

Usability Evaluation in a Remote Access Scenario

Thomas BÜRKLE^{a,1}, Marko MILETIC^a, Kezia LÖFFEL^b, Simone GIUNTA^c, Yaren BAYRAKTAR^c, Patrizia SALZMANN^b, Jürgen HOLM^a

^a*Bern University of Applied Science, Bern, Switzerland*

^b*Swiss Federal University for Vocational Education and Training, Switzerland*

^c*University of Applied Sciences of the Grisons, Chur, Switzerland*

ORCID ID: Thomas Bürkle <https://orcid.org/0000-0002-2936-5375>

Abstract. We present an approach to use the think aloud method for formative usability evaluation for medical software in a remote access scenario which combines researchers in two different locations and test persons distributed across five Swiss hospitals into one study using MS-Teams. We could show that a) these tests can be performed remote in a tight time schedule and b) that results valuable for the formative development of the examined mockup solution could be achieved. Two sessions had to be rescheduled for technical problems but all test persons could complete the think aloud study in 30 minutes at their usual working environment. This offers the opportunity to software providers to develop medical software with better usability in close cooperation with future users.

Keywords. usability evaluation, formative approach, remote access

1. Introduction

Usability of IT systems in medical environments has long been a topic of discussion. Studies e.g. in Germany on a wide variety of different medical IT such as the clinical information system, the radiology, laboratory and intensive care system showed a considerable range in perceived software ergonomics with specialized systems scoring better than the generic clinical system [1]. To improve this situation, the use of formative evaluation methods to develop medical IT software has long been advocated with the goal to use the evaluation results in a feedback cycle to improve the ongoing system development [2,3]. Different tools such as Cognitive Walkthrough [2,4] or Think Aloud [2,5,6,7] have been proposed and successfully used. In our experience, Think Aloud proved better for the evaluation of complex medical software with many windows and different dynamic navigation options [8].

Typically, formative evaluation is carried out bringing the researcher and the test persons physically together in a lab or real world environment. As part of a group of Swiss projects on the influence of digitization we participated in a nursing centered ethnographic study across six Swiss hospitals examining the influence of digitization on the transfer of patient care information [9]. Within the observational study we identified shortcomings in currently used IT applications and discussed those

¹ Thomas Bürkle, Bern University of Applied Sciences, Bern, Switzerland thomas.buerkle@bfh.ch

intensively with the involved nurses and hospitals. One observation was that during or after shift handover nurses compile a paper based document with their patients and personal task list for this shift. During the shift, the nurses will tick off some of the tasks but may also annotate new information gained in the patient room. At the end of the shift some of the information is copied back to the electronic medical record. We developed a programmed mockup to represent a digitized nursing worklist to be configured on PC and used on smartphone.

Faced with researchers at two locations and hospitals spread throughout Switzerland we were tasked to develop this mockup in close feedback with actively working nurses, but with minimal extra effort for those test persons. This prompted the idea to perform a remote access Think Aloud study.

2. Methods

The mockup examined in this evaluation is a web based application with cached data storage to demonstrate the configuration of a personal nursing worklist on PC and later to work with this list on smartphone. It supports several languages and has been programmed as Single Page web Application (SPA) using TypeScript 5.0.4 and the JavaScript framework Vue.js 3.3.4. The mockup consists of two parts. The first part is a desktop GUI for the nurse to individually compose his/her personal worklist for the day. The second part is a simulated mobile phone app with the patient dashboard and worklist to carry throughout the work shift.

The remote think aloud study was conducted by students from the Graubünden University of Applied Sciences (coauthors SG and YB) during the spring semester of 2023 as part of their study curriculum. Test persons were nurses recruited from the Digicare study within the various hospitals. The sample included one nurse from each hospital in the project, 1 nurse covered both Italian speaking hospitals. Remote Sessions were scheduled as 30 minutes remote meetings via MS-Teams with the test persons. For these sessions the link to the web based application was made available to the test person. All sessions were recorded using the MS Teams function for screen recording. Each test person signed a written consent to participate in the study.

The sessions started with a short explanation (5 minutes) of the mockup and the tasks by one of the students who would also lead the conversation whereas the second student would perform the documentation during the session using Mural board from Miro. In addition, either TB or MM participated in each of the 5 remote sessions as silent observers.

The think aloud part was scheduled for 20 minutes. The detailed task list was provided as a document to each test person. It comprised a total of 17 tasks on the desktop GUI and 14 tasks on the smartphone interface. The test person would be asked to start the mockup on his/her own device and then to share the screen. The tasks were on the desktop GUI a) to add respectively remove a patient from the personal worklist for this day, b) to switch to the dashboard view for the first patient (see fig 2), c) to markup items as relevant in the dashboard, d) to add comments for this patient, e) to switch to another patient from the list and to markup further items in that patient's dashboard f) to look at the preview of the worklist on the desktop. Then g) to change to the mobile phone interface, h) to describe what can be seen there, i) to check the dashboard view of the patients and find the events marked as relevant, j) to change to another patient on the mobile, k) to go to the worklist for this patient and find the

annotated comments, l) to tick off a worklist item including a documentation, e.g. vital signs taken with results, m) to filter the worklist to see either performed or open tasks or tasks from another time period. During the session, the test person would be asked to vocalize their thoughts as they occurred.

The last 5 minutes were used for discussion, questions and to gain a short summative impression of the test person for the mockup. Prior to the first real session the whole setup was tested once internally by the team members to ensure feasibility.



Figure 1: Miro Board for test person #4. In blue 3 problems with the Desktop GUI, in lilac 1 problem with the smartphone GUI and in gray one generic comment.

3. Results

We had five test persons with nursing background, actively working as nurses in five different hospitals in German and Italian-speaking Switzerland. In two of the sessions we experienced problems to initialize a stable remote connection and to hear the test persons, potentially due to insufficient hardware or limitations of the hospital network. One session was postponed by one hour, while the other session had to be rescheduled to another day, requiring the use of different hardware in a different location.

We note positively that all participants could solve all tasks in both the desktop GUI and the mobile phone mockup within the scheduled time frame. Several test persons invested additional time to discuss the mockup and its implications in more detail with our student team. Due to flexibility in scheduling from our side, all the tests could be integrated into the daily clinical work routine of the test persons.

Fig 1 shows the test results of one of the tests. For each test person we compiled the results of the test in a Miro board (fig 1), roughly ordered by the test sequence described in the methods. For the complete study results, these individual Miro boards were combined to one document, thus eliminating duplicate problems found by several persons. Thus, we could identify 6 potential improvements in the Desktop GUI in step a) adding / removing patients from the list and 14 potential improvements in the steps c to e compiling the personal worklist. For the smartphone GUI we could identify 6 problems in the patient dashboard (step j) and 13 problems in the patient worklist (steps k to m). In the final questions another 3 problem areas were identified. In a next step

the results were attributed to the respective screenshot of the mockup and positive suggestions for the potential improvement of the GUI were defined (Fig 2).

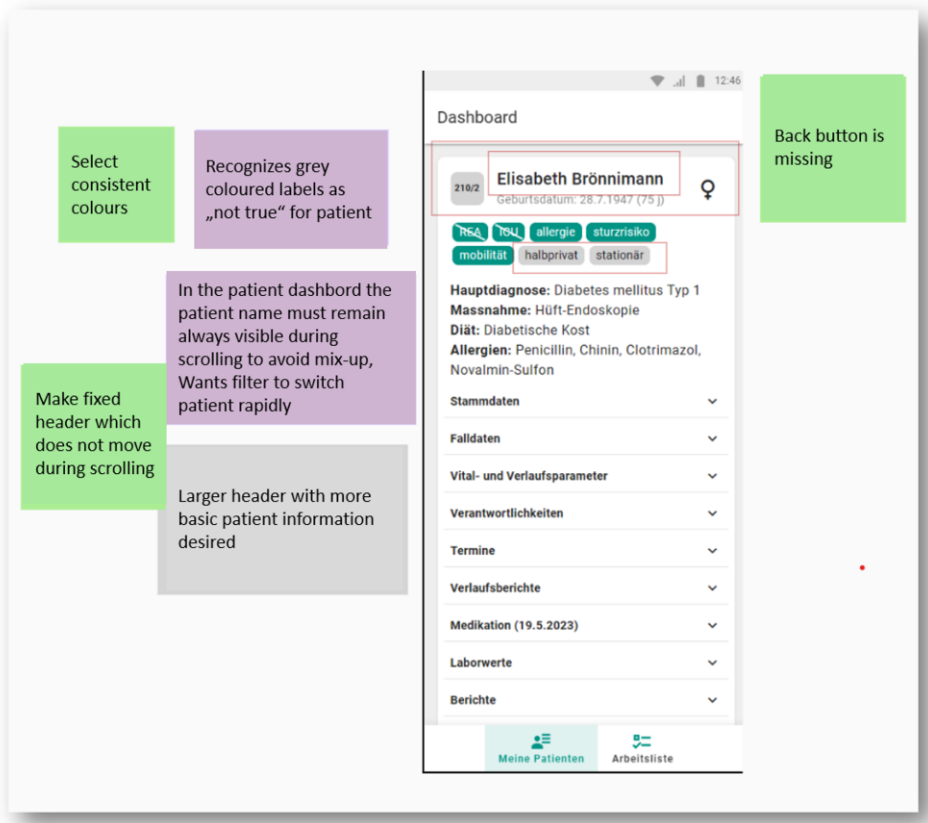


Figure 2: The patient dashboard on the smartphone, in green boxes suggestions for potential improvement

4. Discussion

Conducting a remote think-aloud study was feasible, yielding valuable insights for enhancing the mockup, some of which have already been implemented. This offers the opportunity to perform usability testing a) faster b) with persons working clinically who are otherwise not available c) to reduce cost and effort for the researchers and test persons avoiding traveling and setup of a labor d) with test persons from different countries and backgrounds.

Fundamental technical prerequisites include ensuring high-quality image and sound through the remote connection, along with the capability to procure a recorded video of the session. Furthermore, the examined software must be made available at the place of the test person. In our case, the web based application could be accessed in all five hospital networks. Previous testing is, however, strongly recommended. If local software installation is required, there may be considerable problems with missing user rights in the host (hospital) network. Theoretically, the researcher could also provide

the application on his shared screen and hand over the controls to the test person, but this increases complexity and may result in additional usability problems not connected to the application being tested. In our experience, some test persons were not familiar with MS Teams and found it difficult e.g. to share their screen, although this could be quickly solved. In the five cases, we did not encounter performance problems large enough to hinder the test. Screen size of the remote machine could be an issue, however. When working with individuals in different languages (one of our tests was conducted in Italian), at least one researcher must be fluent in the language and the software able to switch to that language.

We cannot say that our results are fully valid for a genuine smartphone app. In the study, we presented the smartphone interface in the web browser. It permitted e.g. scrolling up and down with the computer mouse, but does not support all options which a touchscreen of a real smartphone would offer. Also, using a mouse to scroll is not the same as just doing this with the fingertips. We thought about using real smartphones but found it too difficult to handle by our test users, who are not necessarily experienced computer users.

5. Conclusions

Remote access usability evaluation seems a promising approach, especially when time and travel options of the test persons are limited. Video recording is a built in function e.g. in MS Teams and thus must not be provided separately. We could prove that a fast formative evaluation for medical software development may be possible. A systematic comparison with the results of a parallel regular lab based Think Aloud study would be desirable to assess the effects and maybe information losses accounted due to the remote situation.

Acknowledgements

We thank all test persons from the participating hospitals. This work was funded by SNF 407740_187503.

References

- [1] Bundschuh BB, Majeed RW, Bürkle T, Kuhn K, Sax U, Seggewies C, Vosseler C, Röhrig R. Quality of human-computer interaction-results of a national usability survey of hospital-IT in Germany. *BMC medical informatics and decision making*. 2011 Dec;11:1-2. <https://doi.org/10.1186/1472-6947-11-69>
- [2] Peute LW, Jaspers MM. Usability evaluation of a laboratory order entry system: cognitive walkthrough and think aloud combined. *Studies in health technology and informatics*. 2005 Jan 1;116:599-604. <https://doi.org/10.1109/icstsn57873.2023.10151607>
- [3] Friedman CP, Wyatt J. Evaluation methods in biomedical informatics. Springer Science & Business Media; 2005 Oct 20.
- [4] Wharton C, Rieman J, Lewis C, Polson P (1994). The Cognitive Walkthrough Method: A Practitioner's Guide. *Usability inspection methods 1994*; 105 – 140
- [5] Nielsen, J (1993). *Usability Engineering*. Boston, MA: AP Professional. ISBN 0-12-518406-9.
- [6] Jaspers MW, Steen T, Van Den Bos C, Geenen M. The think aloud method: a guide to user interface design. *International journal of medical informatics*. 2004 Nov 1;73(11-12):781-95. <https://doi.org/10.1016/j.ijmedinf.2004.08.003>

- [7] Boren T, Ramey J. Thinking aloud: Reconciling theory and practice. *IEEE transactions on professional communication*. 2000 Sep;43(3):261-78. <https://doi.org/10.1109/47.867942>
- [8] Ramming T. Elektronische Untersuchungsanforderungen für Labor und Radiologie: Eine Usability-Studie zur Bewertung klinischer Anwendungen. PhD Thesis University Erlangen 2010.
- [9] Bürkle T, Lueg C, Salzmann P, Salini D, von Kaenel F, Löffel K, Meier L, Moozhiyil SM, Perrini S, Trede I, Volpe A, Holm J. Digi-Care: Exploring the Impacts of Digitization on Nursing Work in Switzerland. *Stud Health Technol Inform*. 2022 May 16;292:57-62. <https://doi.org/10.3233/shti220321>