REFINING THE NEEDS: AN EXPLORATORY STUDY THROUGH USABILITY TESTING

Magnus Åkerman[1], Malin Karlsson[1], Lars-Ola Bligård[2]

Chalmers University of Technology
Product and Production Development, Production Systems[1]
Product and Production Development, Design & Human Factors[2]

magnus.akerman@chalmers.se

Abstract: Factory competitiveness can be enhanced through the use of cognitive support tools at all levels, especially for shop floor operators. To get the most value from implementing a tool it has to fit the needs of tasks, environment and users. The aim of the study presented here was to examine how a usability test of a high fidelity prototype can add value to the need analysis phase. Results show that when using the prototype as a mediating object it can provide new insights to further refine the need analysis, especially when using an environment of high ecological validity.

Keywords: Usability test, ICT, mediating object, production, prototype.

1. INTRODUCTION

To improve competitiveness of factories, cognitive support tools should be utilized to a larger extent in all levels of the organisation (Karlsson, et al., 2013). These tools could be implemented in the form of information and communication technology (ICT). Many companies have invested in ICT for management levels, technicians and others, but among shop floor operators usage of ICT-tools is virtually non existent (Karlsson, et al., 2014). ICT-tools for factory operators would be partly or fully integrated with the company's information technology infrastructure. It will be expensive to invest in the wrong type of technology or overinvest (Fasth-Berglund & Stahre, 2013). Therefore, extensive use of ICT-tools will not guarantee a success. For most gain it is crucial to first find what information or communication support the operators actually need.

The common approach for conducting need analyses, to be used in design, is to conduct interviews and observations (Karlsson, 1996). However, this approach is time consuming and information derived from interviews is usually biased towards current ways of working. It is difficult for many people to imagine, for example, an entirely different information flow. To further refine the needs already found, it would be useful to simplify the observations, by conducting a controlled test in the correct environment. This could be done through usability testing of a prototype, where the prototype functions as a mediating object. A mediating object is something that stimulates discussion, enhances the operator's understanding of a device or concept and simplifies the dialogue between the operators and the developer (Engelbrektsson, 2004). The main idea is that when operators interact with the prototype it will be possible to elicit the needs based on the comments as well as from observing the use. More information may emerge that would be hard to retrieve without the actual interaction in the proper environment.

The aim of the study was to examine how a usability test can add value to the need analysis phase. Results are presented with discussions regarding general usefulness, benefits and issues to consider when conducting this kind of test. The test presented was conducted as part of a larger project with the intention to implement ICT-tools designed for factory operators based on the actual information and communication needs.

2. PROJECT OVERVIEW

Two universities and seven large manufacturing companies are currently involved in a Swedish science project ("Operator of the future," 2012) with the aim to implement and test ICT-tools. The unique feature of this project is that the tools are tailored for the actual needs, regarding information and communication, of the factory operators. The companies each chose a production line or area of focus, in some cases two, where they had either identified problems or deemed interesting in other ways. The first phase of the project was dedicated to a thorough need analysis of the different focus areas with interviews, surveys and observations (Grane et al., n.d.). To further verify and refine the need analysis a case from one of the manufacturing companies was chosen for conducting the usability test. A prototype was implemented with the aim to appear fully functional and integrated, i.e. a high fidelity prototype. The reason for this was to reduce the confusion about the functionality since the purpose was not to test the prototype itself. Instead, the sought after information was previously hidden problems or needs that might appear in the specific situation or after a change of behaviour triggered by interactions with the prototype.

3. THE CASE

The case selected for the test was a department in charge of a heat-treatment process at a large Swedish manufacturing company. Due to the layout of the process, which is quite large, the operators need to climb on steep ladders and crouch under low obstacles and they have to use their hands to stay in balance. The operators in the department are working in five shifts, with five to six operators in each. Each morning any one operator carries out a preventative maintenance round. This round includes checking various machinery and processes all over the department. The operators follow a checklist, printed on two large papers, where most of the points that should be checked are represented. The checkpoints include a short descriptive text and/or a picture as guidance. Some checkpoints also include a brief instruction for what to do if the checked object or process does not appear to be in a correct state. The checklist have some issues: pictures are small and sometimes hard to understand, no statistics regarding the checks or adjustments are captured, with few or no updates the list might include unnecessary checks and new checkpoints might be missed. Because of these problems, together with an easily implemented prototype suggestion, this case was chosen for the usability test and a prototype was implemented and adapted to the specific process.

3.1. The prototype

To replace the static paper checklist a dynamic digital support tool was implemented within the project. The dynamic features of the tool allowed for changing the check frequency using a web interface. The tool also consisted of a client where a user could see, update and interact with the checklist. This client was the target for the test.

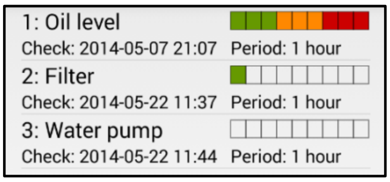

Fig. 1. List view of the prototype checklist tool.

The client is implemented as a mobile application runnable on Android platforms, which provides for a wide selection of hardware. Three different screen sizes were used during the test: 4.3-inch smartphone, 7-inch tablet, and 10-inch tablet. The application displays a list of all the checkpoints (list view). Each checkpoint has a color-coded bar that fills up to red when a new check is needed (Fig. 1). When selecting a checkpoint a detailed view of that checkpoint is displayed. The detailed view (Fig. 2) shows the exact same short instruction and image that can be found in the paper checklist. The user is presented with three alternatives: check with status ok, check with status ok but only after conducting an action, and finally check with status not ok. If the user checked a checkpoint as ok, with or without action, the bar for that checkpoint emptied. A checkpoint with a reported error did not have a time-bar and it was marked with a grey background. The smallest screen shows different views on separate screens while the larger screens showed both the list and details on the same screen.

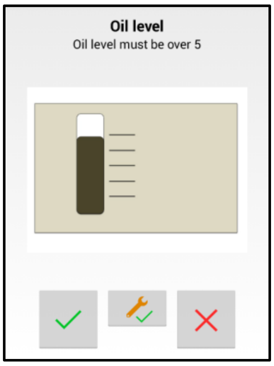

Fig. 2. Detail view in checklist prototype.

3.2. Evaluation method

The test was carried out as an experimental field study in the normal work environment of the operators who participated. The three prototypes of varying sizes were used for the usability study with operators from three different, randomly selected, shifts at the processing facility. The daily preventive maintenance round was the task of focus in the study. The shop floor operators were asked to use the prototype in their normal work environment. The variation in settings between test persons was the size of hardware and ways to carry it. The test was carried out at two separate occasions.

Three ways of collecting data was used during the study: observations of how the operators handled the prototype, a questionnaire and semi-structured interviews. One interview was conducted before the test. After the test the operator took the questionnaire and another, concluding, interview. With their permission the interview was recorded on tape and some of the tests were recorded on video.

Participation was voluntary and the manager of the department supported the test and encouraged participation. New participants were found through snowball sampling (Goodman, 1961) i.e. one participants were asked to recruit a new participant from the shop floor. There was no upper limit for number of participants, everybody that wanted to were able to try the prototype. There were a total of 9 participants in the study of varying ages (Fig. 3) and experience levels. Only male participants were present during the hours of testing.

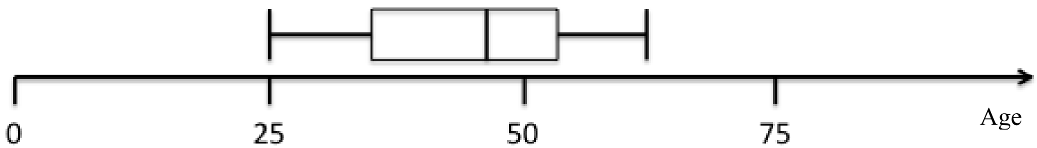

Fig. 3. Boxplot of the age distribution.

Test procedure

The area was prepared with some fake errors that the operators were expected to act on. The types of errors were plausible things that sometimes can be noticed during preventive maintenance rounds. The fake errors were written on paper sheets of various colours next to the spot of a few selected checkpoints, red meaning errors and yellow was quick fix action.

The operators were informed about the test during the morning meeting the same day and were asked to participate. At the start of the test the operator was informed that all participation was optional and that he could abort at any time without consequence.

Before test
A few background questions regarding age, employment time in the department, usage of digital tools, and experience from the maintenance round started of the interview. Then one prototype was given to the test person, in a size not used in the test, without any further introduction to how it worked. The operator was asked to investigate the application and try to figure out what functions it had and how they worked. After a few minutes, if the operator did not grasp all features of the prototype, the test leader interfered and showed on how the prototype works. When the operator felt confident in how to use the application the test in the factory started.

During test
The operator was given one of the three tablets and in some cases aids to carry it with. He was asked to carry out a selected part of the preventative maintenance round (Fig. 4) as they would normally do it but also to act on the errors pre-placed in the factory. During this part of the test the operator could handle, carry and interact with the prototype in any way they wanted and the observers tried to interfere as little as possible.

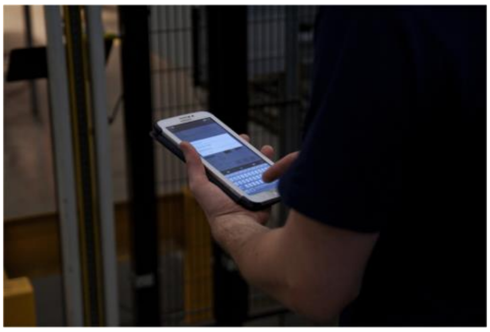

Fig. 4. Prototype used during test.

After test
When returning from the test the operator was asked to fill in a short survey regarding how he felt using the application, if it was easy to understand and if it could be of use in their work. The answers were used as a starting point for the semi-structured interview that followed. The preferred tablet size and way to carry it was discussed as well as issues that could arise using it in a production context.

3.3. Result and analysis of test

The main results from the test were observations of operator behaviour and statements from the interviews. A few noticeable behaviours of participants during the experiment were observed. People with reading glasses needed to use them when using the device. This was by itself quite annoying if interactions with the device were very frequent. Some operators have insensitive fingertips and had problems with small touch buttons, which meant that they often had to retype text. At some physical tasks, when the device needed to be put away, the operators did not always know what to do with it. Finally, at steep ladders or steps, most did not put the device in a pocket or carrying case, this was sometimes borderline dangerous behaviour.

After the test the participants gave feedback and their general thoughts. Most discussions regarded sizes and how to carry the tablets. The operators preferred different sizes of hardware (Figure 5). The motivation for using different sizes varied but could be that it was easier to bring, etc. Another point raised was that different sizes and ways to carrying would be good for different types of tasks and situations. E.g. a larger size would be ok if using it for a limited time or a specific task whereas a smaller hardware would be preferred if used a full work shift. Some operators would prefer putting the tool in their pocket while others wanted to have it visible at all times. Considerations were also expressed regarding usage of gloves and that the factory environment in general was not suitable for sensitive equipment.

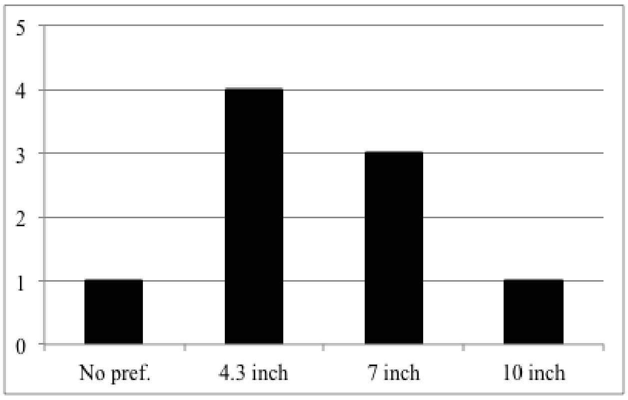

Fig. 5. Preferred size of prototype after test.

The general idea of a dynamic checklist for the maintenance round proved to be successful (Table 1). Most of the problems from the static paper list seem to be solved with this tool: visualizing imminent checks, error states, statistics etc. Furthermore, many expressed that they wanted to add functions to the tool and even use other software in the device since they now had it anyway.

Table 1. Overall perception before and after test.

Nr	Before	After
1	Neutral	Neutral
2	Positive	Positive
3	Positive	Positive
4	Neutral	Positive
5	Neutral	Positive
6	Positive	Positive
7	Negative	Positive
8	Neutral	Positive
9	Neutral	Positive

There are three things to consider regarding results from the study. The first thing is if any new needs regarding information and communication were found. The second consideration is if new results could have been found during the interview and observation phase. Thirdly, if the results contradict previous results, would ignoring this test provide false assumptions?

New results were not surprisingly related to situations with prototype interactions in the specific factory environment. Some dangerous situations were observed because of how the prototype was carried by some people. There were situations when the operator temporarily could not hold on to the prototype e.g. when climbing steep ladders or doing manual work. Interactions with the actual prototype highlighted other problems such as small keyboards for insensitive hands, gloves made for factory environment are not touch screen friendly, oil spills etc.

Many of these problems that arose were possible to foresee even without this test. Observations in the factory environment could have been enough to catch most of the new results, but it would have been difficult to know when and who to observe.

Some interesting results were identified regarding what size of hardware to use for this work. Preferred sizes before testing did not align with the intuitive guess, and many changed their preference after the test. This is interesting because it shows that this information would not have been possible to get with either interviews or observations.

4. DISCUSSION

Even though this was an exploratory study, the approach towards the operators was very systematic. This was probably important since it made them more involved. It also improves the overall credibility of the study. The important results were never with the details in the prototype that was tested, but that was never expressed to the operators. It is important to help the participants to focus on simple details so it is possible to observe their real behaviour. This can be achieved by using a mediating object, in this case, the prototype targeted for usability testing.

Research has shown that mediating objects are useful in eliciting user requirements since it can act as a tool for recall, something to refer to or use in demonstration (Engelbrektsson, 2004). The mediating object becomes an enabler for the user. In this paper´s proposed use, the mediation object transforms the abstract idea of the developers to something concrete for the operators to relate to. It can in the same way be difficult for operators to imagine or express their needs, while an interaction with a prototype makes it easier for the operator to relate and communicate if the prototype helps them or if they have other, earlier unspoken, needs.

The researchers in charge of the test procedure did fully understand the tasks and problems operators could face during the test. They had been partly or fully involved during the entire need collection phase. This understanding was important for the test and the operators' confidence in it. The test procedure was simple and limited so that every operator that participated understood all parts of it. Together, these parts helped to draw the focus towards interesting behaviour and not on details about complicated work tasks. Furthermore, these behaviours would not have been interesting without conducting the test in a valid environment, also known as ecological validity.

The ecologic validity (Howitt, 2010) refers to the extent to which the result of study can be generalized to the settings and situations in which the phenomenon would naturally occur. It is also likely that the operators felt more engaged since it in a clear way related to their everyday work. Whereas a laboratory environment could have made the participants feel like it was only research and not important or even relevant to them.

Often it is possible to get a long way towards a successful product just by guessing what the actual needs are. But if the investment or the consequences of implementing the wrong product is very high, then it becomes important to validate and possibly refine the needs. And as a complementary method, this study shows that a usability test using a prototype as mediating object can be a good approach to validate and refine the needs. However, the usability test shall not be seen as a replacement of initial need analysis, with interviews and observations, rather a complementing approach where the strength comes from using them both together.

Software that is once written can be copied and reused forever even if the final version is very different from the first draft. Because of this, software development often takes an evolutionary approach, where functions are added and tested iteratively (Silva da Silva, et al., 2011). Even if it is often argued that a throw away approach is very useful early in the project even in software development (Kordon, 2002), it is seldom used (Ramesh, et al., 2007). However, this type of test is still not the same as the throwaway prototyping approach. The differences lie in the focus of the analysis. Usually, when implementing through prototyping, the aim of usability testing is to see how well the prototype fit the problems you seek to solve. But maybe it is possible, when searching for new needs and problems, to use different mediating objects or confirm considered problems by implementing a very unlikely or even unfitting prototype.

5. CONCLUSION

It is known that usability tests are important when testing graphical user interfaces or other types of design functionalities. The experiment presented above shows that, when used as an exploratory study, a usability test with a high fidelity prototype can also provide new insights regarding the users and the environment. An unexpected result was the outcome of tablet size preferences. The variations observed did often go against common preconceptions or prejudices. Four general characteristics were observed which should be considered when conducting exploratory usability tests:

- Systematic approach
- Simple mediating objects
- Everyone involved should understand the task and/or problem
- Ecological validity

ACKNOWLEDGEMENT

The authors would like to thank the company and all operators involved in the research. The work has been carried out within the Production Area of Advance at Chalmers and sponsored by VINNOVA. This support is gratefully acknowledged.

REFERENCES

Engelbrektsson, P. (2004). *Enabling the user - Exploring methodological effects on user requirements elicitation*. Chalmers University of Technology.

Fasth-Berglund, Å., & Stahre, J. (2013). Cognitive automation strategy for reconfigurable and sustainable assembly systems. *Assembly Automation*, *33*(3), 294–303. doi:10.1108/AA-12-2013-036

Goodman, L. A. (1961). Snowball Sampling. *The Annals of Mathematical Statistics*, *32*(1), 148–170.

Grane, C., Hansson, I., Karlsson, M., Osvalder, A.-L., Alm, H., & Stahre, J. (n.d.). *Analysis of Future Needs of Supporting Tools for Information Transfer in Industry Based on Field Studies in Mining and Manufacturing*.

Howitt, D. (2010). *Introduction to qualitative methods in psychology*. Pearson Education Limited.

Karlsson, M. (1996). *User requirements elicitation: a framework for the study of the relation between user and artefact*. Chalmers University of Technology.

Karlsson, M., Gustafsson, A., Grane, C., & Stahre, J. (2014). ICT usage and benefits in Swedish manufacturing and processing companies.

Karlsson, M., Mattsson, S., Fast-berglund, Å., & Stahre, J. (2013). Could the use of ICT tools be the way to increase competitiveness in Swedish industry? In *Procedings of the 12th HMS, Analysis, Design, and Evaluation of Human-Machine Systems*.

Kordon, F. (2002). An introduction to rapid system prototyping. *IEEE Transactions on Software Engineering*, *28*(9), 817–821. doi:10.1109/TSE.2002.1033222

Operator of the future. (2012). Retrieved May 15, 2014, from http://www.chalmers.se/en/projects/Pages/Operator of the Future.aspx

Ramesh, B., Cao, L., & Baskerville, R. (2007). Agile requirements engineering practices and challenges: an empirical study. *Information Systems Journal*, *20*(5), 449–480. doi:10.1111/j.1365-2575.2007.00259.x

Silva da Silva, T., Martin, A., Maurer, F., & Silveira, M. (2011). User-Centered Design and Agile Methods: A Systematic Review. *2011 AGILE Conference*, 77–86. doi:10.1109/AGILE.2011.24