technological and usability problems. The company did not use a formal software development method and usability was not prioritized initially in the project. Two stakeholders strongly insisted on taking usability seriously and they gradually succeeded in making usability a significant and comprehensive part of the project. The project management team took a risk by yielding control with the process and allowing anyone involved in the project to have an opinion and express it. The software developer described the space for communication this way: "*We had our arguments and we have been bloody angry at each other, close to physical fights, but it is like that to integrate systems if you ask me, and I find it great that we could ... we really could go directly to each other and say that this is really annoying. Can't we ... I think this is foolish ... but I think this is foolish ... why aren't you done now ... why shall I be done now, and so on. There we really had a very close collaboration.*" So, space was made for rewarding discussions and iterations, but the downside was that much decisions making became very time consuming.

Case 2: Developing a new IT platform for a political organization. This case regards the development of a new IT-platform for a political organization. The IT-

platform was custom-developed by an IT-contractor in close co-operation with the central office in the organization. The co-operation continued over several years where components continuously were delivered and put into use. The project team consisted of six or seven persons from the contractor and the customer's organization. The organizational leaders had strong visions about modernizing the organization and the new IT-platform was a key tool to fulfil this vision. There were strong economic incentives in the project as well. The IT-platform should serve two purposes. First it should replace an existing, but outdated communication platform used by 2.000–3.000 members. Otherwise a costly renewal of the license to the old communication platform was needed, which was not a realistic option. Introducing a new platform should help opening up the organization and make it more attractive to new members. The new platform included an advanced CMS-system available for all members (about 50.000) and specific tools for running effective and professional election campaigns. Second the IT-platform should serve as a new tool for membership administration, which would be decentralized and handed over to the local chapters of the organization. Membership administration includes issues like collection of dues, signing members up for courses and the national congress, and internal polling functionality.

The contractor applied a highly agile and strongly business process oriented approach to the development. This was a key success factor since external events periodically completely did remove the customers' focus from the project and changed the short termed goals. A very special contract was made between the contractor and the customer's organization. No formal requirement specification was agreed upon, but a vision was developed, thoroughly discussed in the management group, and written down. The customers' project manager describes it this way: "*We ended up writing up a two-page contract and some enclosures, which essentially stated that we could put the deliveries into use when we were satisfied, and when we did so we paid. The whole issue of accepting that they had delivered what we needed was handed over to us, by stating 'our experience is, that you only pay if you are satisfied, so let us put that into the contract.' Thus, it was completely up to us to decide when things were approved, but it could not be put into use before it was accepted. This model does not work in all projects, but it was extremely operational in relation to what we were going through.*" The agreement was governed by a fairness principle ensuring that the customer's organization and the contractor treated each other respectfully and this converted potential conflicts to win-win situations. According to one of the key persons, "*enlightened despotism*" dominated within the customer's organization and only three stakeholders were thoroughly involved in the project.

Case 3: Developing a coherent physical and electronic department store. This case regards the development of a new website for a department store with a number of locale houses. The website was developed by an external contractor who were specialized in user centred web development. The customer's organization did only little to involve itself in the project. The customer considered the solution to be a high-class web-solution and it was technically efficient, but it was poorly anchored in the customer's organization. The contractor's information architect experienced the lack of anchoring this way: " *... they might not all have much notion about what this website should be used for, and they also had different positions. The commercial*

manager had another position than the marketing manager, who had another position than the loyalty manager. And then ... they need to clarify it internally, and then they can come to us, because we are going to make something they can use for what they have agreed the system to be used for." The project was completed within five months and five persons from the contractor were core project members. The unique in this case was the idea of creating a coherent solution where the physical and electronic world supplemented each other in order to maintain a leading role in the physical department store market in Denmark and, if possible, also establish a position within the web-shop market. Two different goals were formulated. The first goal was to enable department store customers to buy articles in a traditional web-shop and this was given most attention by the development team. This was a limited success since only about 1 out of 1000 articles from the physical stores were available in the web-shop when it opened. It proved to be a non-trivial task to add articles to the web-shop and to ensure that the organization was able to handle the logistics. The second goal was to present information and to inspire potential customers to buy articles in the physical shops, which was the primary goal according to the business representative. A large effort was put into unifying these two goals. A combined physical and web-based fashion magazine was created and when searching for products at the web site, the search function returned information including the physical placements of the articles in the department stores.

The development process was split into three phases, each sold individually to the customer. This was an efficient way to keep the project on track, but some economic surprises did occur. Most significant was the surprise when the cost of following the strict HTML 1.0 standard was summed up. This standard was not previously followed and the budget was blown for the html-development without adding significant quality to the usability of the solution. Furthermore, the customer did neither want to pay for a thorough analysis of the target group, i.e. department store customers, nor a final user test. These cost savings watered down the user centred process.

3.1 Cross-Analyzing the Cases

Our data suggests three different approaches across the cases, which we use as starting points for analyzing and comparing the cases. Each approach seems to have or could have a significant impact on usability of the end product. The approaches are: (1) The existence or development of living visions or organizational goals in the organizations. (2) The technology used to implement the system and the technical context in which it was implemented. (3) The shaping of the software development process.

Approach 1: The existence or development of living visions or organizational goals in the organizations. All three cases were influenced by visions or organizational goals, but the effect of these was very different.

In case 1, two main goals were important. (1) Tying up customers closer to the company by selling product from more than one branch of the company. This goal was pursued by making the tool customer centred and by making it easy to refer customers to other branches. (2) Following best practises when selling insurance products. This was done by never leaving the customers with obvious needs that were not treated in the sales process. The treatment was documented in the printed policy and signed by the customer. This was done to harmonize the expectations between the

customers and the company and thereby avoiding disappointed and complaining customers when a possible insurance event happened. The redesign of the printed policies introduced a problem with clarity of the policy, since a normal policy that was handed over and signed by the customer was on about 18-20 pages. Since the old tool produced a three-page policy, this change directly influenced the sales process.

In case 2 there was a clear vision about modernizing and opening the organization to make it more attractive to new or younger members. Modernizing included revising the administrative processes in order to save money and strengthen the campaign machinery. For example, the new platform included a web-based publication module where members, from a set of templates, could create folders and posters and send them directly to the printing house without dealing with colour formats and other technical issues. One key to opening the organization was through the design of an individual entry page called 'my page'. My page should give the members easy access to discussion boards, mailing lists, and relevant homepages, but the page suffered from lack of user interface usability. It provided too much information and was difficult to use. This problem could be explained by a significant disagreement among the stakeholders about its purpose, functionality and design.

In case 3 the buyer had a set of visions that was not clearly absorbed in the project team, and some of the project members expressed doubts about the realism of fulfilling the visions. The website should inspire customers and attract them to the physical department stores, and should help building and maintaining customer loyalty. Two means supported this. First, the company developed an electronic and physical fashion magazine, which included various articles about fashion, showed various shopping articles, and linked to other text articles on the website. Second, the buyer introduced a special search concept. When customers were searching for an article or a brand, the search result displayed the various available articles of that brand and where to physically find them in the department stores.

Based on the three cases, we observe how fulfilling visions and goals in a project are strongly influenced by organizational usability. In all three cases the systems were important tools for creating loyalty or solidarity, but different approaches were chosen. In case 1 the utilization of the visions grew out of the comprehensive involvement of the various stakeholders, through workshops and formal or informal evaluations. In case 2 the design of the contract was an important factor for letting the understanding of the organizational usability develop, while the design and redesign of business processes were an important tool to its realization. The small project team with tightly cooperating members was well qualified for the job. In case 3 only one or a few key persons from the customer's organisation understood the concept that was implemented and they did not succeed in making the solution an integral part of the organisation.

Furthermore our data suggests that successful realization of visions and goals depends on thorough and coherent understandings of the users and the situation of use. Thus inadequacy of user interface usability constitutes a significant risk for not fulfilling the visions and goals.

Approach 2: The technology used to implement the system and the technical context in which it is implemented. All three cases relied on web technology and were dependent of the technical context, but the technical impact on usability was very different. One important commonality across the cases was the centralized

architecture that made it easy and relative inexpensive to fix errors and 'roll out' new corrected versions of the software. Compared to traditional software development the test efforts were reduced because of the easy access to fix problems. In case 1 and 2 less attention was directed at the deliveries when they first were put into use, and the organizations thereby failed to profit fully from the centralized architecture.

The tool in case 1 was a Java application running on a number of Citrix servers accessed through a traditional wired network or a high-speed mobile phone connection. On an early workshop the users were asked *"What can we do to make your everyday better?"* This provided important information about the possible improvements of the tool, such as how online access to the national civil registration register could help the users forming the household fast and correct while visiting the customers. The online abilities also made data validation possible through integration to the back-end systems. This drastically reduced the number of errors that required intervention from other employees after the sales were finalized. The wireless setup had a major performance problem and it took up to 17 minutes to print the policy, which preferably should be signed by the customer during the visit.

Case 2 relied on a component based service oriented architecture. This architecture made the solution extremely flexible to expand and modify and supported fast adoptions to changes in the short termed goals of the organization. For example, components of the existing infrastructure was easily integrated into the new solution, which made the solution usable from an early stage in the overall development process, and the ability to fast adoptions to changing goals proved very useful when internal and general elections were announced.

Case 3 took the most conservative approach to technology. The customer's main focus was on getting a stable solution, which they got. The contractor put a lot of effort in delivering a strict html 1.0 compliant solution. This did not have a clear influence on usability of the end product, but increased the cost of the solution significantly. Integration of the web-shop with the existing enterprise resource planner-system was a major issue, which was postponed since the customer's IT-department lacked resources to assist this work. This left the administrative and logistic processes to be carried out more or less manually and thereby exposed to human failures. This caused concerns among the stakeholders and would have been a major problem in the organization had the web-shop been a large success.

The technological comparison suggests a number of things. First, the ability to integrate with other systems can have huge effect on both user interface usability and organizational usability and failing to integrate can have severe consequences for the organization. Our data suggests that successful integration depends on continuously bringing experts together. Second, discovering and utilizing the technological abilities can be a learning process that needs space and time. Relying on well-known technology and solution patterns reduces risks of technical issues, but might also reduce innovation in the solution and in the organization, which can reduce both the user interface usability and the organizational usability. New technology can be used to evolve usability and increase the usefulness of the end product, but with a greater risk. Third, relying on specific technology and standards can introduce limitations, formal and informal. This can be a reasonable overall decision, but the consequences for usability is hard to anticipate.

Approach 3: The shaping of the software development process. In our three cases we see three very different software development processes. Case 1 relied on a human centred development process. The team aimed at putting the customer in the centre in the tool. All possible stakeholders within the company were involved and anyone at the team was entitled to have an opinion and share it. Occasionally this made the process very time consuming and demanding to handle. The result of the development process was a solid all round sales tool, where different orientations of usability were considered. Neither the user interfaces nor the processes were optimized but both were designed well. Through a number of iterations involving various users most parts of the user interface were tested before the final user tests.

Case 2 was a business process centred development process. The main focuses were on identifying important business processes, describing the processes into details, identifying stakeholders in the processes, and then implement the processes. All main design activities started with drawing up and analyzing the involved processes and the project organization saw it as their main task to "*add electric current to the business processes*". The positive outcome of the process-oriented development was a system that supported a variety of processes in the organization and was well integrated with existing and new processes and components. However, it also resulted in a non-optimized user interface with serious flaws.

Case 3 had a user centred development process as starting point. The user centred process was reduced due to economic limitations, since the customer did not want to pay for a target group analysis or a user test. This decision was inconsistent with the contractor's advice. In the development process, focus was on the front-end of the system and the back-end was only minimally adjusted to the customer's organization. The customer took only minimally part in the development project and although the contractor paid some attention to the organizational issues, the integration to the existing business did not work well and introduced a serious risk to the project.

The comparison of the three different development processes suggests two main issues regarding usability. First, a process-oriented approach favours organizational usability while a user centred approach mainly considering direct users, favours user interface usability. The human centred approach of case 1 aiming at considering all possible stakeholders, places it self in between by promoting both organizational usability and user interface usability. Second, the human centred approach required lots of resources because of the broad discussions, which was deliberately avoided in case 2 and 3. In both case 2 and 3 the project managers were clearly aware of the risk of overloading the project and refrained from involving users in specific situations, while the project manager in case 1 aimed at ensuring that 'the user involvement did not get out of hand'.

4 Discussion

We discuss possible means to improve integration of usability work and software development based on the three approaches.

Approach 1: *The existence or development of living visions or organizational goals in the organizations.* We find that the main issues regarding this approach are: (1)

How is a living vision established, evolved, and maintained throughout the development process? (2) How are visions and goals transformed into concrete and usable systems design? (3) How is usability of the systems design evaluated together with the visions? Participatory IT Design [2] and Contextual Design [1] suggest how to develop and utilize visions in systems design, but how to evolve, maintain and evaluate the vision and goals is not discussed. In our cases the visions and goals are initially anchored among the non-technical stakeholders and it becomes their task as vision carriers to maintain and propagate the visions to the entire set of stakeholders, and particularly to anchor the visions and keep them alive together with the key technical stakeholders. This is for example carried out through workshops, and workshops are also used as a place where visions and goals can inform the concrete systems design. Case 1 and 2 include a number of critical decision points, where the intervention by the vision carrying stakeholders was necessary to retained focus on the overall project goals, also in situations where fast and comprehensive reordering of priorities were urgent. Also, we do not see this issue discussed in either the usability literature or the software engineering literature. Since goals and visions seem to have great influence on organizational usability, an iterative process with evaluations and redesigns taking shape in accordance with visions might be a way to better support organizational usability and thereby to better realize the full potential of the solution.

Approach 2: *The technology used to implement the system and the technical context in which it was implemented.* We find that the main issues regarding this approach are: (1) How do we best realize the technological possibilities regarding usability? (2) How do we visualize and evaluate the consequences of the technological choices regarding usability? (3) How do we evaluate the technical implementation regarding usability before it is to late? Both Participatory IT Design [2] and Contextual Design [1] suggest that technology and the technical context are important when planning and designing new IT-systems, but the need for ongoing evaluation during development is not covered. Our cases show that key stakeholders are aware of how technology can support usability work, for example by making it easy and inexpensive to update web-based software on central servers, which should make it possible to fix a number of usability issues with a reasonable cost. Unfortunately, our data also shows that this possibility is not properly utilized, since focus shifts to other important tasks, even though an insufficient or even defective system is put into use. Furthermore, it might be more difficult than anticipated to upgrade the systems after a large number of users have taken the system into use. Also we observe how rigidly relying on standards can introduce new risks, if they are not necessary and coherent with the visions. Adhering to standards can make demand on considerable scarce resources and remove focus from more critical issues.

Approach 3: *The shaping of the software development process.* We find that the main issues regarding this approach are: (1) How is the development process organized? (2) How do the stakeholders stay engaged of the development process? (3) What tools are advantageous and profitable to apply? We have not yet seen a process taking both organizational usability and user interface usability into account in a controlled and efficient manner. This applies to both the involvement of stakeholders and the use of methods and techniques. So far methods and techniques in HCI are primarily backing user interface oriented usability. This is visible for instance in the

many evaluation techniques such as Heuristic Evaluation, Cognitive Walkthrough and Think Aloud Tests. Techniques for uncovering organizational usability issues are far fewer and less commonly used [6].

5 Conclusions

The study reports from three interview-based case studies of software development projects, where important web-based applications were implemented. We have aimed at describing different stakeholders' contributions through cross analysis of the development projects. In all three cases the stakeholders appear as individuals without an archetypical role. They all have positions, interests, and competences that make them important individual contributors. The cases show how end product usability is depending on various factors in the software development project, such as the presence of living visions, the technological choices, and the applied software development processes. Important usability contributors are found both at the user interface usability level and at the organizational level. While many techniques for developing user interface usability are employed, techniques to support the uncovering of organizational usability are lacking. Particularly important are the vision carriers, who are able to keep the project on track with clear focus on the organizational usability issues when plans have to be adjusted. Descriptions of work practises and techniques supporting this task are rare, both in human computer interaction and software engineering.

References

1. Beyer, H., Holtzblatt, K.: Contextual Design, Morgan Kaufmann (1998)
2. Bødker, K., Kensing, F., Simonsen, J.: Participatory IT Design, The MIT Press, Cambridge, Massachusetts (2004)
3. Juristo, N., Windl, H., Constantine, L.: Introducing usability, IEEE Software (2001) 20-21
4. Landauer, T. K.: The trouble with computers, Cambridge, MA: MIT Press (1995)
5. Strauss, A., Corbin, J.: Basics of Qualitative Research, SAGE Publications (1998)
6. Vredenburgh, K., Mao, J., Smith, P., Carey, T.: A survey of user-centered design practice, Proc. CHI 2002, Minneapolis, Minnesota, USA (2002)
7. Wixon, D.: Evaluating usability methods: Why the Current Literature fails the Practitioner, interactions, 10(4) (2003) 29-34

Visions in software development: Achieving value in organizations

Tobias Uldall-Espersen

University of Copenhagen

Department of Computing

Universitetsparken 1

DK-2100 Copenhagen OE

tobiasue@diku.dk

Abstract

This paper reports a case study from a software development project where visions and values among a few key stakeholders in the customer's organization played an important role. The contractor relied on a strongly process-oriented development process, was innovative, and contributed to the development and materialisation of the visions. The result was a new IT infrastructure adding significant value to the organization. However, focusing on processes and key stakeholders' visions and values was not sufficient to achieve the full potential value of the project in the organization.

Introduction

Taking organizational goals and visions into account when developing and deploying new IT systems might be a useful approach to support achieving value in organizations. This paper reports a case study describing the implementation of a new IT infrastructure in a major national political party in Denmark. Different circumstances in and around the party clarified a need for significant changes in the organization and a newly appointed party secretary started this process. The overall vision was to open up the organization and make it more efficient, and at the same time provide the necessary service to stakeholders within and outside the party. The vision grew slowly in a central party office committee and at a certain point of time circumstances made it convenient to act. An external project leader was hired and together with a few key stakeholders, a vision for a new IT infrastructure was developed, important goals specified, and critical success factors defined. During the initial phase of the project a number of external

contractors were consulted and one contractor selected. The selection criteria included traditional matters such as price, cost of ownership, rights to the source code, fulfilment of requirements, technology, the capabilities of the contractor etc, but it also included matters such as chemistry among stakeholders and the ability to work with and further develop the visions. This paper will focus on how the visions evolved and how the focus on visions and organizational values informed the end product and the software development process. A more thorough description of the development project can be found in [3].

Method

The case study is based on two PowerPoint presentations given around the time where the contract was signed and semi-structured interviews [1] with four key stakeholders: (1) the technical project leader, (2) the organizational project leader, (3) a software developer, (4) a user. Each interview was scheduled to last 90 minutes and the same interview guide was used in all four interviews. The interviews were recorded, transcribed, and analyzed using open coding [2]. The interview guide covered four themes: (1) the software development process, (2) software quality, (3) developing usable software, and (4) general experiences with development of usable and useful software products. At the time of the interviews the organizational project leader had left the customer's organization. He was now senior partner in the contractor's organization and there was still a close co-operation between the two organizations.

The case description

This section describes how visions and values were used to define critical success factors, which were directly usable in the software development project. First is described how visions and values among the key stakeholders in the development project were used to formulate a vision for the new IT infrastructure. From this vision desired contributions to the party organization were identified and based on these a number of critical success factors were formulated. Later is described how some of these critical success factors appeared in the software development project and informed the design and the quality of the IT infrastructure.

Materializing the visions

ORGANIZATIONAL VISIONS AMONG KEY STAKEHOLDERS
The overall visions and goals among the key stakeholders in the party was to (1) open up the party, i.e. encourage frankness and transparency when discussing important issues, (2) to make the party more attractive to new and younger members, and (3) to reduce administrative costs. Introducing a new IT infrastructure was seen as a key to achieve this.

VISION FOR THE IT INFRASTRUCTURE
A vision for the new IT infrastructure was formulated in co-operation with the contractor and a contract was made on delivering that vision. The written vision itself was important, but also the process of formulating the vision, which led to a common realization of the core of the vision among the key stakeholders, was an important step in the process. The vision for the IT infrastructure was: '*To provide the best possible digital processes, functions and information to all relevant stakeholders inside and outside the party in the relevant context, at the relevant time, and in the relevant form.*'

DESIRED VALUES
A number of desired values from the new IT infrastructure were identified. According to the key stakeholders the most important values were:

- Increasing the number of party members and the number of voters by strengthening the digital presence and the digital election campaigns.

- Increasing the interest for and use of digital offers among stakeholders outside the party organization.

- Increasing the loyalty among members and voters.

- Strengthening the involvement and the activity in the party through attractive digital offers to members and party organization.

- Increasing the quality and the effect of all digital processes and communications.

- Increasing the income in the party and ensure a scalable effective and efficient administration without increasing costs by increasing the amount of digital processes and publications, and decreasing the amount of physical processes and publications.

IDENTIFICATION OF CRITICAL SUCCESS FACTORS

Based on the desired values from the new IT infrastructure, the following critical success factors were formulated: (1) The individual member must experience (a) clear advantages by using the new digital offers, (b) a closer affiliation to the party through increased insight into own and local conditions, (c) an increased opportunity to participate actively in the political debate, (d) an increased opportunity to inform decisions and to get an overview of the offers from the party. (2) The organization in general must experience the new digital processes as substantial improvements compared to before. There must not be reductions regarding users and processes in the new system compared with the old system. Local chapters must experience a much higher level of service, a better administrative insight, as well as control over their own processes. (3) All critical back office processes must be supported by the new system, and the new system and all processes must be ready and implemented by the end of the year. A demo giving an impression of the new system must be presented at the forthcoming party convention. (4) There must be possibilities for own development and adjustment of all solutions. The solution must be scalable with respect to the number of users and increased frequency of use without increased costs or other investments.

The new IT infrastructure

The new IT Infrastructure offered a significant amount of functionality to the users and maybe even too much. A content management system module was included and all members were offered space for their own homepage and offered functionality for easily creating newsletters and discussion groups etc. Predefined templates were developed and made available to the users in an attempt to make the various homepages more uniform without forcing anyone. Another function added was a web-based publication module. This made it easy to make posters, folders and other campaign material for election campaigns. There were a number of predefined templates available and the system produced a PDF-file ready for delivery at the local printing-house. This was a great help for local candidates.

One of the other central parts of the new IT infrastructure was a webpage called Min Side (English: My page). Min Side was the first page the user met when doing a personal login at the website. The aim of Min Side was to open up the party organization by providing the users with various options and by displaying different information relevant to the user, such as personal information, information from local chapters, and information about debates, meetings, polls etc. Even though the key stakeholders agreed on the goals of Min Side, there were disagreements about the design and the implementation of the page and in the end, Min Side failed its overall purpose. The page was confusing and included an enormous amount of information. Users found it hard to understand some of the terms used, e.g. forum (i.e. a private part of a website typically used for discussions, where people need to sign up to get access), and users had difficulties finding forums they previously had signed up for. When requesting help from the central office users were given workarounds and explained how to complete their task without using Min Side.

The number of employees at the central office had been reduced and there was a need for reducing the work conducted centrally, which led to significant changes in the administrative processes. Member administration was decentralized and handed over to local representatives (i.e. chairmen and treasurers). Work processes were thoroughly analyzed and new processes were designed aiming at simplicity and completeness, i.e. it should be simple to complete a task rather than highly efficient, and users should know that the corresponding work process was completed when a system task was completed. One example of a complete work process was moving members between local chapters. The local representative identified the member in the system and assigned the member to the new chapter. Relevant stakeholders were automatically informed and representatives in the new chapter were asked to confirm the transfer. At the same time the system ensured that all other required actions were carried out. The process of collecting dues was also changed, but this was centralized. Knowing your members and making sure all members paid were political issues with a high level of attention, both within and outside the party. There had been serious legal issues in other organizations because of lack of control with public subsidies given to organizations based on the number of paying members. To avoid this problem the collection of dues was now centralized. Some local representatives noticed that this would inform the social structure in their local organization, since they were used to collect the dues personally and thereby maintained a personal relation with their local members. It was decided to continue to allow this kind of collection of dues and the system was design to support this.

Conclusion

The current case shows that key stakeholders' visions can be applied in software development projects and that doing so can help organizations achieving value when implementing new software systems. By systematically working with visions and materializing these into desired values and concrete success factors, key stakeholders in the two organizations (customer's and contractor's) succeeded in producing an IT infrastructure valuable to the customer from an economic, social, and political perspective. There is a clear main thread from the visions to the end product. The case also suggests that a strong focus on visions, value and processes is not necessarily sufficient to fully realize the potential of an investment. Parts of the investment, e.g. Min Side, did not yield the expected organizational value because of lack of understanding of the users and their ability to use the system.

Furthermore, a number of more general key questions for further research are suggested by the case, e.g. (1) how are realistic and valuable visions regarding IT-usage developed and anchored in different types of organizations and (2) to what degree is a project's success grounded in systematic and methodical work, and to what degree is it grounded in prominent personalities insights and abilities and how can we create the optimal conditions for engaging such personalities in software development projects.

References
[1]   Kvale, S. *InterView*, Hans Reitzels Forlag, 1994.

[2]   Strauss, A. & Corbin, J. *Basics of Qualitative Research*, Sage Publications Inc, 1998.

[3]   Uldall-Espersen, T. & Frøkjær, E. Usability and Software Development: Roles of the Stakeholders, *Proc. HCI International 2007, Beijing, China, 22.-27. July 2007*, Springer-Verlag Berlin Heidelberg (2007), 642-651.

Facilitating usability work using different perspectives

Tobias Uldall-Espersen
Department of Computing,
University of Copenhagen
Universitetsparken 1, DK-2100 Copenhagen
tobiasue@diku.dk

ABSTRACT

Doing usability work is usually considered a multi-disciplinary task, but little is done to advance the multi-disciplinary side when obtaining and transforming usability insights into practical values in software development. This paper argues for a broader conceptualization of usability and suggests a number of usability perspectives. The aim is to support multi-disciplinary considerations among key stakeholders involved in software development projects where practical usability work is conducted. The usability perspectives evolved from an interview study involving 26 respondents in six Danish software development projects. Two examples of multi perspective usability issues are given and it is described how usability issues relating to one or more perspectives under the right circumstances can crystallize in software development projects and have significant impact on end product usability. Finally, it is discussed how to apply the usability perspectives proactively when developing useful and usable software.

Author Keywords

Usability, Software Engineering

ACM Classification Keywords

H5 Information interfaces and presentation; H5.3 Group and Organization Interfaces; K6 Management Of Computing And Information Systems

INTRODUCTION

The aim of this paper is to develop an understanding of usability issues by use of different usability perspectives and to discuss possible, practical implications of such a view. The usability perspectives came into existence when reporting a study [11,12] based on interviews involving 26 respondents from six Danish software development projects. In [12] we argued that usability work in the case studies were oriented towards two different dimensions related to the various goals in the development project, among the stakeholders, and in the organization. The two dimensions found were: (1) usability work oriented towards the user interface or user interests and (2) usability work oriented towards the organization or other stakeholders. In [11] these findings were further analyzed and five usability perspectives were suggested. Based on the five perspectives, this paper analyzes one of the software development projects aiming at enlighten how multi

perspective considerations about usability issues increased the practical value of doing usability work.

The long-term goal of this research is to examine if different usability perspectives can be used to support and facilitate usability work, for example by evaluating software systems using different usability perspectives or by applying different usability perspectives when providing feedback from usability work. It is believed, that relying on different usability perspectives to understand usability issues could help grounding usability work in actual human values and such grounding could be a great advantage [1]. Such grounding could help building relationships with trust and confidence among key stakeholders participating in different way, which seem crucial to establish the proper condition for collaboration and therefore in the long term are important in order to impact software development projects [8,9].

METHOD

Data in this study originates from six case studies of independent Danish software development projects and involves 26 semi structured interviews [6] each of 60-90 minutes. Three cases were considered primary cases and the corresponding 14 interviews were transcribed and analyzed using open coding [10]. The remaining three cases were considered secondary cases and the corresponding 12 interviews were analyzed with the aim of supporting the findings from the primary cases.

The same interview guide was used at all interviews. It had four main themes: (1) The software development process. (2) Software quality. (3) Developing usable software. (4) General experiences with development of usable and useful software products. The interviews used a concrete software development project in the specific organization as a starting point and the respondents were not informed about the interviewer's special interest in usability.

THE FIVE USABILITY PERSPECTIVES

The main observation was that the respondents discussed usability using different perspectives and five perspectives occurred repeatedly across the six software development projects. This section gives a brief introduction to the five perspectives. Further details are given in [11].

1. **The interaction object usability perspective**

Interaction object usability concerns whether users are able to successfully perform isolated interactions with user interface objects in the product.

2. **The task usability perspective**

   Task usability concerns whether the users are able to complete single tasks, i.e. fulfil a (sub) goal through a combination of interactions with user interface objects.

3. **The product usability perspective**

   Product usability concerns whether the product supports the users in reaching a coherent set of goals with effectiveness, efficiency, and satisfaction.

4. **The context of use usability perspective**

   Context of use usability concerns to what extent use of the system, possibly interplaying with other systems, in the actual context of use is effective, efficient, and satisfactory.

5. **The enterprise usability perspective**

   Enterprise usability concerns to what extent goals of the enterprise are fulfilled effectively, efficiently, and satisfactorily through use of the system.

The five perspectives are grounded in data from 26 interviews. It is not believed that the perspectives at this level of maturity constitute a complete and final framework of usability perspectives, i.e. that the definitions of the usability perspectives are fully adequate and no other usability perspectives could be found or added, but the five perspectives played a distinct role in the six studied software development projects.

It is furthermore worth noticing that usability issues can be related to one ore more usability perspectives. When this happens suggested actions can be conflicting or complementing each other. When suggested actions are conflicting there is a risk of sub optimization, i.e. usability using one perspective is optimized at the cost of usability in other perspectives. Thereby the expected impact from addressing the usability issues might not be obtained and the expected value to the system not added. On the other hand, if one or more usability issues related to different usability perspectives are complementing each other, they might add a broader value to the system and in the end help producing a more solid, useful, and usable system.

EXAMPLES

Two examples from one of the primary case studies are given. The case study examines the development of a new internal insurances sales application called *Absalon*. Absalon replaced an old system called *MobilSalg*, which the users were very fond of, but it did not meet modern technical requirements and was replaced.

First example regards the use of a piggy bank icon. This example shows how a simple design issue evolved to a multi perspective usability issue, where different usability

perspectives suggested different actions. It could have added a serious usability problem to the end product, but the users involved in the development project realized the risk before it was too late and the problem was avoided.

Second example regards various data quality issues. This example shows how interplay between different usability issues related to different usability perspectives strengthened the development project and made the system more solid and usable. By continuously involving a multiplicity of key stakeholders, the usability issues crystallized in the project across different usability perspectives and the data quality issues became a core concern in the project.

Example 1: The piggy bank icon
MobilSalg used a lot graphical user interface (GUI) elements and a piggy bank icon was used to symbolize an 'add discount' option in the insurance information sheet. The top of Figure 1 illustrates how this could have looked like.
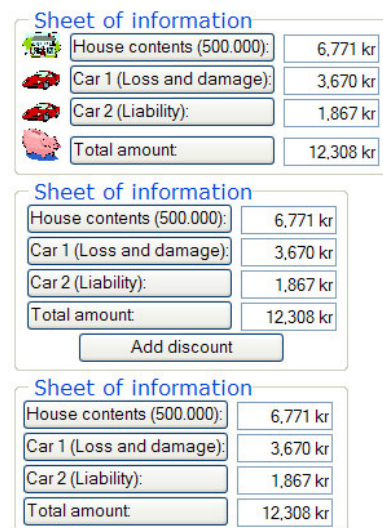

Figure 1 – The figure illustrates three different 'sheet of information' models. In this example of a simple design issue, the use of a piggy bank icon and the 'add discount' option were related to three different usability perspectives.

The initial design suggestion followed an enterprise usability perspective, when the development team decided to remove the piggy bank icon and a number of other GUI elements. They wanted the system to appear professional and found a number of GUI elements childish and inappropriate. The interviewed user did not share this concern. He found the GUI in MobilSalg nice and useful when advising customers and considered the GUI in Absalon usable but improvable.

It was a concern in the development team that the system should be clear, intuitive, and easy to use and as a replacement for the piggy bank icon, the designer added an 'Add discount' button (Middle of Figure 1). Now this

option was clear and easily recognizable for the user and from an interaction object usability perspective, this was a good solution. However, the solution was dismissed by the insurance agents and in the end the option was hidden in the insurance information sheet GUI (Bottom of Figure 1). Now the users had to click on the 'total amount' button to add a discount and there were still no visual clues. The reason for the dismissal was that a strong visual clue would disturb the sales process when sitting together with the customers, as they would ask the insurance agents for a discount. Seen through the context of use usability perspective this issue would have had a significant negative impact on usability.

Example 2: The data quality example
This example shows how a number of usability issues were related across different perspectives and involved different stakeholders. In this example the usability issues seemed to crystallize in the project and added a clear practical value to the end product. The following list describes a number of usability issues related to the data quality work. The list might indicate a possible ordering in time, which is unintended. It seems more likely that the understanding of the usability issues in the example arose and grew more randomly in the project over time and across key stakeholders.

- Following an enterprise usability perspective, data quality in the sales system needed to be improved. In MobilSalg about 8.000 errors were every year investigated and fixed by back office employees adding a significant cost to the sales process. This problem rarely informed the users since the errors occurred after the sales agreement was signed.

- Following a product usability perspective and a context of use usability perspective data quality also needed to be improved. Copies of customer data were held in locale databases and they were not always up to date. This could give faulty or insufficient pictures of the customers and could lead to embarrassing situations and missed sales opportunities.

- To improve product usability online wireless access to backend systems was provided. This removed the locale database problem, but introduced other problems regarding accessibility and performance. The online abilities furthermore enabled closer integration to other systems and other sources of data, and it provided the insurance agents with field access to other tools, such as email and calendar.

- As users now were writing data directly to the main insurance systems, data was validated instantaneously. This informed interaction object usability. Now the users needed to understand and accept rules for data validation of single input fields. The rules were not new, but they had not previously been enforced.

- This informed task usability since task completion time was increased and old habits became obsolete and needed to be replaced. Certain use practices that had developed over time were no longer allowed.

- One way of supporting the habit renewal was to reduce the flexibility of the system. This option was dismissed because it would reduce the context of use usability. The sales situations were very different and it was considered a priority to maintain flexibility with respect to different contexts of use.

- To increase task usability online access to the national civil registration number register was provided. This made it possible to enter a civil registration number and then retrieve the corresponding name and address information. This decreased the time spend on entering data during a sales.

- Since the users now spend less time on entering basic data, focus in the sales situation moved from the system to the customer and the context of use usability was improved.

- Enterprise usability was also improved as basic data was fetched automatically. The availability and consistency of data was increased and it became easier to establish household relations between customers. Combining customer data increased the possibilities for cross sale and the household relations were important for other systems in the organization, such as the data warehouse system and the customer relation management system.

DISCUSSION
The two examples shows how considering usability issues using different perspectives was natural in the software development process. Encouraging such considerations facilitated different stakeholders in contributing important insights and building up a shared understanding of possibilities and challenges in the project. This strengthened the collaboration between usability practitioners, developers, and other stakeholders and increased the practical value of the conducted usability work [4,9]. Furthermore, continuously involving domain/business experts and context of use experts, as well as technical experts and HCI experts, supported focusing on meaningful problems [3] and on maintaining relations to business goals, which previously has been reported lacking in professional usability work [2,7].

Having observed the five usability perspectives in a number of real-life software development projects raises the following questions. (1) Can the five usability perspectives be used to systematically facilitate usability work involving a broader set of stakeholders and (2) would it yield an advantage compared to other usability work approaches?

It is believed that the usability perspectives could be used to facilitate usability work, but this has not yet been tested and

it is not clear how to transfer the usability perspectives to applicable usability work techniques. The five usability perspectives might be naturally combined with many existing UEMs. Most UEMs address mainly the interaction object usability perspective, the task usability perspective and a few UEMs address the context of use usability perspective. The unaddressed perspectives, i.e. the product usability perspective and the enterprise usability perspective, could be seen as supporting perspectives, for example by use of heuristics or guiding questions, which depending on the situation could be more or less meaningful. From our observations, it seems that the usability perspectives most significantly stands out and adds practical value in cases with long-term involvement of various key stakeholders, easy access to key stakeholders, and efficient formal as well as informal routes of communication. This seems hard to reconcile with traditional usability evaluation methods and reported industrial practices [5]. There might be other ways to obtain and utilize insights from various key stakeholders not directly involved in the software development project [13], but whether a synergy effect as in the examples presented here can be obtained, without bringing people together, remains to be shown. Another possibility could be to apply the usability perspectives when working with feedback from usability evaluations in order to evolve a coherent set of redesigns appealing to a broader set of stakeholders. Workshops with various stakeholders were conducted in a number of the studied cases. Applying multiple usability perspectives to analyze feedback from usability evaluations in such multi-disciplinary workshops could be an option. However, if the usability evaluation feedback does not involve multiple usability perspectives, it might be difficult to obtain a valuable multi perspective outcome from the workshop.

A number of advantages could be expected when applying the perspectives and two possible advantages are argued here. First, the perspectives might facilitate the building of realistic expectations among various stakeholders. This is a serious problem, for example reported by Rajanen and Iivari [8] in a recent study where a significant usability investment became a total failure. The study by Rajanen and Iivari is a valuable case study because it shows how usability issues related to specific perspectives and specific stakeholders completely are ignored by other stakeholders using other perspectives. This problem was only realized when it was too late and usability then became a 'curse word' in the organization. Second, the perspectives can be utilized without establishing formal usability work procedures. Stakeholders in a number of the conducted case studies did not work systematically with usability, but still the usability perspectives were observed among the project team members. This suggests an opening for introducing

more systematically work with usability in organizations not yet mature for larger usability investments.

REFERENCES
1. Carter, P. Liberating Usability Testing, *interactions*, *march+april* (2007), 18-22.

2. Dumas, J. S., Molich, R., & Jeffries, R. Describing Usability Problems: Are We Sending the Right Message?, *interactions*, *11*, 4 (2004), 24-29.

3. Følstad, A. Work-Domaine Experts As Evaluators: Usability Inspection of Domain-Specific Work-Support Systems, *IJHCI*, *22*, 3 (2007), 217-245.

4. Hornbæk, K. & Stage, J. Special Issue on The Interplay Between Usability Evaluations and User Interaction Design, *International Journal of Human-Computer Interaction*, *21*, 5 (2006), 117-123.

5. Jerome B. & Kazman, R., Surveying the Solitudes: an Investigation into the Relationships Between Human Computer Interaction and Software Engineering in Practice, in Seffeh, A. *Human-Centered Software Engineering – Integrating Usability in the Development Process,* Springer, 2005, 59-70.

6. Kvale, S. *InterView*, Hans Reitzels Forlag, 1994.

7. Molich, R., Jeffries, R., & Dumas, J. S. Making Usability Recommendations Useful and Usable, *Journal of Usability Studies*, *2*, 4 (2007), 162-179.

8. Rajanen, M. & Iivari, N. Usability Cost-Benefit Analysis: How Usability Became a Curse Word?, *Proc. INTERACT 2007, 11th IFIP TC 13 International Conference*, Springer-Verlag Berlin Heidelberg (2007), 511-524.

9. Redish, J., Bias, R. G., Bailey, R., Molich, R., Dumas, J., & Spool, J. M. Usability in Practice: Formative Usability Evaluations - Evolution and Revolution, *Proc. CHI '02 extended abstracts on Human factors in computing systems*, ACM Press (2002), 885-890.

10. Strauss, A. & Corbin, J. *Basics of Qualitative Research*, Sage Publications Inc, 1998.

11. Uldall-Espersen, T. Exploring Multiple Usability Perspectives, *Proc. INTERACT 2007, 11th IFIP TC 13 International Conference*, Springer-Verlag Berlin Heidelberg (2007), 628-632.

12. Uldall-Espersen, T. & Frøkjær, E. Usability and Software Development: Roles of the Stakeholders, *Proc. HCI International 2007, Beijing, China, 22.-27. juli 2007*, Springer-Verlag Berlin Heidelberg (2007), 642-651.

13. Uldall-Espersen, T., Frøkjær, E., & Hornbæk, K. Tracing Impact in a Usability Improvement Process, *Interacting with Computers*, (2007).

How usability work informed development of an insurance sales system

Tobias Uldall-Espersen
Department of Computing,
University of Copenhagen
Universitetsparken 1, DK-2100 Copenhagen
tobiasue@diku.dk

ABSTRACT

This paper reports a case study of a software development project where an insurance sales system was developed. Two key persons in the project enforced usability work into the development process and usability work became a key success factor. The usability work was comprehensive and became a significant and integrated part of the development project, and it informed both the end product quality and the organization in which the system was implemented. The case study is based on interviews with six key persons in the project.

SETTING THE STAGE

Introduction

This case study describes the development of a new insurance sales system in a Danish insurance company. Usability work was not a part of the formal project development plan or the company strategy, but was forced into the project by two key project members. The work was driven by their professional pride and it proved to be a critical success factor in the project. The development project was influenced by its wide organizational settings and the introduction of new innovative technologies, such as wireless communication, in the organization. The new technology changed current practices both within the development team and among the users. A customer-centered approach was applied to form the new sales processes and the users were heavily involved in the iterative development process.

The description is based on semi-structured interviews [1] with six key persons: (1) the technical project leader, (2) the business project leader, (3) a software developer, (4) a business analyst, (5) the GUI designer and information architect, (6) a user. The business analyst and the GUI designer were responsible for the conducted usability work and special focus is on their experiences in the case. The interviews were conducted within two months after the project ended. Each interview was scheduled to 60-90 minutes, the same interview guide was used too all six interviews. A total of about 8 hours conversation were recorded and analyzed using elements from grounded theory [2].

The system and the users

The system under development was an online sales tool to be used by insurance agents and customer service persons. Insurance agents meet face-to-face with customers in the field and customer service persons provide service to the customers over the phone. About 400-500 people would become users of the system. It replaced a traditional client-based system that was appreciated by the insurance agents, but which did not meet modern technical requirement, i.e. it did not provide online access to customer data and it was hard and expensive to maintain. The aim of the development project was to develop an up-to-date web-based system with the same functionality as the old one.

The development organization

The development project was anchored in the Marketing and Direct Sales (MDS) department, but had threads widely across the organization. MDS was holder of the sales processes and was responsible for the business side of the development project. MDS was tied up to a department in the IT-organization (IT-MDS) that did the technical development. Two other business departments were heavily involved as well: the private general insurance department and the retirement and live insurance department. They were holders of the insurance products and had the product domain knowledge, e.g. how policies were handled throughout and after the sale, how different products could be combined, and legal issues and best practices when advising private customers. Both product departments had comprehensive core IT-systems used for product administration. These served as back-end systems for the sales system and were managed by two other IT-departments. The users of the sales system were organized in five geographical regions independent of the business departments. As the insurance agents were paid on a commission basis and specific requirement to their sales tools where determined by their conditions of appointment, both the local insurance agent association and the internal legal department were involved in the project as well. A number of other IT-departments also took part in the project, such as the departments responsible for the IT-infrastructure, i.e. network, laptops, servers, databases, etc.

The organization did not rely on a standard software development process. The project leaders laid down the process and no formal strategies or procedures were imposed from outside the project. The organization was not experienced with usability work, but a scattered knowledge existed and the consequences of not dealing with usability were well known. A project dealing with usability evaluations and software development was previously conducted and two employees were given formal training. One of these (the GUI-designer) where allocated to the project. Furthermore, a large IT-project implemented in the company three or four years earlier had suffered severely from lack of usability, which seriously influenced the organization.

The development process

The development project was scheduled to about 18 months and at the most 25 people worked on the project. Half way through the project the schedule was slipping. Some participants had difficulty in following the project plan and since other depended on their deliveries the delays spread like ripples in a pond. The participants then focused on the failed deliveries and not on delivering their own work on time, and the project was threatened by disintegration. At this point major changes were introduced in the project. The project leader was replaced with two new project leaders, one responsible for the business side and one responsible for the technical side. All future activities were then broken down into smaller pieces and re-estimated, and 6 weeks time-boxes were introduced in the development process. These changes put the project back on track and the system was delivered with only a minor delay of less than two months.

DESCRIPTION OF THE CASE STUDY

No formal plan for usability work existed, but the need for dealing with usability was recognized by some project members and usability work became a significant part of the project. Two key persons took responsibility for the usability work, one with roots in the business organization (the business analyst, referred to as BA) and one with roots in the technical organization (the GUI designer, referred to as GD). Since usability work was not a formal part of the project, no measurable targets were defined and no formal validations made. Thoroughly understanding and respecting the users, and keeping things simple and reasonable became central concerns in the project.

GD was the main designer of the user interface and the information architect. He had been in the company for some years and had experiences from prior engagements, and he was taking university master classes in usability work. GD's responsibilities were requirement analysis, system and interaction design through development and test of prototypes, and being sparring partner to the software developers throughout the development process. He described his own role as '... being liaise between the business departments and the development team' and other project members referred to him as '... the designer and usability guy'. In the project start he was involved in some workshops with representatives from the user groups and the project team. The aim was to have an open discussion about the shortcomings of the old system and requests for the new system. Identifying what to discard from the old system and what to keep was given special attention and it was discussed what was needed to make the everyday of the users good. Comments were registered and classified as nice to have or need to have. Then GD developed a paper and post-it based prototype to define and test the information architecture. Developing the prototype was an iterative process involving both developers and business. Hand drawn prototypes and later prototypes done in Photoshop and PowerPoint were made in the process. The prototypes were living documents that were continuously assessed by users around the different geographical regions. At some time the Photoshop drawings were given to the developers who started adding business logic to the design. At this point the development process was informed by negotiations and informal user tests, where GD continuously served as sparring partner for the developers and worked on validating various design decisions. 'I try to argue for my view points and they [the developers] argue for their. We might then meet in between ... and then we do a test and ask the users what they want.' This was comprehensive work, but the strategy ensured that most parts of the system were tested with users before the final user test, and continuous involvement of GD ensured consistency across the system.

BA was employed at the general insurance side of the project. She had been in the company for two years. Prior to her engagement she worked with usability in a bank where she was consultant and user test leader. Her responsibilities in the development process were requirement analysis and specification, testing, and user education, and she worked closely together with the users. When doing the requirement analysis and specification, she did contextual inquiries with the insurance agents and customer service persons. This revealed significant differences between the two user groups and resulted in a thorough understanding of their work conditions. After the requirement analysis, BA was made responsible for planning the various tests of the system including a test of a rough paper prototype and two rounds of user tests. The paper prototype was discussed in a focus group with future users. According to BA the user tests were 'not by the book', but conducted '... from the means we had and the resources ... and what we could expect to get out of it and what we would be able to fix in relation to the tight schedule.' The user tests were a mixture of think aloud, focus groups, individual task solving and role-playing. Three test leaders and twelve users spent twice three to four days testing the system. All users were put into a room normally used for training new employees and asked to solve a set of tasks and write down any comment they had or problem they experienced. The test leaders circulated and observed the users, and important

issues were discussed in order to get a thorough understanding of use patterns and potential problems. Problems were finally discussed in plenum in order to identify the root of the problems and the possible needs and desires for changes. Some tasks proved difficult to test realistically as the insurance agents had to read information from the task description instead of talking with the customer. To overcome this challenge the users were set up in a role-play with one user acting as customer and one user acting as sales person. This worked very well. At the end of the project BA was made responsible for educating the users. In her words, 'It was very obvious because you also saw ... there were some things that were difficult for the user, but not necessarily could be changed because it was decided to do it that way – either for business reasons or due to IT limitations, but then we knew what to focus on when educating the users.' A number of instructors (typically users from different regions, who were involved in specification and test) had a week of training in the system and professional communication and were supplied with education material. Then they returned to their local offices and gave a two-day training course to the other users.

The Main Challenges
Key persons in the project identified a number of challenges early. Among the challenges taken most seriously were: (1) Avoiding the system to be a head quarter idea without solid anchoring in the regional organizations' wishes and needs. (2) Taking both user groups into account and building one coherent system. (3) Creating an online sale front-end to an administrative back-end system.

First, users were thoroughly involved and BA noted: 'There are many more things in it than pure usability, i.e. that the system has become better to use.' This referred to the involvement of the users as a political and pedagogical mean in addition to creating a rich understanding of their everyday. As decisions often were made at the head quarter, a gap existed between the head quarter and the employees in the regions. To reduce this gap and to establish a shared responsibility, regional users were involved early in and throughout the process, and the involved users became system ambassadors. This approach worked well and the project team received recognition for that.

Second, about the two groups of users GD notes: 'We made a system that must consider both groups. It was quite a big challenge as one group was ... direct selling and sitting with their laptop in front of them and having an open dialog with the customer. And then you have the customer service persons sitting in a complete different work situation, where they can report very fast on the computer and get a result.... They cannot do the same considerations as a person in the house looking at the two cars in the carport. This is two very different situations.' Fitting the new system to two user groups was an obvious challenge and a number of

differences were identified and addressed in the project. Insurance agents used the old sales system while customer service persons used the administrative back-end system. The old system was mainly mouse controlled and the back-end systems mainly keyboard controlled, and it was considered critical to support both control methods effectively in the new system. When visiting customers, the insurance agents would typically be in control and be able to manage the meeting. They had a fixed sales sequence they explained to the customer. This sequence ensured that they collected all needed information and could deliver a final offer to the customer, but it gave very little room for individuality. Furthermore, the insurance agents typically allowed their customers to look over the shoulder, so information should be structured and displayed with great care and some information should not be directly visible, for example intermediate prices, special discount possibilities, and detailed product conditions which would be distracting in the sales process. On the other hand the customer service people had very limited control of the situation when customers called them with a variety of questions or problems. They were used to enter a minimum of information and then calculate an estimated price. Only when the customer accepted this, then the rest of the information would be entered. The customer service people appreciated flexibility in the system and had a large need for additional information. The solution was to structure information carefully and make it easy accessible in a very flexible system.

Third, another issue was to avoid reflecting the administrative processes from the core insurance systems into the new system. First, two core insurance systems existed, one for administration of general insurance products and one for administration of retirement and life insurance products. The first system was a modern, windows based system and the other was an old mainframe based system, so the core systems were quite different. Second, both systems were administrative systems, whereas it was a sales system that was under development. The processes in the administrative systems were considered to be product centered whereas the sales system should be customer centered. By focusing on making the system customer centered the project team succeeded in creating a suitable sales support system, and adding available cross-organizational banking data, made the sales persons feel more professional and well prepared in the sale process.

The online abilities influenced issues as data quality, validation of data, and online access to historical data and supplementary data sources. This had both organizational and more practical influence. In the old system the insurance agents had a local database containing data from a small subset of the company's customers. Before and after every visit the insurance agent would synchronize data. At the start of the day the agent retrieved the current information about the customers to be visited and at the end of the day changes in this information and new information

were fed back into the core systems. This process had a number of weaknesses. (1) It was not very flexible, i.e. the insurance agents needed to know whom they were going to visit while they were at home, otherwise they could retrieve customer data. This made it very difficult to revise plans during the day. (2) If a new customer were visited, there was a significant work registering the household correct and first then the sales process could start. Sales persons typically want to pay as much as possible attention to the customers and often they entered defective data, which had to be fixed afterwards. About 8.000 incidents a year had to be dealt with manually because of such short cuts. (3) When visiting existing customers there was a risk of having old and incomplete data. If a change in the household or in the product portfolio was not registered correctly, the sales person might look foolish or in worst case lose the customer or miss an extra sale. The majority of these problems were resolved by giving on online access to the back-end systems and by drawing relevant data from other systems, e.g. the national civil registration number database. By utilizing these possibilities, information about a complete household could be inserted correctly or updated within a minute and new data could be validated immediately in the back-end system.

DIFFERENT PEOPLE'S PERSPECTIVES ON THE USABILITY WORK IN THE CASE

No formal evaluation of the usability work and the end product's usability was conducted. As the product only had been used for a short time, the users were still getting acquainted with the new system. Furthermore, a number of technical problems were experienced, which made it difficult to distinguish between problems in the system and problems in the system environment.

The six interviewed persons did all share the understanding, that the usability work was very successful, and one person mentioned how users were applauding after a training session. The key persons agreed that the thorough user involvement and the wide spread discussion and communication was very demanding, but they also agreed that it was worthwhile doing and was a critical success factor in the project.

LESSONS LEARNED

The case was influenced by the lack of traditions for and formal requirements to usability work in the organization, but two key persons managed successfully to put it on the agenda by stubbornly insisting on the necessity of taking usability seriously. It is believed that relying on individuals with a special interest or skills in usability work is a common way to introduce usability work in organizations and introducing usability from within a project might yield some advantages. In this case it seems to have added confidence to the process, since the control of the process remained at the project leader level and the project group was able to follow both the challenges and the progress. The technical project leader supports this interpretation by

stating, that his role in relation to the usability work was to ensure that it '... did not get out of control'. Also, by relying on experienced key persons from within the project, the acceptance and understanding of usability work were able to grow, and the usefulness of the work was widely accepted among the stakeholders.

It seems clear in this case that the close integration of software development and usability work, rooted in a mutual understanding and respect among project participants, was very effective in driving the project in a sound direction. Both persons responsible for usability were centrally placed in the project and took part in the whole software development project, giving a good opportunity to influence all facets of the project. Both persons were aware of the risks of being disqualified when doing usability work and being so close to the project, but none of them experienced problems with this double role. Although the conducted usability work added a significant workload to the development process, the close cooperation seemed to strengthen the overall process and had positive effect on the users, the project team, and the success of the end product.

Making compromises was considered to be necessary by the key persons, both in relation to the level of end product usability and how the various techniques were applied. Given the limited time and resources all usability issues could not be resolved satisfactorily, but when important usability issues were left unresolved the business managers were thoroughly informed about the consequences in a formal way. A clickable prototype would have been useful to test the flow in the application, but in general the applied usability techniques worked well. It was considered desirable to do a regular think aloud test, but there was no clear expectation of what to gain from such a test.

The comprehensive and on-going involvement of the users seemed to have a significant influence on the end product and the reception of the new system among the two user groups. Many users were involved in the process and the way knowledge propagated and was used through the process from requirement analysis and specification to testing and user education, contributed to the necessary anchoring in the organization.

The case also demonstrates how implementing usability work in an organization is hard work. It seems clear that all the informants acknowledged the contribution from the conducted usability work as a critical factor for the success of the project. At the same time people were insecure about whether the successful usability work would be repeated and evolved further in a forthcoming development project, involving a similar sales application for agriculture and industrial insurance products.

REFERENCES

1. Kvale, S. (1994): *InterView*, Hans Reitzels Forlag

2. Strauss, A. & Corbin, J. (1998): *Basics of Qualitative Research*, Second Edition, Sage Publications Inc.

Usability work in practitioners world

Tobias Uldall-Espersen
University of Copenhagen, Dept. of Computer Science
Universitetsparken 1, DK-2100 Copenhagen
tobiasue@diku.dk

ABSTRACT

This study will explore how usability work is conducted among practitioners in organizations as a part of product development and maintenance activities. A qualitative research approach based on case studies, observed practices and interviews with stakeholders will be applied. Existing and new data material will be analyzed with the purpose of identifying important similarities and differences in various cases of applied usability work. The aim of the study is to gain a thorough understanding of usability work conducted under different organizational settings, which could help us increasing the downstream utility of the activities carried out. The research is expected to be useful in industry when applied usability work has limited influence on product design and also, the research is expected to address shortcomings of current usability research.

Author Keywords

Usability, Case Study, HCI, Human Factors, Software Engineering

ACM Classification Keywords

H.5.3 Group and Organization Interfaces, K.6 Management of Computing and Information Systems.

INTRODUCTION

Despite at least 20 years of research into usability engineering, there is still a significant gap between usability evaluation and software product design. Currently, an increasing number of development projects employ usability-engineering techniques in an attempt to improve the quality of software products. For example, the techniques most commonly taken up in the industry are various forms of usability evaluations [12]. While these evaluations often help identifying exhaustive lists of usability problems within a product, they typically have a very limited impact on the subsequent product improvement and current research fails to address why [14].

The current PhD-project is running from May the 1st 2005 until April the 30th 2008 and is part of the Danish USE-Project dealing with Usability Evaluations & Software Design. The PhD-project will try to answer the following research question: *How do practitioners work with usability in organizations and what can we learn from their practices?*

The study is based on three assumptions: (1) Practitioners adapt, improvise and overcome. That is, in order to avoid and overcome critical obstacles in their day-to-day work the practitioners try to adapt to the organizational settings and improvise from their experiences and knowledge. This could introduce important adjustments of the usability improvement process, which could have significant influence on the product improvements. (2) The success of the usability improvement process hangs on more than the usability practitioners' work. Other stakeholders have major influence as well. (3) The stakeholders do not necessarily think in terms of usability. We should be aware of that to avoid being narrow-minded and avoid excluding important observations. With these assumptions in mind, the PhD-project will address individual, organizational and business aspects of practical usability work, and how these aspects are influencing and influenced by current practices. The research approach will be explorative and the project is motivated by its research and commercial importance.

Research motivation

It is an important issue that usability research is rarely carried out within a realistic context of software product development. This issue has been brought to attention a number of times in the last 20 years, for example by Hammond et al. [5], Rosson et al. [10], Buckingham et al. [2], Bellotti et al. [1], and Wixon et al. [14], but few researchers have actually dealt with it [14]. Among the studies listed in comprehensive reviews by Gray & Salzman [1] and Hartson et al. [6], none have been in the context of software product development. Those studies thus fail to take into account the complexity of real-life software development, and the realities in which the results of usability work are to be assessed and used. This raises two issues: (1) We cannot expect results from this type of studies to be reliable descriptions of real-life conditions [8,14]. (2) We have only a limited understanding about how practical work with usability is conducted in different software development contexts [4], for example how and when practitioners adapt to local organizational settings and

how and when other stakeholders influences the results of practical usability work. As a result current usability research might miss factors of importance [14].

Commercial motivation
To serve practitioners, usability evaluation methods (UEM) should have downstream utility. That is, UEMs should have real impact on product design. Recently, Vredenburg et al. [7,12] have documented an increased use of UEMs in industry, but to what degree the applied UEMs actually have downstream utility has not been established by research. Analyses of practical usability work and case studies might help us establish that and by doing so, increase the realism and validity of our claims and recommendations to practitioners. Also, analyzing practitioners' work might help us to understand how practitioners adapt to local organizational settings, how they improvise in their practical work in order to produce results, and how other stakeholders influence these processes. This understanding could help us extract principles that determine the level of product success, which could be of great value to the practitioners and the businesses [14].

RESEARCH APPROACH
The planned study is of an explorative nature and a qualitative research approach has been chosen. Elements from different research traditions like Case Study, Grounded Theory and Action Research will be applied when found useful and appropriate. The object of the analyses is software development projects in which practical usability work plays an important role. The qualitative data from my research will be analyzed together with data from related studies and data from other collection of case studies. The strategy of combining different techniques and cases can, if coherent results are obtained, increase the validity and trustworthiness of the outcome [8,9,15]. The fact that data comes from real-life work can support realism of the results [8] and can increase the legitimacy of claims. The studied projects will be analyzed with the aim of finding important patterns and relations that can be generalized across the cases. Two main activities in the study will be: (1) qualitative and participative studies of practical usability work and (2) analyses of new and existing case studies. These two main activities will be complemented with experiences from participation in national and international research groups.

Qualitative and participative studies of practical usability work
A number of studies of practical usability work in organizations will be conducted. The studies include participation in and observations of usability work, interviews with stakeholders, and analyses of available documentation from the projects. The aim is to gain insight into work processes that lead to usable products and to reflect on how these processes might be improved under the given circumstances. The reflection will be shared with colleagues and practitioners and when possible taken into account and tested in collaboration with the participating organizations.

Analyses of new and existing case studies
A number of case studies dealing with practical usability work will be analyzed. There seems to be a limited number of relevant case studies available. Twenty-five industrial design cases have been published on the CHI 2004 and 2005 conferences, and a few other collections (e.g. [13]) are available. Also, a collection of studies from the NordiChi2006 Industrial Experience Reports might be included. The existing studies are very different regarding details and designs, and only a few have special focus on describing practical usability work as an integrated part of software product development. Hence, an important part of my work will be to take part in a planned collection and publication of a number of new case studies.

Participation in national and international research activities
The involvement in national and international projects like the USE-project and the European MAUSE-project involving more than 20 countries (MAUSE: Towards the Maturation of Information Technology Usability Evaluation, www.cost294.org), yields a unique opportunity to draw on the latest related current research and to cross-examine phenomena of special interest to practical usability work. The USE-project involves a total of nine researchers and PhD-students and during the project period they will conduct related in-depth studies of usability work practices and experiments involving practitioners. So, the USE-project makes it possible to share ideas and experiences and results with other national researchers.

In the MAUSE-project research is aimed at understanding and developing UEMs and understanding the nature of data generated by various UEMs. The aim is to produce results that can be transferred to industry and educators, and the research I participate in is based on a structured framework, which could be complementing the more exploratory part of my study.

STATUS AND FUTURE ACTIVITIES
Based on a case study in a Danish bank, a workshop paper has been published [11] and a journal article is under review. The study describes how usability of an information system was improved and documented. The usability improvements were afterwards traced back to the conducted evaluations in order to estimate the impact from each evaluation. The study suggested a new approach to evaluation of practical usability work. Also, it uncovered new information about the usability improvement process based on real-life experiences, for example how selection of complementing evaluation techniques can influence the result of practical usability work, how different kinds of knowledge are important in the usability improvement process, and how important redesigns can evolve over time based on inputs from multiple evaluations.

I have participated as co-instructor on a workshop for usability practitioners arranged by Rolf Molich in connection to the UI10 conference in Cambridge, Boston 2005. Data from the workshop are to be used in two multi site experiments in the MAUSE-project. Here, the objective is to study traditional feedback from usability evaluations (e.g. problem lists) through coding and merging of problems. Studying feedback from usability evaluations could be useful in order to understand differences in the downstream utility of different UEMs, and to explore the use of usability reports and problem lists as communication and learning tools.

A preliminary study of a limited set of the CHI-cases has been conducted. The study showed only few commonalities between the cases. Further work will be done to develop a framework for analysis of the cases and registration of data.

Currently under planning is a number of interviews with practitioners in the industry. About ten to fifteen interviews will be conducted with different stakeholders in a number of Danish and international companies. Some of the stakeholders will have experiences with usability work and some will not, but all will be key stakeholders in developing usable software products. The aim of the study is to uncover how practitioners develop usable software products with or without specific focus on and formal knowledge about usability.

CONCLUSION

The PhD-project aims at uncovering important issues regarding practical work with usability in a variety of different organizational settings, where local conditions are supposed to influence the usability improvement process. The project is based on independent and original research and research done in cooperation with national and international researchers, and the results of the project will be based on analyses of multiple independent data sources. Special attention is devoted to the issues of relevance, validity and reliability of claims and recommendations in relation to practical usability work. The project aims at contributing to both the research and the commercial communities by addressing a subject that has been underweighted by research, but is important to practitioners, researchers and students in the HCI-communities.

Acknowledgements

REFERENCE LIST

1. Bellotti, V., Buckingham Shum, S., MacLean, A. & Hammond, N. Multidisciplinary modelling in HCI design\…in theory and in practice, *Proc. CHI 1995*, ACM Press (1995), 146-153.

2. Buckingham Shum, S. & Hammond, N. Transferring HCI modelling and design techniques to practitioners: a framework and empirical work, *Proc. BCS HCI'94*, Cambridge University Press (1994), 21-36.

3. Gray, W. D. & Salzman, M. C. Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods, *Human-Computer Interaction 13*, 3 (1998), 203-261.

4. Grudin J., Interactive Systems: Bridging the Gaps Between Developers and Users, in Baecker, R. M., Grudin, J., Buxton, W. A. S., & Greenberg, S. *Readings in Human-Computer Interaction: Towards the Year 2000*, Morgan Kaufmann Publishers, Inc., 1995, 293-303.

5. Hammond, N., Jørgensen, A., MacLean, A., Barnard & P., Long, J. Design practice and interface usability: Evidence from interviews with designers, *Proc. CHI 1983*, ACM Press (1983), 40-44.

6. Hartson, H. R., Andre, T. S., & Williges, R. C. Criteria for Evaluating Usability Evaluation Methods, *International Journal of Human-Computer Interaction 13*, 4 (2001), 373-410.

7. Mao, J.-Y., Vredenburg, K., Smith, P. W., & Carey, T. The State of User-Centred Design Practice, *Commun. ACM 48*, 3 (2005), 105-109.

8. McGrath, J. E. Methodology Matters: Doing Research in the Behavioural and Social Sciences, in Baecker, R. M. et al. (eds.) *Readings in Human-Computer Interaction*, Morgan Kaufmann Publishers Inc., 1995.

9. Naur, P. *Knowing and the Mystique of Logic and Rules*, Kluwer Academic Publishers, Dordrecht, 1995.

10. Rosson, M.B., Kellogg, W. & Maass S. The designer as user: building requirements for design tools from design practice, *Commun. ACM*, ACM Press (1988), 1288-1298

11. Uldall-Espersen, T. Benefits of Usability Work - Does It Pay Off?, *Proc. International COST294 workshop on User Interface Quality Models (UIQM'05)* in Conjunction with INTERACT'05 12.-13. Sept 2005, Rome, Italy, (2005), 7-14.

12. Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. A Survey of User-Centred Design Practice, *Proc. CHI 2002*, ACM Press (2002), 472-478.

13. Wiklund, M. E. *Usability in Practice: How Companies Develop User-Friendly Products*, Academic Press Professional, Inc., 1994.

14. Wixon, D. Evaluating Usability Methods: Why the Current Literature Fails the Practitioner, *interactions, 10*, 4 (2003), 28-34.

15. Yin, R. K. *Case Study Research: Design and Methods*, Sage Pub, Thousand Oaks, CA, 2003.

Benefits of usability work – does it pay off?

Tobias Uldall-Espersen

Datalogisk Institut, Københavns Universitet

Universitetsparken 1, DK-2100 Copenhagen Denmark

tobiasue@diku.dk

ABSTRACT

The value of work to improve the usability of information systems in industrial settings is rarely accounted for. This study reports from an experiment where an administration and risk management system in a bank were improved through three versions during a period of six month. The system has ten users and usability data was collected through questionnaires and logging of data from practical use. The experiment shows how it was possible to improve the system over a broad range of measures covering efficiency, effectiveness and user satisfaction issues. The analysis of return on investment (ROI) of this usability work shows a pay back time of five years, but in this case, the most important improvements were found in issues not accounted for by the ROI. These issues, e.g. expected increased use of the system and a postponement of a replacement of the system alone justified the usability work. This indicates how return on investment analyses are at risk of missing the most important issues.

Author Keywords

Field Experiment, Case study, HCI, Usability engineering, Usability factors, Evaluation Techniques, ISO 9241-11, Return on investment, Cost-Benefits, Metaphors of Human Thinking, Think Aloud, Questionnaire, Logging of usage date, Business process reengineering

ACM Classification Keywords

H.5.2. User Interfaces, D.2. Software Engineering, D.2.9. Management

INTRODUCTION

Traditionally in software development, a number of processes have special focus, e.g. requirement definition, design specification, implementation and test. These processes must be completed and documented consistently and systematically, but this has not sufficiently ensured usability of the developed systems. If usability is to be ensured, the software development processes must be enhanced with further activities. There is a diversity of different techniques, which has been diligently compared, but there seems to be a lack of field studies, documenting the impact of usability work in industrial settings. The purpose of this study is to report how usability has been improved and documented in an industrial software engineering experiment.

The experiment was designed and conducted with the aim of the highest possible realism. All data is collected from real users performing real tasks using a real system in a specific business. Data is based on the usage of an information system used by ten users in a bank in Denmark. Using three versions of the system, data has been collected through six month of logging and two questionnaire surveys have been conducted in the start and at the end of the experiment. The usability work is made by the software engineer who has build the system, which could be a useful approach in future projects.

AIM OF THE EXPERIMENT

The aim of the experiment was to investigate the following:

- Can usability of the system be improved?

- Can changes in usability be identified, documented and measured?

- Does it pay off to identify usability issues and to improve the system?

Can the level of usability be measured, and changes in usability between versions of the software be identified, documented and explained? In the current experiment, this is a precondition for the objective evaluation of the benefits, just as it is necessary to document product performance in an industrial context.

THE SYSTEM AND THE USERS

The evaluated system is a small MS Windows based information system, which was previously developed by the author of this paper. It was implemented in the PowerBuilder programming language and was connected to an Informix database system. The system was used in the bank a couple of years before the experiment was conducted. It is used for administrative purposes, for reporting and for management of risk in relation to a specialized loan department with a total loan amount of about 150 millions euro. The system consists of 10-12

primary windows, where data can be searched, inserted and updated. Further, it has a number of secondary windows and ten reports. The system is the users primary tool for the administration of the loans. In the experiment, only the primary windows were evaluated.

During the test-period ten users had access to the system, and all of them contributed with data. The users were all bank employees with a financial education, and they worked in two different departments. Eight of the users came from the primary department (the primary users), where the system was used mostly. Two secondary users came from a department, which worked with control and risk management for the overall company group. In both departments, the distribution of gender was equal. At the beginning of the experiment the newest employee had been in the department for about 9 month, and hence all of the users where used to internal routines and business rules. All of the primary users had prior to the experiment access to the system, one of them used it only sparsely. Of the secondary users one used the system sparsely and the other got access to the system when the experiment started. The nine users who had access to the system prior to the experiment, all participated in the survey in the start and at the end. They were asked to express their experience about Information Technology (IT) and about the system. Four respondents had middle experience with IT, two had little, and the last three had very great, great, and very little IT experience. Overall a level of experience just below middle. Regarding use of the system, five respondents had much experience, one had middle, two had little and one had very little. The total experience with the system was a little higher than middle, but the distribution was very unequal.

EXPERIMENTAL METHOD

To define usability the ISO 9241-11 standard was used:

"Usability: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."

As suggested by Frøkjær et al. [3] effectiveness, efficiency and satisfaction were a priori considered as being independent factors. Thus, it was necessary to measure all three factors, in order to be able to assess the total change in usability.

The experiment was conducted as a field experiment in an effort to increase its realism compared to laboratory-style studies [8]. During the experiment, two iterations of usability evaluations and system improvements were conducted, and the usage of the initial version as well as the two improved versions were studied. The users worked with all three different versions of the system during the experiment, and all users used the same version at the same time. Every version was used no less than 1½ month. Note that data in this article are limited to the study of the first and the last tested version.

Identification and Processing of Usability Problems

The first part of the experiment consisted in identifying and rectifying potential usability problems. The work was focused with the survey in the start of the experiment, where the users had opportunity to evaluate and comment different parts of the system. The survey was followed by a total of nine usability inspections using both Metaphors of Human Thinking (MOT) based evaluations and think aloud (TA) tests. The MOT evaluation technique is an analytical technique based on five recognized aspects of the human way of thinking [5]. The author conducted two evaluations alone and additionally two pair evaluations with Erik Frøkjær and Kasper Hornbæk – the originators of the MOT-technique. The TA activities were based on the descriptions of Contextual Inquiry [1] pp. 41-66 and TA test [9] pp. 100-115.

The first iteration included the initial questionnaire survey, one MOT based evaluation conducted by the author and three think aloud tests. The survey and the MOT evaluation were link close together, since issues from the survey were given special attention during the MOT evaluation. Both low ratings of parts of the system and the processes as well as user comments were used to point out special issues. This seemed to be an efficient way to bring in some contextual knowledge about the use of the system into the evaluation process. The three TA tests were conducted with three different users, who together used all essential parts of the system. Data from the survey was used for the selection of users and tasks, and the users were ask to collect about two hours work covering the identified tasks, which could be done during the TA sessions.

In the second iteration, two additional users were observed using the improved system for the first time. This served as a special opportunity for evaluating the implemented changes, but the value of the evaluations was limited. More effort should have been put in to identifying tasks covering the improved parts of the system. The iteration was finalized with three MOT evaluations; the first conducted by the author alone and the two following as pair evaluations. All of the nine evaluations were conducted in a very informal way and time was used on both identifying problems and discussing possible solutions.

During the two iterations, 180 intermediate problems were identified. They were consolidated into 99 unique problems from which 40 points of improvement were proposed. The consolidation process was conducted with the main objective to remove clearly identically problems, e.g. problems of the same type regarding the same objects observed during the same processes. Because the author had observed all intermediate problems except 13 originating from the questionnaires, the consolidation could be done without much difficulty, as the context of the problems was clear. The purpose of the proposed points of improvement was to group the consolidated problems into collections that could be fixed together through one proposed redesign. For every point of improvement a

possible redesign was worked out and shortly described, and every redesign were then subsequently estimated and informally prioritized from the following four factors:

- How serious are the covered problems?
- How often do they occur?
- How many users do they affect?
- How much time is needed to implement the solution?

Based on the priority, a plan for implementation was developed. According to the goal of optimizing the return on investment (ROI), the focus was on solving the issues, which gave the users most value for money. At this stage of the usability improvement process, it seemed appropriate to be guided by suppositions about ROI, but it seemed equally important to let broader strategic considerations guide the selection. For example, one of the most expensive improvements of the system was a total reengineering of the user manual and help system. As a single activity this would never pay off in the experiment, but on a longer term it increased the value of the system because the users felt more secure when using the system and the system was easier to deploy in another department, which later was done.

In parallel with the traditionally usability engineering activities, the users had access to a database, where all problems were registered and problems could be added and commented. Additionally the use of the system help function was registered. The users were asked after every access to the help function if they found what they were looking for. None of these techniques had much influence on the result, whereas an informal dialog concerning some of the identified problems and solutions was useful. Of the 180 intermediate problems, 107 were solved.

Measuring Usability

The second part of the experiment consisted in measuring the changes of usability, i.e. improvements of efficiency, effectiveness and satisfaction. For this purpose, logging and two questionnaire surveys were used. The logging was done through the entire six-month experimental period, whereas the surveys were conducted in the start and at the end of the period. All logged data originates from real use of the system in the production environment supporting the highest possible realism. The changes in efficiency were measured by logging, user satisfaction by surveys and effectiveness by both logging and surveys.

Hypotheses

To guide the experiment, a set of 25 hypotheses was put forward, which could help identifying changes in usability. To each hypothesis, a description was made of its relevance and how to clarify it. The most important hypotheses are described below.

For efficiency considerations, six hypotheses were put forward, e.g.:

- *Hypothesis:* Applying usability work, one can reduce the orientation time, i.e. the time from a window is activated until the user interacts with it.

  Relevance: A reduction in orientation time indicates that less effort is required for the user to get a general view of a window, to recall the task and to choose required functionality.

  Clarifying: Through logging it would be possible to measure the time from the window is activated to the first following user-generated event.

- *Hypothesis:* Applying usability work, one can reduce the time from a user starts the program to the first useful windows (e.g. a window with data relevant to the user) is shown.

  Relevance: A reduction in startup time indicates that less effort is required for the user to recall the use of the system.

  Clarifying: Through logging it would be possible to measure the time from the system start event to the activation of the first window with user selected data.

For effectiveness considerations, 12 hypotheses were put forward, e.g.:

- *Hypothesis:* Applying usability work, one can reduce the number of searches for and updates of the same data object.

  Relevance: Repeated search for and updates of the same data object could indicate bad effectiveness e.g. that the user is insecure or makes erroneous updates.

  Clarifying: Through logging, count the number of searches for and updates of each data object.

- *Hypothesis:* Applying usability work, one can reduce the number of interruptions the user experiences.

  Relevance: Each interruption could distract and annoy the user.

  Clarifying: Through logging, count the number of interruptions in each window.

- *Hypothesis:* Applying usability work, one can reduce the number of unsaved changes of data.

  Relevance: Canceled changes indicate that the user starts modifying an object that should not be modified.

  Clarifying: Through logging, count the number of canceled changes in each session.

For satisfaction considerations, seven hypotheses were put forward, e.g.:

- *Hypothesis:* Applying usability work, one can improve the users satisfaction with the system.

  Relevance: Unsatisfied users could be a usability issue.

  Clarifying: Ask the users about their satisfaction in the survey.

- *Hypothesis:* Applying usability work, it is possible to increase the users perception of working efficiently.

  Relevance: Feeling not wasting time could increase user satisfaction.

  Clarifying: Through the survey, ask the user questions about task solving experience.

- *Hypothesis:* Applying usability work, the system can be made more convenient to use.

  Relevance: An inconvenient system could reduce user satisfaction.

  Clarifying: Through the survey, ask the user questions about matters of inconveniency, irritations etc.

Logging

The system logged 31 different types of events generated by the users and their usage of the system, e.g. opening/closing of a window, click with the mouse, change of focus, changes in data and transactions with the database. In the log, it was possible to identify every data object a user viewed or changed.

Questionnaire

The users were asked for background information (age, gender, …), 19 questions about their satisfaction with the system and 20 questions about their experience with the use of the system. All the 39 questions were answered on a five-point Likert scale, and each had a text field for free comments where the users were encouraged to go into details in their answers. The 20 questions about their experience were:

How much do you agree in the following questions?

A1. There are tasks in the system, which are difficult to solve.

A2. There are tasks in the system, which are to time consuming to solve.

A3. There are tasks the system should be able to solve, which cannot be solved.

A4. There are tasks in the system I do not know how to solve.

A5. There are special tasks in the system, which I hand over to others.

A6. There are special tasks in the system, which often are handed over to me.

A7. There are tasks in the system, which I often need help from other in the department to solve.

A8. There are tasks in the system, which I have to help other solving.

A9. There are tasks in the system, which I avoid solving.

A10. There are parts of the system, which annoys me when I use it.

A11. There are parts of the system, which I use without understanding it.

A12. I am insecure about how to solve a task and tries my way.

A13. There are parts of the system, which I feel insecure using.

A14. There are parts of the system, which often give rise to errors.

A15. There are shortcomings in the system.

A16. It is difficult to get around in the system.

A17. I miss feedback from the system.

A18. The system does something different from what I expect.

A19. The system expects me to solve a task in another sequence, than I would have done.

A20. I often make errors in parts of the system.

Seventeen of the questions abut satisfaction dealt with parts of the system and two were general. The non-general questions were:

How satisfied are you with the following parts of the system?

B1. The search function where you can select engagement, investor or guarantor from a list.

B2. Selection of and navigation between windows.

B3. Set up an engagement.

B4. Adding a property.

B5. Adding a tenancy.

B6. Adding a mortgage.

B7. Set up an investor.

B8. Set up a guarantor.

B9. General view of engagements.

B10. General view of properties.

B11. General view of tenancies.

B12. General view of mortgages.

B13. General view of investors.

B14. General view of guarantors.

B15. Reports.

B16. Import of data from data warehouse.

B17. The total general view.

The two general questions were:

B18. How satisfied are you generally with the system?

B19. If a colleague in another department were thinking of starting using the system, would you recommend it?

DATA

About 250.000 log entries originating from about 190 hours of active use of the system were logged. This covers 509 different logins from the users in the period from week 16 to week 44 in year 2003. Totally, the users had been logged on for 1486 hours of which the system was active in 13% of the time. A great deal of the log turned out to be difficult to analyze quantitatively, and was not analyzed. The reason for this was that the usage of the system was more diversified than expected and there were considerable differences in how much the users used the different parts of the system. Nine of the ten users chose to participate in the questionnaire surveys. The surveys were conducted by a third department and were done using a professional web based tool.

In addition to the primary data collection, the use of resources where carefully registered and evaluated, and a final evaluation of the usability improvement process was conducted. To be able to estimate cost-benefits and return on investment later, all time consumption according to relevant activities were collected. This includes time used to study techniques, analyze the system, analyze problems and solutions and the time used to solve the selected problems. The final evaluation was an interview with the responsible department manager, who also was one of the primary users of the system. The purpose of the interview was to get a reaction from the manager to the completed process, the achieved results and the return on investment.

RESULTS

The experiment resulted in important changes of usability in the system and clear changes of efficiency, effectiveness and user satisfaction were documented.

Efficiency

The changes in efficiency were primarily obtained through a better adaptation to users work practice and a reduction in time used to startup and to orientation. The experiment uncovered how the users work practice had changed from the first to the final version of the system. Existent routines evolved, new ones developed and the users had specialized in different parts of the system. This caused a need for adaptation and business process reengineering (BPR). This is shown by the following example: Twice a year the users had to print a summary of every engagement (about 150) and of every investor (about 600). This was done by manually selecting one engagement / investor at a time and then pressing the print icon, which was a rather time consuming process. In the final version of the system there was added a 'print all engagements' and a 'print all investors' function, which saved a great deal of time. In a similar manner, a great part of the data maintenance processes was reengineered. For example the yearly process of registering an engagement account was reengineered from registrations in three general windows to registration in one specialized window, also displaying the last registered account information. This made it easy both to

Changes in Orientation Time
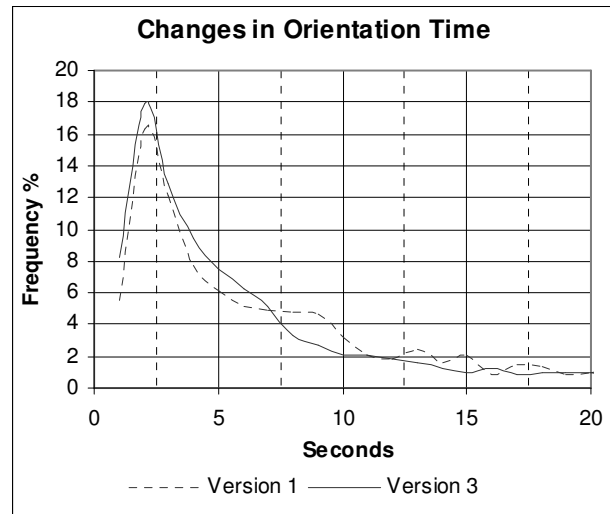
register data and to control the typed in numbers against last registration. The experiment resulted in considerable improvement of the adaptation of the users primary work processes to the system. For a number of important work processes, performance was improved by 40%-76%.

In addition, substantial improvements regarding the startup of the system, navigation in the system and orientation in different windows were documented. The time from starting the system to viewing relevant data was reduced from 43 seconds to 31 (28%). The reduction is even more evident, if we watch the 90% percentile: The time has been reduced from 27.4 to 16.9 seconds (38%). This improvement saves time for the users and may even result in further use of the system because of the faster access to data. The database access time did not change significant during the experiment, and the observed effect can to a certain extent be explained by a centralization of the search facilities to one window and some changes in the navigation. In the first version, the users did all navigation through menu icons or menu entries. In the last version, the navigation was done through buttons placed in every window. Technically, slightly more complex to implement, but it seems to be much more usable. From the measured startup and search times, the experiment shows a reduction of time used to navigation in the order of 60%-70%.

If we look at the time used from a window is opened until the user interacts with it (called *orientation time*), another important result has been pointed out. For the 90%
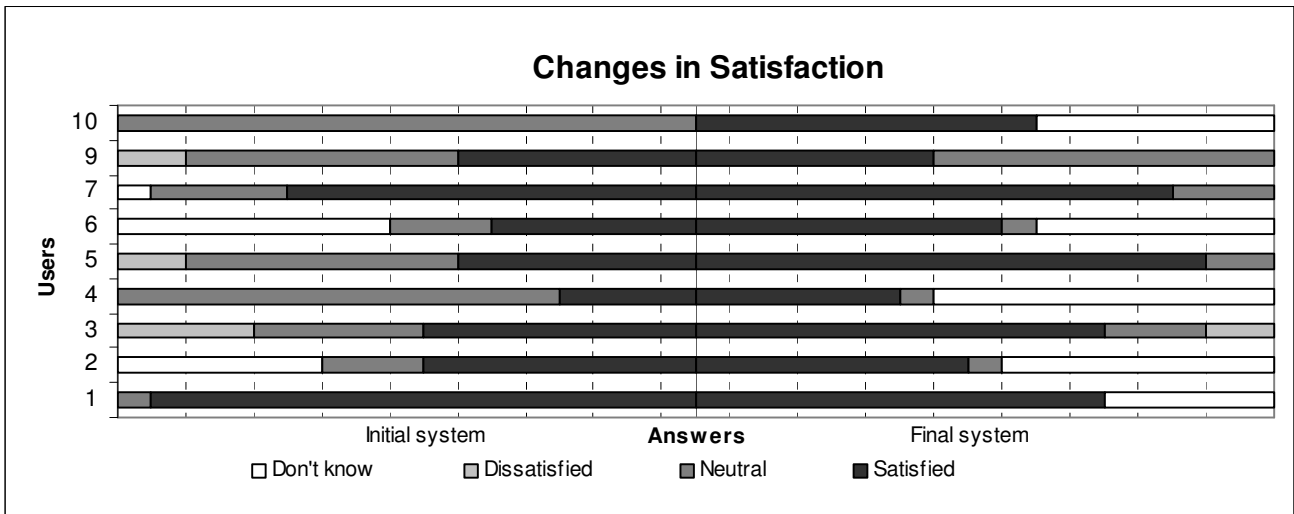
Figure 2 – At the left half is shown the satisfaction with the initial system, and on the right half the satisfaction with the final system. Overall, the users seem less dissatisfied and more satisfied. An increased number of *don't know* could confuse the picture. These have been rejected in the analysis. User 8 did not participate in the questionnaire.

percentile of more than 7.800 observations, the time has been reduced from 8.3 to 5.8 seconds (30%). This is an important result; the fact that it is possible to improve the time used to orientation and forming a general view, could reduce the waste of time in the system. Figure 1 shows a frequency plot of the time used to orientation. In version 1, about 58% of the orientations was done on less than 7.5 second and about 28% between 7.5 and 20 seconds. In version 3, the same numbers where about 68% and 21% which shows, that the users uses less time to orientations. The amount of slow observations (time>20 seconds) seems to bee stable. Some external factors might have influenced the orientation time, e.g. changes in the physical environment or the users learning the interface, but no obvious sources of influence exists. During the experiment there were no changes in the physical environment, and there is no indication of a learning effect within each of the time periods where a single version of the system has been used.

Effectiveness
The changes in effectiveness was documented through both surveys and logging. The users' experience (based on 360 answers from the surveys) with the system did only improve slightly (0.2 points), and only one user expressed a significant change (0.6 point). The improvements documented by the logging were more substantial. A reduction in interruptions on about 88% was observed. To a degree, this fact could be explained by a change in feedback and an improved error handling. The result is fewer disturbances of the users. Likewise, a 79% reduction in the number of times, the users starts changing data without saving were observed. This expresses a more suitable use of the system.

User Satisfaction
The questionnaires gave 138 answers about the initial satisfaction with the system and 128 about the final. On a scale from 1 (very unsatisfied) to 5 (very satisfied) the initial satisfaction was 3.56 and the final was 4.03. In a meta-analysis based on 127 measurements, Nielsen & Levy [11] has shown that 3.5 is a normal level for interfaces before improvement. However, there seems to be considerable variation. On 6 of the 17 questions, there were significant improvements (0.5 – 1 point) and four of nine users expressed significant improvements (0.7 – 0.9 point). Figure 2 shows the change of user satisfaction. Only one user did express dissatisfaction with parts of the improved system. The general questions did not show any significant changes.

Time and Management Consideration
As an important part of the experiment, all relevant time consumption was registered. Overall 15 hours were used to study the techniques, 25 hours to analyze the system, 10 hours to analyze and prioritize problems and solutions and finally 117 hours to implement improvements. This use of time should be seen in lights of limited experience with usability work, but great knowledge about the system.

The final interview with the department manager revealed a general satisfaction with the undertaken process and the achieved results. He emphasized three special issues, to which he felt the work had contributed:

- To give the user a better general view of data.

- To make the system better and easier to use. The typing has been faster and fewer errors are made.

- The users have gained more confidence with the system and they feel less insecure.

The manager expressed that he had difficulty estimating the actual economic benefits. He stressed how the improvement of effectiveness and satisfaction was more important and interesting to him than the gained efficiency.

DISCUSSION

The experiment shows, that important and appreciable improvements of usability can be made through use of existing usability techniques. The fact that the experiment is a field experiment could support its realism, but it may cause that it is hard to generalize from the results [8]. If we look at the improvements two main observations can be made. Waste time can be reduced and work processes can be supported more efficiently e.g. by BPR. This seems to affect all the usability factors, and may be an issue for other systems.

The considered system was originally developed by the users, and the first version served as a prototype. The prototype was prior to the experiment developed further by the author in close cooperation with the users through a couple of iterations. It shows that even when developing software with serious user involvement, important conditions could be overlooked or misunderstood. It also suggest, that the need for BPR will reappear from time to time when the relation between problems, tools and people has changed. A relation Naur has named The Symmetrical Relation [10] p.29.

Should Software Engineers Work With Usability?

Seventy percent of the resources used in the experiment have been used on the implementation of the identified and designed improvements. This makes it appropriate to aim at ensuring usability of the design before the system is implemented. Therefore, usability work and software engineering should be combined in an iterative process. Close integrating of usability work and software development is also described as necessary in e.g. [6]. This could also be a key to weaken the resistance to usability work, which continues to exist among software developers [13], and be a key to support a more holistic understanding and acceptance of the usability field.

In the experiment, a substantial evaluator effect [7] has been established. The evaluations conducted by the author together with two usability experts revealed considerable different problems compared to the problem identified by the author alone. If the software engineers should evaluate their own products, this could be a problem. They are influenced by earlier decisions, known limitations of the tools, and imaginations of the users, surroundings and so on. This could cause a lack of objectivity and may result in that serious problems are overlooked. The software engineers could however strengthen the process of development if these risks are managed. This might be done by involvement of independent experts, by close cooperation with the users and by focusing on the users experience with the system. In some cases, involvement of software engineers could also result in special focus on solvable problems, which could ensure a higher return on investment.

Does it pay off?

The experiment has documented an improvement of the efficiency on about 10% (35-40 hours a year) in relation the total use of the system. This only covers measured improvement and might be bigger, e.g. the effect of the changed *print all* procedure has not been measured, since it only appeared once in the test period. An important question is whether the efficiency improvement together with the changes in effectiveness and satisfaction can justify the costs of the usability work. Totally, the use of time was about 180 hours, which covers studying the techniques (9%), analyzing the system and the possible changes (21%) and traditional software engineering activities (70%). This indicates that the improved efficiency could finance the work over a five-year period, but what pay back time is expected and realistic in industrial settings? Frøkjær & Korsbæk [4] have shown that public information systems tend to have a considerable lifetime and a payback time on five years might not be unrealistic. The value of the effectiveness and the satisfaction is difficult to calculate, but it seems considerable and more important to the company. In the interview, the department manager expressed, that the completed work has extended the lifetime of the system, and a planned replacement has been postponed. Another department has even adopted the system because of the improvements. This is a saving, which immediately seems to have justified the investment.

However, the experiment has shown that we cannot rely on return on investment (ROI) analyses of usability issues. It would be hard, in advance, to calculate ROI because of the complex measurable outcome variables and a ROI analysis might not have been in favor of the conducted work. Afterwards it seems obvious, that it was a reasonable strategic investment, which supports some of the principles discussed by Dray et al. [2]. They argue in favor of using case studies as a tool to document the effect of usability work, in order to convince practitioners, researchers and decision makers.

Are Efficiency, Effectiveness and Satisfaction Equally Important?

In agreement with Frøkjær et al. [3], the experiment has shown that we cannot expect coherence between efficiency, effectiveness and satisfaction. As an illustration from this experiment, the search function was changed significantly. The users got a more transparent search facility with better search options in a centralized search window. This really improved the effectiveness, but the mean search time increased by 15%. The user satisfaction did not chance significantly. This supports the argument, that the three usability parameters must be considered independently and hence that all three parameters must be measured.

The lack of coherence raises another question. We need to control efficiency, effectiveness and satisfaction, but are they necessarily equally important? Many studies focus on efficiency and/or satisfaction, whereas effectiveness seems to be less considered [3]. Does this mean that effectiveness is less important, or could it be due to the fact, that it is harder to measure? In the interview with the department manager, he expressed that in this context effectiveness and user satisfaction were more important to him than efficiency. Would this be a concern, we need to bring in to the work with usability? The priority depends of the context, but it could be important to be aware of, which preferences the users and decision makers have. If the work is guided by these preferences, it may be easier to prioritize the resources and to use them well.

CONCLUSION

The conducted experiment has shown the following:

- It is possible through systematic work to improve efficiency, effectiveness and satisfaction of the system. The combination of analytical and empirical techniques has shown to be very useful, and it seems that the benefits of using expert-reviews justify the extra costs.

- It is possible through systematic work objectively to document changes in efficiency, effectiveness and satisfaction. Questionnaires and logging are useful in this documentation process and complement each other. To limit the use of resources the measurements must be established and focused on the issues of central importance.

- Efficiency improvements were measured and time of return on investment could be estimated to about five years.

- The economic value of the improvements of the effectiveness of the system and the user satisfaction were impossible to make up although the benefits were clearly identified by the direct users and the manager. In the current experiment the usability improvements raised expectations of increased use of the system also in a new department, and a likely replacement of the system was postponed.

- Although the manager recognized the return on investment analysis based on mainly efficiency measures, he made it clear how the identified effectiveness and user satisfaction improvements were the most important result. They alone justified the usability work. This indicates how return on investment analyses are at risk of missing the most important issues.

REFERENCES

1. Beyer, H. & Holtzblatt, K. *Contextual Design*. Morgan Kaufmann, 1998

2. Dray, S., Karat, C.-M., Rosenberg, D., Siegel, D. and Wixon, D. Is ROI an Effective Approach for persuading Decision-Makers of the Value of User-Centered Design? *Panels CHI 2005*, 1168-1169

3. Frøkjær, E., Hertzum, M. and Hornbæk, K. Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? *Proc. CHI 2000*, 345-352

4. Frøkjær, E. & Korsbæk, H. (1992). Informatization Policies in Denmark. In Frissen, Bekkers, Brussaard, Snellen, & Wolters (Eds.), *European Public Administration and Informatization* (pp. 25-47). IOS Press.

5. Hornbæk, K. & Frøkjær, E. Evaluating User Interfaces with metaphors of Human Thinking in *User Interfaces for All*, Lecture Notes in Computer Science 2615, Springer-Verlag, 486-507

6. Hornbæk, K. & Frøkjær, E. Comparing Usability Problems and Redesign Proposals as Input to Practical System Development. *Papers CHI 2005*, pp. 391-400

7. Jacobsen, N.E., Hertzum, M. & John, B.E. The Evaluator Effect in Usability Test. *Late-Breaking Results CHI 1998*

8. McGrath, J. Methodology Matters: Doing Research In The Behavioral And Social Sciences, pp. 152-169 in Baecker, R. M.: *Reading in human-computer interaction: toward the year 2000*, Morgan Kaufmann Publishers, Inc., 1995

9. Molich, R. *Brugervenlige edb-systemer*, Ingeniøren Bøger, 1994, 2. edition 2001

10. Naur, P. Computing: A Human Activity. *ACM Press / Addison-Wesley*, 1992

11. Nielsen, J. & Levy, J. Measuring Usability – Preference vs. Performance. *Communications of the ACM, April 1994/ Vol. 37, No. 4*

12. Uldall-Espersen, Tobias, 2004. Benefits of usability work – does it pay of, *Department of Computer Science, University of Copenhagen*. (Master thesis in Danish)

13. Vredenburg, K., Mao, J., Smith, P. W. and Carey, T. A Survey of User-Centered Design Practice. *Proc. CHI 2002*, 471 - 478

The Usability Perspective Framework

Tobias Uldall-Espersen

University of Copenhagen
Department of Computing
Njalsgade 80
DK-2300 Copenhagen S
tobiasue@diku.dk

Abstract

The usability perspective framework is a tool under development supporting various stakeholders in contributing to software systems usability. This paper describes the work conducted so far and argues how the framework is expected to contribute to both academia and practitioners in industry. The framework's fit to software development systems is discussed and we argue why the framework might succeed where current usability work often fails. Finally the plans and ideas for future work are described.

Keywords

Usability, software development, stakeholders, tools.

ACM Classification Keywords

H5.1. Information systems models and principles,

Introduction

Despite increasing investments in usability work, developing and introducing useful and usable software systems in organizations continues to be a challenge. One problem is that current usability work techniques rarely support stakeholders in developing realistic and coherent understandings of and expectations to development projects. This paper describes the ongoing work on developing a framework supporting stakeholders in developing such understandings and expectations. The work is based on a comprehensive interview study [13,16], a CHI 2007 workshop [15], and a study of practical usability work [17].

The usability perspectives:
1. The interaction object usability perspective
2. The task usability perspective
3. The product usability perspective
4. The context of use usability perspective
5. The enterprise usability perspective

Figure 1 - The five usability perspectives.

Theme: Consistency
- Is the interaction design and visual design of interaction objects consistent?
- Is the task flow design consistent?
- Is the product design consistent across different important parts?
- Is the system consistent with other relevant systems and the actual context of use?
- Is the system technically and organizationally consistent with the enterprise?

Figure 2 - Key questions regarding consistency.

The usability perspective framework

The usability perspective framework is being developed by the author in an attempt to create a tool operational in practical software development. The framework has three core elements: (1) Five usability perspectives (Figure 1), (2) three cross-perspective themes, and (3) 18 key questions relating the usability perspective to the themes (Figure 2 - 4). Each core element is described in the following starting with a description of the completed work the framework is based on. We do not intend to claim that there is a clear cut between the different perspectives, the different themes, and the different key questions. Each element is formed aiming at providing the most valuable insights. However, the elements have a generic style, which in specific situations should be formed and concretized.

Completed work

The completed work focused on building an understanding of practical usability related work in different software development projects. The work gave insights into current practices, important challenges, and useful approaches to ensure usefulness and usability of developed software systems. In [17] we explored how different stakeholders contributed to practical usability work in an organization. The study showed how knowledge from different contexts was important to improve the product. In [16] we explored the role of stakeholders in different organizations developing software. This study showed how different stakeholders addressed usability differently and how they contributed important insights to software development projects, but often in their own way. In [13] we presented five usability perspectives applied by stakeholders in six Danish software development projects (Figure 1). Each perspective makes up a

distinct approach to usability found important in practical software development. In [14] we discuss the relationship between the usability perspectives and practical software development. This work suggests a clear benefit from multi perspective usability considerations.

The work-in-progress

The following section describes the author's ongoing work on developing the framework. The description takes the three elements as starting point.

THE USABILITY PERSPECTIVES

The five usability perspectives (Figure 1) support considerations of usability based on different stakeholders interests and needs. It is believed that such considerations add value to software development projects. In the following we introduce the five perspectives. Further descriptions can be found in [13].

Interaction object usability concerns if users successfully can interact with isolated objects in the product. Interaction objects are often evaluated in isolation, but a reductionistic view can be problematic [14,18]. We suggest considering properties of interaction objects among stakeholders to determine essential requirements to behaviour, design, and functionality with respect to use context and domain.

Task usability concerns whether users are able to complete single tasks through a combination of interactions with available interaction objects. Usability of task design is often measured in terms of task completion time, error rates, and task completion [5], but whether such measures reflects the priorities in organizations can be questioned [17].

Figure 3 - Key questions regarding coherence.

Product usability concerns whether the product supports users in reaching a coherent set of goals with effectiveness, efficiency, and satisfaction. Some stakeholders consider usability of the entire product rather than usability of individual interaction objects or tasks, and hence product usability becomes an important usability perspective. Where task usability concerns having a straight way to reach a specified goal, product usability concerns interplay between different parts of the product allowing numerous roads to reach important goals.

Context of use usability concerns to what extent use of the system, possibly interplaying with other systems, in the actual context of use is effective, efficient, and satisfactory. Understanding the possibilities and limitations of different contexts of use is a key issue. Field studies or contextual interviews can be applied as formal methods for generating insights, but it is rarely realistic to obtain a complete and comprehensive understanding of the context of use through such methods.

Enterprise usability concerns if goals of the enterprise are fulfilled effectively, efficiently, and satisfactorily through use of the system. Usability work should not be limited to only supporting users in reaching own goal; users should also be supported in reaching main goals of the enterprise. Addressing enterprise usability could help increasing the level of integration between business-related stakeholders and IT-related stakeholders, which is a challenge in industrial software development [7]. Furthermore, addressing enterprise usability should make the value of usability work visible to various key stakeholders, which would benefit usability work using other perspectives as well.

To support coherence across the different perspectives three themes are introduced. These are *consistency*, *coherence*, and *fitness*. The themes grew forward during an affinity diagramming activity with the key questions.

Consistency (Figure 2) regards uniformity or agreement among related or similar elements within or across systems, also when considered over time. Consistency increases the reliability of systems and helps subsystems to converge on the same overall goals and visions. Inconsistency increases the risks of contradictive behaviour in or between systems.

Coherence (Figure 3) in a system regards how well different system elements stick together by being logically connected, by being mutually supportive, or by sharing meanings and values across elements. Coherence within and across different perspectives is necessary to holding a system together and making it robust. Coherence across different systems is necessary when systems are related, for example when information is exchanged or shared.

Fitness (Figure 4) concerns the quality of the relationship between system elements. A good fit between stakeholders and tools supports stakeholders in solving relevant problems efficiently and effectively. When unpredictable or uncontrollable variations within a work domain occur fitness becomes a question of flexibility and resilience, and when there are organizational interests involved fitness is a question about how well relationships between stakeholders, tools, and problems fit as an entity in the overall organizational settings.

Theme: Fitness
- Does the applied methods or technology inform the design and implementation of interaction objects?
- Does the applied methods or technology inform the design and implementation of tasks?
- Is the product adaptable to different user behaviour, strategies, and preferences?
- Is the product adaptable to internal errors and faulty or unintentional actions?
- Is the system adaptable to variations in the context of use?
- Is the system adaptable to external circumstance?
- Does the system fit the enterprise in terms of technical and organizational needs, wants, and haves?

Figure 4 - Key questions regarding fitness.

Using key questions in usability work is inspired by [3]. Key questions are easy to apply, are understandable, and can inspire new insights in software development projects. They do not require specific tool knowledge and can be applied by stakeholders independent of their experiences and affiliation in the organization. This decreases the risk when developing software and increases the face validity of the generated insights. Currently we work on developing, explaining, and exemplifying the key questions. Figure 2 - 4 gives an overview of the key questions so far.

The 18 key questions are based on observations from the interviews presented in [13,16] combined with experiences from other research activities (e.g. [15,17]) and various personal experiences. More than 50 questions were generated in a brainstorm and through an iterative process these were generalized and refined using affinity diagramming.

Discussion
In the following the possible academic and practical contribution from the usability perspective framework are discussed and the fit between the framework and software development systems analyzed.

Contribution
A significant amount of research has looked into understanding, developing, and deploying usability work techniques, but a final breakthrough has not yet been made. One flaw is the lacking understanding of the influence organizations and key stakeholders have –and must have– on usability work and software usability. This constitutes a number of shortcomings in current HCI research, which we aim to address.

First, it is not well understood how introducing usability work informs the organizational balance [2] in software development projects. Recent research suggests that introducing usability in an organization imply significant risks [10]. By continuously involving stakeholders in the usability perspective framework, we aim at building coherent understandings and realistic expectations in the organization and thereby reducing such risks.

Second, how to utilize business goals, visions, and values when working with usability are not clear. Business stakeholders are vision carriers and need to be involved. Lack of business insights can lead usability experts to report issues not corresponding with business goals and visions [1,8]. We suggest involving key business stakeholders in the usability perspective framework, both in order to increase the business value of usability work, but also to address the organizational risks mentioned above.

Third, understanding and satisfying developers' needs in order to increase the practical value of usability work have recently been explored [4,6,11]. However such studies are rare and effects in practice have not yet been shown. Addressing the needs of the developers is central in the usability perspective framework to ensure that developers with their technical expertise can contribute to the development process.

Fourth, by maintaining focus on formalized usability work practices, HCI research fails to support organizations with less formal usability procedures even though they still develop useful and usable software products. Parts of our data suggest that more or less justified prejudices about usability work exist, which make organizations refrain from applying formal
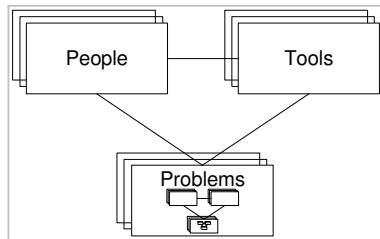
Figure 5 – The Symmetric Relation inspired by [9]. The figure illustrates how a number of people in the project organization system address the problem of designing and developing a new system of people, tools, and problems.
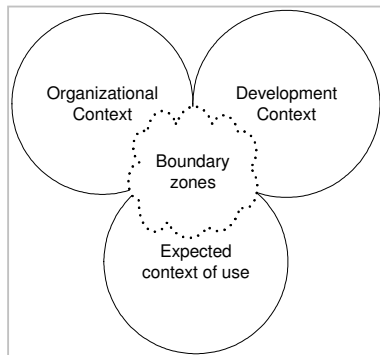

Figure 6 – The boundary zones are zones where stakeholders can develop coherent understandings of and realistic expectations to the development system and the developed system.

usability work. Instead they rely on less formalized work practices. To support both types of organizations the usability perspective framework aims at supporting both formal and informal usability related work.

Software systems development
One of the main challenges in software systems development is creating and maintaining coherence among stakeholders. Lack of coherence can make usability professionals failing to deliver useful feedback or software designers and developers failing to produce software products, with a proper fit to its intended users and expected usage. This challenge can be analyzed using The Symmetric Relation [9]. This relation describes how people, problems, and tools are tied together in a system and that viewing these elements in isolation only makes little sense. People familiar with different tools understand problems and their solutions differently. Developing software involves two sets of symmetric relations: That of the project organization, i.e. the large triangle in Figure 5, and that of the system under development, i.e. the small triangle in Figure 5 corresponding to the problem in the larger triangle. Stakeholders need a coherent understanding of both systems. In the project organization stakeholders come from different contexts and are familiar with different tools. Three contexts seem especially important when developing software: the organizational context, the development context, and the expected or intended context of use [17]. Stakeholders representing different contexts share and build knowledge about problems in different boundary zones [12] (Figure 6). A main challenge is to establish boundary zones where coherent understandings of the project organization and the system under development can be build and where important insights can be

gained. The usability perspective framework aims to facilitate that. The framework makes out a common tool for all stakeholders and thereby simplifies The Symmetric Relation. The five perspectives invite input from the three contexts inspired by the key questions, while the themes supports coherence in the outcome. How stakeholders in practise communicate efficiently within the boundary zones is a topic for further research.

Conclusion and future work
In this paper we have described the work-in-progress on developing a usability perspective framework that encourage contributions from various stakeholders in software development projects. Currently the framework is composed of 18 key questions related to five usability perspectives and three themes. We have described our understanding of software development systems and argued how the framework fits into such systems. Furthermore we have described the expected benefits of adding the usability perspective framework to the toolset practitioners rely on when developing useful and usable software.

At present focus is on finalizing the set of key questions and on exemplifying them in order to support practitioners' future use. To increase validity of the framework the examples will be based on everyday observations, which are independent of the research conducted so far. In the future we aim at conducting studies of practical use of the framework in order to validate our expectations and to further improve and develop it. Case studies and action research are considered as methods for that research. Quantitative studies of use of the framework are also considered, but a suitable research design has not yet been found.

References

[1]   Dumas, J. S., Molich, R., & Jeffries, R. Describing Usability Problems: Are We Sending the Right Message?, interactions, 11, 4 (2004), 24-29.

[2]   Gallivan, M. J. Changes in the Management of the Information Systems Organization: an Exploratory Study, Proc. the 1994 computer personnel research conference on Reinventing IS : managing information technology in changing organizations: managing information technology in changing organizations, ACM Press (1994), 65-77.

[3]   Hornbæk, K. & Frøkjær, E. Evaluating User Interfaces With Metaphors of Human Thinking, LNCS 2615, Springer-Verlag (2002), 486-507.

[4]   Hornbæk, K. & Frøkjær, E. Comparing Usability Problems and Redesign Proposals As Input to Practical Systems Development, Proc. CHI '05, ACM Press (2005), 391-400.

[5]   Hornbæk, K. & Law, E. L.-C. Meta-Analysis of Correlations Among Usability Measures, Proc. CHI '07, ACM Press (2007), 617-626.

[6]   Hornbæk, K. & Stage, J. Special Issue on The Interplay Between Usability Evaluations and User Interaction Design, International Journal of Human-Computer Interaction, 21, 5 (2006), 117-123.

[7]   IBM whitepaper. Igniting Innovation Through Business and IT Fusion, (2006).

[8]   Molich, R., Jeffries, R., & Dumas, J. S. Making Usability Recommendations Useful and Usable, Journal of Usability Studies, 2, 4 (2007), 162-179.

[9]   Naur, P. Computing: A Human Activity., ACM Press / Addison-Wesley, 1992.

[10] Rajanen, M. & Iivari, N. Usability Cost-Benefit Analysis: How Usability Became a Curse Word?, Proc. INTERACT 2007, Springer-Verlag Berlin Heidelberg (2007), 511-524.

[11] Redish, J., Bias, R. G., Bailey, R., Molich, R., Dumas, J., & Spool, J. M. Usability in Practice: Formative Usability Evaluations - Evolution and Revolution, Proc. CHI '02, ACM Press (2002), 885-890.

[12] Riisgaard Hansen, T. Strings of Experiments: Looking at the Design Process As a Set of Socio-Technical Experiments, *Proc. of the 9th conf. on Participatory design,* ACM (2006), 1-10.

[13] Uldall-Espersen, T. Exploring Multiple Usability Perspectives, Proc. INTERACT 2007, Springer-Verlag Berlin Heidelberg (2007), 628-632.

[14] Uldall-Espersen, T. Facilitating Usability Work Using Different Perspectives, Proc. the 4th COST294-MAUSE Open workshop on Downstream Utility (November 6th, 2007)

[15] Uldall-Espersen, T., Blandford, A., Jokela, T., & Frøkjær, E. Increasing the Impact of Usability Work in Software Development, *Proc. CHI '07*, ACM Press (2007), 2873-2876.

[16] Uldall-Espersen, T. & Frøkjær, E. Usability and Software Development: Roles of the Stakeholders, Proc. HCI International 2007, Springer-Verlag Berlin Heidelberg (2007), 642-651.

[17] Uldall-Espersen, T., Frøkjær, E., & Hornbæk, K. Tracing Impact in a Usability Improvement Process, Interacting with Computers, 20, 1 (2008), 48 - 63

[18] Vicente, K. The Human Factor: Revolutionizing the Way People Live With Technology, Routledge, 2004.

Increasing the Impact of Usability Work in Software Development

Tobias Uldall-Espersen

Department of Computing
University of Copenhagen
Universitetsparken 1
DK-2100 Copenhagen
Denmark
tobiasue@diku.dk

Timo Jokela

Department of Information
Processing Science,
University of Oulu
P.O. Box 3000
90014 University of Oulu
Finland
timo.jokela@oulu.fi

Ann Blandford

UCL Interaction Centre
University College London
Remax House, 31-32 Alfred Place
London WC1E 7DP,
United Kingdom
A.Blandford@ucl.ac.uk

Erik Frøkjær

Department of Computing
University of Copenhagen
Universitetsparken 1
DK-2100 Copenhagen
Denmark
erikf@diku.dk

Abstract

A key challenge when producing usable and useful software is the lack of impact of usability work on software development. We aim at developing a more coherent and realistic understanding of this challenge and the possibilities of how to increase the impact from usability work when developing high quality software products. We present a workshop gathering usability practitioners and researchers in order to establish and thoroughly discuss a corpus of case studies covering a broad range of usability practices in software development. The result of the workshop will be summarized in a conference paper and an international recognized scientific publisher will publish the case studies.

Keywords

Usability, Case Study, Software Engineering, Software Quality, Organizational influence, Usability Requirement Management

ACM Classification Keywords

D2 Software Engineering - D2.1 Requirements/ Specifications, D2.9 Management; H5 Information Interfaces And Presentation - H5.2 User Interfaces, H5.3 Group and Organization Interfaces; K6 Management Of Computing And Information Systems - K6.3 Software Management, K6.4 System Management

Introduction

Despite at least 20 years of research into usability engineering, a significant gap between usability work and software development persists, causing a lack of impact from usability work on software development. This lack of impact is a key challenge when developing high quality software; a challenge that current research has failed to address. HCI-researchers have frequently suggested that research should take into account the work of usability practitioners when doing applied usability research ([8][12][3][2][14][4]). As Cooke & Mings put it: "Ideally, academic researchers should inform workplace practices and workplace practices should inform academic research and teaching" ([4] p. 296). The suggestions seem to have had limited impact on the research communities. For example, among the usability studies listed in comprehensive reviews like Gray & Salzman [6] and Hartson et al. [9], none is rooted in real-life practices.

As the use of technology has spread an increasing number of people suffer from using products with poor usability. Information technologies have through Internet and information appliances (e.g. mobile phones and hand held computers) become important and sophisticated daily tools for hundreds of millions of people. As a consequence the scope of usability work is changing to cope with these new technologies and the conditions under which they are used. Practitioners rely on traditional and new usability work practices when trying to impact software development, and an increasing number of software development projects make use of usability work in an attempt to improve the quality of the software produced. Vredenburg et al. reports that the techniques most commonly taken up in industry are various forms of usability evaluation [13],

and Forrester Research reports that the highest single priority of a sample of 212 e-commerce vendors was improved usability [7]. Attempts to integrate usability work into software development is difficult and challenging [10] and in many cases suffering from a lack of impact on software product usability. This lack of impact affects an increasing number of people and has severe consequences including unsuitable user interface designs and limited support for the core work activities of users, and lack of productivity increases despite heavy investments in information technology ([5][1][11]).

We present a workshop bringing practitioners and researchers together providing an opportunity to analyse and bridge the gap between usability work and software development. The workshop will be based on experiences from practical usability work and analysing the conditions that influences the impact on software development will be the main focus.

Workshop goals

The workshop's goals are:

(1) Establishing and thoroughly discussing a corpus of case studies covering a broad range of experiences with usability work in software development, where a significant impact on software product usability has been made, or where an expected impact was not realised. The corpus should serve to develop a more coherent and realistic understanding of the conditions that influence the impact from usability work on software development and support exploration of the possibilities to increase the impact when developing high quality software products. The workshop is intended organized as a mutual learning experience,

supporting participants in developing ideas and thoughts, both within the industry and the research community.

(2) Collecting and publishing a corpus of high quality case studies of rewarding usability work to be used in future research, in teaching and as inspiration to practitioners. The corpus of the case studies is expected to be useful to (a) usability practitioners and software developers as a source of ideas of how to integrate usability work effectively in software development, (b) researchers who seek a more realistic understanding of outstanding research questions and key challenges in usability work in industry, (c) students of HCI courses who want to gain a more complete understanding of how to do usability work in practice.

Workshop issues
We propose four themes and a case study template as a starting point when describing the cases and the participants are encouraged to add additional topics that are relevant to understanding the cases. The goal is not to come up with a simple prescription ("Given challenge X, solution Y is recommended") but rather to provide enough structure that contrasting approaches can be explored side by side, while at the same time providing the reader with enough detail to let the reader think about how the particular details of the case influenced its dynamics.

The four themes
We suggest the following four themes as starting points for the case studies. Each theme points to conditions that could influence the achieved impact on software development and will form a basis for discussions.

Theme 1: "Handling usability as a quality issue in a software development process". Usability of software is one of a number of software quality attributes. Therefore, it seems useful to explore how usability is handled as a part of the software quality management process when designing software and how that influences software development. *Theme 2: "Integrating usability work in a software development process"*. Two often-mentioned concerns are that usability work is conducted too late in the software development process and that the results from usability evaluations are hard to utilize in software development. Integrating usability work further in software development could diminish such concerns and increase the subsequent impact on software development. *Theme 3: "Situating usability work in an organizational context"*. Usability work requires organizational knowledge, can depend on organizational involvement, and can motivate organizational changes. Therefore it seems likely that the organizational context could influence how usability work is conducted and how it impact software development. *Theme 4: "Planning, conducting, utilizing and evaluating usability work"*. Usability work could be organized as a coherent process were requirements are specified, usability evaluations are planned and carried out, evaluation results are utilized in software development, and fulfilment of the requirements validated. Such a structured process could signal seriousness and imply acceptance, and influence the impact on ongoing software development.

The case study template
The template identifies what central issues any case study should address. The participants are encouraged to add additional topics that are relevant to understanding the case. The full template is available

on www.diku.dk/infosys/chiworkshop/. The structure of the template looks like this: Section 1: "Setting the stage". This section should describe the industrial context of the software development project, the motivation for adopting usability work and the expected outcome. Section 2: "Description of the case study". The description of the case study should include a description of the plan and the conducted activities. Furthermore it should describe measurement procedures, contributions from key persons, realized impact on the software development project, and the key challenges in the conducted work. Section 3: "Different peoples perspective on the usability work in the case". This section should encourage the participants to explore and describe how other key persons experienced the outcome of the conducted usability work. Section 4: "Lessons learned". This section should describe the important lessons learn including which activities that would be repeated in a future project and what would be done differently.

References

[1]   Brooks, F. P. (1987), "No Silver Bullet: Essence and Accidents of Software Engineering", IEEE Computer, 20, 4, 10-19.

[2]   Bellotti, V., Buckingham Shum, S., MacLean, A., & Hammond, N. (1995). Multidisciplinary modelling in HCI design, in theory and in practice. Proceedings of CHI 1995. ACM Press. 429 – 436.

[3]   Buckingham Shum, S. & Hammond, N. (1994). Transferring HCI Modelling and Design Techniques to Practitioners: A Framework and Empirical Work. Proceedings of BCS HCI'94, 21 – 36.

[4]   Cooke, L. and Mings, S. "Connecting Usability Education and Research With Industry Needs and Practices." IEEE Transactions on Professional Communication 48.3 (2005): 296-312

[5]   Gould, J. D. & Lewis, C. (1985), "Design for usability: Key principles and what designers think", Communications of the ACM, 28, 3, 300-311.

[6]   Gray, W. D. & Salzman, M. C. (1998), "Damaged merchandise? A review of experiments that compare usability evaluation methods", Human-Computer Interaction, 13, 3, 203-261.

[7]   Forrester Research (2003), "2004 Outlook For eCommerce Tech Spending", Business View and IT View Data Overview, www.forrester.com.

[8]   Hammond, N., Jorgensen, A., MacLean, A., Barnard, P. & Long, J. (1983). Design Practice and Interface Usability: Evidence from Interviews with Designers. Proceedings of the SIGCHI, 40 – 44.

[9]   Hartson, H. R., Andre, T. S., & Williges, R. C. (2001), "Criteria for evaluating usability evaluation methods", International Journal of Human-Computer Interaction, 13, 4, 373-410.

[10] Juristo, N., Windl, H., & Constantine, L. (2001), "Introducing usability", IEEE Software, 20-21

[11] Landauer, T. K. (1995), The trouble with computers, Cambridge, MA: MIT Press.

[12] Rosson, M., Maass, S. & Kellogg, W. (1988). The Designer as User: Building requirements for the design tools from design practice. Communications of the ACM, 31, 11, 1288 – 1298.

[13] Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. (2002), "A Survey of User-Centered Design Practice", In Proceedings of CHI 2002 Proceedings of ACM Conference on Human Factors in Computing Systems, 472-478, New York NY:ACM Press.

[14] Wixon, D. (2003), "Evaluating usability methods: Why the Current Literature fails the Practitioner", interactions, 10(4), 29-34.

Book proposal
Real life usability work
Tobias Uldall-Espersen, Ann Blandford, Erik Frøkjær, and Timo Jokela (Editors)

Suggested titles:

Real life usability work
Increasing the impact of usability work in software development.
Usability work in software development
User-Centred Design in real life

Introduction to the subject area

Researchers and practitioners concur that a key challenge in developing high-quality software is to design user interfaces that support users in accomplishing their work in an effective, time-efficient, and satisfactory manner (Gould & Lewis 1985; Boehm 1991). This challenge has implied an increasing focus on human-computer interaction and usability engineering. For the past 20 years, these disciplines have refined techniques for assessing and improving the usability of software (Lewis 1982; Nielsen 1993; Uldall-Espersen et al. 2008).

Despite several years of research into usability engineering, there is still a lack of understanding of how to efficiently conduct practical usability work. Currently, an increasing number of development projects employ usability-engineering techniques in an attempt to improve the quality of the software produced. As examples, improved usability was reported as the highest single priority of a sample of 212 e-commerce vendors (Forrester Research 2004). The techniques most commonly taken up in industry are various forms of usability evaluation (Vredenburg et al. 2002). The cost-benefit arguments for including usability evaluation in software development activities are strong (Bias & Mayhew 1994, 2005), and notable products have benefited from extensive usability evaluations. While these evaluations often help identifying an exhaustive list of usability problems with the software, they typically have a very limited impact on the subsequent development activities. This lack of impact has severe consequences including unsuitable user interface designs, limited support for the core work activities, and lack of productivity gains despite heavy investments in information technology (Gould & Lewis 1985; Brooks 1987; Landauer 1995).

Current research fails to address why practical usability work only have a limited impact on real-life software development. There are three main reasons for this. First, most research studies focus on generating reports that list usability problems with a certain piece of software. Recently, Wixon (2003) argued that such studies have limited relevance in practical software development, because 'it ignores that problems should be fixed, not just found'. Second, only very few studies of usability evaluation are conducted in real industrial settings. Among the studies listed in comprehensive reviews (e.g. Gray & Salzman 1998; Hartson et al. 2001), none have been conducted in the context of software systems development. Thus, those studies fail to take into account the complexity of real-life software development and the practical realities in which the results of usability evaluations are to be assessed and used. Third, research is only beginning to address how stakeholders understand and contributes to usability work (Hornbæk & Frøkjær 2004; Dumas et al. 2004; Uldall-Espersen & Frøkjær 2007). Therefore, it is not clear how different stakeholders should be involved in usability work when developing software, nor is it clear which is the most effective content to present to stakeholders so as to facilitate timely and accurate amelioration of usability issues.

Overall objectives and mission of the book

The ability to produce rapid and lasting impacts from usability work on software development processes, products, and organizations are crucial when developing high quality software products. Through a corpus of practical case studies presented at a recent workshop, this book offers a broad set of experiences from people who aim at creating such impacts. Each story in the book contains experiences important to the involved stakeholders such as usability experts, software engineers, or project managers.

Book proposal
Real life usability work
Tobias Uldall-Espersen, Ann Blandford, Erik Frøkjær, and Timo Jokela (Editors)

The core of this book is a set of rich descriptions of case studies from practical software development projects, including description of contexts, organizations, and processes. Both researchers and practitioners authored the case studies. The case studies describe both novel and more traditional approaches to usability related work. For example, can increasing the salary of a project team meeting usability goals increase usability and how can human centred design inform the design of technical infrastructures? How can usability be introduced bottom up, can a lasting impact from bottom up usability work be expected, and how can real-life problem solving processes be transferred to the software under development?

Among the cases we also have reports from usability work in large-scale IT-systems development projects. The cases cover systems development in many different settings under many different organizational conditions, both industrial and scientific. One may assume the work conditions across such projects are really different, but is this in fact true? Another important theme covered in the book is usability and offshore software development. We present two cases where challenges and possible solutions regarding offshore development are discussed. Furthermore we present some cases where new methodological approaches are introduced to practitioners.

This book is not just a book sharing HCI-practitioners experiences with other HCI-practitioners. Usability work concerns all stakeholders in the software development processes and other stakeholders have also shared their experiences with usability work in practical software development. The book is partly based on the result of a workshop at the CHI 2007 conference in San Jose, California. The title of the workshop was 'Increasing the impact of usability work in software development.' and more information can be found on http://www.diku.dk/infosys/chiworkshop/.

Audience

By creating a corpus of case studies we aim at describing real-life situations where usability work produced a significant impact on software or where an expected impact did not occur. It is believed that such a set of case studies, through their rich descriptions covering a broad range of usability work processes, will be useful to usability practitioners, HCI researchers and HCI students as well as to software developers and managers of IT projects who are concerned about making usable IT solutions for their customers.

Contribution

The book will include a set of rich case studies describing practical usability work in a variety of different software development contexts and from various perspectives. The cases cover a wide range of approaches to usability work at different levels of maturity. It includes both traditional and emerging questions and solutions regarding development of technology used every day. At present, such a collection does not exist.

The book will present a set of case studies enabling the reader to reflect and learn from the various cases. We do not see the set of cases as successes or failures, but we believe that every case has important and realistic experiences to pass on to the readers. This includes both experiences with different usability work techniques and experiences with different stakeholders reactions to the conducted usability work.

Existing publications (Competitors):

Michael E. Wiklund, Usability in Practice, Academic Press, 1994, ISBN 0-12-751250-0, 610 pages, sold out
Strengths:
This is probably the book coming closest to the book we propose here. The book contains 17 cases describing usability work conducted in a number of different organisations. Among the organizations are

a number of large multinational companies, some developing consumer hardware and software products, and some developing and utilizing major IT-solutions to support their core businesses.

Weaknesses:
The weakness of the book is its age and the fact that it is sold out. The book is 200 – 300 pages longer than we expect ours to be and it is not very structured. It clearly needs an update reflecting the large changes in IT-usage and the progress of usability work and usability research within the last 15 years.

Randolph G. Bias & Deborah J. Mayhew, Cost-Justifying Usability, Elsevier, 2005, ISBN:0-12-095811-2, 660 pages, £38
Strengths:
By focusing on cost-justification, this book presents a clear perception of a measurable impact from usability work in software development projects. Furthermore, it exemplifies how to advocate for usability work in organizations when facing reservations from management. The book has 22 chapters and most of them are written by establish experts within the practical HCI field. The book describes various cases and other relevant aspects of practical and theoretical usability work.

Weaknesses:
By maintaining a strong focus on cost-justification, the book pays less attention to other highly relevant considerations when doing usability work, such as the development of organizations and the creation of new possibilities through innovative IT usage. Furthermore, usability work processes seem to be influenced by a significant set of individual differences, which makes it hard to make valid generalizations. A number of the presented cases are only briefly described and a lot more details are needed to really understand and learn from the conducted usability work of the cases.

Carol Righi & Janice James, User-Centered Design Stories, Elsevier, 2007, 0-12-370608-4, 535 pages, £35
Strengths:
This book contains 22 easy-to-read User-Centred Design Stories following the Harvard Case study method. The majority of the authors are experienced practitioners. Using the Harvard Case study method gives good insight into the interplay between the various stakeholders.

Weaknesses:
Following the Harvard Case study method makes the book longer than necessary and the chapters less focused on what really matters to us – the experience with practical usability work. The primary audience is new practitioners; and more experienced practitioners and researchers will have a limited benefit from reading the book due to the general lack of technical details in the cases.

Tentative table of contents:

Introduction and overview:
1. Introduction
2. Usability practitioners' experiences of the impact of usability work in software development

User centred design
3. Money for better usability. A case study on the development of user interface for a mobile phone
4. Innovation in testing; Innovation in design
5. The Conversion Challenge; Fix flat sales
6. Integrating usability work in a software development process: A case study on Claims Analysis
7. Impacts of Classification of Usability Problems (CUP) on system redesign
8. Has Usability Become a Curse Word? - An Interpretive Case Study on Usability Cost-Benefit Considerations

9. Effects-Driven IT Development: Managing Change by User-Centred Design

UCD in large enterprise systems
10. User centred design and development of financial operational software tools
11. How usability work informed development of an insurance sales system
12. The impact of usability on large science projects
13. Usability in e-Science: The DiaMoND case study

UCD and the development of technical platforms
14. The Cross-Functional Challenge of Usability Work in New Packaged Software Development
15. Extending human-centred design to the design of technical infra structures

UCD in offshore development
16. HCI + SE Integration Case Studies from Offshore development Projects
17. Arriving at Shared Perspectives on Software through User-Centred Design Processes
18. Offshoring Usability

Summary
19. Summery chapter

About the authors:

A preliminary synopsis for the book is attached.

Timetable:

Conducted activities:
- Holding the CHI workshop 'Increasing the impact of usability work in software development.'
- Preparing an initial book proposal
- Initial commitment from the authors
- Receiving and reviewing workshop papers and abstracts
- Submitting the book proposal to IGI

Forthcoming activities:

Relative timeline / Activity dead line	Activity duration	Activity
Starting at week 0		Proposal accepted by IGI Global
By end of week 2	2 weeks	Obtain final commitment from the authors
By end of week 6	4 weeks	Writing initial versions of full chapters
By end of week 10	6 weeks	Reviews of full chapters
By end of week 16	4 weeks	Writing final versions of full chapters
By end of week 20	4 weeks	Finalizing the introduction and summary chapter and putting the book together.

Book proposal
Real life usability work
Tobias Uldall-Espersen, Ann Blandford, Erik Frøkjær, and Timo Jokela (Editors)

Editors:

Tobias Uldall-Espersen (Corresponding Editor)
Department of Computing
University of Copenhagen
Universitetsparken 1
DK-2100 Copenhagen Ø
Denmark
tobiasue@diku.dk
Phone/Fax: +45 35321451/ +45 35321401

Professor Ann Blandford
UCL Interaction Centre
University College London
Remax House, 31-32 Alfred Place
London WC1E 7DP,
United Kingdom
A.Blandford@ucl.ac.uk
Phone/Fax: +44 7679 5288 / +44 7679 5295

Associate professor Erik Frøkjær
Department of Computing
University of Copenhagen
Universitetsparken 1
DK-2100 Copenhagen Ø
Denmark
erikf@diku.dk
Phone/Fax: +45 35321456/+45 35321401

Timo Jokela, PhD
Department of Information Processing Science,
University of Oulu & Joticon Ltd.
P.O. Box 3000
90014 University of Oulu
Finland
timo.jokela@oulu.fi, timo.jokela@joticon.fi
Phone/Fax: +358 40 5118250, +358 8 5531890

The editor's CVs are attached.

References:

Bias, R. G. & Mayhew, D. J. (1994), *Cost-justifying usability,* San Francisco, CA: Morgan Kaufman.

Bias, R. G. & Mayhew, D. J. (2005), *Cost-justifying usability,* San Francisco, CA: Morgan Kaufman, 2nd edition.

Boehm, B. (1991), "Software Risk management: Principles and Practices", *IEEE Software,* January, 1991, 32-41.

Brooks, F. P. (1987), "No Silver Bullet: Essence and Accidents of Software Engineering", *IEEE Computer,* 20, 4, 10-19.

Dumas, J., Molich, R., & Jefferies, R. (2004), "Describing Usability Problems: Are we sending the right message?", *interactions,* 4, 24-29.

Forrester Research (2004), "2004 Outlook For eCommerce Tech Spending".

Gould, J. D. & Lewis, C. (1985), "Design for usability: Key principles and what designers think", *Communications of the ACM,* 28, 3, 300-311.

Gray, W. D. & Salzman, M. C. (1998), "Damaged merchandise? A review of experiments that compare usability evaluation methods", *Human-Computer Interaction,* 13, 3, 203-261.

Hartson, H. R., Andre, T. S., & Williges, R. C. (2001), "Criteria for evaluating usability evaluation methods", *International Journal of Human-Computer Interaction,* 13, 4, 373-410.

Hornbæk, K. & Frøkjær, E. (2004), "Usability Inspection by Metaphors of Human Thinking Compared to Heuristic Evaluation", *International Journal of Human-Computer Interaction,* 17, 3, 357-374.

Landauer, T. K. (1995), *The trouble with computers,* Cambridge, MA: MIT Press.

Lewis, C. (1982), "Using the "thinking-aloud" method in cognitive interface design", Research Report RC9265.

Nielsen, J. (1993), *Usability Engineering,* San Diego CA: Academic Press.

Uldall-Espersen, T. & Frøkjær, E. (2007), Usability and Software Development: Roles of the Stakeholders, In J. Jacko (Ed.): *Human-Computer Interaction*, Part I, HCII 2007, LNCS 4550, pp. 642–651, Springer-Verlag Berlin Heidelberg.

Uldall-Espersen, T., Frøkjær, E., & Hornbæk, K. (2008) "Tracing Impact in a Usability Improvement Process", *Interacting with Computers*, 20, 1, 48-63.

Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. (2002), "A Survey of User-Centered Design Practice", *In Proceedings of CHI 2002 Proceedings of ACM Conference on Human Factors in Computing Systems,* 472-478, New York NY:ACM Press.

Wixon, D. (2003), "Evaluating usability methods: Why the Current Literature fails the Practitioner", *interactions,* 10(4), 29-34.

Book synopsis
Real life usability work
Tobias Uldall-Espersen, Ann Blandford, Erik Frøkjær, and Timo Jokela (Editors)

Table of contents

Introduction and overview:

User centred design

UCD in large enterprise systems

UCD and the development of technical platforms

UCD in offshore development

Summary

About the contributors

Book synopsis
Real life usability work
Tobias Uldall-Espersen, Ann Blandford, Erik Frøkjær, and Timo Jokela (Editors)

Introduction and overview:
1. Introduction
Tobias Uldall-Espersen, Ann Blandford, Erik Frøkjær and Timo Jokela

2. Usability practitioners' experiences of the impact of usability work in software development
Jenni Anttonen & Hannakaisa Isomäki

A survey measuring the impact of usability work in practice revealed that usability practitioners' methodological competence and education are on a good level. However, nearly half of the respondents were not happy with their opportunities to have an impact. The greatest factors impeding their work were that there is too little time and other resources to do the work, usability activities are integrated into projects too late, the significance and scope of usability work is not understood, and the status of usability work as a part of software projects is weak. The respondents felt that to increase their possibilities to make an impact, they need to develop their own skills in communication and management, and to gain more technical and economic knowledge. Also, increasing management support was perceived to be essential.

User centred design
3. Money for better usability. A case study on the development of user interface for a mobile phone
Timo Jokela

A user interface development project is described where usability had an impact on the amount of project incentives (extra salary for a successful project) that was paid to the project team. The extent to which the (quantitative) targets for usability were achieved formed one basis for calculating the amount of incentives. This approach - the better the usability targets are achieved, the more money is paid - made the project team search for a user interface design solution of high-level usability. As a result, the target levels were achieved and even exceeded, and incentives paid to the project team accordingly. The conclusion is that 'money for usability' is an effective means for increasing the impact of usability work.

4. Innovation in testing; Innovation in design
Barbara Hernandez

After adopting usability testing and experience design several years ago, TechSmith and product team members have come to realize how user experience impacts sales and adoption rates. User Experience team members are now typically embedded in agile product development teams at the company. In looking at how people might work with new functionality, questions arose; the answers would drive design. Existing software did not hold the answers and has existing usability issues we wanted to avoid. Existing software also lacked the focus we needed for the product. The User Experience team felt that understanding how people dealt with the problems in the real world, instead of on the computer, would show how to best solve the problem. This case study outlines how the experience test was constructed, the tools that were used and how the team incorporated those results into design to create an innovative solution.

5. **The Conversion Challenge; Fix flat sales**
   Barbara Hernandez
   TechSmith, a "trial ware" software company was faced with a problem – flat sales for a mature product. All we heard from our thousands and thousands of users is how they couldn't live without it and how it has changed the way they communicate with people. So, why were sales flat? To answer this question, the focus became the user experience. This case study looks at how a product's user experience, not just practical usability, can have a profound effect on both the adoption and sustained use of a software application. A full developer was dedicated to this goal. Marketing and business objectives were considered along with experience goals when designing the new experience for the product. Our approach included user experience testing – rather than pure usability testing to ensure we were achieving those goals. The end result was a 40% increase in sales.

6. **Integrating usability work in a software development process: A case study on Claims Analysis**
   Ann Blandford and Suzette Keith
   We present a case study in which we set out to adapt Claims Analysis specifically to suit the design of Digital Library systems. The work involved embellishing CA with an explicit interaction model and information seeking scenarios and expressing it in a form that was accessible to software developers with no background in usability work. While scenarios were valued by the development teams we worked with, Claims were not, being considered "too academic". The work highlighted different value systems and priorities of human factors specialists and software developers, including prioritizing user problems or design solutions, and different ways of thinking in terms of interactions and functions.

7. **Impacts of Classification of Usability Problems (CUP) on system redesign**
   Effie Lai-Chong Law, Sigurbjörg Gróa Vilbergsdóttir and Ebba Thora Hvannberg
   To enhance impacts of usability evaluation outcomes on system redesign, it is imperative to derive from such outcomes some systematic information that enables a development team to understand what problems are, to analyze why they occur, and to gain insights how to fix them - this is exactly the main function of a tool known as Classification of Usability Problems (CUP) scheme. Basically, CUP can be applied to outcomes of any empirical or analytic usability evaluation method (UEM). Presumably, results of CUP can not only foster developers' understanding about individual UPs, but also help them prioritize UPs for fixing as well as inspire them to create ideas for redesigning the system in question, and prevent the reoccurrence of UPs in other future projects. In this chapter, we first describe the CUP in detail, which comprise two major parts, Pre-CUP and Post-CUP, with each consisting of a set of attributes and associated values. Then, we present a review of existing defect classification systems (DCSs) being deployed in the field of software engineering and HCI, including User Action Framework (UAF), Orthogonal Defect Classification (ODC), Root Cause Analysis (RAC) and others, to compare and contrast their commonalities and differences and to discuss how CUP is distinct from them. Then we report several case studies about the application of CUP in different contexts to illustrate its scoping, usefulness, usability, reliability, and validity.

8. **Has Usability Become a Curse Word? - An Interpretive Case Study on Usability Cost-Benefit Considerations**
   Mikko Rajanen and Netta Iivari

This article contrasts usability cost-benefit analysis literature with an empirical case in industrial setting, in which usability cost-benefit considerations (along with other usability activities) resulted in usability becoming a curse word. An interpretive case study was carried out in a small-to-medium sized software development organization. The article describes how the company was introduced with usability activities (including usability cost-benefit considerations), and was thus expected to change their current practice. Empirical analysis reveals that clearly divergent meanings and motives were attached to usability and its cost-benefit analysis in the organization. Increased sales and reduced development costs were strongly emphasized as benefits of better usability. However, very surprising meanings were attached to them both. Furthermore, the development costs associated with better usability were the main failure factor of the whole usability improvement effort. Implications both for usability cost-benefit analysis theory and practice, and more generally for improving the position of usability work in software development organizations are discussed.

9. **Effects-Driven IT Development: Managing Change by User-Centred Design**
Morten Hertzum and Jesper Simonsen
To customers IT systems are means to an end. However, most IT projects are organized as though the IT system is an end in itself in that they are successfully completed when the vendor delivers an IT system with a specified functionality. This dissociates technical implementation from organizational implementation and implies that IT projects can be successful even if the developed systems never become used in ways that produce the ends desired by customers. Effects-driven IT development is an emerging UCD approach that strives to avoid this dissociation by a sustained focus on the effects desired from using a system and measurements that demonstrate the presence or absence of these effects during actual system use. The premise of effects-driven IT development is to establish a strategic partnership in which customer and vendor share the responsibility of providing IT solutions that provide specified usage effects. The intension is to capture the purpose of a system in terms of effects that are both measurable and meaningful to the customer. This can be seen as a supplement to, or as opposed to, a focus on IT functionality.
The paper argues the need for a UCD approach that joins technical and organizational implementation through a focus on the effects obtained by actually using a system, describes our emerging approach to effects-driven IT development, and reports our experiences from an evaluation in which parts of this approach were used in the configuration, trial use, and assessment of an electronic patient record (EPR) system for a stroke unit.

UCD in large enterprise systems
10. **User centred design and development of financial operational software tools**
Joachim Sander
This article describes a case study how to integrate a User Centered Design approach (UCD) most effectively into a complex software development procedure, like software development of Enterprise Resource Planning software (ERP). The process description and later analysis will contain a discussion of the UCD process. Furthermore it will focus on End User Involvement activities (EUI), how these needed to evolve to be most beneficially integrated into the project to support the project progress. Overall a special focus will be given to the iterative manner of the UCD Process and its resulting consequences.

11. How usability work informed development of an insurance sales system

Tobias Uldall-Espersen

Introducing usability work in an organization is not a trivial task, but can imply a number of short-term and long-term challenges. This interview-based case study describes the development of a new insurance sales application. Usability work was not a part of the formal project development plan or the company strategy, but was forced into the project by two key project members. The work was driven by their professional pride and it proved to be a critical success factor in the project. The development project was influenced by its wide organizational settings and the introduction of new innovative technologies, such as wireless communication, in the organization. The new technology changed current practices both within the development team and among the users. A customer-centered approach was applied to form the new sales processes and the users were heavily involved in the iterative development process. Despite the relative success of the conducted usability work, only a very limited impact on future software development projects was obtained.

12. The impact of usability on large science projects

Cecilia Aragon & Sarah Poon

Much of the discussion of the importance of usability to software development has been focused on commercial software. However, large scientific software projects can also greatly benefit from the application of usability engineering principles. This chapter describes software developed for astrophysicists studying supernovae with the goal of measuring the expansion history of the universe. By performing iterative software design and other usability engineering techniques throughout the project, we were successful in developing a supernova data catalog and workflow management tool that improved scientists' efficiency, situational awareness, and productivity. Special care was taken to involve the scientist users in all aspects of and at all stages of the design, implementation, and testing. Integrating usability design throughout the project had a significant impact on its success.

13. Usability in e-Science: The DiaMoND case study

Andrew Warr, Grace de la Flor, Marina Jirotka and Sharon Lloyd

E-Science, or cyber-infrastructure as it is known in the US, promotes a vision of large-scale, collaborative and multi-disciplinary research. It is believed that this vision will give rise to new forms of multi-institutional and multi-disciplinary science that will allow new and more complex scientific questions to be answered. Not only will this vision create a significant difference to the practices and products of quantitative science, but also to qualitative scientific approaches and practices. However, a number of challenges have emerged from researchers attempts to turn the vision into reality. Though e-Science technologies offer significant opportunities and benefits, the complex nature of the technologies developed and the corresponding infrastructure often create considerable challenges for researchers and practitioners who wish to take advantage of e-Science technologies in their working environment. If prospective e-Science tools and technologies are to be deployed into researchers' environments to support their everyday work, e-Science development must address future users' and organisations' needs and requirements if they are to realise their full potential. Thus, a key concern that we address in this chapter is that of usability in e-Science; facilitating the establishment and use of e-Science infrastructure and technologies in practitioners' and researchers' everyday work practices and organisational contexts. We will outline some of the challenges emerging from the experiences of participants attempting

to develop working and usable grid enabled systems. Drawing upon our analyses of in depth interviews and ethnographic observations we present the impact of usability in our case-study project, namely eDiaMoND. We discuss the major issues that have emerged when the eDiaMoND project stakeholders attempted to address users' needs and organisational contexts in development. We then consider these issues within the wider debate about uptake of e-Science technologies and conclude with some lessons learned and recommendations for future development.

UCD and the development of technical platforms

14. The Cross-Functional Challenge of Usability Work in New Packaged Software Development

Tonja Molin-Juustila

This chapter presents a case study of usability work in the context of developing packaged software applications. It will be shown how - within such a context - especially during the uncertain and dynamic early phases of very new software product, the impact of usability work faced the organizational challenge of cross-functional interaction. The paper provides practical experience from a concrete case of improving the status of usability work in one international software product company. The case provides better understanding of how usability work – within the context of packaged software development – is clearly a cross-functional issue. With a new product and new emerging business area the user needs and requirements for the new product are not very accurate. Unique requirements of specific pilot customers need to be generalized and negotiated together with many different interest groups (e.g. marketing, sales, support, consultants, even partner companies). In addition to software engineering - in order to impact new product development - usability work needs to be better integrated to the release-independent and on-going activities of other organizational functions as well.

15. Extending human-centred design to the design of technical infrastructures

Eija Kaasinen and Marketta Niemelä

Human-centred design as defined by the ISO 13407 standard is a well-established practise to the design of individual software applications. Many application features are, however, not defined only in the application itself but by the underlying technical infrastructures such as device platforms, databases, middleware and interface libraries that are typically fixed before the application development takes place. To increase the impact of usability work, human-centred design should be extended to the design of those technical infrastructures. In this chapter we introduce this design challenge and describe a case study where we have applied human-centred design to the design of a ubiquitous computing architecture. Key elements of the design approach are extensive scenario work that aims to foresee future applications, user evaluation of scenarios and proof-of-concept prototypes, identification of typical usage patterns that repeat from one application to another as well as thorough, multidisciplinary analysis of usage patterns and user evaluation results to identify implications to the technical infrastructures.

UCD in offshore development

16. HCI + SE Integration Case Studies from Offshore development Projects

Anirudha Josh

Though the processes of human-computer interaction (HCI) design and Software Engineering (SE) affect each other deeply, enough has not been done to integrate them. We

reviewed and participated in several case studies from the Indian IT industry to study the integration of human-computer interaction (HCI) design into software development by process-conscious Indian software vendors. Several problems seem to occur because HCI skills are either not used, or are not used early enough in a project or when the HCI professional lacked process support to carry out all HCI activities in the project. In the one case where HCI professionals were indeed used early and with a multi-disciplinary team, the results were positive. The case studies point to a greater need to integrate HCI into existing SE process models and establishing benchmarks that are widely acceptable. We propose modified SE process models by integrating important HCI design activities and deliverables while retaining the essential structure of the established SE models.

17. Arriving at Shared Perspectives on Software through User-Centred Design Processes
Robert Gillham
User-centred design is often little more than an afterthought in traditional software development processes. Development projects meanwhile often suffer from poor communication between stakeholders and a lack of shared vision. This paper describe a case study where user centred design was brought to the forefront of a development effort to address both issues.

18. Offshore Usability
Eric Schaffer
This chapter gives insights into strategies, challenges, and tactics for completing usability work in an offshore environment. We will share HFI's long experience in applying offshore staff from India to the development of applications for Western countries. We have proven that this can be quite successful; in fact we feel this is the likely best-practice model for getting usability work done on a large scale. But the road to success is long and exacting. This article will help companies assess their opportunities in offshore usability work, or ask the right questions if they are offered an offshore facility.

Summary
19. Summary chapter
An invitation has been sent to Dominic Furniss (UCL) to co-author a summary chapter.

About the contributors

How usability work informed development of an insurance sales system

Tobias Uldall-Espersen

Department of Computing,

University of Copenhagen

Universitetsparken 1, DK-2100 Copenhagen

tobiasue@diku.dk

ABSTRACT

This paper reports a case study of a software development project where an insurance sales system was developed. The organization under study had only limited experience with usability, and usability work was introduced bottom up in the project. Two key people forced usability work into the development process and usability work became a key success factor. The usability work was comprehensive and became a significant and integrated part of the practical development work, and it informed both the end product quality and the organization in which the system was implemented. The case study is based on interviews with six key stakeholders in the project.

SETTING THE STAGE

Introduction

This case study describes the implementation of a new insurance sales system called Absalon in a Danish insurance company. Usability work was not a part of the formal project development plan or the company strategy, but was forced into the project by two key project members. The work was driven by their professional pride and it proved to be a critical success factor in the project. The development project was influenced by the facts that it concerned a wide range of the organization and that it introduced new innovative technologies, such as wireless communication, in the organization. The new technology changed current practices both within the development team and among the users. A customer-centered approach was applied to form the new sales processes and the users and other stakeholders were heavily involved in the iterative development process.

The interviews

This case description is based on semi-structured interviews [3] with six key stakeholders. The interviews were conducted within two months after the project ended. Each interview was scheduled to last 60-90 minutes; the same interview guide was used at all six interviews. A total of about 8 hours conversation were recorded and analyzed using open coding [7]. The transcripts were coded and emergent themes were found and will be presented in the following. A number of important themes are illustrated using quotations from the interviews. The quotations were originally in Danish and we have not tried to make an exact translation word for word. Instead our aim has been to translate the meaning of what was said, while maintaining the living atmosphere of the interview and the details in the discussions. The remaining part of this section gives a short introduction to each of the six interviewed stakeholders.

- *The project leader.* The project leader was brought in from outside the organization about half way into the project. At that time he had more than ten years experience with development of insurance sales and service systems. His main responsibility was to manage the IT-resources and to report to the steering committee. He worked closely with the business project leader and a technical coordinator, both being experienced in the organization.

- *The business project leader.* The business project leader was responsible for controlling the business resources, for the deliveries of business relevant input from the organization to the project team, such as insurance product descriptions and specifications, and at the end of the project for the final test of the system. Throughout the project a large number of business stakeholder from various parts of the organization were involved and the business project leader was responsible for keeping them up to date on the project. Furthermore he represented the department holding the sales processes and thereby the department who would be the future owner of the system.

- *A software developer.* This software developer also served as secondary IT-architect in the project. He was an experienced Java programmer and he was employed around the time when the programming work was starting up. He considered his own role as partly being the "coding guy", but he also took part in various important discussions about how specific parts of the system should be designed and implemented.

- *A business analyst.* This business analyst was initially representing the general insurance department, but throughout the project she became more involved in the development project and she was very

much the person driving the user involvement in the project. She was first involved in some basic analyses and in the requirement specification process. Later she was asked to design the system tests and the user tests, and she considered the user tests as being her "darling". At the end of the project she was made responsible for the user education program as well.

- *The designer.* The designer was the person responsible for designing the user interface and the information architecture. Among his colleagues he was referred to as the "usability guy", the "designer", the "information architect" and so on, but he described himself as the person being the liaison between the business organization, the development team, and individual users.

- *A user.* This user was involved in the specification phase, the test work, and the ongoing user education. He had participated in the development of the previous sales system in the organization called MobilSalg (English: MobileSales), and he had a thorough knowledge about the users' work practices. He was continuously involved in testing different parts of the system, and he found it important that he was able to stand up for the system when educating new users.

System, Users and Goals

The system under development was an online sales tool built on a new wireless IT-platform utilizing 3G mobile technologies. The project leader described the system this way:

Project leader: Absalon is simply a sales application build on our existing policy systems. We have our TIA system [an application for insurance administration] handling general insurance and a system running at an AS400 platform handling retirement insurance. Absalon is a sales tool; built on these systems using its own Java platform in such a way that it reuses all business logic from the existing systems. The system has about 500 users, typically insurance agents working with laptops and wireless connections at the customers' homes, people in the central customer service centers in Denmark, and a group of retirement specialists working in the local branches around the country. (**Quotation 1**)

The three groups of users were very different and both the contexts of use and the sales processes needed to be understood well. Furthermore there was a shift in roles among the users; from being perceived as regular sales people to being perceived as responsible customer advisers offering a broader and more integrated portfolio of financial services:

Developer: The background is that we want a more integrated sale. We have three product lines in the company. One is general insurance, which is the largest. Another one is bank and the third is retirement and life insurance. You want to encourage insurance agents, customer service personnel, and bank employees to sell products not only from their own product line. (**Quotation 2**)

This shift was based on two conditions. First, it was well known that the customers' loyalty increased when they had products from more than one product line, and increasing the number of such double

customers would increase the revenue. Second, there was an increasing focus on being responsible in the sector, that is: not leaving the customers with obvious need for special insurances, such as life or retirement insurance.

The system replaced a traditional client-based stand-alone system – MobilSalg – that was appreciated by its users but did not meet modern technical requirement such as providing online access to customer data. The system was also hard and expensive to maintain. The aim of the development project was to develop a modern system with the same functionality as MobilSalg. According to the project leader, seeing it as a replacement process was central in the project, and MobilSalg was an important starting point:

Project leader: An overall project statement was that with Absalon we should build a new system just as good as MobilSalg, which they were really satisfied with; yet measured on the whole sales process from the beginning to the end, that is: including aftercare [Aftercare refers to the considerable work done to identify and fix errors that prevent sold insurance policies in being created in the main insurance systems]. MobilSalg was very flexible and they could type in whatever they wanted; problems first occurred when data was uploaded to the core systems. Now we build on another structure and the users get lots of problems and things they need to decide on. It is obvious that some of this work will be experienced as being more difficult. On the other hand you get a higher quality in the end; that was the overall judgment. But MobilSalg has been important, you must realize that they were very fond of it out there, and it has also served as a frame of reference, both when choosing things to carry on from MobilSalg and when choosing things not to carry on. On the other hand, with MobilSalg as frame of reference we were also less apt to add new functionality. Then we were just as well off – or not more badly off than today. In that way MobilSalg meant a lot. (**Quotation 3**)

The above quotation also indicates how organizational issues such as "looking at the whole sales process" to some extent overrules the fact that the user are satisfied with the system. The developer follows up on some of these thoughts when explaining how the new system is expected to benefit the overall organization:

Developer: Today we have an off-line application where the user uploads data to the backend-systems. We are making that upload unnecessary. Errors happen and people have to check what is going on and find the cause of the error. We remove that work in Absalon, because they get the error right away. I do not think we will get rid of the problems, but we will reduce them so the insurance agents and the customer service personnel can deal with them. (**Quotation 4**)

About 8,000 of such problems occurred every year and were dealt with manually in the backend of the organization: these constituted a significant cost.

The organizational context

The development project was anchored in the Marketing and Direct Sales (MDS) department, but had threads widely across the organization. MDS was holder of the sales processes and was responsible for

the business side of the development project. MDS was tied up to a department in the IT-organization (IT-MDS) that did the technical development. Two other business departments were heavily involved as well: the general insurance department and the retirement and live insurance department. They were holders of the insurance products and had the product domain knowledge, for example how policies were handled throughout and after the sale, how different products could be combined, and legal issues and best practices when advising private customers. Both product departments had comprehensive core IT-systems used for product administration managed and maintained by two other IT-departments. The users of the sales system were organized in five geographical regions independent of the business departments. As specific requirements of the sales tools were specified in the insurance agents' employment contracts, both the local insurance agent association and the internal legal department were involved in the project as well. A number of other IT-departments also took part in the project, such as the departments responsible for the IT-infrastructure, such as network, laptops, servers, databases, etc.

The organization did not rely on a standard software development process. The project leaders laid down the process and no formal strategies or procedures were imposed from outside. The organization was not experienced with usability work, but scattered knowledge existed and the consequences of not dealing with usability were well known. A research project dealing with usability evaluations and software development had previously been conducted in the organization and the designer had been given formal training in usability work. Furthermore, a large IT-project implemented in the company three or four years earlier had suffered severely from lack of usability.

The development process

The development project was scheduled to last about 18 months and at the most 25-30 people worked on the project. Half way through the project the schedule was slipping. The delays spread like ripples in a pond and the project was threatened by disintegration. Major changes were introduced in the project and two new project leaders replaced the original project leader. Future activities were broken down into smaller pieces and re-estimated, and 6-week time-boxes were introduced. These changes put the project back on track and the system was delivered with a delay of less than two months. Furthermore, the users were heavily involved. The project leader gives the following description of the development process and he emphasized how the users had been involved from an early stage:

Project leader: There was a wish of working with iterative development, but you could question if we succeeded. Sometimes we did, sometimes we did not; it was a little different how it came out. We broke things down and a few pieces were developed at the same time. Then we ran through it with the users and got it tested or at least we ran our own tests

continuously. Having said that, we did a lot of testing at the end of the project as well; and you can recognize traditional phases of software development in the project, such as analysis, development and test. I would especially like to comment on the user involvement; it was important because it was a sales application and because it spreads across a number of business areas, so the users were involved from an early stage. It was very comprehensive. They were involved as soon as we had something to show, for example paper sketches and clickable prototypes. (**Quotation 5**)

Stakeholders were heavy involved in the project and comprehensive usability work was conducted. However, in the eyes of the project leader this effort was not a natural part of the development project:

Project leader: No, I would not say so. I would probably say, that we have had some going back there as well. That is the danger ... do not misunderstand me, but when you have a lot of user involvement, as we had, then you have to listen to them. You get some feedback and you have to change some things. We experienced that as well. Sometimes we were saying Oops! This is not good enough. We have to fix that. We then had to go back, but at the end of the day that would be seen as quality or usability. It gave us extra work but we took that work. (**Quotation 6**)

Another central theme in the development process was making it customer centered or putting the customer in the focus in the application. In a sales situation this would increase the support given to the sales people and make them appear more professional. The user gives this description about why it matters and how it relates to the organizational goals:

User: The most important thing is putting the customer in focus. When the customer calls and we cannot see ... Oh, it is in your husband's name, ok, what is his name and his civil registration number. Oh, you do not remember. We cannot look it up then. We need to do something different. Now you can see the entire household at once, it is really customer friendly; it is the customer in the centre. Now you do not have to guess.... Oh you have other insurances? I cannot see that. Oh you have bank or life and retirement, I cannot see that either. Absalon elegantly indicates that you are a life and retirement insurance customer or a bank customer. It is much better. (**Quotation 7**)

The development process was informed by a number of stakeholders and different organizational interests making the decision process blurry and also stressing the need for compromises and strategic alliances, as important decisions were taken at a higher level without involving the users. This rich quotation gives a great insight into how the development project evolved from the developer's point of view, how decisions were taken, and some of the dangers when involving a broad range of stakeholders.

Interviewer: You had some considerations about implementing a shopping cart or a product list. Could you develop that point further?

Developer: In addition to the users we have two or three other interest groups. There were great discussions because they use different systems, such as our TIA system and our Java platform. We were aiming at making a sales application that was easy and intuitive to use and required a minimum of data entering, but we had this core system requiring a lot of data to be entered in a specified flow. Following that flow, you would sell one product at a time, but we wanted to sell say five products at a time and only enter common data once. The result was a compromise. The administrative stakeholders and the selling

stakeholders had very opposite interests. The result was this shopping cart solution, because we are not doing insurance policy administration. We are selling insurance policies, we are selling retirement products, and we are selling bank products.

Interviewer: So the TIA people were pulling one way and you...?

Developer: No, it was on a higher level. It was project leaders and ... I went there with the main architect, but we were the only developers ... maybe there was a TIA developer as well ... and then the people who made MobilSalg telling us this is a good idea and that is a bad idea. You can do this in MobilSalg, but you cannot do that. Often myths arose about what MobilSalg could, so it was at an overall level. We argued for what we meant was right, but we did not take the final decision, altogether.

Interviewer: Who took the final decision then?

Developer: It was taken by the back door. We made a proof of concept and people told us it looked well. Then the business described it ... give and take a little. So the business decided it ... the business resources. But I do not know who took that decision. (**Quotation 8**)

So far we have described how the development project was formed, and some of the organizational issues that informed the project and its outcome. We have highlighted different stakeholders' thoughts about project goals and organizational goals, the decision-making process, quality of the end product, and the consequences of involving a lot of stakeholders in the development project. Now we turn to the core of the case study, which is the description and analysis of the usability work that was conducted.

DESCRIPTION OF THE CASE STUDY

As a starting point we look at the roles of three of the main stakeholders and then discuss some of the main challenges that influenced the project. Since usability work was not a formal part of the project, no measurable targets were defined and no formal validations made. Thorough efforts of understanding and respecting stakeholders, focusing on business goals, and keeping things simple and reasonable became central concerns in the project.

The role of the stakeholders

No formal plan for usability work existed, but the need for dealing with usability was recognized by some project members and usability work became a significant part of the project. Two key people took responsibility for the usability work: the business analyst rooted in the business organization and the designer rooted in the technical organization. They worked closely together with other stakeholders and are in the interviews generally presented as people bridging gaps between more established groups of stakeholders, such as IT stakeholders, business stakeholders, and users. In their work they needed to gain an understanding of and to respect the limitation of their own roles and resources available in the organization, and at the same time identify and make use of every opportunity and seek influence whenever possible. The business analyst describes the cooperation this way:

Business analyst: I believe it was very important that we were physically placed together; not like you were sitting in this room and the developers were sitting in another room three floors away. You could just turn on your chair and say No! It cannot be true that it cannot be done! We must be able to figure something out because it is very important to the users. In that way the developers contributed by taking us seriously and bothered to do something. You also had to be awake when they fixed other problems and then ask them to fix that and that as well. And the common understanding is rooted in ... with my education I have look at some code in the early days. Then you can read over the shoulder and that understanding across the different professions has also been very important. (**Quotation 9**)

The following three sections will give a detailed description of the role of the designer, the business analyst, and the user.

The role of the designer

The designer had been in the company for some years, had experiences from prior engagements, and was taking university master classes in usability work. His responsibilities were requirement analysis, system and interaction design through development and test of prototypes, and being sparring partner to the software developers throughout the development process. Other project members referred to him as *'... the designer and usability guy'*, but his own description of his role was much more nuanced.

The designer: I had some different roles. We have a business department drawing up the system requirements. How to reach a 1.000 dollars premium is their job to figure out, I do not need to know. And then we have the developers. These two groups could work it out, but usually they have not considered how to put things together, in what order things should be presented, and so on. Then I become a liaison between the two groups, sparring with both of them, trying to make some requirements that in the end meet our users' demands. And how is this done? We have this experience database ... we have developed a number of systems; we do not rely on that alone, but it helps us recognize what works in this branch, such as terms and terminologies, and how to do things. We try creating some standards, so when switching from one system to another, things are called the same and done in the same way. Obviously, we do things in a new way if they originally were badly designed; we do not want bad things to be inherited, but we like inheriting some of the good things. My role is being the devil's advocate regarding the users; trying to propagate their wishes and I have done so in various ways. I arranged an early workshop where the users could speak freely; what was bad about MobilSalg, what would be nice in Absalon ... and I then did a relevance analysis; is this relevant, do we need it, is it interesting? If it was up to me, the developers should also have participated ... a subset of the developers, a business representative etc, so all experienced what the users wanted. (**Quotation 10**)

Based on the requirements the designer developed a paper and post-it based prototype to define and test the information architecture. The business project leader supplemented the above description of how specifications were evolved:

Business project leader: Such a specification process would typically have two or three iterations where it was reviewed. Maybe the developer made a prototype asking: Would it work this way? Could that be an idea? Or our graphic designer could

make a mock-up ... draw it up in Photoshop asking: Is this what you imagined? At the end we had a specification on which we signed off and froze for the moment. (**Quotation 11**)

Developing the prototype was an iterative process involving both developers, context of use experts, e.g. the different types of sales people, and domain experts, e.g. people with knowledge about specific products, the administrative processes and the administrative backend systems. Hand drawn prototypes and prototypes done in Photoshop and PowerPoint were developed. The prototypes were living documents continuously assessed by users around the different geographical regions. At some stage the Photoshop drawings were given to the developers who started adding business logic to the design.

The development process was informed by negotiations and informal user tests, where the designer continuously served as sparring partner for the developers and worked on validating various design decisions. *'I try to argue for my view points and they [the developers] argue for theirs. We might then meet in between ... and then we do a test and ask the users what they want.'* This was comprehensive work, but the strategy ensured that most of the system was tested with users before the final user test, and continuously involving the designer ensured consistency across the system.

The business analyst

The business analyst had been in the company for two years. Prior to her engagement she worked with usability in a bank where she was consultant and user test leader. Her responsibilities in the development process were requirement analysis and specification, testing, and user education, and she worked closely with the users. A number of contextual inquiries with insurance agents and customer service personnel were conducted as part of the requirement analysis and specification work. These revealed significant differences between the two user groups and contributed a thorough understanding of the context of use. The business analyst describes her work like this:

The business analyst: I represented the general insurance department. Other people represented retirement insurance, IT and many others, but my job was to define the requirements for the general insurance products. It was in the beginning. It was basic analysis and requirement specification; there was nothing in the project about process improvements compared to MobilSalg. MobilSalg should be replaced one-to-one. After the requirement specification I became responsible for the tests; first describing the different types of tests we needed. It was both the developers' tests and system tests verifying that data was saved correctly in the policy system. For example, when we tick here, is it then handled correctly, are the right fields opened and closed in the system? Very system specific! And then there was the user test, which became my darling. Because of lack of time and resources we had to make some compromises regarding the methods. It was a mixture ... I do not know if you would call it focus groups, think aloud, or task solving. It was not done by the book, but we did not have enough people with usability experience and we did not have much time. After the requirement specifications we made some rough paper mock-ups. The users reviewed the mock-ups, and after the system development ... the development project was not over, but

some parts of the system were finished, we held two user tests lasting three or four days. We tested with insurance agents, customer service personnel, and retirement specialists. They got some tasks ... I think there were 12 test users, and we were three people circulating and observing and registering how they performed. It was not really think aloud tests and it was not focus groups. It was a little pseudo... (**Quotation 12**)

As illustrated in this quotation the user tests were adjusted according to the available time and space, but the team still gained important insights. Neither time nor space allowed for individual user tests and all users were put into a room normally used for training new employees and asked to solve a set of tasks. Therefore, the user tests were '*not by the book*', but conducted '*... from the means we had and the resources ... and what we could expect to get out of it and what we would be able to fix in relation to the tight schedule.*' Problems were considered in open discussions in order to identify the root of the problems and the possible needs and desires for changes. Some tasks proved difficult to test realistically as the insurance agents had to read information from the task description instead of talking with the customer. To overcome this challenge test participants were set up in a role-play with one person acting as the customer and one person acting as the user, i.e. the sales person. This worked very well. At the end of the project the business analyst was made responsible for educating the users. In her words:

The business analyst: It was very obvious because you also saw ... there were some things that were difficult for the user, but could not necessarily be changed because it was decided to do it that way – either for business reasons or due to IT limitations, but then we knew what to focus on when educating the users. (**Quotation 13**)

A number of instructors (typically users from different regions, who were involved in specification and testing) had a week of training in the system and professional communication and were supplied with education material. Then they returned to their local offices and gave a two-day training course to the other users.

The user

The interviewed user was one of a number of users involved in the process. He was an insurance agent, had worked in the organization for more than 15 years, and was very experienced. He was very satisfied with the main thread running through the project and the influence he felt the users had on the product:

User: You can see, it is like that, and it is going to look like that. Is that okay? And then we expressed our opinions. I would say there has been a main thread through it all. And they were fast. We had noticed some icons were missing. It was fixed over the night and we could look at it the next day. That was cool and everybody was filled with enthusiasm. (**Quotation 14**)

His positive involvement made him a strong and proud ambassador for the project. He contributed early his comprehensive insight into the sales situations and sales processes. The contribution included recommendations about selected default values, recommendations about how to design various flows,

descriptions of different sales situation, and recommendations based on his contextual knowledge. He was involved in the project for more than half a year and took part in the user tests and in the user education as well. He was very fond of the test work and had this comment about the tests:

Interviewer: How were the tests conducted? Did you get a piece of paper and follow that or?

User: Yes, and then there was some free playing or what you would call it. When we had been going through some tasks and then were permitted to play with the system on our own. It is important that you are not completely controlled. You also want free playing; now I just look up my self, how is that working? If I want dog insurance, how does that work? There have been time for both, but there have also been a sort of a main thread through it. We were given 10 tasks with the message: This is typically how the tasks will look like when you are going to teach. Try to go through them and see if it makes any sense. (**Quotation 15**)

During the interview, the user pointed out some important conflicts related to both the organization and the technical environment. Most importantly he discussed his thoughts of being either a sales person or an adviser to the customer, and how that dilemma influenced the design of the system.

The Main Challenges

The informants pointed out a number of challenges in the project. These challenges introduced both new possibilities and significant risks in the project. The challenges were related to three different conditions: organizational conditions, technical conditions, and the conditions of context of use.

The organizational conditions

There was a significant interplay between the system under development, the development process, and organizational issues, such as organizational goals and organizational changes. The main organizational goals were to start following a set of best practices defined by the financial sector, to increase focus on cross-selling products from different product lines, for example by referring customers to colleagues in other branches, and to reduce the resources used on sales aftercare. The aim of fulfilling these goals influenced the design of the sales process and the efficiency of the system, as described by the user:

Interviewer: Do you know how many sales you can make at one day with Absalon compared with MobilSalg?

User: It is difficult. I think it is difficult because MobilSalg worked real well, I must say that. To be honest, I do not think it is more efficient right now, but it is probably more effective. We are going through the retirement part, which now is automatically attached, and we generate a lot of references to and from the bank and life insurance branches. We get a lot of references from out customer service personnel only doing general insurances. They refer customers to the bank ... I think we sell more, but I do not think we make more customers, because we have to go through more with each customer. We have saved something by making the system more efficient. We do not have to enter so much information and we do not need to write anything when referring a customer. References are automatically created with the customer data entered. I do not think we visit more customers, but I think we generate additional sales more effectively, and thereby get more double customers

[customers having products from more than one product line]. We better work with the existing customers and make them double customers ... and then get a system that can handle that. That is the idea.

Interviewer: I assume that insurance agents are paid on a commission basis?

User: Yes, to a very high degree.

Interviewer: How do you handle a situation where you are told to do something in a way that takes longer time? Is that not controversial?

User: Yes, but that is counterbalanced ... I do not think it takes longer time. There are two different types of customer groups or employees. The customer service personnel at the offices. They probably get more ...(pause)... they are measured on their references; so they have to ...(pause)... their job description will be a little different when referring people. They might say oh, are we now going to refer customers, but they are measured at it. Maybe a customer talk will take more time ... obviously it will when they have to talk about life and retirement insurance. Regarding the insurance agents, I do not think it would take much longer. We were always obliged to talk life and retirement insurance with the customer, so there is no shortcut there. Now we just cannot get around it. [...] In my eyes we are being more professional in out relations with the customers, we act as we are expected to and in fact a little more. (**Quotation 16**)

Quotation 16 makes it clear that the result of the development project was not a system supporting faster sales and larger commissions to the users. Rather the developed system supported an evolution in the organization by breaking down barriers between product lines and evolving sales people to customer advisors.

When implementing changes in an organization there is a risk of generating uncertainty among the affected people, which might lead to resistance or absence of engagement. In the case under study, special focus was on avoiding the system being and idea from the head office without solid anchoring in the regional organizations' wishes and needs. One way to avoid this perception was by thoroughly involving a number of regional users. The business analyst noted: '*There are many more things in it than pure usability, i.e. that the system has become better to use.*' This referred to the involvement of the users as political and pedagogical instruments in addition to creating a rich understanding of the users daily work. According to Quotation 14 the users appreciated having influence on the product design through their thorough involvement and the users' satisfaction with the development process was further reinforced by the development team's ability to realize and making use of opportunities that cropped up:

The business analyst: Then there was the pedagogical angle especially in relation to the insurance agents saying: It is so nice that we are finally heard, that you finally consult us! Overnight they could see changes. They asked for some small changes ... could we have an icon there and so on ... things we already had decided on. Then we went to the designer and when they arrived the next day it was implemented in their demo version. They were crazy about it. So, especially in this company, it was a pedagogic/ political signal stating, that it was not a head office system. (**Quotation 17**)

The comprehensive involvement of users resulted in a strong, shared responsibility and the users became system ambassadors and instructors in the final user education program. To further harmonize expectations with the future users, a tour around the country was made, where the project, its scope, and its time schedule were presented and discussed. This approach worked well and the project team received recognition for that. At the end of the project, the extensive user education also served as a tool for anchoring the system in the organization. Twenty instructors from around the sales organization educated 500 users in two-day courses and this work was considered especially important for the adoption of the system.

The technical conditions

The majority of the technical challenges originated in the new wireless platform and the abilities to integrate systems as the new platform brought about. One issue was to avoid reflecting the administrative processes from the core insurance systems into the new system (Quotation 8). The online abilities made it tempting to simply reflect the established administrative processes in the new web-based user interface and organizational pressure enhanced this idea. The processes in the administrative systems were product centered whereas the sales processes should be centered on the customer. By focusing on the customer (Quotation 7) the project team succeeded in creating a suitable sales support system, and adding available cross-organizational banking data made the sales people feel more professional (Quotation 16) and well prepared in the sale process.

The online abilities influenced issues such as data quality, validation of data, and online access to historical data and supplementary data sources. This had both organizational and more practical influence:

The business project leader: If I have to emphasize one thing ... it could be the customer part. We aimed at making it easy to deal with the customers, for example to enter customers and find customers. We implemented a look-up function to the national civil registration number database. If you are visiting a new customer and making an offer, you only need to ask for their national civil registration number and enter that into the system. Then we fetch their names and addresses and create the relations between the customers. All things they used to do manually. (**Quotation 18**)

In MobilSalg the insurance agents had a local database containing data from a small subset of the company's customers. Before and after every visit the insurance agents synchronized data. At the start of the day the agent retrieved the current information about the customers to be visited and at the end of the day updated information was fed back into the core systems. This process had a number of weaknesses, which the new technology could reduce (Quotation 3, Quotation 4). (1) It was not very flexible, i.e. the insurance agents needed to know who they were going to visit while they were at home,

otherwise they could not retrieve customer data. This made it very difficult to revise plans during the day. (2) If a new customer was visited, there was significant work registering the household correctly before the sale could start. Sales people typically want to pay attention as much as possible to the customers and often they entered defective data, which had to be fixed afterwards adding a significant amount of aftercare to the sale. (3) When visiting existing customers there was a risk of having old and incomplete data. If a change in the household or in the product portfolio was not registered correctly, the sales person might look foolish or in worst case lose the customer or miss an extra sale.

The majority of these problems were resolved by giving online access to the back-end systems while visiting the customers and by drawing relevant data from other systems, e.g. the national civil registration number database. By utilizing these possibilities, information about a complete household could be inserted correctly (Quotation 7) or automatically updated within a minute, new data could be validated immediately in the back-end systems (Quotation 3), and historical data displayed to support the sales people in their work.

The context of use

Workshops were an important tool in order to understand the context of use and to evaluate early design decisions, in relation to both usability and business goals. An initial workshop with 80 stakeholders from all around the organization became the starting point scoping the project, and a number of smaller workshops were later held as well:

Business project manager: At the initial workshop we got the users' wishes on functionality. Then we started drawing solutions, first line drawings at paper and then some Photoshop drawing, and we held some workshops where we got together with 5-7 future users at a time and ran through a customer case on paper and asked the users to use the application. (**Quotation 19**)

Early in the project it was realized that the groups of users had significant differences and it was considered especially important to gain an understanding of the different contexts of use and how the sales processes differed with these contexts. It was not considered realistic to build different user interfaces and understanding the sales processes thoroughly became a key activity:

Project leader: Basically we need to start from the users' work situations. We have two main groups: the insurance agents visiting the customer and looking him in the eyes, and the customer service personnel using the phone. They are very different, but they will need to use the same system. My reason for emphasizing this is, that it is exactly the understanding of the different sales processes ... they are not completely alike ... that understanding is totally fundamental and I believe that is the essence of the matter. Without that understanding you will not get a usable system and that is what we have taken as the starting point. We have been very close to these people and interviewed them and held workshops for them and user tests and

anything continuously, but the fundamental issue is still, how do you work today and how will you work with this new system. (**Quotation 20**)

Also the designer was aware of how understanding the context of use was a key issue in order to make the users accept the system:

Designer: We made a system that must consider both groups. It was quite a big challenge as one group was direct selling, sitting with their laptop in front of them and having an open dialog with the customer. And then you have the customer service personnel sitting in a complete different work situation. They can report very fast on the computer and get a result, but they cannot do the same considerations as a person in the house looking at the two cars in the carport. These are two very different situations. (**Quotation 21**)

Fitting the new system to two user groups was an obvious challenge and a number of differences were identified and addressed in the project. Insurance agents used the old sales system while customer service personnel used the administrative back-end system. The old system was mainly mouse controlled and the back-end systems mainly keyboard controlled, and it was considered critical to support both control methods effectively in the new system. And then there was an issue about controlling the sales processes, which influenced the way the insurance agents worked:

Business project leader: We made one system and we did not make a system with two user interfaces, but we aimed at making the system so flexible that both user groups could use it. This means that the process ... the sales process in the system is less rigid than it was in the old system. This is done mostly by considering the customer service personnel, because they work over the phone. They cannot control the conversation in the same way the insurance agents can. The customer service personnel need full control of the sequence and how to enter the system. The insurance agents needed to get used to that. Subsequently the insurance agents have realized it was an idea that they had the sales processes in their head. It is the insurance agent who controls the sequence. It is not the system that controls the sequence, and they seem to have realized that now. (**Quotation 22**)

Furthermore, the insurance agents typically allowed their customers to look over the shoulder, so information should be structured and displayed with great care and some information should not be directly visible, for example intermediate prices, special discount possibilities, and detailed product conditions which could be distracting in the sales process. On the other hand the customer service personnel had very limited control of the situation when customers called them with a variety of questions or problems. They were used to enter a minimum of information and then calculate an estimated price. Only after the customer accepted this would the rest of the information be entered. The customer service personnel appreciated flexibility in the system and had a large need for additional information. The solution was to structure information carefully and make it easily accessible in a very flexible system.

DIFFERENT PEOPLE'S PERSPECTIVES ON THE USABILITY WORK IN THE CASE

No formal evaluation of the usability work and the end product's usability was conducted. As the product had only been used for a short time when the interviews were conducted, the users were still getting acquainted with the new system. Furthermore, a number of technical problems were experienced, which made it difficult to distinguish between problems in the system and problems in the system environment.

The six interviewees all shared the understanding that the usability work was very successful, and one person mentioned how users were applauding after a training session. The key people agreed that the thorough user involvement and the wide spread discussion and communication had been very demanding, but they also agreed that it was worthwhile doing and was a critical success factor in the project. The business project leader expressed it this way: *'I feel we got a good solution; one of the reasons being the significant amount of discussions about how things should be done.'* Despite his initial reservation, the project leader expresses his acknowledgement this way:

Project leader: I think this user involvement worked well. I really think so. There is an incredible response. They appreciate it out there. It is clearly positive and it goes well together with the education where we also are in dialog with them out there. And using them as instructors and letting them educate their local users has worked in an exemplary fashion. It has been super. (**Quotation 23**)

LESSONS LEARNED

The case was influenced by the lack of traditions and formal requirements for usability work in the organization, but two key people managed successfully to put usability and direct user involvement on the agenda by stubbornly insisting on the necessity of taking usability seriously. It was not a giant leap into a user-centered design process, but a number of small steps navigating between different stakeholders:

Business analyst: We wanted to bring in all these people because they were so different, but we did not have time for these one-on-one user tests. Some of us had all the wishes and others managed the time schedule and the money and ... because it [usability work] has not been in the company before it was a small step at a time. (**Quotation 24**)

The starting point was to make a one-to-one transformation of functionality from the old system to the new one. The informants were never asked when this requirement came into existence, but it seems clear that it was watered-down gradually, when the organizational goals were introduced and the technological abilities recognized. Yet, it seemed as a useful strategy to use the old system, both as a

starting point for designing the new system, and as a baseline throughout the development project, especially when building a shared understanding of the potential of the new system with the users.

It is believed that relying on individuals with a special interest or skills in usability work is a common way to introduce usability work in organizations and introducing usability from within a project might yield some advantages. In this case it seems to have added confidence to the process, since the control of the process remained at the project leader level and the project team was able to follow both the challenges and the progress. Stone et al. [6] (p.589) suggest that concerns about usability are grounded in unawareness and fear of extra work. Such concerns are also found in this case, but it seems that the community of stakeholders and the shared responsibility seen in this study counteracts the concerns and makes them controllable. The technical project leader supports this interpretation by stating, that his role in relation to the usability work was to ensure that it '... *did not get out of control*'. Also, by relying on experienced key people from within the project, the acceptance and understanding of usability work were able to grow, and the usefulness of the work was widely accepted among the stakeholders.

It seems clear in this case that the close integration of software development and usability work, rooted in a close collaboration and a mutual understanding and respect among project participants, was very effective in driving the project in a sound direction. This is consistent with Dumas [5], who argues that the most important factor for software developers responding positively to usability findings is their relationship with the usability practitioners, and with Hornbæk and Stage [2], who argue for the importance of usability practitioners' understanding of the needs and wants of the development team. The two people responsible for usability were centrally placed in the project and took part in the whole software development project, giving a good opportunity to influence all facets of the project, that is from the initial analysis to the final user education. Both people were aware of the risks of being disqualified when doing usability work and being so close to the project, but neither of them experienced problems with this double role. Although the usability work added a significant workload to the development process, the close cooperation within the project team seemed to strengthen the overall process and had positive effect on the users, the project team, and the success of the end product. This is for example recognized in Quotation 23.

Adapting to the realities was crucial and making compromises was considered to be necessary by the key people, both in relation to the level of end product usability and how the various techniques were applied. The available opportunities for doing usability work were exploited in the best possible way and the at times novel and exploratory approach to usability work did reveal important insights. A clickable

prototype without underlying functionality would have been useful to test the flow in the application, but in general the applied usability techniques worked well. It was considered desirable to do a regular think aloud test, but there was no clear expectation of what would be gained from such a test. Given the limited time and resources not all usability issues could be resolved satisfactorily, but when important usability issues were left unresolved the business managers were thoroughly informed about the consequences in a formal way.

The comprehensive and on-going involvement of the users seems to have had a significant influence on the end product and the reception of the new system among the users. Many users were involved in the process and the way knowledge propagated and was used through the process from requirement analysis and specification to testing and user education, contributed the necessary anchoring in the organization and helped maintain focus on business relevant issues [1,4]. Furthermore realizing and taking advantage of opportunities when fixing errors and presenting progress did strengthen the process.

In the end, this case demonstrates how implementing usability work in an organization is hard work. It seems clear that all the informants acknowledged the contribution from the usability work as a critical factor for the success of the project. At the same time the project leader did not consider it a natural part of the development project and people were insecure about whether the conducted usability work would be continued in future projects. However, the case also demonstrates that a few committed people successfully can introduce usability work bottom up in organizations, if room is given for such work when developing software. The key then seems to be to ensure that usability work develops into an integrated part of the software development project. Not necessarily in the eyes of the project leaders and the managers, but most importantly in the eyes of the hand-on practitioners, that is: the people producing the software and the people who are going to use it.

Reference List

1. Følstad, A. Work-Domaine Experts As Evaluators: Usability Inspection of Domain-Specific Work-Support Systems, *IJHCI*, *22*, 3 (2007), 217-245.

2. Hornbæk, K. & Stage, J. Special Issue on The Interplay Between Usability Evaluations and User Interaction Design, *International Journal of Human-Computer Interaction*, *21*, 5 (2006), 117-123.

3. Kvale, S. *InterView*, Hans Reitzels Forlag, 1994.

4. Molich, R., Jeffries, R., & Dumas, J. S. Making Usability Recommendations Useful and Usable, *Journal of Usability Studies*, *2*, 4 (2007), 162-179.

5. Redish, J., Bias, R. G., Bailey, R., Molich, R., Dumas, J., & Spool, J. M. Usability in Practice: Formative Usability Evaluations - Evolution and Revolution, *Proc. CHI '02 extended abstracts on Human factors in computing systems*, ACM Press (2002), 885-890.

6. Stone, D. J. C., Woodroffe, M., & Minocha, S. *User Interface Design and Evaluation*, Morgan Kaufmann, 2005.

7. Strauss, A. & Corbin, J. *Basics of Qualitative Research*, Sage Publications Inc, 1998.

A usability toolkit for industrial systems development

The aim of this research project is to develop a usability toolkit, which can support industrial systems development projects in producing more useful and usable IT-systems and in keeping IT-development projects on track. The work will benefit the industry developing IT-systems and organizations and people depending on their use. So, the project aims at both social and economic contributions to society. The research project will be conducted in close collaboration with industrial partners and ongoing international research projects, yielding a unique opportunity to let the applicant connect with international researchers in different research environments.

Background

The usability of an IT-system concerns the users' ability to solve specific tasks and thereby reach specific goals through use of the system. Usability is defined in the ISO 9241-11 [1] standard and in the ISO 9126 standard [3]; and the ISO 13407 standard on Human Centred Design [2] gives some details on how to conduct human centred systems development. However, the standards are difficult to apply and the actual industrial practises vary a lot [28].

In industrial IT-systems development, development teams use various tools to identify and ameliorate usability problems, such as usability testing [7] or heuristic evaluation [25], but using these tools is difficult. Most tools are designed aiming at finding a maximum of problems rather than with the aim of fitting the development process or supporting the organizations implementing the IT-system. Consequently, usability professionals might care more about collecting data rather than systematically analyzing and making use of them [28]. For example, a common challenge is that extensive usability problems are pointed out too late in the systems development process. Fixing such problems adds significant costs to the development project and delays the product delivery. Another challenge is that even though comprehensive lists of usability problems in a IT-system are produced, the problems might have limited relevance [6] and only lead to few new insights [8,16,17,23]. Thus, usability work becomes an unnecessary risk and resistance towards usability can grow in the development team and in the organization implementing the IT-system [29].

The consequences of introducing IT-systems with poor usability can be severe. People interact with and rely on IT-systems in various sectors, for example in the areas of energy and supply services, healthcare, communications, transportations, production, finance, sales and service, and administration. Depending on the sector, poor usability can cause losses of life, physical and mental industrial injuries, social imbalance and isolation, reduced production and waste of resources, as well as dissatisfied employees and customers and lost sales [10,26,37,38]. Therefore, there is a strong social and economic potential in supporting the industry in increasing the usefulness and usability of developed and applied IT-systems.

The usability toolkit

This research project will develop a usability toolkit that can enhance systems development methods as used in industry. The usability toolkit will comprise a collection of usability tools that support each other, for example by building on common definitions, data and measures, coherent procedures, and shared values. Each tool will be designed as a component that can be plugged-in to the systems development process, i.e., it can be used when appropriate in the systems development process without specific requirements and preconditions. Such a toolkit does not currently exist, but would be valuable in industrial systems development [13].

The basic process when developing IT-systems is shown in Figure 1. Key stakeholders, for example developers, designers, users, managers, and usability professionals, work in a systems development context with a common objective of designing and creating an envisioned use context, which the IT-system will be a part of. In their work they rely on different tools and these specific tools informs their understanding of the work objective (i.e. the problem) [24].
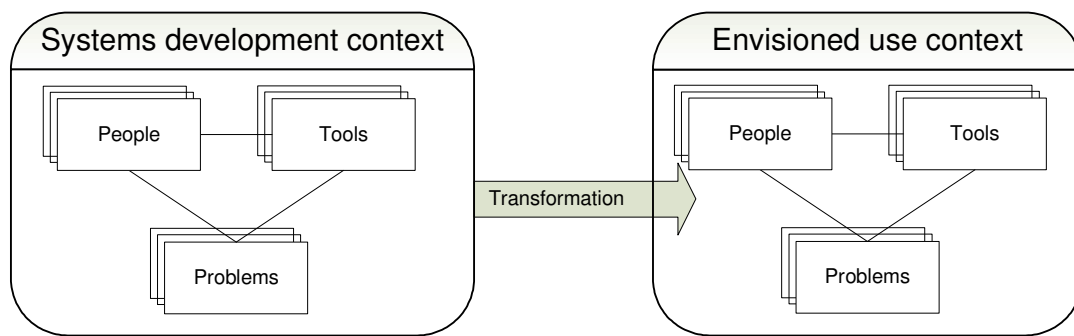

Figure 1 – The usability toolkit facilitates key stakeholders (1) in collaborating in the systems development context and (2) in maintaining focus on the important usability issues when transforming goals and vision from the development context to the use context. The figure builds on The Symmetrical Relation [24].

In this development process the usability toolkit will facilitate (1) collaboration among key stakeholders and (2) maintaining focus on the important usability issues. *First*, the toolkit will support key stakeholders in collaborating effectively in order to grow adequate and convenient solutions. The key stakeholders need to develop a coherent understanding of possibilities and limitations in the development context and in the envisioned use context [5,29,36]. Non-technical stakeholders, such as users and business managers, are often involved in early stages (e.g. in the analysis or specification stage) and in final stages (e.g. in the test stage) of development projects. Their involvement in the middle stages (e.g. the design or implementation stage), where the system is build, can be limited or non-existent. This makes it hard to maintain a main thread among key stakeholders throughout the development project. The usability toolkit will help keeping the project on track by continuously involving the necessary stakeholders. *Second*, the toolkit will help key stakeholders in obtaining needed and relevant information regarding usability issues and help

maintaining focus on those issues, while transforming goals and visions from the development context to the use context. This is a challenge since the envisioned use context might change during the process [30] and since usability issues might emerge and disappear throughout the project. Thus, the usability toolkit must have downstream utility i.e., it must direct attention towards key usability issues from the preliminary problem formulation stages, through the early sketching or explorative prototype activities, through the analysis and design activities, and onto the implementation and realization activities.

Relation to international and Danish research

The two foci above have only recently been targeted directly by the HCI research community; and they both seem promising in order to improve the practical impact of usability work. Thus, providing information on usability issues directly useful for collaboration among key stakeholders are shown effective in e.g. [8,17,18,20,31]. Likewise has downstream utility been researched in e.g. [4,9,15,19,22,34,36] and in the European MAUSE-project (MAUSE: towards the MAturation of Information Technology USability Evaluation, www.cost294.org) in which the HCI-group at DIKU is partner. The aim of the MAUSE-project is to bring more science to bear on Usability Evaluation Methods development, evaluation, and comparison, aiming for results that can be transferred to industry and educators, thus leading to increased competitiveness of European industry and benefit to the public. The international collaboration established in the MAUSE-project will be a cornerstone in the current research project. With partners from more than 20 countries we have an international established network, which will be drawn on and involved throughout the research project. Short-termed stays at different partners will be arranged when found useful, for example when doing experimental work as described in the research plan.

The current research project builds directly on usability research where the HCI-group at DIKU including the applicant are highly active and have demonstrated their ability to raise important research questions and find successful solutions. Furthermore, the research project is a natural successor to the USE-project (Usability Evaluation & Software Design) funded by the Danish Research Agency through the NABIIT Programme Committee (Grant no. 2106-04-0022). Here the HCI-group at DIKU collaborates with the HCI-group at Aalborg University. An important lesson learned from the USE project is, that we need to consider usability work an inter-human activity rather than a question about just identifying usability problems [27].

However, industrial usability work is still often based on traditional academic research on isolated usability work tools; think-aloud usability tests and different types of expert reviews are among the most popular tools in industry [14,21]. These tools are often used without coherence to other systems development activities. Gulliksen et al. [14] find that usability professionals in Sweden prefer using techniques, such as think-aloud, prototyping, interviews, field studies, and scenarios. They also find that the key success factors when doing usability work is (1) that the work is in the project plan from the start, and (2) that the work gets support from the key stakeholders. Thus, the

research by Gulliksen et al. suggests that the industrial practice not necessarily coheres with the key success factors.

At present, research conducted at DIKU and by other MAUSE partners have provided strong empirical results suggesting the usefulness of the proposed usability toolkit [8,11,12,17,18,36]. However, the results need to be brought together, described as a whole, and evaluated in collaboration with industrial partners. By doing so we aim at reaching a coherent, simple, yet powerful toolkit focusing on addressing usability issues through collaboration and generating relevant insights throughout the systems developing process.

The applicant's background

The applicant is currently writing up his Ph.D. thesis, with the working title "*On the understanding of usability work in systems development*". Deadline for this work is April 30th 2008. During his Ph.D. research, the applicant has focused on building a broader and more coherent understanding of usability work as conducted in industry. This research has generated new insights regarding different stakeholders interest in and contributions to industrial usability work [32,35], and suggested new approaches to combining usability work activities and evaluating such work [36]. Furthermore the applicant has shown a strong commitment to international research activities. He is an active partner in the European MAUSE-project and was the main organizer of an international workshop on "Increasing the Impact of Usability Work in Software Development" at the CHI 2007 conference [34]. A book proposal based on this workshop has been submitted to an international publisher and is currently under review.

As a part of his Ph.D. study the applicant developed The Usability Perspective Framework (UPF) [33]. This is based on five perspectives on usability found important by key stakeholders in industrial systems development [32]. The UPF is designed to support coherent usability work activities involving the key stakeholders throughout the systems development process. Through a set of key questions the UPF will inspire activities such as early workshops, design meetings, and different types of usability evaluations, and form a main thread in the usability toolkit.

Research plan

The research plan includes six main activities as shown in the table below. Each activity will be reported either in form of scientific publications or on the project homepage. The usability toolkit will be made accessible on the web, possibly in collaboration with the Usability Professionals' Association (UPA, www.upassoc.org).

	Activity	Month
1	Validate the UPF	1 – 6
2	Analyze, classify, and select candidate tool for the usability toolkit	1 – 9
3	Re-engineer candidate tools in order to tie them together under the UPF	6 – 12
4	Evaluation Phase 1: Evaluate individual candidate tools	9 – 15

| 5 | Evaluation Phase 2: Evaluate usability toolkit through supplementary training courses | 15 - 24 |
| 6 | Evaluation Phase 3: Evaluate usability toolkit under realistic industrial settings | 24 - 36 |

Table 1 – Main activities in the research plan

The *first* activity aims at validating the usefulness and face validity of the UPF. The UPF is based on observations and experiences from industrial systems development and will be used to frame the usability toolkit. The validation process will include both studying related work and letting industrial practitioners review and comment the framework.

The *second* activity aims at establishing a set of candidate tools for the usability toolkit. This activity will include a thorough review of existing usability tools. Under the guidance of the UPA a usability body of knowledge (BOK) is under development – a development project we have been encouraged to take part of. This BOK makes up a collection of independent tools used in industrial usability work. Complementing the BOK, the MAUSE-project's Working Group (WG) 1 has produced another collection of usability evaluation methods (UEMs) – a work we have been involved in since 2005. The BOK and the MAUSE UEM collection will, together with other recent scientific contributions serve as starting point for this activity. From these collections a set of candidate tools will be selected and their relevance and expected usefulness and utility will be considered.

The *third* activity aims at re-engineering the candidate tools in order to create a coherent usability toolkit. Each candidate tool will be reviewed and the necessary adjustments will be developed in order to tie the tools together under the UPF.

The *fourth* activity will be to evaluate the candidate tools under experimental conditions to make probable that the adjusted candidate tools will be useful and usable. Two examples of such an experiments are [8,18]. Partners from the MAUSE-project will be invited to take part in this activity, for example through collaborate or multi-site experiments.

The *fifth* activity will be to evaluate the relevance and usefulness of the candidate tools through workshops and supplementary training courses for industrial practitioners including usability professionals, managers, users, and developers. This work will both target Danish and international practitioners and will be conducted in collaboration with partners from the MAUSE-project.

The *sixth* activity will be to evaluate and reiterate the usability toolkit based on realistic industrial usage. Contacts with at least two industrial partners will be established and we aim at testing the toolkit in four industrial projects. The industrial partners will preferably have some significant differences, for instance with respect to applied development methods, internal vs. external systems development, and branches.

Post. Doc. Application
A usability toolkit for industrial systems development

References

1. ISO 9241-11, Ergonomic Requirements for Office Work With Visual Display Terminals (VDT)s - Part 11 Guidance on Usability, (1998).
2. ISO 13407, Human-Centred Design Processes for Interactive Systems, (1999).
3. ISO 9126, Software Engineering - Product Quality, (2001).
4. Blandford, A., Adams, A., Attfield, S., Buchanan, G., Gow, J., Makri, S., Rimmer, J., & Warwick, C. The PRET A Rapporter Framework: Evaluating Digital Libraries From the Perspective of Information Work, *Inf.Process.Manage.*, *44*, 1 (2008), 4-21.
5. Bødker, K., Kensing, F., & Simonsen, J. *Participatory IT Design*, The MIT Press, Cambridge, Massachusetts, 2004.
6. Dumas, J. S., Molich, R., & Jeffries, R. Describing Usability Problems: Are We Sending the Right Message?, *interactions*, *11*, 4 (2004), 24-29.
7. Dumas, J. S. & Redish, J. *A Practical Guide to Usability Testing*, Intellect, Exeter, UK, 1999.
8. Følstad, A. Work-Domaine Experts As Evaluators: Usability Inspection of Domain-Specific Work-Support Systems, *IJHCI*, *22*, 3 (2007), 217-245.
9. Følstad, A., Bark, I., & Gulliksen, J. How HCI-Practitioners Want to Evaluate Their Own Practice, *Proc. of the 4th Nordic conference on Human-computer interaction - changing roles*, ACM (2006), 417-420.
10. Forrester Research 2004 Outlook for ECommerce Tech Spending, (2004).
11. Frøkjær, E. & Hornbæk, K. Cooperative Usability Testing: Complementing Usability Tests With User-Supported Interpretation Sessions, *Proc. CHI '05 extended abstracts on Human factors in computing systems*, ACM Press (2005), 1383-1386.
12. Frøkjær, E. & Hornbæk, K. Metaphors of Human Thinking for Usability Inspection and Design, *ACM Trans.Comput.-Hum.Interact.*, *14*, 4 (2008), 1-33.
13. Greenberg, S. & Buxton, W. A. S. Usability Evaluation Considered Harmful (Some of the Time), *Proc. of the SIGCHI conference on Human factors in computing systems*, ACM Press (2008)
14. Gulliksen, J., Boivie, I., Persson, J., Hektor, A., & Herulf, L. Making a Difference: a Survey of the Usability Profession in Sweden, *Proc. of the third Nordic conference on Human-computer interaction*, ACM (2004), 207-215.
15. Hartson, H. R., Andre, T. S., & Williges, R. C. Criteria for Evaluation Usability Evaluation Methods, *IJHCI*, *13* (2001), 373-410.
16. Høegh, R. Usability Problems: Do Software Developers Already Know?, *Proc. of the 20th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction design - activities, artefacts and environments*, ACM (2006), 425-428.
17. Hornbæk, K. & Frøkjær, E. Comparing Usability Problems and Redesign Proposals As Input to Practical Systems Development, *Proc. the SIGCHI conference on Human factors in computing systems*, ACM Press (2005), 391-400.
18. Hornbæk, K. & Frøkjær, E. Making Use of Business Goals in Usability Evaluation: An Experiment With Novice Evaluators, *Proc. of the SIGCHI conference on Human factors in computing systems*, ACM Press (2008)
19. John, B. E. & Marks, S. J. Tracking the Effectiveness of Usability Evaluation Methods, *BIT*, *16* (1997), 188-202.
20. Law, E. L.-C., Lárusdóttir, M., & Nørgaard, M. COST294-MAUSE Open Workshop on Downstream Utility: The Good, the Bad, and the Utterly Useless Usability Evaluation Feedback. November 6th 2007, Toulouse, France, (2007).
21. Mao, J.-Y., Vredenburg, K., Smith, P. W., & Carey, T. The State of User-Centered Design Practice, *Commun.ACM*, *48*, 3 (2005), 105-109.
22. McDonald, S., Monahan, K., & Cockton, G. Modified Contextual Design As a Field Evaluation Method, *Proc. of the 4th Nordic conference on Human-computer interaction - changing roles*, ACM (2006), 437-440.
23. Molich, R. & Dumas, J. S. Comparative Usability Evaluation (CUE-4), *BIT*, (2007).
24. Naur, P. *Computing: A Human Activity.*, ACM Press / Addison-Wesley, 1992.
25. Nielsen, J. Finding Usability Problems Through Heuristic Evaluation, *Proc. of the SIGCHI conference on Human factors in computing systems*, ACM Press (1992), 373-380.
26. Nielsen, J. Return on Investment for Usability (Http://Www.Useit.Com/Alertbox/20030107.Html), (2003).
27. Nørgaard, M. Usability Work: A Human Activity, *Proc. the 4th COST294-MAUSE Open workshop on Downstream Utility (November 6th, 2007)*, (2007)

28. Nørgaard, M. & Hornbæk, K. What Do Usability Evaluators Do in Practice?: an Explorative Study of Think-Aloud Testing, *Proc. of the 6th conference on Designing Interactive systems*, ACM (2006), 209-218.

29. Rajanen, M. & Iivari, N. Usability Cost-Benefit Analysis: How Usability Became a Curse Word?, *Proc. INTERACT 2007, 11th IFIP TC 13 International Conference*, Springer-Verlag Berlin Heidelberg (2007), 511-524.

30. Riisgaard Hansen, T. Strings of Experiments: Looking at the Design Process As a Set of Socio-Technical Experiments, *Proc. of the 9th conf. on Participatory Design*, ACM (2006), 1-10.

31. Tohidi, M., Buxton, W., Baecker, R., & Sellen, A. Getting the Right Design and the Design Right, *Proc. of the SIGCHI conference on Human Factors in computing systems*, ACM (2006), 1243-1252.

32. Uldall-Espersen, T. Exploring Multiple Usability Perspectives, *Proc. INTERACT 2007, 11th IFIP TC 13 International Conference*, Springer-Verlag Berlin Heidelberg (2007), 628-632.

33. Uldall-Espersen, T. The Usability Perspective Framework, *Proc. of the SIGCHI conference on Human factors in computing systems*, ACM Press (2008)

34. Uldall-Espersen, T., Blandford, A., Jokela, T., & Frøkjær, E. Increasing the Impact of Usability Work in Software Development, *Proc. CHI '07 extended abstracts on Human factors in computing systems*, ACM Press (2007), 2873-2876.

35. Uldall-Espersen, T. & Frøjkær, E. Usability and Software Development: Roles of the Stakeholders, *Proc. HCI International 2007, Beijing, China, 22.-27. July 2007*, Springer-Verlag Berlin Heidelberg (2007), 642-651.

36. Uldall-Espersen, T., Frøkjær, E., & Hornbæk, K. Tracing Impact in a Usability Improvement Process, *Interacting with Computers*, *20*, 1 (2008), 48-63.

37. Vicente, K. *The Human Factor: Revolutionizing the Way People Live With Technology*, Routledge, 2004.

38. Watson, S. The Business of Customer Experience: Lessons Learned at Wells Fargo, *interactions*, *January + February* (2008), 38-43.