Instant Data Analysis: Evaluating Usability in a Day

Kjeldskov, Jesper; Skov, Mikael B.; Stage, Jan

Published in:
Proceedings of NordiCHI 2004

DOI (link to publication from Publisher):
10.1145/1028014.1028050

Publication date:
2004

Document Version
Early version, also known as pre-print

Instant Data Analysis: Conducting Usability Evaluations in a Day

Jesper Kjeldskov, Mikael B. Skov and Jan Stage
Department of Computer Science, Aalborg University
Fredrik Bajers Vej 7, 9220 Aalborg East
{jesper, dubois, jans}@cs.auc.dk
+45 9635 8080

ABSTRACT
When designing a usability evaluation, key decisions must be made regarding methods and techniques for data collection and analysis. Although there is a strong body of research within human-computer interaction regarding the appropriate choices of data collection methods and techniques, much less research has been conducted examining and comparing methods and techniques for analyzing the collected data. This paper presents a data analysis technique which allows usability evaluations to be conducted, analyzed and documented in a day; Instant Data Analysis (IDA). The use of this technique is exemplified through a usability evaluation of a software product for a large hospital for which traditional video data analysis and Instant Data Analysis were applied independently through a controlled experiment. Among our key findings, the experiment revealed that in only 10% of the time required to do the video data analysis, Instant Data Analysis identified 85% of the critical usability problems in the system being evaluated. At the same time, the noise of unique usability problems usually characterizing video data analysis was significantly reduced.

Author Keywords
Usability evaluation, data analysis, discount usability

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION
Usability evaluations have proven to be invaluable tools for assessing the quality of computer systems. However, evaluating usability is often difficult, time consuming and expensive. Hence, with the purpose of reducing effort and increasing return of investment, the last decade of usability research has investigated extensively into questions such as which experimental methods to apply [3], how many evaluators to involve in expert- and user-based studies [2], how many test subjects to use for user-based studies [8, 17, 19], and in which settings to conduct the evaluations [4, 5, 9]. Providing tools for conducting usability evaluations 'at a discount' [9, 10, 11, 12], this research has greatly improved software designers' and researchers' ability to select the most appropriate methods and techniques for usability evaluation within the specific constraints of their individual software development projects. Thus, for example, it is generally acknowledged that heuristic evaluation is a very cost-effective technique for quick-and-dirty evaluation of usability, see e.g. [13]. But it is also well-known that while being more time consuming, user-based evaluations (e.g. using the think-aloud protocol) on the other hand tend to facilitate identification of a higher number of problems and more relevant problems, see e.g. [3].

One of the most resource demanding activities in a usability evaluation is the analysis of collected empirical data. Not only is it time consuming, but data analysis is also very vital as it extracts key findings of the usability evaluation. Furthermore, it is a key activity in usability evaluations as evaluators may find themselves influencing the findings through different interpretations [1].

The data analysis activity is, however, vaguely described in the mainstream usability literature like e.g. [10, 16, 17]. Many methods and techniques exists for analyzing the empirical data from usability evaluations like, for example, grounded analysis [18], video data analysis [7], cued-recall [14], and expert analysis [13], etc. However, instrumentation in data analysis of usability evaluations is often poorly discussed [1] and the relative value of applying these methods and techniques to analysis of usability is still largely unidentified. Of special interest, it seems implicitly assumed that thorough video analysis with detailed log-files and transcriptions of usability evaluation sessions is the most optimal way to analyze usability evaluation data. However, the added value of spending large amounts of

NordiCHI '04, October 23-27, 2004 Tampere, Finland

time on video analysis in relation to the results subsequently produced is still questionable [9].

This paper extends Nielsen's discount usability tradition [9, 10, 11, 12] by exploring how much the required effort during the data analysis activity can be reduced. It challenges the implicit assumption that video analysis is the most optimal way to analyze usability evaluation data. Alternatively, we present a novel technique for increasing the speed of data analysis and improving its focus: Instant Data Analysis (IDA). Used in combination with the think-aloud protocol for user-based evaluations, this technique makes it possible to conduct a complete usability evaluation in a day.

The paper is structured in the following way. First we outline the basic idea behind the Instant Data Analysis technique IDA. We then present an experimental study in which we applied the IDA technique to a usability evaluation of a commercial software product. For comparison, the video data from this evaluation was also analyzed using a conventional technique for identifying usability problems based on log-file transcription. Based on our experiment, the findings from the Instant Data Analysis and the video data analysis are outlined and gaps and overlaps are discussed. Finally, we conclude and point out avenues for further research.

INSTANT DATA ANALYSIS

Motivated by the challenges discussed above, we have developed a technique for reducing the efforts spent on analyzing data from usability evaluations. We named the technique Instant Data Analysis (IDA). The aim of applying this technique is to make it possible to conduct an entire usability evaluation in one single day. Consistent with the discussion in [9], it is not the aim of the IDA technique to be the best technique for analyzing usability data, but to be effective in supporting fast identification of the most critical usability problems of a software system.

Background

The IDA technique grew out of experiences from previous research on developing new methods for usability evaluation [15]. In the study by Pedell et al., the researchers tried to capture additional data about the usability evaluations they were conducting over a number of days by ending every day with a 1 hour debriefing session. While initially looking for 'metadata' on the experimental methods applied, the study revealed that many of the usability problems later identified through thorough analysis of video recordings had actually already been brought out into the open during the researcher's debriefing sessions. On the basis of this finding, we decided to refine the debriefing approach to focus explicitly on the identification of usability problems.

The IDA technique adopts the wide-spread assumption that identifying the highest number of critical usability problems of a software product can lead to improved quality through redesign. Thus, the IDA technique does not aim at producing an elaborated theoretical understanding of the use of the system being evaluated. Rather it aims directly at producing a ranked list of usability problems for the design team to take into consideration when redesigning the system.

The IDA technique is designed to be combined with the use of the well-established think-aloud protocol for user-based usability evaluations as described in for example [16, 17]. The technique can be applied to both laboratory-based and field-based think-aloud evaluations. The IDA technique exploits the fact that think-aloud usability evaluations typically involve a test monitor and a data logger with high level usability expertise. When conducting the evaluation, the test monitor and the data logger typically gain a strong insight into the evaluated system's key usability problems very quickly. While some of these problems may be captured by taking notes, much of this insight is often lost and needs to be reconstructed during later video data analysis. Rather than loosing this valuable moment of insight, the IDA technique extends the think-aloud sessions with a joint data analysis session.

Key literature on usability evaluation suggests that 4 to 6 test subjects are sufficient to gain an overall idea about the usability of a software product [8, 19]. Spending roughly 45 minutes per test subject on the data collection and 15 minutes on changeover, this makes it possible to conduct the whole evaluation in only one day while still leaving room for analyzing data and writing up the findings.

Procedure

The use of the IDA technique follows immediately after the actual conduction of the think-aloud usability test sessions. Aiming at conducting the entire evaluation in one single day, 4 to 6 think-aloud sessions should provide a proper foundation for the analysis. During the usability test sessions, the data logger records incidents or problems. This will be used for the later problem identification and categorization. After the think-aloud sessions, the test monitor and the data logger conduct a one hour brainstorming and analysis session. The purpose of this session is to produce a list of usability problems as experienced by the 4 to 6 test subjects.

The roles of the test monitor and data logger during the data analysis are to articulate and discuss the most critical usability problems that they have identified during the think-aloud sessions. Also, they should rate the severity of each problem stating if it is, for example, critical, serious or cosmetic [6]. Assisting the brainstorming and analysis process, the test monitor and data logger may use printed screenshots of the system and notes taken by the data logger during the think-aloud sessions. The aim of the process is not to identify as many usability problems as possible, but to identify the most critical ones.

The analysis session is assisted by a facilitator, who may or may not have been present during the think-aloud sessions. The role of the facilitator is to manage the brainstorming and analysis session, asking questions for clarification and writing all identified usability problems on a whiteboard/flip-over as they are presented by the test monitor and data logger. The facilitator should also make sure to keep an overview of the identified problems as the session progresses, categorizing them in themes, avoiding redundancy etc.

After the one hour brainstorm and analysis, the facilitator spends 1-1½ hour on his own writing up the contents of the whiteboard/flip-over into a ranked list of usability problems with short descriptions and clear references the system. Finally, the test monitor, data logger and facilitator run through the problem list together to ensure consensus.

METHOD

We have evaluated the use of the proposed technique for Instant Data Analysis through a usability evaluation of a commercial software product prior to its implementation at a large regional hospital.

System

The system evaluated in our study is a resource booking system for the healthcare sector. The idea of the system is that hospital staff can schedule upcoming operations, specify overall details and book the necessary resources e.g. doctors and operation rooms. The system runs on a standard desktop or laptop PC connected to a central server via network. The prospective target users of the booking system include doctors, nurses, and secretaries.

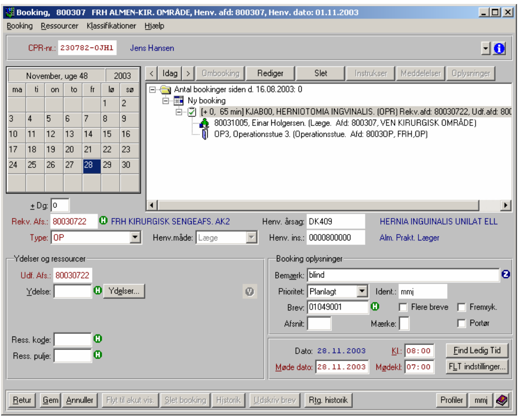

Figure 1. The Booking System (Courtesy of B-Data).

Figure 1 illustrates one of the main interfaces in the booking system where the user can assign staff and other resources to an upcoming operation.

Participants

The study included a number of test subjects acting as users of the evaluated system and a number of researchers taking the roles outlined by IDA and by traditional video data analysis.

Five test subjects (four females and one male) between 25 and 64 years of age participated in the usability evaluation of the Booking System. They were all staff at the Hospital of Frederikshavn and had diverse practical experience with hospital work and medical procedures ranging from 1 year to 37 years. The test subjects had all received training in the booking system but none of them were using the system on a regular basis at the time of the study. All of the test subjects were either experienced or advanced experienced users of computers and information technologies.

In addition, four trained usability researchers participated in different roles on evaluating the use of the IDA technique. All evaluators had significant previous experience with the conduction of usability evaluations and had managed or participated in more than 40 usability sessions. One researcher acted as test monitor during the test sessions with the five test subjects. A second researcher acted as data logger during the sessions writing down as much as possible during the tests. This researcher also operated the video equipment. A third researcher observed the sessions and also logged data for supporting a later video analysis. Finally, a researcher observed the sessions and acted as facilitator in the following IDA session.

Setting

The five usability evaluation sessions were conducted at a state-of-the-art usability laboratory at Aalborg University facilitating close-up observation of the test subject's interaction with the evaluated system. As outlined in figure 2, this involved the use of three different rooms: a subject room, an observation room and a control room.

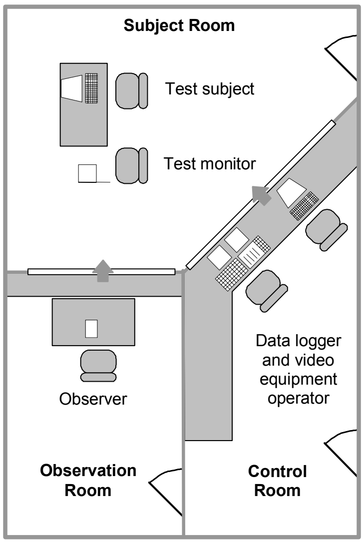

Figure 2. Physical layout of the usability laboratory.

The subject room contained two chairs for the test subject and the test monitor and a computer table with a standard

PC with 19 inch monitor, a conventional keyboard and a mouse. The PC had the Booking System pre-installed and was connected to a central booking server through a wide area network via a VPN client.

Three motorized remote-controlled cameras were mounted in the ceiling of the subject room. One camera was placed directly above the table allowing a view of the subject's interaction with the keyboard and the mouse. The second and third camera provided close-up and general overviews of the test subject and test monitor (figure 3). A scan-converter transformed the image on the computer screen to a composite video signal. The four video signals were merged to one video feed and recorded digitally. For sound recording, we used a single table microphone. In addition, the test monitor had a wireless in-ear monitor system, which allowed the data logger in the control room to speak back to him during the evaluation if necessary.

Figure 3. Evaluation setup in the subject room: test monitor (left) and test subject (right).

The control room allowed the data logger to survey the subject room through one-way mirrors and by means of the motorized cameras. During the evaluation, the data logger controlled the video recording equipment, took notes and created a preliminary, overall log-file.

Figure 4. Control room: equipment for operating the motorized video cameras in the subject room.

The observation room allowed two researchers to observe the evaluation through one-way mirrors. Like in the control room, the observation room also facilitated a relayed screen image from the test PC and a large monitor with the merged video signals from the cameras. During the evaluations, the researcher who was going to do the video analysis took notes and created a preliminary, detailed log-file. The other researcher, who was going to facilitate the Instant Data Analysis session observed the evaluation sessions in order to generate an overview of the use of the system and did not take notes of any kind.

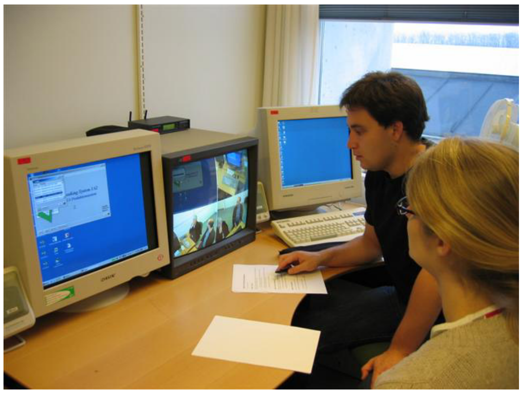

Figure 5. Observation room: relayed screen signal from test PC and video signals from cameras in the subject room-

Additionally, two members of the software development team monitored the evaluation in a second observation room and took notes. The outcome of this is, however, not touched upon in this paper.

Procedure

The usability evaluation was conducted in one day (five hours) using the following procedure.

Briefing and introduction. First, the test monitor introduced the test subjects to the evaluation procedure. They then filled out a brief questionnaire about their demographic profile. Following this, they were handed the first task.

Task Solving. The evaluation sessions were structured by three tasks assignments. The tasks were given to the test subjects one at a time by the test monitor. The test subject worked on solving the tasks until they felt that they had completed it or until the test monitor asked them to move on to the next one.

Thinking-aloud. The test-subjects were asked to think-aloud during the evaluation sessions, explaining their interaction with the system and articulating their comprehension of the design. The test monitor did not provide help but asked questions for clarification

Debriefing and questionnaire. Following each evaluation session, the test subjects filled out a questionnaire and were debriefed about their experience of the evaluation.

Data Analysis

The data from the usability evaluation sessions was analyzed independently by two teams of researchers applying a traditional video data analysis technique and the Instant Data Analysis technique respectively.

Instant Data Analysis (IDA)

The Instant Data Analysis followed the procedure described above. The test monitor, data logger and facilitator spent 1 hour on brainstorming and analyzing the usability of the evaluated system assisted by the notes of the logger and printed screenshots. During this time, the facilitator listed and organized the identified usability problems on two large whiteboards (figure 6) and asked questions for clarification. Severity was ranked in a collaborative effort between the test monitor and data logger. Following the joint session, the facilitator spent 1 hour on compiling a list of usability problems.

Figure 6. Instant Data Analysis: Facilitator listing usability problems identified by test monitor and data logger on whiteboard immediately after the evaluation sessions.

The Instant Data Analysis produced a list of usability problems ranked as critical, severe or cosmetic with approximately 2 lines of explanation. The total time spent using the traditional Instant Data Analysis technique amounted to 4 man-hours

Video Data Analysis (VDA)

The analysis of the video data followed a standard approach to identifying usability problems. First, the preliminary log-files for each of the five test subjects created during the evaluation sessions were completed by looking through all videos. Following this, the video tapes were then examined thoroughly for identification of usability problems assisted by the log file and each usability problem was described in detail and ranked in relation to its severity.

The Video Data Analysis produced a detailed log file of the five evaluation sessions and a list of usability problems ranked as critical, severe or cosmetic with approximately 5-7 lines of explanation. The total time spent using the traditional Video Data Analysis technique amounted to approximately 40 man-hours.

Following the Instant Data Analysis and the video data analysis, the two lists of usability problems were merged in a collaborative effort. As a part of this, small variations in severity ratings were discussed until consensus had been reached.

FINDINGS

The instant data analysis (IDA) technique generated a list of 41 different usability problems. From these 41 problems, we classified 11 problems as being critical, 15 problems as being serious, and 15 problems as being cosmetic. The identified critical usability problems related to various issues of the interaction between the system and the user, for instance a couple of problems related to difficulties in solving key tasks with the booking system e.g. like rescheduling a booking or changing the severity status of a booking from normal to emergency. Other critical problems were triggered by slow response times where subjects by mistake would activate functions on different interfaces. Finally, a few critical problems were results of misunderstanding of concepts and words used in the interface. Thus, on its own, using the IDA technique seemed to be suitable for identifying usability problems in a collaborative effort.

	IDA	VDA	TOTAL
Critical	11	12	13
Serious	15	15	22
Cosmetic	15	19	27
Total	41	46	62

Table 1. Numbers of usability problems identified using the Instant Data Analysis technique (IDA) and using the Video Data Analysis technique (VDA).

Comparing the IDA results with the results of the video data analysis approach, we found that the latter identified a total of 46 different usability problems where 12 were critical, 15 were serious, and 19 were cosmetic (see table 1). The table also outlines the total number of identified problems by the two approaches. In total, the two techniques identified a list of 62 different usability problems where 13 were critical problems, 22 were serious problems, and 27 were cosmetic problems. The distribution of the identified problems across the two analysis techniques is illustrated in figure 7.

Considering the identified problems, we found that both approaches assisted in identifying nearly all critical problems where IDA identified 11 of the 13 (85%) critical usability problems whereas video data analysis identified 12 of the 13 (92%) critical usability problems. The two problems not identified by IDA related to general irritation of the participating subjects on the slow response time of the system and secondly, a software failure of the system during the test. However, in summary, the two approaches produced very similar results concerning the most critical usability problems of the tests.

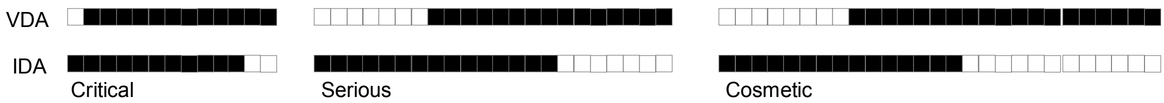

Figure 7. Distribution of usability problems identified using the Video Data Analysis (VDA) and Instant Data Analysis (IDA) techniques. Each column represents a usability problem. A black square indicates that a problem was identified using that specific technique. A white square indicates that a problem was not identified using that specific technique but was found using the other technique.

The serious and cosmetic usability problems exhibited a different distribution between the two analysis techniques. Where the IDA technique identified 15 serious problems, a total of 22 serious problems were identified by the two approaches together. Thus, the IDA approach identified 68% of the serious problems found in total. On the other hand, the video data analysis also identified 15 serious problems (68%) meaning that eight serious problems were identified by both approaches. Considering the cosmetic problems, we found that the IDA technique identified 15 of the total 27 problems (56%). The 12 remaining cosmetic problems unidentified by IDA related primarily to specific interaction problems for the subjects typically only experienced by one of the five subjects. As examples, the following two cosmetic problems were only identified by the video data analysis "*the subject would prefer to specify search criteria in terms of weeks instead of days*" and "*the subject realizes that the busy pointer is only active when inside the current interface*". A total of 7 out of the 27 cosmetic problems (26%) were identified by both analysis approaches.

A high number of the usability problems identified in the video data analysis approach were experienced by only one subject test subject (26 problems of the total 46). As considered in e.g. [3], it can be discussed whether these are really problems at all, or if they are noise added to the picture by non-generalizable subjective experiences of interaction with the system. Information about how many test subjects experienced the different usability problems was not included in the problem list generated from the IDA technique. But some of these 26 problems were also identified by the IDA approach. However, the majority of problems experienced only by one single test subject (16 of the 26) were *only* identified in the video data analysis and not in the instant data analysis. Thus, the use of the IDA approach allowed for the omission of a significant part of this noise.

When merging the two lists of problems, a number of the usability problems identified by the video data analysis were combined in one single usability problem as identified by the IDA technique. As an illustrative example, the IDA approach identified the following serious problem "the *subject has problems with the logical order of sequence in the booking interface*". This particular problem was also identified in the video data analysis, but it was described in four different, separate problems, e.g. "*the subject attempts to attach a service before a ward in the booking interface*" or "*the subject is unable to recognize whether information is missing in the booking interface*". Thus, the video data analysis was able to provide a richer and more diverse snapshot of some of the identified problems. On the other hand the use if the IDA approach generated more general and high-level abstractions of the identified problems.

DISCUSSION

The findings of our experiment presented above raises a number of interesting issues for discussion.

Firstly, it is interesting that the Instant Data Analysis technique was actually able to assist the identification of most of the critical and serious usability problems also found using the video data analysis approach. This indicates that thorough analysis of video recordings from usability evaluations may not be necessary in order to produce a useful outcome, and that reducing the effort spent on data analysis may be a viable way of releasing time and resources for other activities, such as an additional evaluation at a later stage of the project. On the other hand, however, missing out on a number of critical and serious problems may not be acceptable, in which case the IDA technique should be used with caution.

Secondly, it is noteworthy that the Instant Data Analysis technique assisted in identifying a relatively large number of serious problems, which did not come out clearly from the video data analysis. Qualitatively, these problems were often related to more general issues of the interface and interaction design of the system than the serious problems identified by the video data analysis approach. When applying the results to redesign, this may help designers to focus on overcoming overall problems rather than patching up a number of smaller ones. This quality of the problems identified by the IDA technique may be attributed to the fact that the test monitor and data logger did not have direct access to data about the specific user interactions with the system. Thus they were forced to do the analysis on a higher level of abstraction and not distracted by specific instances of a larger problem.

Thirdly, it was interesting to find that even though the Instant Data Analysis technique identified a similar number of cosmetic problems as the video data analysis technique, the problems identified by the two approaches were very different. While the problem list produced by the Instant Data Analysis technique does not include information about how many test subjects experienced each of the problems, it is noteworthy that out of the 12 cosmetic usability problems only found through the video data analysis, only one problem was experienced by more than one user (it was experienced by two). Viewing such unique problems as noise rather than 'real' usability problems, this finding suggests that the IDA technique assists in identifying the 'right' cosmetic problems.

This discussion can also be generalized to the serious usability problems identified with the two techniques. Among these problems, 4 of the 7 problems identified only by the video data analysis were experienced only by one test subject. Similarly, one of the two critical problems only identified in the video data analysis (50% of the added value in this category of problems) was also only experienced by one user. This specific problem was related to a bug in the system. In the light of these findings, it thus seems that the use of the IDA technique assists in suppressing much of the noise from 'irrelevant usability problems' only experienced by one test subject.

Finally, it is relevant to discuss the findings from the experiment in the light of the time spent on applying the two techniques (4 man-hours spent on Instant Data Analysis versus 40 man-hours spent on video data analysis). While the differences in required efforts in terms of man-hours speaks for itself in terms of the cost-benefit ratio of identified usability problems, it should, however, also be noted that the video data analysis approach produced a more detailed description of each of the identified usability problems with references to the video recordings through the log file. The Instant Data Analysis approach did not provide this detailed information but merely produced a list of shortly described problems with clear pointers to screenshots. Whether or not this added outcome of the video data analysis is relevant for the redesign of the evaluated system is questionable and depends on the specific system development project. If the team conducting the evaluation is also going to be involved in its subsequent redesign, the log file may be redundant and not worth the hassle. If, on the other hand, a usability evaluation report is handed over to a team of designers who did not participate in the evaluation, the log file may provide valuable cues for the interpretation of some problems. However, it should at the same time be noted, that if the designers involved in the redesign of a system on the basis of a usability evaluation have to go back to the original empirical data themselves anyway to make sense of the identified usability problems, the conducted data analysis was probably not good enough in the first place.

CONCLUSIONS

This paper has presented a new technique for analyzing empirical data from user-based usability evaluations: Instant Data Analysis (IDA). Combined with 4 to 6 think-aloud sessions, this technique makes it possible to conduct, analyze, and report from a usability evaluation in a day.

The presented IDA technique has been evaluated through the usability evaluation of a commercial software product and compared to traditional video data analysis through a controlled experiment. The key findings from this experiment are that the use of the IDA technique assists usability researchers in identifying as much as 85% of the critical problems and 68% of the serious problems of the evaluated system while at the same time reducing the noise of unique cosmetic and serious usability problems significantly. The time required for identifying these usability problems using the Instant Data Analysis technique amounted to only 10% of the time required to do a traditional video data analysis. On the basis of these findings, the Instant Data Analysis technique is a promising approach to reducing analysis effort within the tradition of discount usability.

FURTHER WORK

The findings presented in this paper leave a lot of room for further research. For example, it would be interesting to see if the presented findings are generalizable to data analysis conducted by non usability experts – or if the video data analysis technique produces better results compared to the Instant Data Analysis technique in this situation. Also, in the light of the time pressure characterizing industrial usability studies, it would be interesting to know if it is easier to teach usability novices to sufficiently master the IDA technique compared to teaching them traditional video data analysis methods. In extension of the last issue in the discussion section, it would also be interesting to investigate further into to use of usability evaluation results in subsequent redesign.

ACKNOWLEDGMENTS

The authors thank B-Data, the Hospital of Frederikshavn, the test subjects who participated in the evaluation and Mette Mullerup for introduction to the Booking System and assistance with planning the evaluations. We also thank Rune T. Høegh and Benedikte S. Als who participated in conducting the usability evaluation and analyzing the data.

REFERENCES

1. Gray, W. D. and Saltzman, M. C. Damaged Merchandise? A review of experiments that compare usability evaluation methods. *Human-Computer Interaction 13*, 3 (1998), 203-261.
2. Jacobsen, N. E., Hertzum, M. and John, B. E. The Evaluator Effect in Usability Tests. *Conference summary on Human Factors in Computing Systems, CHI 1998,* ACM Press.
3. Karat, C. M., Campbell, R., and Fiegel, T. Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation. *Proceedings of CHI 1992,* ACM Press, 397-404.
4. Kjeldskov, J. and Skov, M. B. Creating a Realistic Laboratory Setting: A Comparative Study of Three Think-Aloud Usability Evaluations of a Mobile System. *Proceedings of Interact 2003,* IOS Press, 663-670.

5. Kjeldskov, J. and Stage, J. New Techniques for Usability Evaluation of Mobile Systems. *International Journal of Human-Computer Studies 60* (2004), 599-620.
6. Molich, R. *Usable Web Design* (In Danish). Ingeniøren|bøger, 2000.
7. Nayak, N. P. Mrazek D., Smith R.D. Analysing and Communicating Usability Data. *SIGCHI Bulletin* 27, 1 (1995).
8. Nielsen, J. Why you only need to test with 5 users, http://www.useit.com/alertbox/20000319.html (2000).
9. Nielsen, J. Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. In R. G. Bias and D. J. Mayhew (eds.). *Cost-Justifying Usability.* Academic Press Professional, 1994.
10. Nielsen, J. *Usability Engineering.* Boston: Academic Press, 1993.
11. Nielsen, J. Big paybacks from 'discount' usability engineering, *IEEE Software 7,* 3 (1990), 107-108.
12. Nielsen, J. Usability engineering at a discount. In Salvendy, G., and Smith, M.J. (Eds.), *Designing and Using Human-Computer Interfaces and Knowledge Based Systems.* Elsevier Science Publishers, Amsterdam, 1989, 394-401.
13. Nielsen, J. and Molich, R. Heuristic evaluation of user interfaces. *Proceedings of CHI 1990,* ACM Press, 249-256.
14. Omodei, M. M., Wearing, A. J., and McLennan. Head-Mounted Video and Cued Recall: A Minimally Reactive Methodology for Understanding, Detecting and Preventing Error in the Control of Complex Systems. *Proceedings of the 21st European Annual Conference of Human Decision Making and Control.* University of Glasgow, Scotland (2002).
15. Pedell, S., Graham, C., Kjeldskov, J. and Davies, J. Mobile Evaluation: What the Metadata and the Data Told us. *Proceedings of OzCHI 2003,* Australia.
16. Preece J., Y. Rogers H., Sharp D., Benyon D., Holland S. and Carey T. *Human-Computer Interaction.* Workingham, Addison-Wesley, 1994.
17. Rubin J. *Handbook of Usability Testing.* Wiley, 1994.
18. Strauss, A. and Corbin, J. *Grounded Theory in Practice.* Sage Publications, 1998.
19. Virzi, R. A. Refining the test phase of usability evaluation: how many subjects is enough? *Human Factors, 34,* 4 (1992), 457-468.