SINTEF REPORT

SINTEF ICT

Address: NO-7465 Trondheim, NORWAY
Location: Forskningsveien 1
Telephone: +47 22 06 73 00
Fax: +47 22 06 73 50

Enterprise No.: NO 948 007 029 MVA

TITLE	

Evaluating user-centred methods: Results from a survey study among Nordic HCI-practitioners

AUTHOR(S)

Ida Bark, Asbjørn Følstad, Jan Gulliksen

CLIENT(S)

EFFIN research project, financed by the Norwegian Research Council through the FIFOS program

REPORT NO.	CLASSIFICATION	CLIENTS REF.	
STF90 A05129	Open	NRC reference number: 153048/530	
CLASS. THIS PAGE	ISBN	PROJECT NO.	NO. OF PAGES/APPENDICES
Open	82-14-03810-3	403318.04	25
ELECTRONIC FILE CODE		PROJECT MANAGER (NAME, SIGN.)	CHECKED BY (NAME, SIGN.)
eRoom/EFFIN/WP-04/Survey among HCI practitioners		John Krogstie	Jan Heim
FILE CODE	DATE	APPROVED BY (NAME, POSITION, SIGN.)	
	2005-06-29	Bjørn Skjellaug, Research Director	

ABSTRACT

One goal of the research project Efficiency through user involvement (EFFIN) is to develop a procedure to evaluate user-centred activities with regard to their efficiency in a system development process. In order to develop an evaluation procedure that is relevant for the practitioners of HCI (Human-Computer Interaction) we need to know more about practitioners' views on: (1) desired outcomes of user-centred activities; (2) ideal and feasible ways to evaluate user-centred activities, and (3) Usefulness and use of existing methods and practices.

The study was conducted as a questionnaire survey. The survey included 277 respondents, mainly from the Nordic countries.

The HCI-practitioners were found to focus more on getting high-quality results from their user centred activities; and less focused on the integration of the results in the development process. This implies a need for an evaluation procedure that (1) allows the practitioner to get feedback on the quality of their results, while (2) the practitioner is motivated to focus more strongly on the utilization of the results in the development process. High quality results are not useful unless they are used.

In the earlier phases of development projects the HCI practitioners are not always able to use the methods they find most adequate. Field studies are conducted far less, and low-cost evaluation methods are conducted more, than their perceived usefulness should account for. An evaluation procedure should enable the practitioner to investigate and explicate the assumed gain in usefulness when substituting a low-cost method with a more resource demanding (and more adequate) method.

The survey results also show a fairly high maturity in among HCI practitioners with regard to evaluation of user-centred activities. This will hopefully facilitate the dissemination of the evaluation procedure to be developed in the EFFIN-project.

KEYWORDS	ENGLISH	NORWEGIAN
GROUP 1	ICT	IKT
GROUP 2	HCI	Menneske-maskin interaksjon
SELECTED BY AUTHOR	Survey	Spørreskjemaundersøkelse
	Method evaluation	Metodeevaluering

TABLE OF CONTENTS

1 Introduction

In the research project Efficiency through user involvement (EFFIN) we aim to develop a procedure to evaluate user-centred activities with regard to their efficiency in a system development process. In order to develop an evaluation procedure that is relevant for the practitioners of HCI (Human-Computer Interaction) we need to know more about practitioners' views on:

(1) Desired outcomes of user-centred activities
(2) Ideal and feasible ways to evaluate user-centred activities
(3) Usefulness and use of existing methods and practices.

The study was conducted as a questionnaire survey. The survey included 277 respondents, mainly from the Nordic countries.

2 Previous research

2.1 Desired outcomes of user-centred activities

"Cost-justifying usability" was one of the earliest approaches to the evaluation of user-centred methods and work practices (e.g. Bias & Mayhew, 1993; Karat, 1997; Mantei & Teorei, 1988). This tradition has been important in shaping common knowledge in the HCI field with regard to beneficial results of usability. Typical beneficial results advocated in this tradition include:

- Increased end-product usability
  - o Increased quality in use
  - o Reduction of user error
  - o Increased work effectiveness and efficiency
  - o Ergonomic improvement
  - o Lowered costs associated with product introduction, customer support, and system maintenance
  - o Increased sales
- Increased development process efficiency
  - o Early identification of errors and usability problems
  - o Common view of the product development across users and stakeholders
  - o Reduced development time and cost
  - o Increased user and customer satisfaction

2.2 Evaluating user-centred activities

The "cost-justifying approach" has been instrumental in developing the view that user-centred activities may be evaluated as a function of their associated costs and benefits; calculated either by estimation or empirical studies. Basic costs and benefits can be translated into return on investment figures so that the value of usability engineering can be compared with other activities.

The cost-justifying approach provides knowledge that may be valuable when "selling in" user-centred activities to a reluctant project manager or customer. However, cost-benefit ratios at times indicate almost ridiculously high return on investments (see e.g. Nielsen & Gilutz, 2003), and there is little precision with regard to which usability methods that are used. Therefore it may be difficult to utilize literature on cost-justification as support for choosing between different user-centred methods.

The methods of user-centred activities has been evaluated in a more systematic and less biased manner in what may be termed the "reductionist approach" (e.g. Gray and Saltzmann, 1998; Law

and Hvannberg, 2002; 2004). According to the researchers that may be labelled "reductionist", user-centred methods should be evaluated according to the quality of the results generated from the method. Quality of the results may be evaluated with regard to:

- Reliability; the method produces the same results every time it is used, all relevant aspects of the context unchanged
- Thoroughness; the number of identified issues (e.g. predicted user problems) in relation to the number of real issues (e.g. real-life user problems)
- Validity; the number of identified real issues (e.g. identified real-life user problems) in relation to the number of false positives (e.g. predicted user problems that does not correspond with real-life problems)
- Effectiveness; the combination of thoroughness and validity (since high thoroughness alone may result in a method that produces a high number of non-real issues, whereas high validity alone may result in a method identifying only a low proportion of real and important issues)
- Cost-effectiveness; the combination of effectiveness and cost
- Downstream utility; the utilization of the results from user centred methods in the overall development process.

The reductionist approach has been criticized by e.g. Wixon (2003) for not being sufficiently oriented towards evaluating user-centred methods as part of a development process.

2.3 Usefulness and use of user-centred methods

Existing survey studies have provided interesting insights with regard to the use and usefulness of user-centred methods and work practices.

Integration of user-centred design in the industry was investigated in a survey by Venturi and Troost (2004). They found that the most frequently used methods were interviews, high and low fidelity prototyping, expert evaluation, "quick and dirty" usability test and observation of real usage.

Similar results were found in a survey on usability professionals in Sweden (Gulliksen et al., 2004). In this study the highest rated methods (1) involved users, (2) were comparatively informal, and (3) were concerned with design issues. High-rated methods included "Think-aloud", low fidelity prototyping, interviews, field studies and scenarios. The study also indicated that successful usability work is characterized by being …

- part of the project plan from the start
- supported from the project management, overall management, and users
- accepted by the software developers.

A survey study by Vredenburg et al. (2002) conclude that informal and less structured methods tend to be used more widely than formal and structured methods. In the study, informal usability testing, low-fidelity prototyping and heuristics received the highest rankings. Formal methods like cognitive walkthrough received the lowest ranking.

3 Objectives

The main objectives of the present survey study were to gain knowledge of typical perceptions within the community of HCI-practitioners with regard to the following three issues:

1) Desired outcomes of user-centred activities

User-centred activities may result in many beneficial results. In order to have the right focus when investigating the effectiveness of user-centred methods, it is important to know which results the

practitioners see as important. Activities based on analysis methods, design-input methods, and evaluation methods will have different desired outcomes. Thus, in the survey we differentiated between these three method-categories.

2) Ideal and feasible ways to evaluate user-centred activities
Practitioners' perceptions on ideal ways to evaluate user-centred activities are important to get an adequate vision for the development of an evaluation procedure. At the same time we need to know the existing maturity level with regard to evaluations of user-centred activities in the practitioner population, in order to develop an evaluation procedure that will be judged as feasible.

3) Usefulness and use of user-centred methods
- Which user-centred methods are perceived to be the most useful by practitioners?
- To what extent are the different methods actually used by the practitioners?
- Are HCI-practitioners "optimally" involved in development projects?
  - o Are the practitioners involved in the project phases they regard to be most crucial?
  - o Are the practitioners actually using those methods perceived to be the most useful?

Knowledge of use and perceived usefulness will enable us to establish a baseline for use and perceived usefulness of existing methods, and may be helpful to practitioners when choosing methods for their own development projects. If we know how useful a method is perceived to be and also know how much the same method is used, it is possible to say something about practitioners' level of freedom in development projects. If practitioners have only a limited degree of influence and freedom in different project phases, this will be visible in a lack of match between how much methods are used and their perceived usefulness. If practitioners are in full control of their project activities, the match between reported use and perceived usefulness should be almost perfect.

4   Method

A questionnaire survey was chosen as method in order to enable data collection from a large sample of the population of HCI-practitioners. The population was restricted to include mostly practitioners from the Nordic countries.

4.1 Participants

The precise extent of the HCI-population is not known, thus a random sample of participants is difficult to achieve. Convenience sampling from HCI conferences and HCI organization mailing lists were chosen as a suitable means to get a reasonably representative sample. To be included in the survey analysis the participant had to reported that (1) his role in projects was usually "Usability professional" and/or "UI designer" and that (2) he had two or more years of user-centred work experience. The main target of the survey was the Nordic population of HCI-practitioners, but a minor number of other nationalities were included in the final sample.

4.2 Recruitment

The survey was distributed at three conferences: Yggdrasil'04 (Lillehammer, Norway. 56 respondents), NordiCHI'04 (Tampere, Finland. 86 respondents), and STIMDI'04 (Gothenburg, Sweden. 24 respondents). The survey was also distributed as a web-survey (112 respondents). Invitations to the web-survey were sent via e-mail to persons recommended by the respondents at the three conferences, and to the mailing lists of the HCI organizations STIMDI (Sweden), Danish SIGCHI, and Icelandic SIGCHI.

As incentive the participants were offered to be included in a raffle with two Apple iPod mini as prizes. The participants were also offered information on the survey results via the project's home page.

4.3 Material

The survey consisted of 27 items. It was distributed as pen-and-paper questionnaires at the conferences, and as a web-based questionnaire to the other participants. The questionnaire items were developed on basis of a pre-study consisting of semi-structured interviews with eight HCI-practitioners and piloted on a small number of respondents.

4.4 Analysis

The first step of the analysis was to filter out unwanted respondents. All respondents (1) reporting less than two years user-centred work experience, or (2) not reporting to be working mostly as usability specialists or UI designers, were filtered out. Participants providing only partial responses were also filtered out. Analyses of differences between groups or relationships between variables were conducted as visual inspections of graphical representation of descriptive analyses. To illustrate the magnitude of some of the relationships bivariate rank correlation analyses (Spearman's rho) were conducted.

4.5 Start, mid and end-phases of development projects

In the pre-survey interviews with HCI-practitioners the interviewees expressed that they were not comfortable with assigning the user-centred methods they used to phases as described in e.g. ISO 13407 (1999). Rather, they preferred too discuss user-centred methods as belonging to either the start, mid, or end phase of a project. Because of this, we chose to categorize the methods included in the questionnaire according to one or more of three different phases.

- Start phase - including analysis methods. Important activities include project planning, analysis, and specification.
- Mid phase - including design-input methods. Important activities include design, implementation, and early evaluation.
- End phase - including evaluation methods. Important activities include evaluation of the late and final versions of the system, deployment, and acceptance of delivery.

4.6 User-centred methods included in the survey

The methods included in the survey were selected based on the sources of Maguire (2002) and Usabilitynet (2005). The list was completed based on the HCI-practitioner interviews conducted as part of the survey development. The methods are presented in Table 1.

	Analysis	Design input	Evaluation
	(Start phase)	(Mid phase)	(Mid-and end phase)
Questionnaire surveys	X		X
Interviews	X		X
Field studies	X		
Workshops	X		
Personas	X		
Scenarios	X		
Use cases	X		
Focus groups	X	X	
Task analysis	X		
Rapid prototyping		X	
Advanced prototyping		X	
User tests			X
Design patterns		X	
Card sorting		X	
Storyboarding		X	
Guidelines		X	
Heuristic evaluation			X
Cognitive walkthrough			X
Expert evaluation			X
Evaluation workshop			X

Table 1. User-centred methods included in the survey

No detailed method descriptions were provided in the questionnaires. It was assumed that the HCI community has a fairly consolidated understanding of the different methods. (Possible exceptions from this may be the methods of focus groups and workshops, which, from the authors' experience, may be interpreted somewhat differently among different HCI-practitioners).

5 Results

A total 277 responses were received. Only participants that answered that their project role usually was "Usability professional" and/or "UI designer" and that their user-centred working experience was more than two years were considered. Responses of 179 participants made it through this filter. All results presented below are from this sample of 179 responses.

5.1 Respondents' Profile

The participants were asked which role they usually have in a project. The different alternatives they could chose from were; "Usability professional", "UI designer", "Project manager", "Software developer", and "Other" (free text). In the selected sample population 78 percent answered that their role in projects was "Usability professional" and 66 percent answered "UI designer". Everyone in the sample had answered one or both of these choices, but since it was possible to enter more than one alternative on this question there was in addition 18 percent who also had entered project managers, 11 percent software developer and 13 percent who reported "other".

The user-centred work experience among the participants was between 2 and 35 years with a median value of 5 years.

53 percent of the participants reported to be in-house experts, 48 percent consultants, 19 percent researchers and 3 percent "other" (multiple responses possible).

The respondents were, with a few exceptions, divided among the Nordic countries. 33 percent were from Sweden, 27 percent from Finland, 23 percent from Norway, 12 percent from Denmark, 2 percent from Iceland and 4 percent from other countries. The amount of participants from each Nordic country were corresponding fairly well with the number of inhabitants of each country; only Denmark somewhat underrepresented).
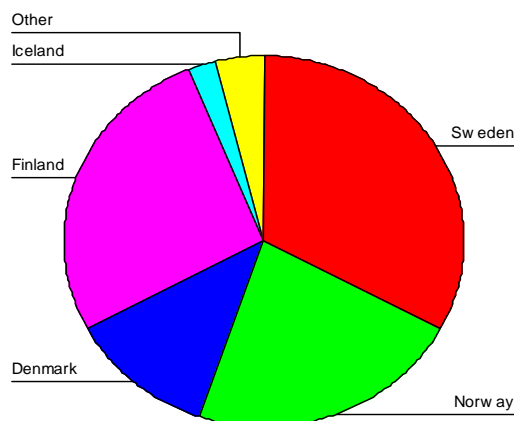
Figure 1. Participants' country affiliation

The respondents' academic background were reported as one of the following: "Psychology", "Informatics", "Technology", "Design" and "Other" (free text). (Multiple responses possible.) Distribution of respondents background: 46 percent technology, 24 percent design, 16 percent informatics, 10 percent psychology and 26 percent had a background in other areas (fig. 3). Common examples of "other areas" included business, cognitive science, ergonomics, social science, and communications.
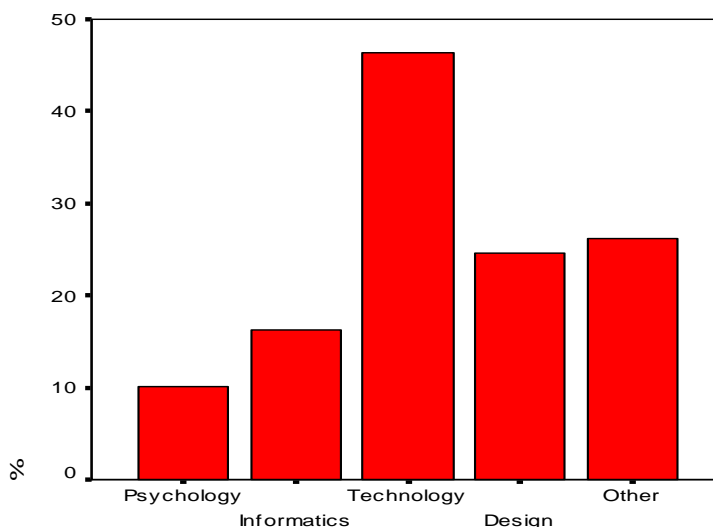
Figure 2. Participants' academic background

The respondents were asked what kind of systems they usually worked on. Answer alternatives included "Web-based applications", "Business applications", "Big complex systems (for example control rooms)", "Mobile ICT", "Forms and documentation" and "Other" (free text). (Multiple responses possible). 67 percent answered web-based applications, 45 percent business applications, 41 percent mobile ICT, 22 percent forms and documentation, 15 percent big complex systems, and 23 percent other types of systems.
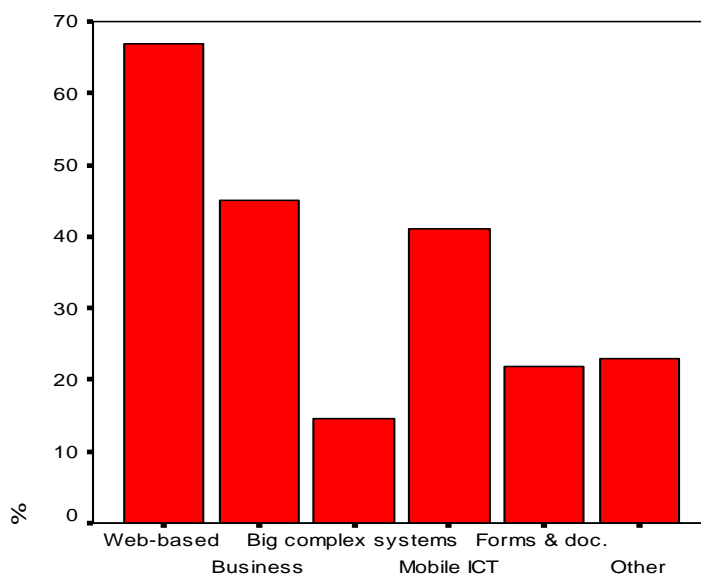
Figure 3. Participants' typical development project

The respondents working situation were reported as "Consultant", "In-house expert", "Researcher" or "Other" (free text). (Multiple responses possible.) 53 percent of the respondents reported to be in-house experts, 48 percent consultants, 19 percent researchers and 3 percent "other".
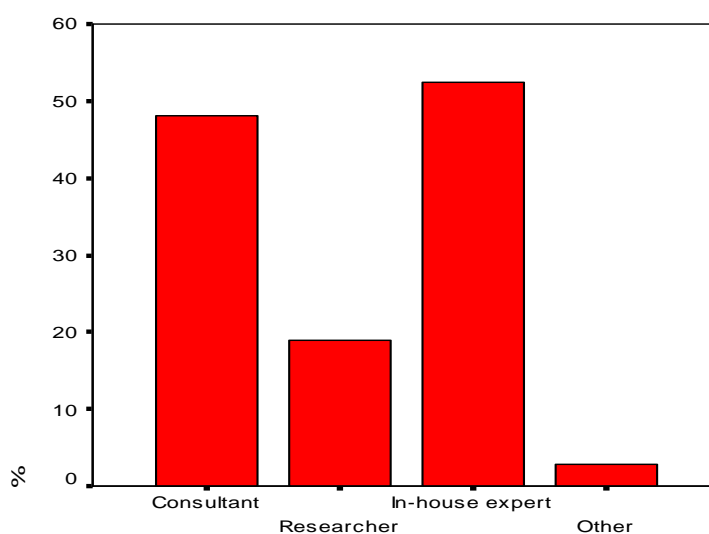
Figure 4. Respondents' working situation

The participants reported their professional titles in free text. This in order to enable judgements with regard to the nature of the survey sample, as well as to get an indication on how well defined

the HCI design profession actually is. 83 percent of the respondents provided their working title. The majority of the reported titles are listed in the frame below.

Reported professional titles
The following list summarizes the bulk of the reported professional titles that were indicating user-centred activities. The word in bold in each paragraph is to be put together with the other words of the same paragraph:

Usability Architect / Consultant / Specialist / Analyst / Designer / Engineer / Expert / Manager

Concept Developer / Designer

UI Designer / Specialist / Design Specialist / Design Coordinator

User Interface Designer / Specialist

GUI Designer

Interaction Designer / Design Specialist / Architect / Specialist

Information Architect / Designer

Design / Function Development / Human Factors / Software / Systems / Product Design / Usability **Engineer**

(Graphic) **Designer**

Business / HCI / Quality / Senior / User Experience **Consultant**

User Experience specialist

5.2 Desired outcomes of user-centred activities

The respondents were asked to indicate which potential outcomes of user-centred activities are most important given that one is to examine the usefulness of these activities. The respondents were to choose three outcomes from the following list: "Customer satisfaction", "Project leader satisfaction", "End-user engagement", "Identified usability problems", Impact on the development process", "New design suggestions", "New understanding of users, task and context of use", "Usability of the developed system", and "Usability awareness in the development team".

The respondents were to rank the potential outcomes from 1 to 3 (3 being the most important). Initial analyses, however, showed that differentiation between the three highest ranked results gave little added insight in spite of increased complexity. Thus, this differentiation was not included in the final analysis. In the final analysis, importance-scores were calculated as the percentage of respondents choosing a potential outcome as belonging to the three most important.

Most important outcomes from user-centred analysis
The respondents reported that with regard to user-centred analysis, it is important to give particular attention to "new understanding of users, tasks and context of use" (58 percent). Details are presented in Figure 5.
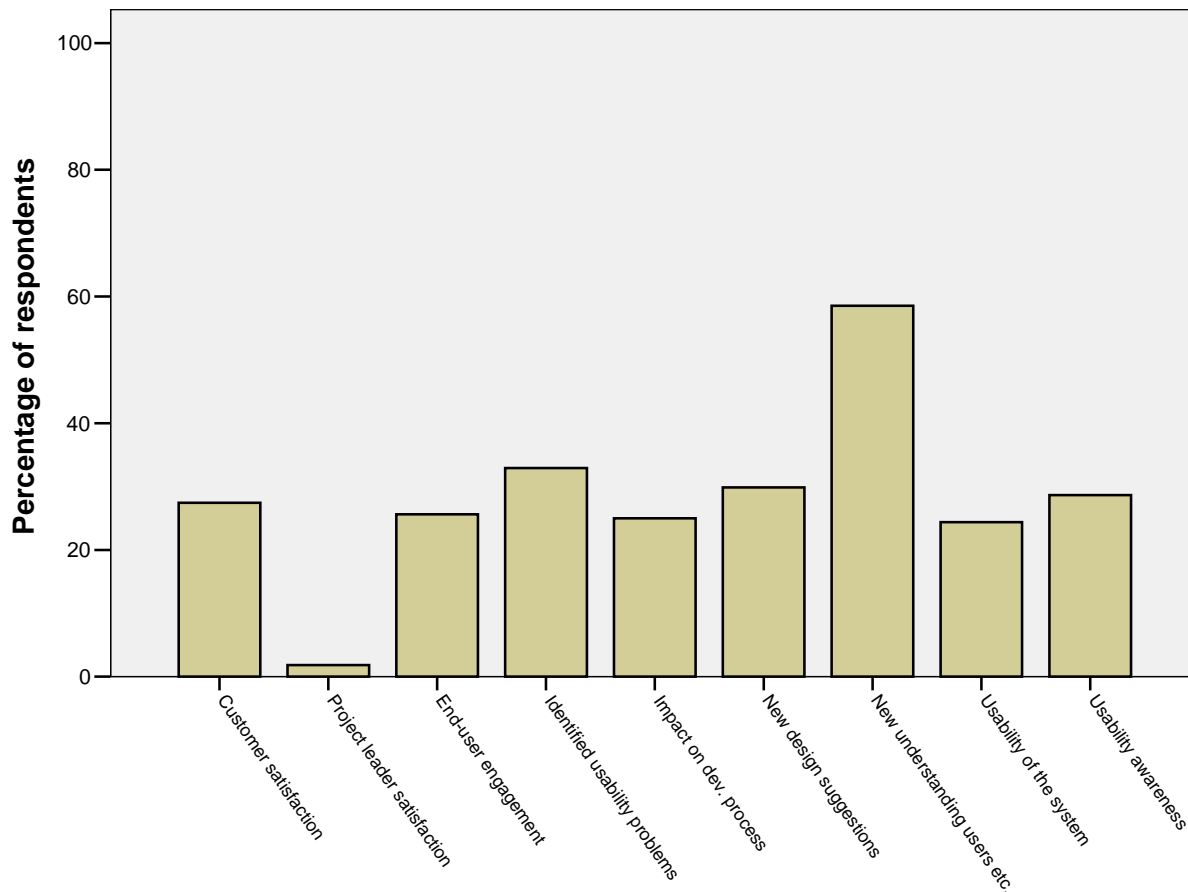

Figure 5. Percentage of respondents indicating the possible results of analysis methods as among the three most important to examine in a method evaluation (n=164)

Most important outcomes from design-input activities
The most important outcomes with regard to design-input activities were reported to be "New design suggestions" (45 percent), "New understanding of users, task and context of use" (36 percent) and the "Usability of the developed system" (36 percent). Details are presented in Figure 6.
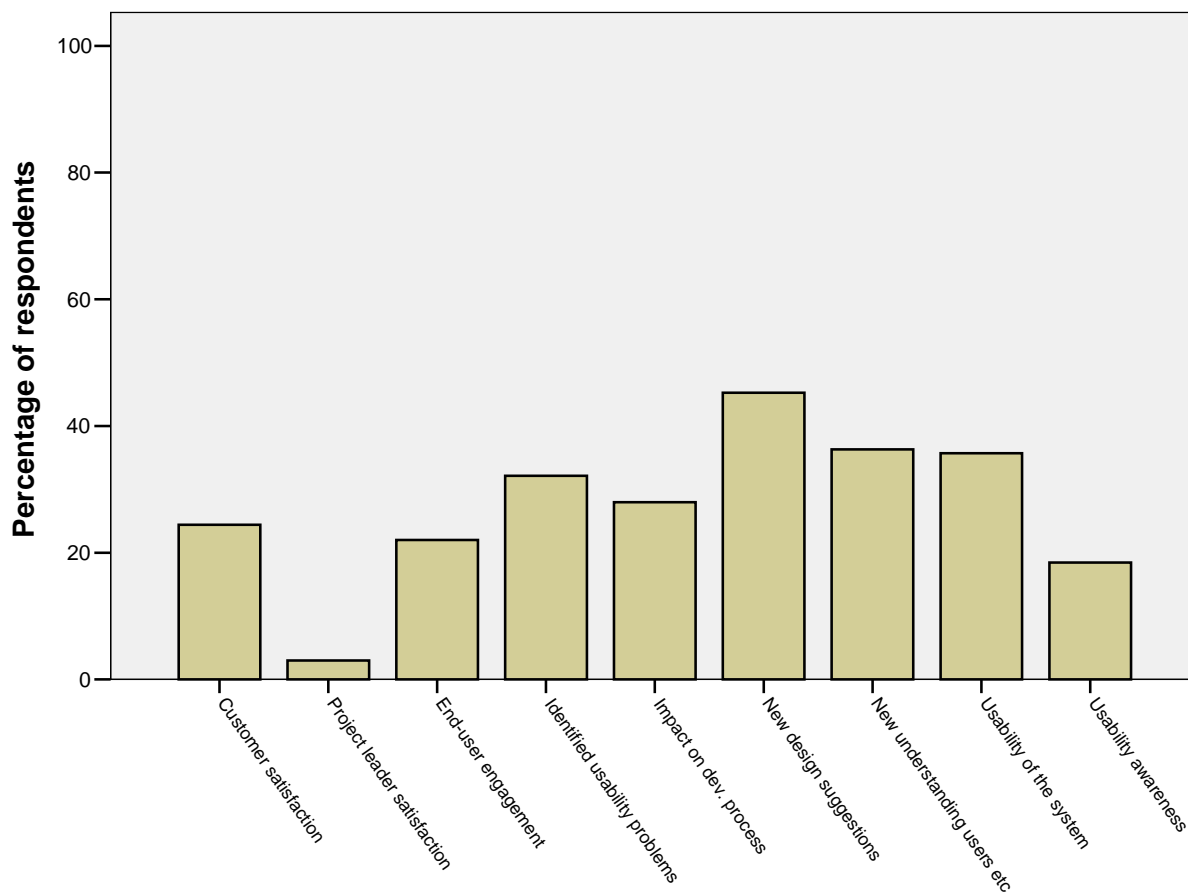

Figure 6. Percentage of respondents indicating the possible results of design-input methods as among the three most important to examine in a method evaluation (n=168)

Most important outcomes from user-centred evaluation
Regarding user-centred evaluation, the three highest ranked outcomes were "Identification of usability problems" (63 percent), "Usability of the developed system" (56 percent) and also "Customer satisfaction" (39 percent). Details are presented in Figure 7.
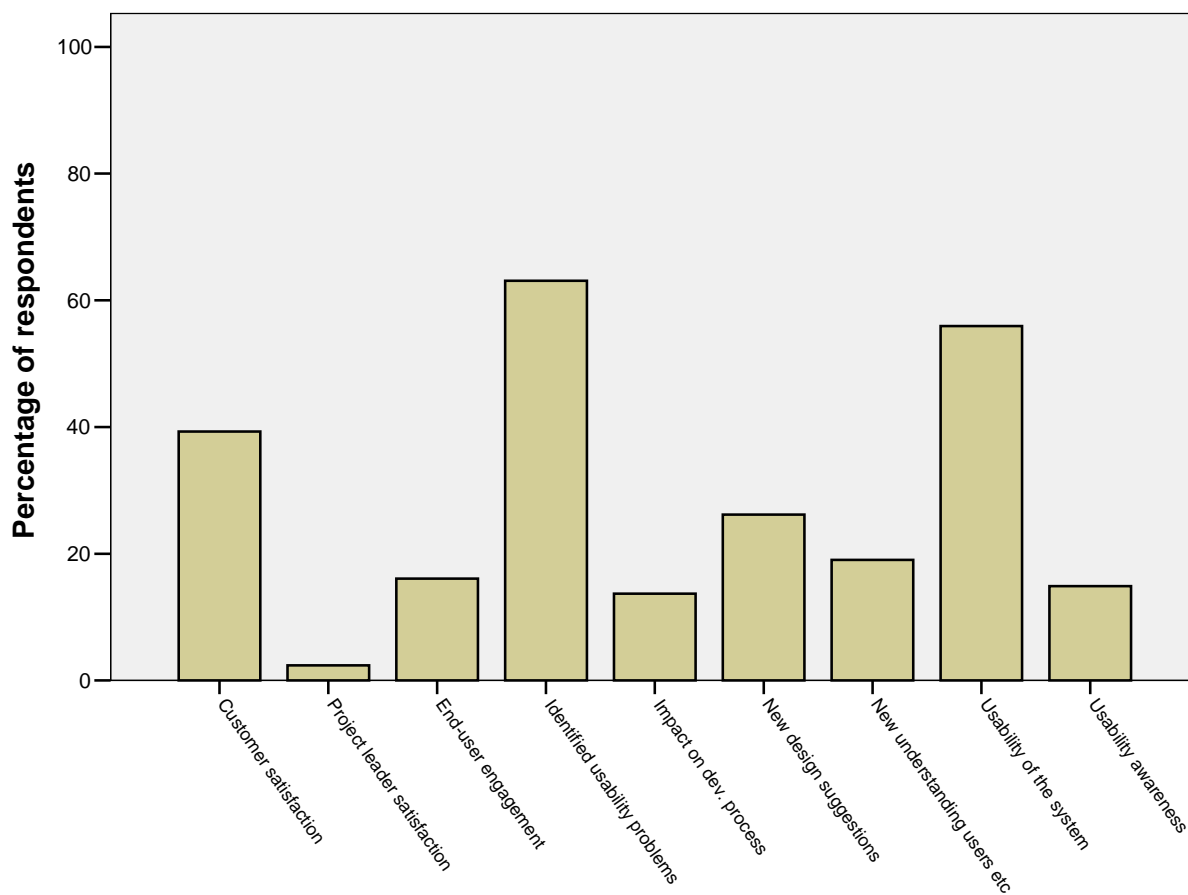

Figure 7. Percentage of respondents indicating the possible results of analysis methods as among the three most important to examine in a method evaluation (n=168)

Most important outcomes all over – total importance-score across the three method categories
The percentage of respondents choosing a potential result for any of the three categories of methods was calculated as a total importance-score.

Across all methods, the potential outcomes given the highest importance-score were "Identified usability problems" (69 percent) and "New understanding of users, tasks and context of use" (69 percent). "Usability of the developed system", "New design suggestions", and "Customer satisfaction" received the 3., 4. and 5. highest scores.

The lowest importance-score was received by the potential result "Project leader satisfaction" (5 percent), followed by "End-user engagement", "Usability awareness in the development team", and "Impact on development process".
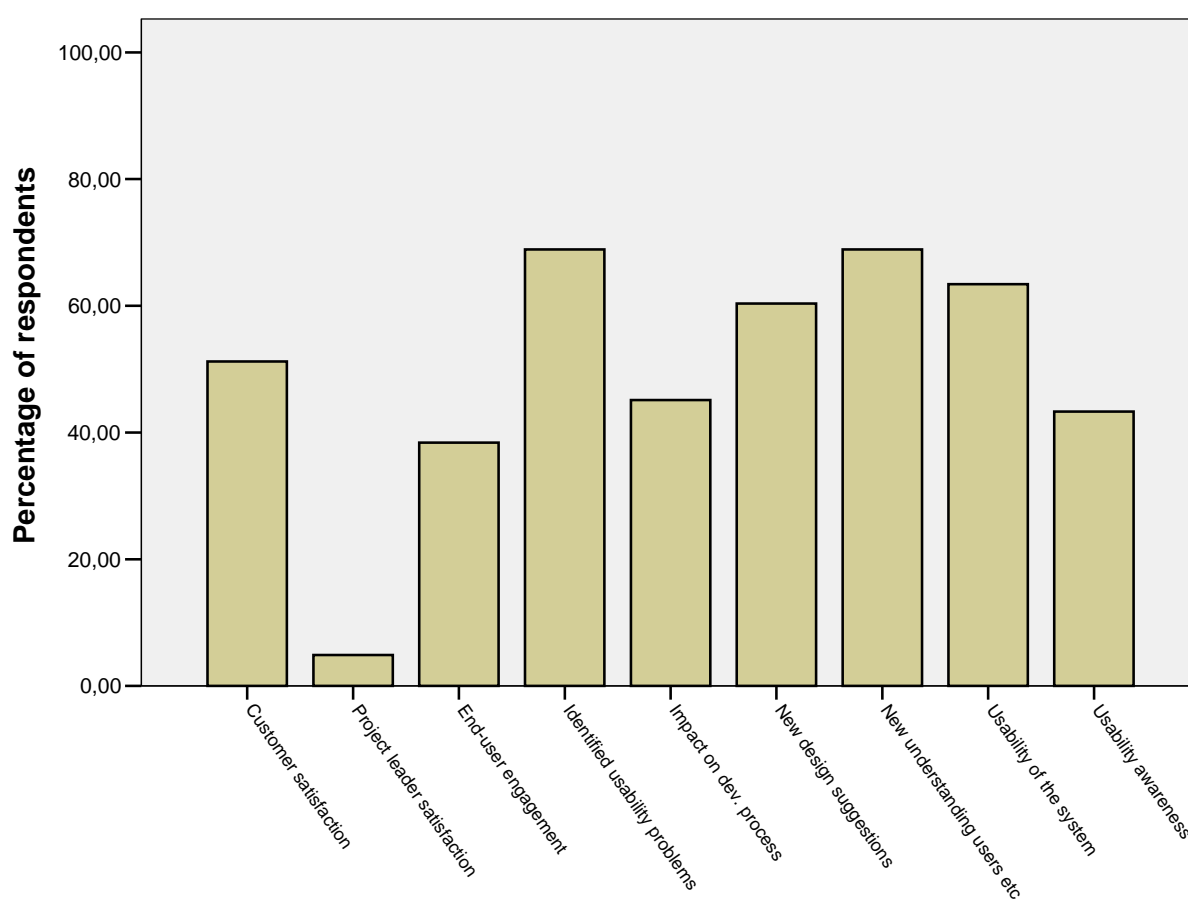
Figure 8. Total importance-score across the three method categories (n=168)

5.3 How to evaluate user-centred activities?

How would HCI-practitioners ideally evaluate their own user-centred practice?
The participants were asked how they ideally would like to evaluate the usefulness of their own user-centred practice. Answer alternatives included: "Repeated user tests", "Evaluation meeting", "Surveys", "Customer satisfaction reports", "Subjective evaluation", "Trust gut feeling", "Not necessary", and "Other"

56 percent answered that they would like to use repeated user tests, 41 percent reported that they want to use customer satisfaction meetings. "Subjective evaluation" and "Trust gut feeling" were

among the three least wanted ideal approaches to evaluation of own practice. And only 1 percent of the participants answered that it wasn't necessary to evaluate the usefulness of ones own user-centred practice. Details are presented in Figure 8.
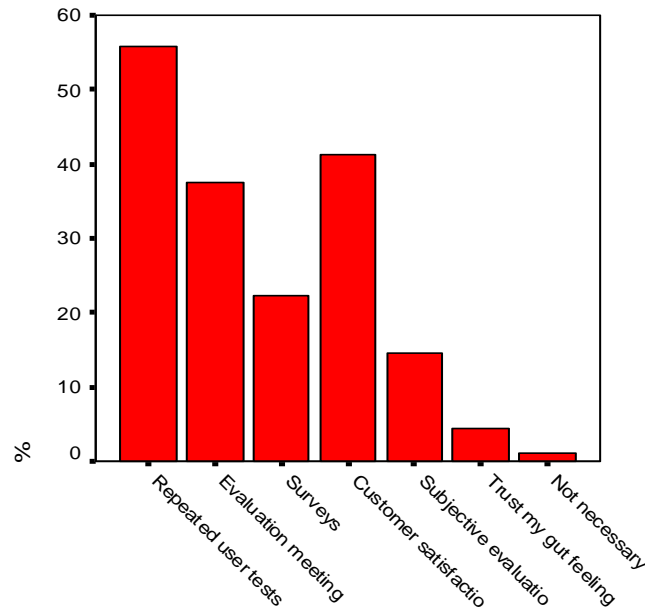
Figure 8. The respondents' ideal evaluation of user-centred practice (percentages)

How are HCI-practitioners evaluating their own practice as of today?
The participants were asked how they usually evaluate the usefulness of their own user-centred practice. The participants could chose between the same alternatives as above.

37 percent of the participants answered that they used repeated user tests, 27 percent answered evaluation meeting. As little as 5 percent of the participants answered that they do not evaluate the usefulness of their user-centred practices.
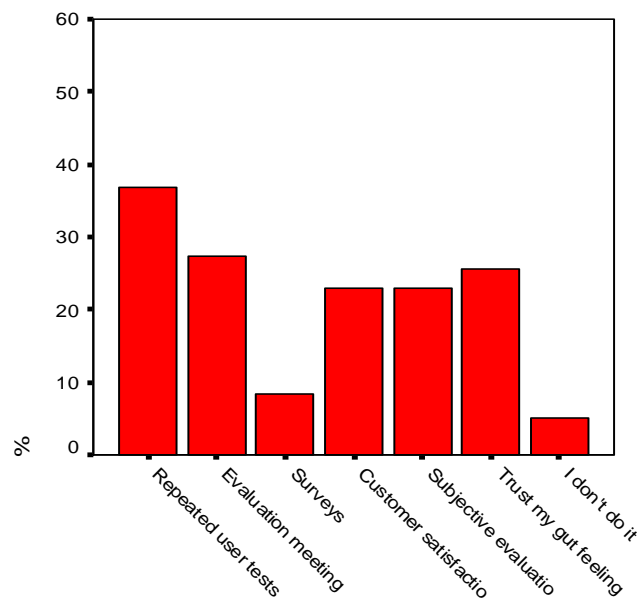

Figure 9. The respondents' actual evaluation of user-centred practice (percentages)

5.4 Usefulness and use of user-centred methods

The usefulness and use of user-centred were investigated as:
- Perceived usefulness of user-centred methods
- Actual use of user-centred methods
- Whether HCI-practitioners are "optimally" involved in development projects

5.4.1 Perceived usefulness of user-centred methods

The participants were asked to rate the usefulness of a number of methods applied in the different phases of the development process. Rating was done on a scale from 1-5, where 1 represented "Not at all useful" and 5 represented "Very useful".

Start phase. The methods to be rated on usefulness in the start phase of projects were Surveys, Interviews, Field studies, Workshops, Personas, Scenarios, Use cases, Focus groups, and Task analysis. The three methods that were given the highest mean rating were Field studies (4,3), Interviews (4,2) and Scenarios (4,0). The three lowest ranked methods were, from the bottom, Focus groups (3,3), Surveys (3,4) and Personas (3,5).

Mid phase. The methods to be rated on usefulness in the mid phase were Rapid prototyping, Advanced prototyping, User tests, Card sorting, Storyboarding, Guidelines, Heuristic evaluation, Focus groups, Design patterns, Cognitive walkthrough, Expert evaluation, and Evaluation workshop. The three methods that were given the highest mean rating were User tests (4,6), Rapid prototyping (4,5) and Expert evaluation (4,0). The three lowest ranked methods were, from the bottom, Focus groups (2,9), Design patterns (3,2) and Card sorting (3,2).

End phase. The methods to be rated on usefulness in the end phase were User tests, Heuristic evaluation, Expert evaluation, Surveys, Evaluation workshop, Interviews, and Cognitive walkthrough. The three methods that were given the highest mean rating were User tests (4,5), Expert evaluation (3,9) and Heuristic evaluation (3,7). The three lowest ranked methods were, from the bottom, Surveys (3,0), Cognitive walkthrough (3,3) and Evaluation workshop (3,5)

Perceived usefulness across different project types
It is interesting to note that there was very little difference between practitioners working with different kinds of systems, with regard to how useful they perceive the different user-centred methods to be.

As an example, Table 2 presents the mean usefulness ratings for the methods allocated to the mid phase across practitioners working with different kinds of systems. The mid phase was chosen as example since this phase encompasses the greatest number of methods.

	All	Web based applic. (n=111)	Business applic. (n=75)	Big compl. systems (n=26)	Mobile ICT (n=59)	Forms & document. (n=38)
User tests	4,6 (0,7)	4,6 (0,6)	4,6 (0,6)	4,4 (0,9)	4,6 (0,6)	4,5 (0,7)
Rapid prototyping	4,5 (0,7)	4,5 (0,7)	4,6 (0,6)	4,6 (0,6)	4,5 (0,5)	4,4 (0,8)
Expert evaluation	4,0 (0,8)	4,1 (0,8)	4,1 (0,8)	4,0 (0,7)	4,0 (0,9)	4,3 (0,7)
Advanced prototyping	3,7 (1,0)	3,8 (1,0)	3,8 (1,1)	3,9 (0,8)	3,9 (1,0)	3,6 (1,2)
Heuristic evaluation	3,7 (0,9)	3,8 (0,9)	3,8 (0,9)	3,6 (1,0)	3,7 (0,8)	3,8 (0,9)
Evaluation workshop	3,5 (1,0)	3,6 (1,0)	3,5 (1,1)	3,2 (0,9)	3,5 (1,1)	3,6 (1,1)
Story-boarding	3,5 (1,0)	3,5 (1,0)	3,6 (1,1)	3,9 (0,9)	3,7 (1,1)	3,5 (1,1)
Cognitive walkthr.	3,5 (1,0)	3,5 (1,0)	3,5 (1,1)	3,3 (1,0)	3,5 (1,0)	3,5 (1,1)
Guidelines	3,3 (0,9)	3,4 (0,9)	3,3 (1,1)	3,0 (0,8)	3,3 (1,0)	3,5 (0,9)
Card sorting	3,2 (1,1)	3,5 (1,1)	3,2 (1,2)	3,1 (0,9)	3,0 (0,9)	3,3 (1,1)
Design patterns	3,1 (0,9)	3,2 (0,9)	3,1 (1,0)	3,0 (0,7)	2,9 (0,9)	3,1 (1,0)
Focus groups	3,0 (1,1)	3,1 (1,1)	2,9 (1,0)	2,8 (0,9)	3,0 (1,0)	2,8 (1,0)

Table 2. Mean and standard deviations of User-centred method usefulness ratings in the mid phase; across groups of practitioners working with different kinds of systems.

5.5 Actual use of user-centred methods

The participants were asked to report which user-centred methods they use for each of the three phases of the development process. In the questionnaire the participants were asked to report their use of the different methods as "yes, sometimes", "yes, often", or "no, not at all".

When analyzing the data, differentiating between "yes, sometimes" and "yes, often" was not judged to provide much added insight, only complexity in reporting. Thus the results will be presented as belonging to the categories "yes" or "no".

Start phase. The methods reported to be most used by the participants in the start phase of projects were Interviews (92 percent), Scenarios (89 percent), and Use cases (89 percent). The least used methods were Focus groups (64 percent), Personas (65 percent) and Surveys (70 percent).

Mid phase. The methods reported to be most used by the participants in the mid phase were User tests and Rapid prototyping (both 95 percent), and Expert evaluation (93 percent). The least used methods were Card sorting (43 percent), Design patterns (48 percent), and Focus groups (53 percent).

End phase. The methods reported to be most used by the participants in the end phase were User tests (97 percent), Expert evaluation (88 percent), and Heuristic evaluation (75 percent). The least used methods were Cognitive walkthrough (47 percent), Evaluation workshop (52 percent), and Surveys (53 percent).

Reported use of methods across project types

With regard to reported use across different development projects, there were only minor differences between practitioners working with different kinds of systems. Most differences were found with regard to practitioners working with big complex systems and forms and documentation. As an example, Table 3 presents the reported usage of methods in the mid phase; across practitioners working with different kinds of systems.

	All	Web based applic. (n=111)	Business applic. (n=75)	Big compl. systems (n=26)	Mob. ICT (n=59)	Forms & doc. (n=38)
User tests	95 %	97 %	97 %	84 %	96 %	97 %
Expert eval.	96 %	97 %	97 %	100 %	96 %	100 %
Rapid prototyp.	92 %	92 %	94 %	91 %	91 %	85 %
Heuristic eval.	82 %	84 %	85 %	92 %	92 %	81 %
Adv. prototyp.	74 %	74 %	71 %	92 %	78 %	63 %
Guidelines	73 %	76 %	80 %	73 %	79 %	79 %
Cogn. walkthr.	62 %	63 %	71 %	63 %	73 %	62 %
Eval. Workshop	58 %	57 %	60 %	52 %	57 %	69 %
Storyboarding	56 %	60 %	62 %	71 %	62 %	61 %
Focus groups	53 %	55 %	52 %	50 %	61 %	47 %
Design patterns	47 %	55 %	54 %	52 %	38 %	47 %
Card sorting	43 %	53 %	47 %	50 %	38 %	53 %

Table 3. Percentage of respondents that reported to use the different user-centred methods in the mid phase; across groups of practitioners working with different kinds of systems

5.6 Are HCI-practitioners "optimally" involved in development projects?

To investigate whether the HCI-practitioners are optimally involved in development projects, the participants usefulness-ratings of the methods were compared with the participants reported actual use of the methods.

The comparison was conducted by (1) sorting the methods according to mean perceived usefulness, and (2) sorting the methods according to reported usage of the methods. Deviations between the two diagrams indicate methods that are used more or less than their perceived usefulness accounts for. Pair wise bar diagrams for the three project phases are presented below.
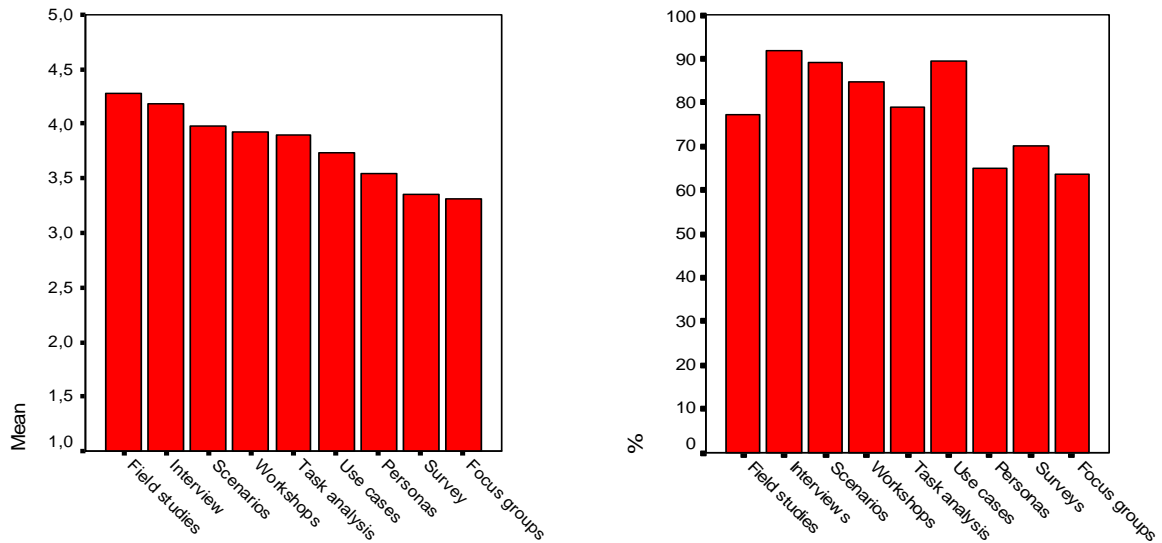
Start phase

Figure 10. Left: Mean ratings of usefulness of methods used in the start phase. Right: Usage of methods in the start phase.

The magnitude of the difference between actual use and perceived usefulness of the methods in the start phase may be illustrated by using Spearman's rho to calculate their rank correlation coefficient, with regard to the rank order of the different methods. $r_s=0{,}72$ (p=0,03). Field studies seem to be used less than their perceived usefulness should indicate. Use cases seem to be used more than their perceived usefulness indicate.

Mid phase

Figure 11: Left: Mean ratings of usefulness of methods used in the mid phase. Right: Usage of methods in the mid phase

Rank correlation coefficient calculated by Spearman's rho: r_s =0,90 (p=0,00). In the mid phase expert evaluation, heuristic evaluation, guidelines, and focus groups seem to be used more than these methods perceived usefulness should indicate. Advanced prototyping, evaluation workshops

and card sorting may seem to be used somewhat less than their perceived usefulness seem to indicate.
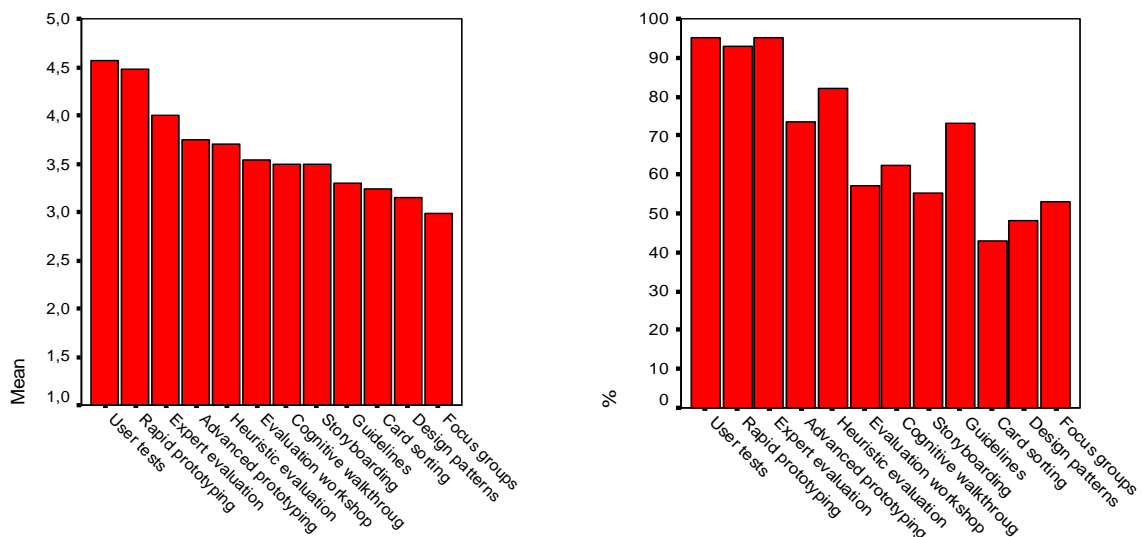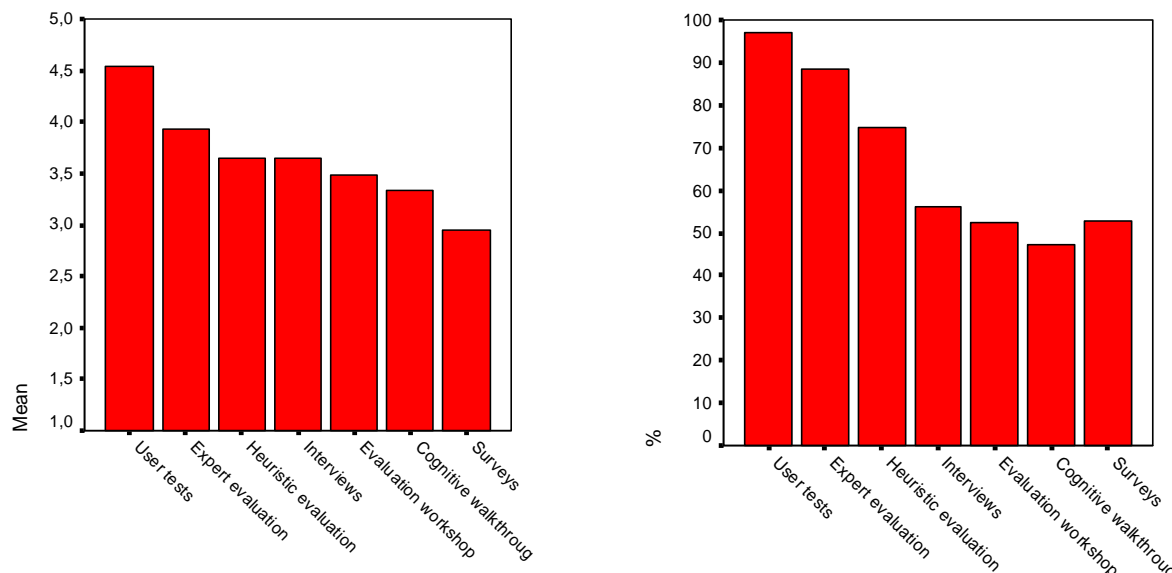
End phase

Figure 12: Left: Mean ratings of usefulness of methods used in the end phase. Right: Usage of methods in the end phase

Rank correlation coefficient calculated by Spearman's rho: rs=0,96 (p=0,00).In the end phase it mostly seems like the methods perceived as the most useful are also the most used, and vice versa.

6 Discussion

The presented results provide insights in work practices and opinions of 179 HCI-practitioners. The participants were sampled by convenience. Even so, the participants' profiles agree fairly well with at least the authors' intuitions with regard to the population of HCI-practitioners at large. There was a good spread of participants between the Nordic countries.

The discussion of the results' implications, with regard to the research questions will be structured according to the three main issues addressed in this report. Following this, the validity and generality of the results will be discussed.

6.1 Desired outcomes of user-centred activities

Across all methods, the outcomes given the highest importance-score were "Identified usability problems" and "New understanding of users, tasks and context of use". Followed by "Usability of the developed system", "New design suggestions" and "Customer satisfaction".

The most interesting results were the low importance-scores received by "Project leader satisfaction", "Usability awareness in the development team", "End-user engagement", and "Impact on development process". This is very surprising given that the full utilisation of results from user-centred activities depends on a good relation with the team that builds the system.

The low importance-scores received by "Project leader satisfaction", and "Usability awareness in the development team" may be an indication of the "us-them" attitude HCI-practitioners sometimes displays when describing their relation to software developers – an attitude that not

necessarily serves the purpose of full integration of HCI in the software development process. Quite possibly HCI practitioners may benefit from increased insight with regard to whether or not their user-centred work practices are sufficiently integrated in the development process at large; indicating the importance of including impact as a dimension of an evaluation process for user-centred activities.

The low importance-scores received by "Impact on development process" may (similarly to the low scores received by "Project leader satisfaction" and "Usability awareness in the development team") indicate that the HCI-practitioners do not necessarily see themselves as empowered actors in the development process; leaving the responsibility of utilising results from user-centred activities to other actors.

Contrary to what might be expected, "End-user involvement" received relatively low importance-scores. This may indicate that for many HCI-practitioners, end-user involvement is regarded more as a means towards usable systems rather than a goal in itself.

6.2 Ideal and feasible ways to evaluate user-centred activities

The three most popular means of evaluating ones own practices were repeated user tests, customer satisfaction reports, and evaluation meetings. However, in real life the evaluation activities of "subjective evaluation " and "trust gut feeling" were more used than they would be in an ideal world. Very few of the respondents reported that the ideal way of evaluating ones own practice was subjective evaluation, to trust ones gut feeling, or don't do it at all.

It is interesting to note that HCI-practitioners in general seem to experience the need for procedures to evaluate their own practices, and that gut feeling or subjective evaluations are not sufficient in this regard. The results indicate a fairly high maturity with regard to evaluation of user-centred activities.

6.3 Usefulness and use of user-centred methods

Which user-centred methods are perceived to be most useful? Start phase
Field studies, interviews, use of scenarios, workshops and task analysis were among the methods rated most useful in a projects' start phase. Focus groups, surveys and personas were rated the least useful. Great differences in perceived usefulness exist between data collection methods e.g. field studies, interviews, surveys and focus groups. Field studies and interviews were rated highest, surveys and focus groups were rated lowest.

One reason for the low rating of focus groups may be widespread familiarity with reported weaknesses in the method (e.g. risks of biased results due to uncontrolled group dynamics).

The survey method's low usefulness-score may indicate that the information gathered through surveys is regarded as too shallow to work properly as input to a project. In field studies and interviews (two methods that are often combined) the HCI-practitioner may gain a deeper and more unbiased understanding of the relevant needs and requirements for the project.

Which user-centred methods are perceived to be most useful? Mid phase
The two methods reported to be most useful in the mid-phase were no surprise: User tests and rapid prototyping. These are two strongly advocated methods throughout the HCI literature. It was however a bit surprising that in spite of a steady trickle of studies with negative findings regarding the reliability and validity of heuristic evaluation (see e.g. Law and Hvannberg, 2004) this method still is given a fairly high usefulness rating. The reason for this may be that heuristic evaluation is cheap and easily applicable, and that the individual HCI-practitioner feels that heuristic

evaluations may provide sufficiently high quality results provided that the evaluator has sufficient expertise.

Guidelines and design patterns received low usefulness-scores in the mid phase. This may seem a bit odd, given that one of the more important design inputs for any practitioner would be good design examples and principles of the past. The reason for these methods' low rating may indicate that the existing body of design patterns and guidelines does not serve the purpose for HCI-practitioners.

Which user-centred methods are perceived to be most useful? End phase
In the end phase no other method even comes close to the usefulness-score of user tests. Expert evaluation was rated second highest, probably due to the low-cost and versatile nature of this method.

Cognitive walkthrough rates surprisingly low on perceived usefulness, given that this is a low-cost task oriented method. Perhaps the low usefulness score of Cognitive Walkthrough is due to the method's cumbersome evaluation procedure, sometimes making cognitive walkthroughs uninspiring and drawn out.

The survey method received the lowest usefulness-score in the end phase. This is somewhat surprising given the convenience and ease with which a survey evaluation may be carried out: The results may indicate that subjective feedback on user satisfaction (traditionally being a main focus of evaluation surveys) is not regarded as interesting as objective feedback on user problems.

Are HCI-practitioners optimally involved in projects?
The issue of HCI-practitioners' influence and freedom in projects may be investigated as the match between perceived usefulness of methods and frequency of use. The different project phases are discussed individually.

For the end phase, the comparative analysis showed an almost perfect match between the methods most frequently used and the methods perceived to be most useful. This probably indicates that HCI-practitioners are given relatively great degrees of freedom and influence in the end phase. This is in line with the general opinion within the field; HCI-practitioners often claim to be involved a little later in the project than they would really like to have been.

In the mid phase, several methods are more used than their usefulness scores should indicate. In particular this goes for expert evaluations, heuristic evaluations, and guidelines, and to a certain degree focus groups. This finding may indicate that other forces than the best judgment of the HCI-practitioners are at play in the mid phase of projects. The reason for the popularity of expert evaluations and heuristic evaluations may be that they are extremely low-cost methods, and may be easily fitted to a tight project cost- and timeframe. With regard to guidelines and focus groups, these methods may be requested by other persons than the HCI-practitioner.

Field studies are far less used in the start phase than one should expect on basis of the reported usefulness of this method. This probably is due to the cost demanding nature of this method. The usefulness of the method may be known to the HCI-practitioners, but the practitioner may still not have the opportunity to use it because of budget and time constraints. Use cases are somewhat more used than their perceived usefulness may account for. This probably is due to the need for using use cases as a description format to communicate findings with others in the development project.

Small differences between practitioners working with different kinds of systems
There was very little difference between practitioners working with different kinds of systems, with regard to usefulness-scores of the different user-centred methods. The between-groups differences in reported use were somewhat larger than the between-groups differences in usefulness-scores, but the overall impression is that the groups of practitioners working with different kinds of systems are more similar than different also with regard to their use of the different methods.

This finding is not at all according to the intuitions of the authors, given the great differences in both development process and process requirements for projects of mobile ICT, big complex systems, web-based applications and business applications. One explanation for this lack of differences may be that HCI-practitioners look to the general HCI-tradition, rather then their own work particular situation, when evaluating and choosing methods for their own work practice. One less inspiring explanation may be that HCI-practitioners perceive all systems like nails, for which the ultimate hammer is user tests. It would be interesting to pursue these explanations in future research.

6.4 Validity and generality

The validity of the present study greatly depends on the quality of the questionnaire and whether the sample of participants is representative of the HCI-community at large. The questionnaire was developed on basis of a series of HCI-practitioners, interviews and adjusted based on early piloting. Also the face validity of the items of the questionnaires was quite high, in that the questions explicitly focused on the information sought. (The intentions of the items were not hidden from the participants as may be the case in other questionnaires.) These factors should contribute to a confidence in the quality of the questionnaire.

With regard to the sample, the procedure for sampling was a convenience sample. This sampling method was chosen since the exact extent of the population of HCI-practitioners is not known. This is an obvious weakness of the study. However, the participants' profiles were quite in line with the authors' intuitions with regard to the characteristics of the HCI-population at large. Hopefully the participants' profile is as convincing to the readers of this paper as well. The sample of participants is mainly from the Nordic countries why the conclusions may not be generalized to the HCI-population at large. However, many of the findings presented in the present paper are in line with results from other surveys (e.g. Vredenburg et al. (2002)). Depending on whether a HCI-community may be seen as sufficiently similar to the participants' profile of this survey, some conclusions in this paper may be applicable also outside the Nordic countries.

7   Developing an evaluation procedure for user-centred activities; lessons learnt from the survey

In the EFFIN project, one main ambition is to develop a procedure to evaluate the efficiency of user-centred activities in a development process. The results from the present survey will be used as a basis for orienting this development.

7.1 Existing basis for evaluation of user-centred practices

The survey results indicate that the maturity for evaluating user-centred methods and practices seem to be fairly high in the HCI-community. Particularly in association with repeated user tests, evaluation meetings, and customer satisfaction reports. This is promising with regard to dissemination and uptake of an evaluation procedure.

Repeated user tests are an accepted way of conducting method evaluation, but may be too resource demanding. However, as part of a general method evaluation effort (e.g. with regard to strategic decisions with regard to what kind of methods a HCI group should use) the use of repeated user tests may be advisable.

Existing utilization of evaluation meetings and customer satisfaction reports may represent a useful low-cost vehicle for evaluation activities, e.g. to gain information on (1) the quality of the results of the user-centred activity and (2) the actual utilisation of the results of the user-centred activity.

7.2 Objective of an evaluation of user-centred practice

Quality of the results of user-centred activities
The survey indicates that HCI-practitioners particularly are preoccupied with "Identified usability problems", "New understanding of users, tasks and context of use", and "New design suggestions" as focus of attention for an evaluation of user-centred work practice. This clearly indicates an interest in the evaluation of the quality of the results from user-centred activities. An evaluation procedure should include facilities for getting feedback on quality of results.

Impact of the results of user-centred activities
"Usability of the developed system" was given high importance-score in the survey, indicating that an evaluation procedure should provide information on the quality of the end-result of the development process is important. This means that the HCI-practitioners impact on the final product of the development process is regarded as important, even though the low importance-scores of "Project leader satisfaction" and "Impact on development process" may indicate that the average HCI-practitioner does not see herself as a fully integrated part of the development team.

In order to have impact on the "usability of the developed system" it is necessary to make sure that the results of the user-centred activities actually are utilised; not just forgotten and lost in the remainder development process. Thus, the HCI-practitioner's actual impact on the development process is a necessary prerequisite in order for a user-centred activity to lead to increased usability of the end product.

Objectives of the evaluation - summarized
The most important objectives of an evaluation procedure, as indicated by the survey results, may be summarized as the following:
- Evaluate the quality of the results of user-centred activities
- Evaluate the actual impact of the results of user-centred activities
- Evaluate the usability of the developed system

7.3 Background knowledge for evaluation of the developed tools and methods for evaluation of user-centred practice

The information on perceived usefulness and actual use of user-centred methods provides important knowledge for future evaluation of the tools and methods to be developed in the EFFIN project.

In particular it should be noted that the later project phases is when the HCI-practitioner seems most at liberty to actually use the methods she finds most useful. This may indicate that trial of developed tools and methods may easiest be conducted in later project phases rather than earlier, to avoid any bias introduced by the HCI-practitioner not being in sufficient control over her own activities.

Further it should be noted how the different user-centred methods are rated with regard to usefulness. In the case where an evaluation of the developed tools and methods involves comparative analysis of methods, it should be expected that any ranking of the usefulness of the methods reflect the usefulness-scores of the present survey. In case of deviation from the survey results, this needs to be explained.

Which kind of system or product that is developed does not seem to be important when trying out an evaluation procedure. The survey indicate little variation between different categories of development projects with regard to perceived usefulness and actual use of user-centred methods.

References

Andre, T. S. (2000). "Determining the effectiveness of the usability problem inspector: A theory-based model and tool for finding usability problems." PhD Thesis. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Bias R.G., Mayhew D.J., Cost-Justifying Usability, Academic press, Inc., London, 1994.

Gray & Saltzman (1998). "Damaged merchandise? A review of experiments that compare usability evaluation methods." Human Computer Interaction. 13(3):203-261

Gulliksen J., Boive I., Persson J., Hektor A., Herulf L., Making a difference - a survey of the usability profession in Sweden, Proc. NordiCHI 2004, 207–215

Hartson, H. R., Andre, T. S., Williges, R., C. (2001). "Ultimate criterion for UEM effectiveness - Finding real usability problems" International Journal of Human-Computer Interaction, December 2001, Vol. 13: 373-410

ISO 13407, (1999), Human-Centered Design Processes for Interactive Systems, International Organization for Standardization, Geneva.

Karat, C-M. (1997). "Cost-justifying usability engineering in the software life cycle". In Helander, Landauer & Prabhu (Eds.). Handbook of human-computer interaction. Amsterdam: Elsevier Science.

Law, E. L.-C., Hvannberg, E. T. (2002) "Complementarity and convergence of heuristic evaluation and usability test: a case study of universal brokerage platform". Proc. NordiCHI 2002, Aarhus, Denmark

Law, E. L.-C, Hvannberg, E. T. (2004) "Analysis of strategies for improving and estimating the effectiveness of heuristic evaluation". Proc. NordiCHI 2004, Tampere, Finland

Maguire M., "Methods to Support Human-Centered Design", Int. J. Human-Computer Studies, 2002, 587–634

Mantei, M. M. & Teorey, T. J. (1988). "Cost/benefit analysis for incorporating human factors in the software lifecycle", ACM, April 1988, 31, 4: 428-439.

Nielsen, J. & Gilutz, S. (2003) "Usability Return on Investment", Nielsen Norman Group

Usabilitynet (2005). Available online at www.usabilitynet.org

Venturi G., Troost J., Survey on the UCD integration in the industry, Proc. NordiCHI 2004, 449–452.

Vredenburg K., Mao J., Smith P.W., Carey T., A Survey of User-Centered Design Practice, Proc. CHI 2002, CHI Letters 4(1), 471–478.

Wixon, D. (2003). "Evaluating usability methods". Interaction, July-August: pp. 29-34.