Active Collaborative Learning: Supporting Software Developers in Creating Redesign Proposals

Anders Bruun, Janne Jensen, Mikael Skov, Jan Stage

Active Collaborative Learning: Supporting Software Developers in Creating Redesign Proposals

Anders Bruun[1], Janne Juul Jensen[1,2], Mikael B. Skov[1] and Jan Stage[1]

[1]Aalborg Universitity, Department of Computer Science, DK-9220 Aalborg, Denmark
[2]Trifork A/S, DK-8000 Århus C, Denmark
bruun@cs.aau.dk, jjj@trifork.com, {dubois, jans}@cs.aau.dk

Abstract. Redesign proposals have been suggested as means to improve the feedback from usability evaluation to software development. Yet redesign proposals are usually created by usability specialists without any involvement of the software developers who will implement the proposals. This paper reports from an exploratory study where redesign proposals were created in an active and collaborative learning process that involved both software developers and usability specialists. The focus was on the support that the developers needed in order to contribute constructively to improve the usability of the system. The findings show that this process had a considerable impact on the developers' understanding of the usability problems, especially the weaknesses of the system. They were able to contribute constructively to create redesign proposals, and they found the workshop very useful for their future efforts to eliminate the usability problems that have been identified.

Keywords: Usability evaluation, usability problem, redesign proposal, developer involvement, active collaborative learning, exploratory study.

1 Introduction

A usability evaluation is conducted to assess the usability of an interactive software system. Usability is the system's ability to help specified users achieve specified goals in a particular environment in an effective, efficient and satisfying way [10]. A formative usability evaluation is conducted to improve the interaction design. Formative evaluations are carried out during development, often iteratively, with the goal of detecting and eliminating usability problems [1]. A formative usability evaluation establishes a strong and useful basis for understanding and improving the design of a software system. Exploiting this requires feedback that significantly impacts how developers understand the usability of the system [7].

The classical feedback from a usability evaluation is a report that describes the usability problems that have been identified [18]. The aim is to present the results of the evaluation in a format that is useful for the developers that are going to eliminate the identified usability problems. Unfortunately, a substantial body of research documents that usability reports are of limited utility and effect, e.g. [5, 7, 15]. To overcome this, it has been argued that evaluators who conduct a usability evaluation should not only

adfa, p. 1, 2011.

describe what the usability problems are, but also suggest how they could be resolved, e.g. [11]. Such suggestions have been denoted as redesign proposals, and their advantages have been studied empirically, e.g. [6].

Redesign proposals are a great step forward in providing useful feedback about usability to developers of interactive systems. However, redesign proposals are usually created solely by evaluators who are usability specialist and subsequently handed over to the developers in an old-fashioned presentation style involving only one-way communication. This division in roles for evaluators and developers is similar to a traditional teaching approach.

In education, many authors have questioned the qualities of a traditional teaching style. Dewey's theory on education was based on the idea that education consists primarily in transmission through communication; a process of sharing experience until it becomes a common possession [4]. He emphasized that the traditional teaching's concern with delivering knowledge, needed to be balanced with a much greater focus on the students' actual experiences and active learning. The teacher should be acknowledged as the intellectual leader of a social group, not by virtue of official position, but because of wider and deeper knowledge and matured experience [17]. Dewey was a key proponent of experiential education with learning through action that embody teaching methods such as cooperative learning and active learning.

This paper explores how software developers can become actively involved in creating redesign proposals to resolve identified usability problems. We explore different patterns of collaboration between usability specialists and software developers based on Dewey's theory on education combined with the idea that redesign proposals should be a key vehicle for providing feedback from a usability evaluation. In section 2, we present related work on feedback from usability evaluations. In section 3, we describe our exploratory study where active collaborative learning was employed to understand the results of a usability evaluation and create redesign proposals. Section 4 describes our findings from the study. In section 5, we discuss the implications of the findings for software development practice. Finally, section 6 provides the conclusion.

2 Related Work

There is a rich body of research on feedback from usability evaluations. The classical approach from the first half of the 1990s suggested that the results of a usability evaluation should be provided in a written usability report where the key element is a list of the identified usability problems and a detailed description of each [18].

Research from the same period demonstrated significant difficulties with usability reports. An early paper on this topic was based on analysis of real usability reports. Many of these reports did not describe usability problems but redesign solutions; it was emphasized that it is extremely rare for an evaluator to simply describe a problem, unless it is deemed as not having an apparent solution. These observations led to a proposal for a general format for usability problem lists [11]. Other publications from the same time also deal with the format of the usability problem list [13, 16].

There is also a considerable number of empirical studies that deal with feedback from usability evaluations. The early works are primarily concerned with evaluation methods and only secondarily with report formats, e.g. [1]. More recent research has focussed on a more rigorous problem reporting format [14] and practical guidelines for making problem descriptions [3]. There has also been work on different types of support for structured usability problem reporting [9].

A different stream of work has documented the problems with the classical usability problem report. A comparative study showed that the usability report had a strong impact on the developers' understanding of specific usability problems and supported a systematic approach to deal effectively with problems, whereas observation of user tests facilitated a rich understanding of usability problems and created empathy with the users and their work [8]. Two other studies analysed in more detail the utility of problem descriptions for the developers who should eliminate the problems [15, 22].

A number of studies have dealt with redesign proposals as opposed to mere problem descriptions. A study on the impact of inspections on software development concluded that specific recommendations to fix specific problems had a considerable positive effect [19]. Another study tracked how attempts to resolve usability problems influenced the usability of the system. Their feedback included redesign proposals to individual usability problems [12]. A more recent study has developers assess the relevance of usability problem descriptions compared to redesign proposals, and it was concluded that developers assessed redesign proposals as having higher utility in their work than usability problem descriptions but no developers, however, wanted to receive only problems or redesigns [6]. This also led to a set of recommendations for describing usability problems in practical usability work. For an overview of suggestions for other kinds of feedback, see [8].

The research on usability problem descriptions and alternatives has been questioned by authors with a background in usability practice. For example, it has been argued that in much research on usability evaluation methods, the effectiveness of the different methods has been compared in terms of the usability problems identified with an assumption of a direct link to design improvements [20]. It has also been claimed that the literature on usability evaluation is fundamentally flawed by its lack of relevance to applied usability work, and a key flaw is the focus on finding, rather than fixing, usability problems [22]. There are also practitioners who have been more constructive by emphasizing that redesign proposals are too often quick fixes, and they are only as brief as many of the problem descriptions. Therefore, they suggest that a more developed form of redesign proposals is defined [5].

The usability reports and the variety of alternatives that have been suggested share a common feature. They are almost exclusively based on usability specialists acting as evaluators who are being active in describing problems and creating redesign proposals. The developers, on the other hand, are passively receiving the descriptions and proposals. This is not a very effective way of establishing a common ground for improving the system, as emphasized in the introduction with reference to key education theories.

3 Exploratory Study

Our exploratory study aimed to investigate the effects of integrating usability evaluation feedback into a tailored design activity in a workshop setting where the developers were actively involved. Thereby, we aimed to alleviate some of the inherent problems of merely presenting usability problem lists.

3.1 System

The evaluated system is a web-based building permit system that citizens and companies use to apply for building permits, e.g. for construction of new houses or factories and refurbishing of existing houses. The existing system was an interactive form that was directly inspired by a previous paper-based form, which was usually used by Danish municipalities. The form is an integrated part of a collection of web-based services that is provided for municipalities in Denmark.

3.2 Usability Evaluation

We conducted a classic think-aloud usability evaluation, cf. [18], of the building permission system to generate a usability problem list. This was done in a traditional manner using lab evaluations with pre-assigned tasks.

3.3 Participants

10 subjects (3 females and 7 males) participated in the evaluation of the system. The participants were from 29 to 64 years old. They all had experience with the Internet and web-based services in general, but had varied levels of formal training and education (from a self-taught to a PhD) and had varied experiences with filling in online municipality forms and also varied experience with building and refurbishing houses. None of the participants had any prior experience with the system that was evaluated.

3.4 Procedure

From understandings of the prospective users and system functionality, the evaluation team constructed usage scenarios and task assignments for the usability evaluation. In total, three tasks were defined.

The four authors of this paper conducted the evaluation. The ten participants were scheduled for participation in the evaluation over two days, with five participants each day. Before the sessions, the participants were given a written test instruction outlining the purpose and the procedure for the evaluation. Furthermore, the participants were asked to fill in a demographic questionnaire. The sessions were conducted in a classical usability lab with one-way mirrors, where the participants were placed in a test room with the test monitor. A data logger in an adjacent observation room observed the evaluation while making notes. The participants were asked to follow the

think-aloud protocol during the test. After the test sessions, the participants were asked to fill in a NASA TLX test [24, 25] to assess how they perceived the mental workload.

Each day, after all evaluation sessions, we conducted a problem identification session using the Instant Data Analysis method [26]. This method was used to facilitate quick and efficient identification of the usability problems, which were needed in the subsequent design workshop. The usability problem identification procedure for both days were as follows 1) Problems observed from memory by the test monitor, 2) problems observed from memory by the test monitor or the logger, 3) problems identified from systematic walk-through of assigned tasks, 4) problems identified from systematic walk-through of logger's notes, and 5) severity categorization of each identified usability problem.

3.5 Result

We identified a total of 75 usability problems of which 7 were categorized as critical, 38 as serious, and 38 as cosmetic. Furthermore, we classified 4 of the problems as incomplete because we did not have adequate knowledge of the application domain to classify these problems according to the scale.

Considering the identified problems, we found that several of them were related to the basic concepts where the participants found it difficult to distinguish between the different kinds of permissions and applications. Furthermore, none of the participants could figure out which fields were mandatory, causing them to be uncertain about the completeness of the submitted form. Also, the division between building applicant and owner caused several problems for most of the participants. Finally, nearly none of the participants were able to attach appendices and drawings to their application. Accordingly, the NASA TLX test showed that the participants assessed mental workload and frustration as the most important contributors to the experienced workload.

3.6 Redesign Workshop

Previous research has demonstrated that usability evaluation feedback inherently involves several challenges and obstacles, as emphasized above in the overview of related work. To address some of these issues, we designed a redesign workshop with the aim of facilitating feedback through active developer participation and collaborative activities between developers and the usability specialists who had acted as evaluators.

The fundamental idea of the redesign workshop was to "look forward" through integrated design activities rather than having a retrospective approach only focusing on identified problems and, thereby, not focusing on overcoming the limitations of the current system. In the workshop, the software developers should become active in the evaluation of their own system. To achieve this, we integrated three components: 1) active involvement of the developers, 2) focus on the future system, and 3) support from usability specialists.

3.7 Participants

The participants in the redesign workshop were developers from the software company Dafolo A/S and usability specialists from Aalborg University. The participating developers represented different skills and job positions at the software company ranging from programmers over project leaders to a development manager. We will refer to them collectively as developers. In total, five developers participated in the workshop.

The usability specialists were the authors of this paper and are all researchers within interaction design and human-computer interaction and have extensive experience with usability evaluation. They held either a PhD or Master degree within human-computer interaction.

We divided the participants into three groups:

- Group 1: 3 Developers
- Group 2: 2 Developers + 1 usability specialist
- Group 3: 3 usability specialists

The overall aim was to explore to what extent the developers could contribute actively to the creation of redesign proposals and how much support they would need to do so. We also observed the group dynamics and assessed the redesigns they.

3.8 Preparation

We prepared the redesign workshop in the following way. We reduced the original list of usability problems from 75 to a short-list of 25 problems. For the reduced list we selected problems that were associated with one particular use scenario on applying for construction permit. Thus problems that were irrelevant to this particular scenario were omitted. Also, the reduction was done in order to create a basis for the redesign activity where the participants would be able to address all or nearly all problems. The reduced problem list contained 2 critical problems, 2 serious problems, 7 cosmetic problems, and 14 incomplete problems.

We also formed the three groups in advance. We divided the five developers and the four usability specialists into the three groups, where we strived for diversity in each groups with both junior and senior participants. Group 1 was formed with three developers only, and without any usability support. Group 2 was formed by combining two developers and a usability specialist who should provide support in the process. Group 3 was formed to provide product support for the other two groups; they conducted a redesign activity in advance of the workshop.

The participating developers received no information or results from the usability evaluation prior to the workshop as our experience was that this was counter-productive, because the developers would go into a defensive mode..

3.9 Activities

The design workshop consisted of three major activities (or elements):

1. Problem presentation and illustration
2. Redesign session
3. Plenary presentation and discussion

The first activity encompassed introduction to the workshop and presentation of the usability evaluation results. We scheduled this activity for an hour and it served the purpose of introducing the evaluation to the developers and more importantly presentation of the short-list of usability problems. The selected 25 problems were presented and illustrated in the interface and their severity was motivated. The usability specialists were leading this activity. All participants received a copy of the problem list.

The second activity was the redesign sessions. We gave the groups the task that they should produce a redesign of the system that addressed the shortlist of usability problems. Group 1 and 2 were placed in two separate rooms in our usability lab, allowing us to monitor them while working on the redesign. Two of the usability specialists from Group 3 acted as data loggers for each of the groups. The groups were instructed to 1) produce a written description of the redesign proposal, 2) illustrate the new interface layout and flow on a flip-over, and 3) note which problems they had addressed. Furthermore, they were instructed to prepare an oral presentation in the plenum session.

The third activity was a plenary session where group 1 and 2 presented their redesigns. After that, Group 3 presented the redesign they had prepared in advance. Each group presented their solutions for app. 15 minutes followed by questions and discussion for another 15 minutes. After the individual presentations, the solutions and design strategies were compared and discussed.

We measured the developers' attitudes to or perceptions of the strengths and weaknesses of the current system. This was done three times during the workshop: before they were introduced to the problem list, after the design session, and after the entire workshop. Each time, the participants were instructed to list 5 weaknesses and 5 strengths of the existing system.

4 Findings

In this section we present the finding from the explorative study with focus on the redesign proposals that were produced and the usefulness of the workshop.

4.1 Redesign Strategies and Design Proposals

This subsection describes the three redesign strategies and the design proposals that were developed during the design workshop.

Redesign strategies. The two developer groups (1 and 2) adopted a bottom-up strategy where they mainly addressed usability problems at the low-level description. We identified no visible effect of introducing a usability specialist into a developer group (#2) compared to the group consisting solely of developers (#1).

The adopted strategies also implied that Group 1 and 2 addressed the usability problems rather explicitly and had no problems in assessing which usability problems their solution addressed or solved.

The lack of an overall design strategy caused problems for both groups, as they several times had to change or alter their solutions, e.g. order of data input.

The developer group (#1) addressed nine usability problems in total. They primarily focused on the critical and serious problems and intentionally left out cosmetic problems in the beginning of the activity. Of the nine problems addressed, three problems were critical, two were serious, and four were cosmetic. As stated above, their overall strategy was a bottom-up approach focusing on the problems one by one.

The developers and usability specialist group (#2) addressed a total of 12 usability problems. Of these, three were critical, two were serious and three were cosmetic. Group 2 also adopted the bottom-up approach where they addressed problems as they were listed in the problem list instead of focusing on solving critical issues first.

When the usability specialist group (#3) created their redesign proposal in advance of the workshop, they adopted a more top-down strategy where they started to discuss the overall future solution, e.g. the presentation order of the different elements of form filling. After creating the redesign, they retrospectively traversed the whole problem list to mark which problems they had resolved. while doing this, it seemed that they utilized their knowledge and experience from the actual usability evaluation. Sometimes they were uncertain if they had solved a usability problem because they had less knowledge about the application domain. They addressed a total of 16 problems, where four were critical, five were serious and seven were cosmetic.

Design proposals. Group 1 presented the most technically advanced redesign proposal in which the main theme regarded process automation in a wizard based solution, see Figure 1. The wizard consisted of the following steps:

1. Type of construction work
2. Owner/Applicant and property information
3. Type of construction work (further questions)
4. Submission overview/receipt.

To illustrate this, we first need to review one of the experienced usability problems regarding users not knowing whether to submit a permit application for the construction work or a more simple notification. A permit application is needed in case the construction work conflicts with district plans, in which case the municipality must grant permission before construction may begin. A notification is used when the construction work obeys district plans. The users (citizens) in our evaluation were unable to decide, whether they were required to submit a “permist application” or a “notification”, which was selected in the pdf form via radio buttons. Group 1 proposed an

advanced solution for this, in which users would answer a set of questions regarding the type and size of the construction work. They also proposed to incorporate a Geographic Information System (GIS) component in which users could draw sketches of the buildings. By using this type of information, the system could potentially decide if the construction work would need permission or just a notification, hereby automating the process.

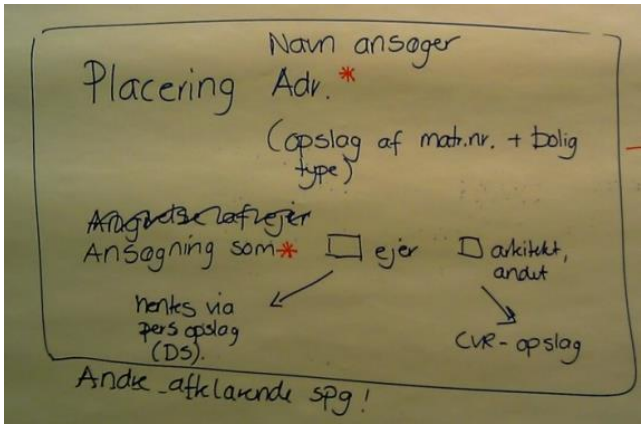

Fig. 1. A window from the redesign proposal from Group 1.

Another automation was suggested in the case of finding information about land register, which was another problem encountered by users in the evaluation. Users did not know where to find this information and the format of a land register was also unclear. Group 1 suggested that this problem could be overcome by automatically looking up the land register based on the given address of the construction work.

Group 1 also addressed a problem with irrelevant input fields. As an example, some users were confused about the information required in input fields for "Applicant address", because there also were input fields for "Owner address". The users were confused about whether to fill in the same information in both cases, as they were applicants and owners at the same time. Group 1 suggested a solution where the system leaves out irrelevant fields based on previous selections in the wizard.

Group 2 also suggested a wizard based design solution, in their case with the following steps:

1. Owner/Applicant information
2. Information about property
3. Type of construction work.

The main theme for this design proposal evolved around contextual help providing more clear and simple examples of when, for instance, users should select "permit application" or "notification". This addressed problems concerning complex terminology. Another problem addressed was also identified during the usability evaluation and regarded some users who did not notice the "Save" button, which allowed them to save the pdf form on their computer. Group 2 addressed this by placing all buttons consistently in a toolbar at the same location of each page (at the bottom). This group

also focused on a problem regarding the format for typing in addresses in input fields. In the pdf form, one field about street address was split into two parts: One field for street name and another for the house number in the street. Other fields in the form, however, required the user to type in street name and number in the same field. This inconsistency confused some of the users. Group 2 suggested a solution in which all address fields were of the same format, more specifically they designed for one input field containing both street name and number. As an extra feature this group also added a progress bar at the top of the layout to indicate the current state of the submission process.

Group 3 also proposed a wizard-based redesign. This group chose a top-down approach which made them emphasize a complete reordering of the workflow in the following steps:

1. Owner/Applicant information
2. Information about property
3. Type of construction work.

These steps are similar to the ones defined by Group 2, but the main theme for Group 3's design was simplicity for the citizen. For instance, they focused on letting office clerks decide whether or not a submitted application should be categorized as "permit application" or "notification" such that the citizen would not need to decide. This contrasts with the solutions suggested by groups 1 and 2, which suggested that the user (citizen) should decide.

Summary. From the overview of the design process, we see that the redesign strategies varied between top-down and bottom-up approaches and that the three groups addressed a different number of usability problems. The main themes of the redesign proposals also differed between the three groups, but all applied a wizard approach of three to four steps. The wizard steps of Group 2 and 3 were similar while Group 1 selected an alternative order and also designed an extra step. All steps in the redesign solutions deviated from the order in the original pdf form.

4.2 Problem Understanding

This subsection describes our findings on the developers' perception of the problems of the system. In the following we describe the numbers and categories of identified strengths and weaknesses and the collective list of these as prioritized by the five participants.

Categories of Strengths and Weaknesses. Using grounded theory [21] we identified a set of categories to describe and compare the contents of the perceived strengths and weaknesses before and after the redesign exercise. The identified categories are shown in the leftmost column of Table 1 which shows the number of mentioned strengths and weaknesses distributed according to these categories.

We see that the participants were able to identify 20 strengths in total before the redesign task and 13 after and 20 weaknesses in total before the group work and 20 after. Thus, the number of strengths perceived by the participants was considerably reduced after the redesign task, whereas the number of weaknesses remained the same.

Table 1. Number of strengths and weaknesses distributed according to identified categories. * denotes the categories with a change of at least 2 in either strength or weakness.

	Strengths before	Strengths after	Weaknesses before	Weaknesses after
Barcode	2	2		
Breadth	1			
Save	1			
Consistency		1		
Digital signature *	6	4	1	
Recognizability *	5	3		
Clarity *	2	2	5	1
Simplicity	1		1	
Instructions	1		5	4
Attachment	1	1	1	
Excess				1
Structure *				3
Terminology *			1	5
Visibility			2	1
Workflow			1	
Land register			2	1
Transparency *				3
Total	**20**	**13**	**20**	**20**

It is also interesting to note that the categories instructions, simplicity, breadth and save were mentioned as strengths before the exercise, but not after. Only a single new category emerged after (consistency).

We observed more shifts concerning perceived weaknesses compared to strengths. In Table 1 it is shown that three new categories (excess, structure and transparency) emerged after the redesign exercise, while four others disappeared compared to before (digital signature, simplicity, attachment, workflow).

For strengths, two categories were rated more than 2 higher after the exercise than before, and none were rated more than 2 lower after the exercises, while for weaknesses, three categories were rated more than 2 higher after the exercise than before, and one category was rated more than 2 lower after the exercises.

Based on the above observations we see that the redesign exercise lead the participants to change their perception about the system strengths, while for weaknesses it is more a change in type according to our findings from the usability evaluation.

Collective Strengths and Weaknesses. Table 2 shows the five strengths and weaknesses, which the participants collectively agreed on at the end of the workshop, in prioritized order.

For the five collective strengths it is interesting to note the connection to the result found in Table 1, where the participants initially perceived the categories of digital signature and recognisability as the foremost strengths but after the redesign exercise the excitement of these seemed reduced. The possible consequence of this may be what we see in Table 2 where digital signature is unmentioned and recognisability has the lowest priority of the five strengths.

We see a similar connection considering the collective list of five weaknesses in the categories instructions, excess, terminology, instructions and structure (in prioritized order). In Table 1 we see that instructions and terminology are the primary perceived weaknesses before and after the group work which correspond well to the priorities in Table 2. The two remaining categories of excess and structure were perceived important by the participants after the redesign exercise, which is also reflected in Table 2.

Thus, we see a connection in categories of strengths and weaknesses perceived after the exercise and the collective prioritized list.

Table 2. Five collective strengths and weaknesses (prioritized).

	Collective strengths	**Collective weaknesses**
1	Barcode useful for office clerks [**Barcode**]	Instructions provide no overview [**Instructions**]
2	Pdf form provides a final overview before submission [**Clarity**]	Many input fields seem unnecessary [**Excess**]
3	Pdf form provides an overview of requirements [**Clarity**]	Terminology in form difficult to understand [**Terminology**]
4	Provides ability to save temporarily [**Save**]	Required level of detail on provided information is unclear [**Instructions**]
5	Pdf form similar to paper version – well known to office clerks [**Recognisability**]	Layout of pdf form is not intuitive [**Structure**]

Strengths and Weaknesses after Reading the Usability Report. Table 3 shows the perceived strengths after reading the usability report. From this we see a total of 14 strengths, which corresponds well to the total of 13 found after the redesign task, see Table 1. Most of the categories after reading the report are the same as the strengths

identified in Table 1. Recognisability has, however, decreased considerably, and two new categories; time and integration have emerged. We also see similarities to table 2 showing the list of collective strengths, of which all categories are represented in table 3. Thus, the number and types of strengths after reading the report correspond well to the findings identified previously.

Table 3. Perceived strengths after reading the usability report distributed according to identified categories.

	Strengths before report	Weaknesses after report
Digital signature	5	1
Clarity	3	3
Recognizability *	1	
Barcode	1	
Save	1	
Time	1	
Integration *	2	
Structure		2
Terminology *		2
Instructions *		7
Visibility		1
Transparency *		1
User errors *		3
Functionality		1
Affordance		1
Total	**14**	**22**

Table 3 presents the perceived weaknesses after reading the usability report. The number of identified weaknesses is 22, which corresponds to the total of 20 identified after the group work, shown in [8]. Considering the categories observed after reading the report we see that the majority of these are the same as those identified after the redesign task and in the collective list of weaknesses shown in Table 2. The categories user errors, functionality and affordance, however, were not identified during the workshop and are new. Regarding the weaknesses we also see a strong connection to the observations done during the workshop.

4.3 Usefulness of Workshop and Usability Report

After reading the usability report we asked participants to anonymously respond to a survey asking questions about the usefulness of the workshop and the usability report. The majority of questions required participants to answer a five point Likert scale where 1 = Strongly disagree and 5 = Strongly agree.

Usefulness of Workshop. All participants rated the general usefulness of the workshop as 5, which show clear indication of a positive perceived value.

We were also interested in examining the understandability of the presented usability problems before and after the redesign exercise, that is, whether or not the exercise helped them gain a more thorough understanding of the problems. Based on the presentation of the problems (just before the exercise) the average rating of the clarity was 4.6, which clearly indicates that the participants found the problem descriptions during the presentation to be understandable.

When asked if the redesign exercise provided a more clear understanding, the participants on average rated 3.2, which indicate that the exercise did not increase or decrease the clarity of the presented usability problems.

Usefulness of Usability Report. The usefulness of the report was rated as 4.8 on average. Although this rating is a bit lower than the workshop rating, we see a clear indication that participants also found the report valuable. We also asked participants to rate the understandability of problem descriptions in the report, which resulted in an average rating of 4.4. This result is also a bit lower compared to the workshop.

In addition we asked participants to openly comment on their experience in reading the report and all found the report positive. As an example one of the participants mentioned that "I think that the report is very good. It provides a good explanation of the approach and it works fine as an encyclopaedia of problems in the current system. The individual problems are easy to understand and it is very good that the critical issues are described in further detail".

Finally we asked participants to openly comment on both the workshop and the usability report. Four of these explicitly mentioned the workshop, where one answered that "The report combined with the workshop was very constructive because we discussed specific design solutions to overcome the problems" and another stated that "I hope that we will see more of this. It is very insightful to discuss different ideas on how to improve a system. It is also fun to disagree on, what the most important strengths and weaknesses are and then listening to the arguments".

5 Discussion

Our exploratory study with the redesign workshop originated from the experience that usability evaluation results and especially feedback have proved to have limited effects on software development in practice [5, 7, 15]. From an educational point of view, the transfer of usability evaluation knowledge seems to be ineffective and have very limited success.

In designing the workshop, we were inspired by Dewey's theory on education, which is seen as being constituted by transmission through communication. It reflects a sharing experience process until it becomes a common possession [4]. Dewey emphasizes that traditional teaching's concern with delivering knowledge need to be balanced with a much greater focus on the students' actual experiences and active learning [17].

We attempted to integrate cooperative and active learning through actual participation of both evaluators and developers in the process of conceptualizing the usability of the system and the redesigning activity. Thus, we explored how Dewey's theory on education can be combined with the idea that redesign proposals should be a key vehicle for providing feedback from a usability evaluation. It is interesting that all the developers rated the workshop very high, and their understanding of the problems identified in the usability evaluation clearly improved.

Dewey's educational theory focusses on learning by doing. But Dewey not only reimagined the way that the learning process should take place, but also the role of the teacher in the learning process. The teacher should not stand at the front of the room doling out bits of information to be absorbed by passive students. Instead, the teacher's role should be that of facilitator and guide. Instead, the teacher should be a partner in the learning process, guiding students to independently discover meaning within the subject area. He expressed it this way: "Were all instructors to realize that the quality of mental process, not the production of correct answers, is the measure of educative growth something hardly less than a revolution in teaching would be worked."

In our study, the developers grew and improved their understanding of the usability problems that had been identified even if they did not produce the best overall solution. Yet they were able to contribute constructively to create redesign proposals and in particular to flesh out the concrete details of the overall pattern that was created by the usability specialists.

In our study, we had two groups of developers where one of them had process support from a usability specialist while the other only included developers. The experiences from the process and the redesign proposals created demonstrate very little difference between those two groups. Thus the inclusion of a usability specialist in one of the groups did not seem to make an important difference on the process. Group 1 and 2 both chose a low level bottom-up approach to redesign. The proposed designs evolved around three main themes, where Group 1 suggested the most technical advanced solution, while Group 2 focused on contextual help. We expected Group 1 to come up with a simpler and more pragmatic solution than the one they chose, since this would ease their workload on the upcoming system considerably. This indicates that they were dedicated in solving the usability problems and hereby to create an improved user interface.

After the workshop, the developers designed and implemented a new version of the system. This version was based on the overall ideas from Group 3, filled in with details from the redesigns created by Group 1 and 2. While the inclusion of a usability specialist in Group 2 made little difference compared to Group 1, the basic ideas in the solution from Group 3 was very useful as a basic pattern for the actual redesign of the system. Thus the influence from the usability specialists was more on the product and a pattern for that, than on the redesign process itself.

The aim of the workshop was to involve the developers actively in a collaborative process of redesigning the existing system and thereby improve the developers' understanding of the usability problems that had been identified. We measured the success in terms of the perceived strengths and weaknesses of the system as it was ex-

pressed by the developers. We observed a substantial reduction in the number of identified strengths after the redesign task. This shows that participants viewed the system in a more realistic perspective, which may have been caused by them improving their understanding of the usability problems through the active involvement in the redesign exercise.

We observed that the categories instructions, simplicity, breadth and save were mentioned as strengths before the exercise, but not after. This indicates that the participants have been influenced by the identified usability problems. Høegh et al. [8] conducted a study with two developers, who were asked to list five strengths and weaknesses before and after reading a usability report. That study showed that the developers did not change their perceived strengths noticeably, which contrasts with our findings. This difference can be explained in terms of Dewey's theory on education because they only passively read the usability report.

Considering the perceived weaknesses we observed that participants identified the same number before and after the redesign exercise. This is a bit surprising since we saw a considerable change in the strengths before and after the redesign task. However, we observed considerable changes in the categories of clarity, structure, terminology and transparency. The categories digital signature, simplicity, attachment and workflow were identified as weaknesses before the exercise but not after, and we saw that participants who perceived these categories as problems shifted towards the new categories of structure, excess and terminology. Thus we see the participants being influenced by the group work in terms of perceived weaknesses. This corresponds to the findings of Høegh et al. [8], which show a developer completely altering his perceived weaknesses after reading the usability reports.

In our study, the developers had not received prior training in user interface or interaction design. It would be interesting to see if they could do better if they had been trained before the redesign process.

Training relates to the more general idea of qualifying barefoot practitioners in usability engineering. Bruun and Stage [27] have discussed this in relation to usability evaluation where they trained practitioners in that with very positive results. When software developers can be trained to conduct usability evaluations, it should be possible to achieve the same in redesign. This is important in practice, because the amount of usability specialists is limited, and many smaller development companies have no possibility of involving external specialists.

In this collaboration, a usability specialist may become a leader and facilitator of design, rather than the sole source of it. It is often argued that generating many alternatives is a key part of good design. In that case, it seems to be an advantage to have a facilitator for a larger group of barefoot practitioners rather than requiring this individual to be solely responsible for generating design ideas.

6 Conclusion

This paper has presented an exploratory study where a usability evaluation was followed by a redesign workshop. In this workshop, evaluators and developers worked

together in an active collaborative learning process inspired from Dewey's educational theory. The aim of the workshop was that the developers should understand and try to resolve the identified usability problems. The workshop was a success in the sense that the developers' understanding of the problems with the system changed considerably, and they expressed a high level of satisfaction with the process. They did not produce the redesign proposal that eventually was implemented, but they were able to choose that as the best redesign, and they could fill in the details of this overall design pattern to arrive at a considerably improved system. The study indicates that usability specialist support to the redesign process is less important than having usability specialists generate a redesign pattern where the details can be filled in by the developers.

We have conducted an exploratory study. It is not possible to make any definite conclusions on this qualitative basis, but the study is useful for defining more quantitative hypotheses for a further study.

The basis for assessing the success of the workshop has been the developers' understanding of the strengths and weaknesses of the system and their statements about the process. It would be interesting with a follow-up study to assess the downstream utility of the workshop in the developers' work to eliminate the usability problems that have been identified and to measure the utility of the redesign proposals that were produced in the workshop.

References

1. Bailey, R. W., Allan, R. W., & Raiello, P. (1992) Usability Testing vs. Heuristic Evaluation: A Head-to-Head Comparison, Proc. Human Factors Society 36th Annual Meeting, 409-413.
2. Bevan, N., Singhal, N., Werner, B., Degler, D. and Wilson, C. Formative Evaluation. In Usability Body of Knowledge.
   http://www.usabilitybok.org/methods/formative-evaluation
3. Capra, M. G. (2007) Comparing Usability Problem Identification and Description by Practitioners and Students. In Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting (pp. 474-478). Santa Monica, CA: HFES.
4. Dewey, J. (1916) Democracy and Education: an introduction to the philosophy of education.
   http://en.wikisource.org/wiki/Democracy_and_Education
5. Dumas, J., Molich, R., & Jefferies, R. (2004). Describing Usability Problems: Are we sending the right message? Interactions, 4, 24-29.
6. Hornbæk, K., and Frøkjær, E. (2005) Comparing usability problems and redesign proposals as input to practical systems development. Proceedings of CHI 2005, pp. 391-400.
7. Hornbæk, K., Frøkjær, E. (2006) What kind of usability-problem description are useful for developers?, HFES 2006, pp. 2523-2527.
8. Høegh, R. T., Nielsen, C. M., Overgaard, M., Pedersen, M. B. and Stage, J. (2006) The Impact of Usability Reports and User Test Observations on Developers' Understanding of Usability Data: An Exploratory Study. International Journal of Human-Computer Interaction, 21(2):173-196.

9. Hvannberg, E.T., Law, E. L.-C., and Lárusdóttir. M. K. (2007) Heuristic evaluation: Comparing ways of finding and reporting usability problems. Interacting with Computers 19(2): 225-240.
10. ISO (1998). ISO 9241-11 Ergonomic Requirement for Office Work with Visual Display Terminals (VDTs) – part 11: Guidance on Usability. Switzerland, International Organization for Standardization.
11. Jeffries, R. (1994). Usability Problem Reports: helping Evaluators Communicate Effectively with Developers. In J. Nielsen & R. L. Mack (Eds.), Usability Inspection Methods (pp. 273-294). New York, NY: John Wiley.
12. John, B. E. & Marks, S. J. (1997) Tracking the Effectiveness of Usability Evaluation Methods, Behaviour and Information Technology, 16, 4/5, 188-202.
13. John, B. E. and Packer, H. (1995) Learning and using the cognitive walkthrough method: a case study approach, in I. Katz, R. Mack and L. Marks (eds), Proceedings of ACM CHI`95 Conference on Human Factors in Computing Systems (ACM, New York), 429-436.
14. Lavery, D., Cockton, G., & Atkinson, M. P. (1997) Comparison of evaluation methods using structured usability problem reports. Behaviour & Information Technology, 16, 4/5, 246-266.
15. Law, E. L-C. (2006). Evaluating the downstream utility of user tests and examining the developer effect: A case study. International Journal of Human-Computer Interaction, 21 (2), 147-172.
16. Mack, R. and Montaniz, F. (1994) Observing, predicting and analyzing usability problems, in J. Nielsen and R. L. Mack (eds), Usability Inspection Methods (John Wiley and Sons, New York), 295-339.
17. Polito, T. (2005) Educational Theory as Theory of Culture: A Vichian perspective on the educational theories of John Dewey and Kieran Egan. Educational Philosophy and Theory, 37(4):475-494.
18. Rubin, J. (1994) Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. New York: John Wiley & Sons.
19. Sawyer, P., Flanders, A., & Wixon, D. (1996) Making a Difference - The Impact of Inspections, Proc. CHI'96, ACM Press, 376-382.
20. Smith, A. and Dunckley, L. (2002) Prototype Evaluation and Redesign: Structuring the Design Space through Contextual Techniques, Interacting with Computers, 14, 821-843.
21. Strauss, A., & Corbin, J. (1998) Basics of qualitative research. Techniques and procedures for developing grounded theory (2nd ed.). Thousand Oaks, CA: Sage.
22. Uldall-Espersen, T., Frøkjær, E., Hornbæk, K. (2008) "Tracing Impact in a Usability Improvement Process", Interacting with Computers, vol. 20, issue 1, pp. 48-63.
23. Wixon, D. (2003) Evaluating Usability Methods: Why the Current Literature Fails the Practitioner, interactions, 10, 4, 29-34.
24. Hart, S. G. Nasa-task load index (nasa-tlx); 20 years later. NASA-Ames Research, 2006.
25. Hart, S. G. and Staveland, L. E. Developing the NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload. In P.A. Hancock and N. Meshkati (Eds.), *Human Mental Workload*, Amsterdam: North-Holland (1998), 239-250.
26. Kjeldskov, J., Skov, M. B. and Stage, J. (2004) Instant Data Analysis: Evaluating Usability in a Day. *Proceedings of NordiCHI 2004*, pp. 233-240. ACM.
27. Bruun, A. and Stage, J. (2014) Barefoot Usability Evaluations. *Behaviour and Information Technology*, DOI: 10.1080/0144929X.2014.883552.
28. Gothelf, J. (2013) *Lean UX*. O'Reilly.