VI

A Practical Method for Evaluation of Human-Computer Interfaces

CARL ÅBORG [a, b]

BENGT SANDBLAD [c]

MAGNUS LIF [c]

[a] *Futura, Statshälsan Research and Development Occupational Health, St. Olofsg. 9, S-75321 Uppsala, Sweden.*

[b] *Luleå University of Technology, Dept. of Human Work Science, Luleå, Sweden.*

[c] *Center for Human-Computer Studies (CMD), Uppsala University, Sweden.*

Abstract

This paper introduces a method, the ADA-method, for usability evaluation of information systems, used by skilled professionals. It is intended to be used by occupational health care organisations, as a complement to their more traditional methods for investigation of the physical and psycho-social work environment.

The objectives of the developed ADA-method are (i) to identify usability and cognitive work environment problems in computer supported work, and (ii) to be a basis for further analysis and discussions concerning improvements of the system. The intention is to identify major usability problems related to the user's cognitive work environment.

An evaluation with the ADA-method is performed as an observation interview, during which the observer documents potential usability problems.

1. Introduction

1.1. Health issues

Computer technology has a great impact on working conditions, as well as on health and well-being of individuals. A rapidly increasing number of people is using computers in their daily work. In several organisations the operation of computer support systems has become a full-time job for many employees. In the 1970s, reports began to appear about adverse health effects of computerisation, and since then numerous studies have shown that poorly designed VDU-work is associated with a variety of physical and psychological problems. In the 1970s and 1980s the primary emphasis in examining Human-Computer Interaction (HCI) at work was on physical ergonomic aspects and technology design. The studies mainly concerned physical problems, such as eye strain, visual fatigue and musculoskeletal symptoms. During the 1980s also skin problems (e.g., the prevalence of dry skin), stress responses and psychological complaints were studied and discussed (Bergqvist, 1993; Bauer, Aronsson, Åborg, & Örelius, 1990).

The Swedish Foundation for Occupational Health and Safety for State Employees is the organiser of occupational health services for state employees in Sweden. The foundation conducted a study in co-operation with National Institute of Occupational Health, where a total number of 5.771 state-employed VDU-users were examined. Standardised questionnaires were used, containing questions on job content, physical and psycho-social work environment, and mental and somatic health symptoms (Åborg, Aronsson, & Dallner, 1993; Aronsson, Åborg, & Örelius, 1988). The VDU-users were investigated during two periods of two years each. The results showed that a variety of health complaints, especially eye strain, and neck/shoulder problems, were common among VDU-users, and that some of these problems were clearly stress-related. 70% of all subjects claimed that they did not have enough influence on decisions made concerning computerisation (e.g., purchase of equipment or computer programs affecting their own work). In the study the type of VDU-work was classified into five different categories: Data entry, Data acquisition, Interactive communication (involves both data entry and data acquisition), Word processing and Programming/Computer-aided design.

The proportion of people considering the demands on attention and concentration too high were largest in the group working with interactive communication. The personnel in this group who worked for 6 hours or more per day at a computer terminal also reported high frequencies of psychological complaints (45%). These complaints were categorised as follows: (1) Often lacking in concentration; (2) Often restless or tense; (3) Often irritated or impatient; (4) Often anxious, uneasy or nervous; (5) Often depressed, dejected or sad.

According to a transactional stress model, the stress reactions can be caused by an imbalance between job demands and the opportunities for individuals to control and cope with these demands. This model has shown to be useful for understanding and predicting stress reactions and stress-related diseases (Karasek & Theorell, 1990).

The use of computers at work has often increased workload, work demands and the risk of loosing jobs, and decreased personal control and social support (Aronsson, Dallner, & Åborg, 1994; Smith & Carayon, 1993).

There is also evidence that the stress associated with VDU-work may contribute to musculoskeletal problems (Smith & Carayon, 1993). Psychological stress can lead to an increased physiological susceptibility by effecting hormonal and circulatory responses, and to behaviour that increases the risk of musculoskeletal disorders. Mental stress is probably the underlying cause of many of the health-related symptoms suffered by VDU-users. Inappropriately designed computer support creates tension, irritation and aversion. The human-computer interface is one important factor influencing user stress, and subsequently user health.

1.2. Usability

An important concept within HCI is *usability*. In Draft for International Standard ISO/DIS 9241-11 (1995), usability is defined 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". Here, the effectiveness of a system relates to the work objectives (goals), the efficiency relates to effectiveness in relation to the resources needed to perform the tasks. Satisfaction concerns *acceptability* and *comfort.*

Nielsen (1993) has proposed the following definitions:

- *Usefulness* (related to work objectives)
  This relates to the work processes, and how well the system contributes to a good work result (e.g., in quality terms). Usefulness can be divided into:

- *Utility* (related to functionality of the application)
  This relates to if the desired and needed functionality is at hand. In other words: "Is this the correct tool for the work procedures?".

- *Usability* (related to the user)
  This relates to if the user can interact efficiently with the system through its user interface.

Most of the mentioned definitions are intended to be used in the system development process (i.e., by the system developers as a guidance for design and construction of the user interface).

This paper focuses upon usability evaluation of information systems used by skilled professionals in working life. The introduced evaluation method is to be used in occupational health care organisations, as a complement to their traditional methods for investigation of physical and psycho-social work environment. The method is not intended to be used during system development, but for evaluation of running information systems in everyday use. However, the result of such an evaluation may lead to changes in the existing system.

What we need is a definition, and evaluation methods, that are more related to the user's *cognitive work environment* (Lind, Nygren, & Sandblad, 1991). Cognitive work environment problems are caused by limitations in the work environment that hinder the users from efficiently using their skills. Such hindrances are often associated with the human-computer interface. Our definition of usability includes the extent to which the user can interact efficiently with the system without unnecessary mental effort, caused by cognitive work environment problems. Such problems can lead to inefficient work procedures, bad performance and low user acceptance as well as to somatic and mental health symptoms.

1.3. Existing evaluation methods

Methods for evaluating usability can be separated into usability testing methods, that is, where users are involved, and usability inspection methods, that is, where users are not involved.

A traditional method of usability testing is *performance measurement*. The purpose is here to measure whether a usability goal is reached or not. User performance is almost always measured by having a group of test users perform a pre-defined set of tasks while collecting data on errors and times. The tests are usually carried out in a laboratory. With such a test many usability problems will be found. One advantage of this test method is that the result is given in hard numbers which makes comparison of different design solutions easy. Unfortunately there are seldom enough time, money or laboratory expertise available to use this kind of method (Nielsen, 1993). Difficulties in sampling, methodological problems in planning, validity and reliability of obtained measures are other pitfalls in usability testing (Holleran, 1991).

Questionnaires are useful for issues concerning users' subjective satisfaction and possible anxieties, but are less useful for other usability aspects (Nielsen, 1993).

Questionnaires may be distributed to many users and is an inexpensive survey method.

Thinking aloud (Lewis, 1982) is a method where the users verbalise their thoughts while using the system. Through this test, users let the usability expert understand how they view the computer system. This is an inexpensive test that identifies users' misconceptions of the system. Drawbacks with this method include that it is not very natural for users to think out loud and it is difficult for expert users to verbalise their decision process. Expert users execute part of their work automatically (Schneider & Shiffrin, 1977; Shiffrin & Dumais, 1981). Therefore, it is difficult to capture usability problems concerning efficiency in daily use.

One method that includes users, developers and usability experts, and may be carried out early in the design process is *pluralistic walkthrough* (Bias, 1991). Representatives from the three categories meet and discuss usability problems that are associated with the dialogue elements in different scenario steps. The main focus is on how users react in different situations (e.g., a user may claim that, in a certain situation, he or she would "Hold down the shift key while pressing Enter"). Pluralistic walkthrough is an effective method for evaluating the learnability of a user interface. However, it is not as effective for evaluation of interfaces in daily use since the users are not able to predict how they will interact with the system when they become skilled.

There are also several different inspection methods available. One such method is *cognitive walkthrough* (Polson, Lewis, Rieman, & Wharton, 1992). With this method an evaluator examines each action in a solution path and tries to tell a credible story describing why the expected user would choose a certain action. The story is based on assumptions about the users' background, knowledge and goals, and on understanding the problem solving process that enables a user to guess the correct action. The method focuses on evaluating ease of learning, particularly by exploration. Therefore, this method is not applicable regarding inspection of interfaces for skilled users.

Another inspection method is *heuristic evaluation* (Nielsen & Molich, 1990). The evaluator uses sets of guidelines (i.e., heuristics) and compares those with the interface. The heuristics form a checklist that the evaluator uses during his work. It is easy to learn and inexpensive to use. A drawback is that evaluators using this method seldom manage to identify domain specific usability problems due to lack of domain knowledge.

Recently a series of methods for measuring usability has been developed in the ESPRIT MUSIC project (Corbett, Macleod, & Kelly, 1993). The usability of a product is defined through analytic measures, performance measures, cognitive

workload measures and user attitude measures. Analytic measurements are performed early and are based on a dynamic model of the user interface and on the user tasks. It estimates performance parameters for human interaction dependent on the use of specific interface objects. Performance measurement can be enhanced by using the DRUM tool for analysis of video recording. Cognitive workload is measured through heart rate variability and respiration and subjectively by the use of questionnaires. Questionnaires are also used to measure the user attitude. This is an extensive method that can be used for a number of evaluations. However, these methods are intended to be used by human factor experts and not by Occupational health experts with limited knowledge in HCI.

The *software checker* (TCO, 1992) is a rather simple method for evaluation of software. The method includes a number of questionnaires to be answered by the user. With this checker the user is able to examine the capabilities of the software to achieve intended goals, the effects the software will have on work routines and on the organisation, the ergonomic qualities and the training. The result of the evaluation is the basis for choosing a certain software or not. This checker is easy to use and helps the users to identify shortcomings of different software. A drawback is that many questions are difficult to answer with a simple yes or no, since they often are general to their nature.

Other methods for evaluation of office work are built on German theories on human activity and the design of work tasks (Bokander, 1992). These "Action Regulation Theories" describe the basic characteristics of human activity and derive from them a number of principles and criteria for designing work tasks. To analyse work tasks on the basis of these principles a number of methods and instruments have been developed.

RHIA-instruments measure regulation hindrances leading to negative mental load. The most well known method based on Action Regulation Theory is the *VERA-instrument*. It is used to determine the "scope of action" or the regulation requirements for a specific task, the degree to which a worker can make autonomous plans and decisions at the work place. Still another instrument with the same theoretical background has been developed to give guidance when deciding which part of a work activity that should be computerised and which should not. It is called the *KABA-method* ("Contrastive Task Analysis") (Dunckel, 1989).

These methods are of great help when identifying shortcomings in workers' situation from a psychological point of view. They do not, however, give much support for identifying shortcomings in the user interface.

1.4. The need for new evaluation methods in occupational health

The method introduced here is intended to be used by occupational health care organisations (OHC). Almost all employees in Sweden are entitled to occupational health care services, through local health care centres all over the country. At the centres there are medical, ergonomic, technical and psycho-social experts (e.g., psychologists). OHC-personnel have a long tradition of investigating and describing work environment, health and safety factors. Often the aim of the investigation is to find indications of causal relationships between work environment factors and health and well-being reactions. Sometimes the purpose is to produce a general description of the health and safety situation in a company or an organisation.

A main goal to all OHC-organisations is to prevent health and safety problems and improve health and well-being. Therefore it is necessary to identify risk-factors and risk-situations at an early stage, before having caused accidents or medical disorders. The quality of life, health, well-being and satisfaction of the employees is closely related to the effectiveness and productivity of the organisation. In a work environment investigation it is also important to identify factors leading to reduced effectiveness. Investigation methods are primarily based on questionnaires, interviews and observations. Different methods are often combined. There are several well tested and frequently used tools and methods for eveluating both physical and psychological work environment factors. However, practical methods to be used by OHC-experts for evaluation of human-computer interfaces are still lacking.

The ADA-method, presented in this paper, follows this well established tradition of occupational health and safety work, with the aim of adding a new, important issue. There are a number of situations where the ADA method can be applicable. A general work environment investigation can lead to hypotheses of connections between the human-computer interface and observed health reactions. The ADA-method can be then used in a second study where these hypotheses are tested. The ADA evaluation complements the earlier findings concerning work environment, work organisation and somatic and mental health complaints. A more specific study that aims at investigating the relations between work organisation, work content and stress reactions can benefit from using the ADA-method if operating computers is part of the work. The ADA-method can also be used to identify and analyse users problems with a specific computer system.

2. The ADA method

2.1. Objectives

The objectives of methods for evaluation of usability are to make sure that the application under study allow the users to perform their work efficiently and that the cognitive work environment is acceptable. The specific objectives of the ADA-method are to:

- identify problems in computer supported work that are related to bad functionality and to the cognitive work environment,
- be a basis for further analysis and discussions concerning improvements of the information system,
- identify the most important problems, not necessarily all possible ones.

The ADA-method is based on a mixture of observations, interviews and questionnaires (Figure 1).

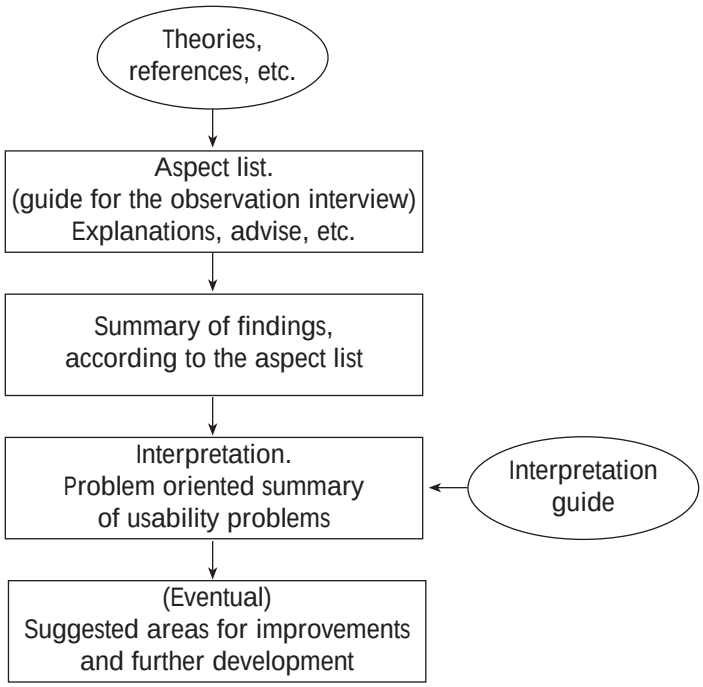

Figure 1. The general structure of the ADA-method.

2.2. The evaluation procedure

The evaluator performs observation interviews with users during their ongoing work with the information system. If necessary the interview is completed after the observation period. The interview is based on an *interview guide* which contains a list of usability aspects and advises for how to perform the observation interviews. The findings from the observation interview are interpreted using an *interpretation guide*. The conclusions are later presented for both the users and the management, as a basis for a dialogue concerning future improvements of the computer system.

The evaluators (i.e., the experienced OHC experts), should be able to use the method after a two days tutorial and some practical experiences.

2.3. The aspect list

A central part of the method is the aspect list, including the usability aspects that the method is intended to cover. The aspects and the explanations are mainly a collection of existing material, found in literature and in our own earlier research. Important bases are studies of work conditions and health of VDU-users (Aronsson et al., 1994), standards (e.g. ISO, 1995), psychological controlled experiments (e.g. Nygren, Allard, & Lind, In press), field studies of work activities (e.g. Nygren & Henriksson, 1992) and participation in development projects (e.g. Borälv, Göransson, Olsson, & Sandblad, 1994). Headlines of the aspect list are:

1 The role of the interviewed/observed person

2 Work tasks and work organisation

3 Functionality of the information system

4 Structure and technology of the computer system

5 Competence and rules for usage

6 Accessibility and authority

7 Training, introduction and changes

8 Manuals, help, support and guidance

9 System functions:

  9.1 Response times

  9.2 Control

  9.3 Error controls and tolerance

10 User interface:

- 10.1 Type of interface
- 10.2 Disposition of screen area
- 10.3 Menus, levels
- 10.4 Orientation
- 10.5 Parallel (simultaneous) presentation of information
- 10.6 Input functions
- 10.7 Control
- 10.8 Form, font, etc.
- 10.9 Use of colours
- 10.10 Icons
- 10.11 Feedback functions

11 Subjective judgements

12 Others

2.4. The interview guide

The interview guide contains more detailed information on the aspects to be observed and discussed. The aspects are divided into sub-areas. Each sub-area is further explained through a set of questions to be answered by the observer and/or by the observed person. Finally, some general remarks and practical advises are given. Two examples of such aspect specifications are given below:

Example: 10.5 Interface. Parallel information presentation.

- Is all information required to perform a task simultaneously available?

  Is enough information always provided simultaneously to successfully manage the complete task?

  Does the user have to switch between different views or windows?

  Are there many windows related to one task?

  How is the switching between sequential windows performed? Simple or demanding manipulation?

It is important that all information required to accomplish a task is simultaneously presented on the screen. Having to consider which information that is needed, switching between different windows, and finally integrating the information is one of the most common sources for unnecessary cognitive load. It often leads to overload of the short term memory, slow performance and a high error rate.

Example: 10.6 Interface. Input/editing functions

- How is information entered into the system?

  Which technology is used? Keyboard? Mouse?

  Switching between different technologies?

  Automatic? Keystrokes? Arrow keys? Function keys? Pointing and clicking with the mouse?

- How much information is entered into the system?

- What kind of information is entered?

  Text, numeric values, selection between pre-defined values?

- Are there any limitations in what kind of and how much information that can be entered?

  Are the limitations distressing? Do they bring about additional work?

- Are there different kinds of validations integrated in the system?

  Are the rules automatically verified? What kind of feedback is provided in an error situation? Is the feedback plain and obvious? Is it easy to make corrections? The questions here concern the ability of the interface to utilise such possibilities.

In many situations the work task consists of entering larger or smaller amounts of information. Such functions should be flexible, with a minimum of typing and cognitive load. The concentration should be focused on the task contents and the accuracy of the information.

2.5. The evaluation report

The findings from the observation interview are summarised and documented with support from the interpretation guide. This guide helps to structure the text material according to the following head-lines:

- Work task aspects, related to the functionality of the system.
- System aspects which can cause usability problems.
- Cognitive load of different nature.
- The user's control and possibilities to influence changes.
- Knowledge and competencies of the user.
- Subjective experiences and problems.

The findings in the summary from the observation interview are analysed, and the corresponding interpretation in problem terms are noted for each head-line. The result is the interpretation report of possible usability problems.

2.6. How to use the method

The ADA-method is to be used by OHC personnel, after a short period of training, as a part of their investigations of work environment and health in VDU-work. An investigation by an OHC-unit normally follows a plan which includes the following processes:

1. Clarifying and reaching agreement on the purpose and the overall plan of the investigation.
2. Planning the activities and the time schedule.
3. Informing employees, managers and other interested parties.
4. Data collection.
5. Analysing the results.
6. Presentation of results to those involved.
7. Evaluation.

When applying the ADA-method an overall plan should be clarified. The information activities should include both written and oral information to all those involved. The data is collected during the observation interviews. Before the results are presented, the findings are discussed with the individuals interviewed and their supervisors. This can be performed in a group discussion to obtain the users spontaneous reactions and ideas, get feedback on the results and gather supplementary data. A written report is produced according to the ADA report model. In this report the findings and results of the observation-interviews are summarised in a problem-oriented way. The purpose of the report is to provide a basis for a dialogue with employees, employers and project managers on problems and possible remedies in order to improve the computer system. The emphasis is on describing problems which may lead to unnecessary cognitive load and cognitive work environment problems. Possible remedies are not described in detail. The purpose of this method is not to define how to redesign the systems. That will be a separate project, after the ADA-evaluation.

The evaluation is performed at the users' workplace, and it considers not only the software, but also the specific tasks and the organisational context. The time required for one observation-interview is approximately 2 hours, and 2 more hours are needed to summarise the interview. Analysing the results and writing the report will take 3-4 hours. This means that approximately 2 working days are needed to conduct an ADA-evaluation including 3 interviews. Since the method is rather time consuming it is seldom possible to involve every individual in a user group. Consequently, interviewees has to be selected that represent different categories of users, e.g. users with different expertise. Users without prior computer experience should not be selected, since they have problems and needs not related to a specific computer system.

The observation interviews may be complemented with a questionnaire dealing with some of the aspects in the ADA-method to find out which problems that are common among several users. Such a questionnaire can be distributed to large numbers of users; to entire user populations, to big samples, or to many different groups of users. Questionnaires distributed to a large number of users, and interviews with a small sample of that population is often a fruitful combination.

3. Evaluation of the method

Today the method has been used at more than ten different workplaces. The evaluation described here is based on three cases: one new information system used by the Swedish National Tax Board, one system for telephone-booking of tickets at an airline company and, finally, one system for appointment-booking at an OHC centre.

The purpose of this evaluation was to investigate if two evaluators get similar results when evaluating the same system with the ADA-method.

3.1. Procedure

Three different systems were evaluated. The ADA-method was used by two evaluators for each system. The evaluators performed three observation interviews each at the different offices. All observation interviews were made on different individual users of the analysed information system. The evaluators were instructed to document the findings according to the ADA-method. These findings were then analysed by a human-factors expert. All findings (potential usability problems) were categorised and given a weight; 0 for not relevant, 1 for important and 2 for very important.

In total there were four evaluators performing the evaluations. Two of these were Occupational Health experts and two were usability experts. They evaluated the systems according to the following schedule (Table 1):

	Evaluator 1	Evaluator 2	Evaluator 3	Evaluator 4
Office 1	X	X		
Office 2	X		X	
Office 3			X	X

Table 1. Two evaluators evaluated each system. In total there were four evaluators performing the evaluations.

3.2. Results

Table 2 shows the number of findings identified by the two evaluators at each workplace. The number of findings given a weight 1 and a weight 2 is presented. The findings given a weight 0 are not listed. The table shows the total number of findings identified by the two evaluators and the number of findings that were discovered by both evaluators. "Total, weight 1" means the total number of findings with weight 1 identified in the evaluation. "Same, weight 1" means the number of findings with weight 1 that were identified by both evaluators.

	Office 1	Office 2	Office 3
Total, weight 2	13	19	19
Same, weight 2	11	17	15
Total, weight 1	20	18	10
Same, weight 1	6	6	3

Table 2. Number of weighted findings identified during the evaluations.

Table 3 shows the number of findings identified by both evaluators divided by the total number of findings identified at each workplace in percentage. This has been derived for both findings with weight 1 and weight 2. The table also shows the mean value and the standard deviation when comparing the result from the three evaluations.

	Office 1	Office 2	Office 3	Mean	Std.dev.
Same/total, weight 2 (%)	84	89	79	84	5,27
Same/total, weight 1 (%)	30	33	30	31	1,92

Table 3. Percentage of total number of findings that were identified by both evaluators in each evaluation.

Table 4 shows the frequencies of findings with weight 2. "Same, weight 2" means the number of findings of weight 2 that were identified by both evaluators. "Different, weight 2" means the number of findings of weight 2 that only were identified by one of the evaluators. Table 5 shows the frequencies of findings with weight 1.

	Office 1	Office 3	Total
Same, weight 2	11	15	26
Different, weight 2	2	4	6
Total	13	19	32

Table 4. Number of findings with weight 2.

	Office 1	Office 3	Total
Same, weight 1	6	3	9
Different, weight 1	14	7	21
Total	20	10	30

Table 5. Number of findings with weight 1.

In a χ^2-test the results from office 1 and office 3 was compared. The 0-hypothesis was that the degree of findings discovered by both evaluators in each evaluation should be different when using the ADA-method for different systems. The test for findings with weight 2 was not significant (χ^2 = 0,163, df=1). The test for findings with weight 1 was also not significant (χ^2 = 0, df=1).

Utilisation

It is rather difficult to measure the actual validity of the method since that would require some kind of "correct" answers. However, what we can do is to judge the response on the methods' utility, from the "consumers" of the method, i.e. decision-makers and users at the studied work places. The method is supposed to lead to

usable knowledge, new insights and, if necessary, to specific actions to solve particular problems.

The manager responsible for the development of the studied system at the three different offices were also the persons who had ordered the ADA-evaluation. These persons were questioned about the utilisation of the results of the ADA evaluation.

- They all found the method usable in practice
- They all found the results of the evaluations helpful in solving specific problems

All of the respondents gave several examples of new understanding and specific corrective actions based on the "ADA findings". Some examples:

- the need for better printout facilities was emphasised. This was fulfilled in the new version of the program.
- problems with the overview were identified. This was regarded in the new system.
- the suggestions on how to use fonts and colours to highlight information was used in the new version.
- aspects of the software and cognitive load were considered when an investigation was performed concerning the users' eye problems

Some examples of quotes from the three clients:

"We used the results from the ADA evaluation when we planned improvements of the new version of the software...The risk of building a system that includes shortages for the user that we are not aware of decreases. "

"First of all, it feels good to let someone who has not been involved in the project judge the system...Secondly, I found it useful to get your opinion about what you regarded as major and minor problems in the system"

"The wrong causes for problems are identified to often. When Occupational Health experts have studied health problems, they usually find causes in the physical environment. The ADA method can help identifying causes in the software product. Just the fact that the users' are able to speak to someone from the outside can help solving the problems."

The utility of the method has also been discussed with the users and the people responsible for systems development at the organisations where the observations were made. The results of these discussions indicated that the observation interviews pointed out a number of new aspects concerning usability problems not previously known. It is our experience that such discussions after the summary of the interviews are very important. In this dialogue a more complete list of problems that should be solved can be obtained, and the priorities of the findings can be defined.

4. Discussion

Table 2 and 3 shows that a large number of the findings discovered during the evaluations were identified by both evaluators. This is especially true for the findings that were regarded as very important. Table 2 shows that 84 % (Std.dev. = 5,27 %) of the total number of very important findings were identified by both evaluators. 30 % (Std.dev. = 1,92 %) of the important findings were discovered by both evaluators.

According to the χ^2-test the degree of findings discovered by both evaluators was more or less the same when using the ADA-method at the different workplaces. The test indicates that the reliability of the method is good.

The results show a difference in "user agreement" for findings with weight 1 as compared to findings with weight 2. One reason for this could be that the evaluators have given priority to very important findings and not documented all findings of less importance. Another reason could be that the number of potential findings of less importance is much larger than the number of potential findings that are very important.

Different observers will to some extent identify different findings, but the very important findings are likely to be found. Better results when using the ADA-method can probably be obtained if more than two persons perform each evaluation.

The method has been taught to OHC experts during some tutorials. Primary tests indicate that novice users of the method are able to identify most of the major problems discovered by experts. However, novice users are more likely to document a larger number of less important findings. In the future we will perform studies to test these preliminary results.

The study of the utilisation of the method indicates that the ADA-method is a useful input for improving the evaluated information systems.

The presented method for usability evaluation has a specific purpose and must be judged in relation that. The ADA method is developed for use in connection with OHC investigations. It is only intended to identify major usability problems and can not be compared to more comprehensive investigations performed by HCI experts. However, not all computer applications in working life can be evaluated by highly qualified usability experts. Since computer support is becoming more and more common, and since occupational health care investigations are rather common, at least in Scandinavia, this method fulfils an important purpose.

Acknowledgements

Financial support was provided by the Swedish Council for Work Life Research.

References

ARONSSON, G., DALLNER, M., & ÅBORG, C. (1994). Winners and losers from computerisation. A study of the psychosocial work conditions and health of Swedish state employees. *International journal of human-computer interaction, 6* (1), 17-35.

ARONSSON, G., ÅBORG, C., & ÖRELIUS, M. (1988). *Datoriseringens vinnare och förlorare* [Winners and losers from computerisation. Summary in English]. Arbete och hälsa. Arbetsmiljöinstitutet, 27.

BAUER, H.G., ARONSSON, G., ÅBORG, C., & ÖRELIUS, M. (1990). Health aspects of work with VDU-s. In L. Berlinguet, & D. Berthelette (Eds.), *Work with display units -89* (pp. 283-290). Amsterdam: Elsevier Science Publishers.

BERGQUIST, U. (1993). *Health problems during work with visual display terminals.* Arbete och hälsa, Arbetsmiljöinstitutet 1993:28.

BIAS, R.C. (1991). Walkthroughs: Efficient Collaborative Testing. *IEEE Software, 8* (5), 94-95.

BOKANDER, I. (1992). *Arbetsanalys av kontorsarbete* [Analysis of office work]. Lund, Sweden: Psykologiska Institutionen, Lunds universitet.

BORÄLV, E., GÖRANSSON, B., OLSSON, E., & SANDBLAD, B. (1994). Usability and Efficiency. The HELIOS Approach to Development of User Interfaces. In U. Engelmann, F.C. Jean, & P. Degoulet (Eds.), *The HELIOS Software Engineering Environment, Supplement to Computer Methods and Programs in Biomedicine, 45*, 63-76.

CORBETT, M., MACLEOD, M., & KELLY, M. (1993). Quantitative Usability Evaluation – The ESPRIT MUSiC Project, In G. Salvendy, & M.J. Smith (Eds.), *Proceedings of the Fifth International Conference on Human-Computer Interaction, HCI International' 93* (pp. 313-318). Amsterdam: Elsevier Science Publishers.

DUNCKEL, H. (1989). Contrastive task analysis. In K. Landau, & W. Romert (Eds.), *Recent developments in job analysis* (pp. 125-136). London: Taylor and Francis.

HOLLERAN, P.A. (1991). A methodological note on pitfalls in usability testing. *Behaviour & information technology, 10* (5), 345-357.

ISO/DIS 9241 (Draft). Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 10: Dialogue Principles, Part 11: Guidance on Usability, Part 13: User Guidance (1995). Geneva, Switzerland: International Organization for Standardization.

KARASEK, R. & THEORELL, T. (1990). *Healthy work: Stress, productivity and reconstruction of working life.* New York: Basic books.

LEWIS, C. (1982). *Using the 'thinking-aloud' method in cognitive interface design.* (IBM Research Report RC 9265, 2/17/82). Yorktown Heights, NY: IBM T.J. Watson Research Center.

LIND, M., NYGREN, E., & SANDBLAD, B. (1991). Kognitiva arbetsmiljöproblem och gränssnittsdesign [Cognitive work environment problems and design of user interfaces] (Rep. No. 20, CMD). Uppsala, Sweden: Uppsala University.

NIELSEN, J. (1993). *Usability Engineering.* San Diego: Academic Press, Inc.

NIELSEN, J., & MOLICH, R. (1990). Heuristic evaluation of user interfaces In J.C. Chew, & J. Whiteside (Eds.), *Proceedings of Human Factors in Computing Systems, CHI' 90* (pp. 249-256). New York, NJ: ACM.

NYGREN, E., ALLARD, A., & LIND, M. (In press). Skilled Users' Interpretation of Visual Displays. *Human Computer Interaction.*

NYGREN, E., & HENRIKSSON, P. (1992). Reading the Medical Record 1. Analysis of Physicians Ways of Reading the Medical Record. *Computer Methods and Programs in Biomedicine, 39*, 1-12.

POLSON, P.G., LEWIS, C., RIEMAN, J., & WHARTON, C. (1992). Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces. *International Journal of Man-Machine Studies, 36* (5), 741-773.

SCHNEIDER, W., & SHIFFRIN, R.M. (1977). Controlled and Automatic Human Information Processing I, *Psychological Rev., 84,* 1-66.

SHIFFRIN, R.M., & DUMAIS, S.T. (1981). The Development of Automatism. In J.R. Anderson (Ed.), *Cognitive Skills and their Acquisition.* Hillsdale, NJ: Erlbaum.

SMITH, M., & CARAYON, P. (1993). A balance model for examining psychological stress in VDU work. In H. Luczak, A. Cakir, & G. Cakir (Eds.), *Work with display units -92* (pp. 35-39). Amsterdam: Elsevier Science Publisher.

TCO, THE SWEDISH CONFEDERATION OF PROFESSIONAL EMPLOYEES. (1992). *Software checker. An aid to the critical examination of the ergonomic properties of software. Version 2.0.* Stockholm, Sweden: Informgruppen AB.

ÅBORG, C., ARONSSON, G., & DALLNER, M. (1993). Work organisation and health aspects of work with VDU:s. In Luczak, H., Cakir, A. Cakir, G. (Eds.), *Work with display units -92* (pp. 359-363). Amsterdam: Elsevier Science Publisher.