DEVELOPMENT OF A TEAM-LEVEL USABILITY TESTING METHOD OF GROUPWARE

Thesis booklet

Dalma Geszten

Supervisor: Dr. Károly Hercegfi
Mentor: Dr. Balázs Péter Hámornik

Budapest
2021
**Introduction and main objectives**

Current topics of work psychology are team processes (Marks, Mathieu & Zaccaro, 2016) and team interactions (Juhász, 2015), since a common form of work is teamwork: people are able to achieve better results in a team than individually.

Collaborative software became widespread in the 1990s as a combined effect of the spread of the World Wide Web and the fact that teamwork became a popular research topic (Schmidt & Bannon, 2013). The main goal of collaborative software is to support teamwork, and since it has become an integral part of everyday work, the usability of software is a key issue.

The scientific discourse on evaluating the usability of collaborative software began in 1988, when Grudin summarized the difficulties of collaborative software evaluation (Grudin, 1988). It was followed by an intense study of the topic, in which researchers experimented with various methods. Pinelle's (2000) article summarizes these “early” studies that evaluate collaborative software, but at the end of the article the author expresses criticism of previous studies: the methods are not well documented or not documented at all. Pinelle pointed out that new, time and cost effective collaborative software evaluation methods are needed (Pinelle, 2000).

The evaluation of the usability of collaborative software has been investigated from several directions. Analytical methods evaluate the software based on expert knowledge but without the involvement of actual or potential users. These methods usually suggest software design directions based on task models. A common factor of these methods is that their goal is to create the most appropriate software for a specific organization (Herskovic, Ochoa & Pino, 2009; Pinelle, Gutwin & Greenberg, 2003; Veer & Welie, 2000). Analytical methods include research based on expert analysis utilising mechanics of collaboration and heuristics based on this framework (Baker, Greenberg & Gutwin, 2002; Pinelle & Gutwin, 2008). By developing the theory of the mechanics of collaboration, the authors attempted to create a framework that includes basic collaborative actions that could be the foundation for future cost-effective evaluation methods (discount evaluation methods) (Gutwin & Greenberg, 2000).

Empirical methods represent another direction in the usability evaluation of collaborative software, in which the evaluation of the software is performed with the involvement of users, most often under real, everyday working conditions. Field studies typically investigate how software affects collaboration in the workplace (Gumienny, Gericke, Wenzel & Meinel, 2013; Tang, Isaacs & Rua, 1994), and what makes a software successful within the organization (Pipek & Wulf, 1999). Various methods are used in field studies: observation, questionnaire,
interview, usability situations (scenario) and log-file analysis can also be found (Christensen & Ellingsen, 2016; Gumienny et al., 2013; Haynes et al., 2005; Marlow et al., 2016).

Although, it has been two decades since the publication of Pinelle’s (2000) article, the gap of well-documented and rapid methods still exists. Thus, the fast spread of collaborative software was not followed by an explosion in usability assessment methods. This served as a motivation for my doctoral research: the creation and validation of a new empirical method to investigate the usability of collaborative software at a team level.

Method

The process of developing the team-level usability testing method consisted of four stages. As a first step, I conducted a laboratory study involving teams collaborating at the same time and at different locations. The main objective of the First Laboratory Study was to assess the types of usability problems the team-level usability test can reveal. In addition, it was also important to investigate the extent of useful data the different data collection methods can explore. As a next step, before the Second Laboratory Study, I was interested in what kind of problems would occur under real, workplace conditions compared to controlled laboratory conditions. Therefore, a field study was conducted before the Second Laboratory Study. Summarizing the results and lessons learned from the Field Study and the First Laboratory Study, the Second Laboratory Test followed, during which I analyzed a different software than in the First Laboratory Study. To investigate the relationship between usability problems and team communication patterns, sequence analysis was performed using a self-developed code system (based on the communication transcripts of the First and Second Laboratory Studies). As a final step of my doctoral research, I tested how results of experts analysis of given collaborative software compares to results of a laboratory study. Therefore, as the last part of my doctoral research, a heuristic analysis was performed.
During the development of the team-level usability test, my research questions were the followings:

- What factors influencing software usability can the team-level usability test reveal?
- What types of team usability problems can the team-level usability test reveal?
- What factors influencing software usability can the field study used in the development of the method reveal compared to the laboratory study?
- What types of team usability problems can the laboratory study reveal compared to heuristic analysis?
- What is the relationship between usability problems and a team’s communication patterns?

New scientific results

Thesis point 1.

The team-level usability test developed during my doctoral work is able to explore team usability problems. Content analysis of communication transcripts, interview and questionnaire data revealed different types of problems and are therefore crucial elements of the developed method.

In the First and Second Laboratory Studies, I analyzed the communication and interview transcripts as well as the text answers of the questionnaires based on the theory of the mechanics of collaboration (Pinelle et al., 2003) using content analysis method (Geszten, Hámornik & Hercegfi, 2020). (I have successfully used content analysis method to explore usability problems earlier. I was able to explore the usability problems of the examined software by the content analysis of interview transcripts (Geszten et al., 2018).) Team usability problems identified during doctoral research are usability problems that arise during teamwork and affect team collaboration while using the software. The team-level usability test was able to explore team usability problems in laboratory situations. Team usability problems in the First Laboratory Study were related to following mechanics: explicit communication, basic awareness, and work protection, while in the Second Laboratory Study, to basic awareness and work protection.
In both laboratory studies, the problems with basic awareness were related to specific collaborative functions of the investigated software (avatar, synchronization, save, zoom). Problems with these features have a negative impact on collaboration. The support of basic awareness is crucial in collaborative software, as evidenced by several studies (Gutwin & Greenberg, 1996, 1998; Gutwin, Roseman & Greenberg, 2004; Ignat, Oster, Fox & Shalin, 2015; Lopez & Guerrero, 2017).

The overwriting problem related to the work protection mechanic also appeared during the First and Second Laboratory Study. According to the work protection mechanic, a software must prevent team members from accidentally deleting or overwriting each other's work (Pinelle et al., 2003). Overwriting is the most serious usability problem, with serious negative impact on collaboration, which is confirmed by earlier studies (Dew, Turner, Desai, Martin & Kirchhoff, 2015; Pinelle & Gutwin, 2008).

The difference between the problems identified in the First and Second Laboratory Study is that team usability problems related to explicit communication only appeared in the First Laboratory Study. This is not surprising, since in the First Laboratory Study, the participants worked together in different locations to simulate virtual teamwork, while in the Second Laboratory Study, they worked together in person, in the same room. Problems related to explicit communication negatively affected the cooperation in line with previous results in the literature (Geszten et al., 2020; Pinelle et al., 2003). Majority of the team usability problems in the First and Second Laboratory Study were identified in communication transcripts (55% and 77%) and interview transcripts (34% and 21%), while questionnaire data identified most of the contextual factors influencing collaboration (56% and 59%). In addition, some problems only appeared in the communication transcripts (overwriting) or in the questionnaires (team mood), so different types of data are crucial elements of the team-level usability testing method.

To summarize, team-level usability test is able to explore team usability problems, therefore has a valuable role among existing collaborative software evaluation methods.
Thesis point 2.

The team usability problems and contextual factors influencing the collaboration revealed by the team-level usability test developed during my doctoral work are real, valid aspects of software usability in the field.

According to literature, several methods should be used in a research process to support the authenticity of data analysis and interpretation (Szokolszky, 2020; Thurmond, 2001). Accordingly, during my doctoral research, as part of the method development, a field study was conducted following the laboratory study. Usability of synchron collaborative software was evaluated both in the First Laboratory Study and in the Field Study. Using the team-level usability testing method, two types of factors influencing collaboration were identified in the First Laboratory Study: team usability problems and contextual factors. While team usability problems are related to the mechanics of collaboration (problems are most often caused by the inadequate support of the mechanics), contextual factors refer to social and environmental factors that influence collaboration.

Similar to the laboratory study, team usability problems and contextual factors influencing the collaboration also emerged in the Field Study. The team usability problems observed in the Field Study were the following: the lack of the possibility to alternate between different parts of the workspace and the visibility of comments. Contextual factors referred to the physical artefacts that supported collaboration, i.e.: the importance of using whiteboard and notebook, which played a significant role in understanding the process of collaboration. This is consistent with previous studies, stating that collaborative software can be characterized by usability problems and contextual problems resulting from inadequate support of the mechanics of collaboration (Steves, Morse, Gutwin & Greenberg, 2001).

Thus, the results of the Field Study support that the team usability problems and contextual factors influencing collaboration are relevant aspects of software usability that also appear in the field.
Thesis point 3.

The team-level usability test developed during my doctoral work can reveal additional results compared to the efficiency-optimized Nielsen heuristic analysis: team usability problems that heuristic analysis cannot.

As a final step of the development of the method, a Nielsen heuristic analysis (based on the involvement of experts) was performed. The subject of the analysis was the Miro collaborative whiteboard software, which was also examined in the Second Laboratory Study. With the help of the comparison of the laboratory study and heuristic analysis, I examined the types of problems the my method can explore compared to heuristic analysis. According to the literature, heuristic analysis can reveal different types of problems than a usability study with real users, which is also supported by the results of my research. While in heuristic analysis, experts tend to evaluate the software more comprehensively, real users are able to explore usability problems related to their daily tasks (Nielsen, 1994; Steves et al., 2001). It depends on the purpose of research which method should be used, but if possible, experts recommend the combined use of the two methods (Lazar et al., 2017; Rubin & Chisnell, 2008). During the heuristic analysis, the experts evaluated the usability of the software based on the collaborative heuristics related to the mechanics of collaboration (Baker, Greenberg & Gutwin, 2001).

The main difference is that while problems occurred related to basic awareness, work protection, and verbal communication during both the laboratory study and the heuristic evaluation, the heuristic analysis did not reveal the most serious problem encountered during the laboratory test: the overwriting. There were several overwriting problems in the shared workspace in the laboratory study, when one participant accidentally modified or deleted the work of the other. This type of problem was identified only by the laboratory study.

Thus, the team-level usability testing method can reveal different types of results based on the experience of real or potential users than a heuristic analysis.

However, heuristic analysis also has several advantages. It examines all the collaborative features of the software, rather than focusing on a single task, providing an overview of the usability of the features.

The results of the two methods complement each other well, in practice we can get a complete picture of the usability of the examined collaborative software by examining the opinion of expert’s and the experience of users together. This result confirms previous research that
analytical and empirical methods are both important when examining the usability of collaborative software, but neither method can be substituted for the other (Steves et al., 2001). Overall, the team-level usability test was able to identify additional results compared to the heuristic analysis, therefore I recommend its use in collaborative software evaluation scenarios.

**Thesis point 4.**

<table>
<thead>
<tr>
<th>The team-level usability test developed in my doctoral research works for different teams with different dimensions:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. It works for teams whose members know each other well, but also for teams whose members did not know each other.</td>
</tr>
<tr>
<td>b. It works for teams that rated collaboration with a high score, but also for teams that rated it low.</td>
</tr>
</tbody>
</table>

4. a.

When evaluating the usability of collaborative software, participants are typically team members, who are familiar to each other (Christensen & Ellingsen, 2014; Gumienny et al., 2013; Marlow et al., 2016; Pipek & Wulf, 1999; Tang et al., 1994). In the First Laboratory Study, the participants were members of the same university group, so I did not examine the degree of familiarity here. However, for the Second Laboratory Study, I recruited participants from a larger and more heterogeneous group (students from a given university faculty), so I considered it important to assess the extent of familiarity of the participants in a separate questionnaire question as an element of method development. According to the results, team usability problems also occurred in teams whose members knew each other to varying degrees. It means that the team-level usability testing method is able to reveal usability problems regardless of the degree of familiarity of the participants. These results are in line with previous research, the results of the usability test are not affected by the degree of familiarity of the participants (Gutwin & Greenberg, 1996; Gutwin et al., 2004). It is important in the practical application of the method, as it makes the recruitment of participants easier.
The number or type of problems encountered in each team was independent of the degree of familiarity of the participants, so the team-level usability testing method works for teams whose members know each other well, but also for teams whose members did not know each other before.

4. b.

In most cases, the final stage of usability testing is a questionnaire or interview that investigates the participant's subjective experiences of the usability test situation and software usability (Lazar et al., 2017; Rubin & Chisnell, 2008). The subjective feeling of appropriate support of teamwork and collaboration is a part of collaborative software usability (Pinelle & Gutwin, 2002; Pinelle et al., 2003). The subjective evaluation of the success of the collaboration has also appeared in other collaborative software usability studies (Geszten et al., 2015; Sutcliffe & Alrayes, 2012). During the post-experiment questionnaire, team members were asked to rate the collaboration: “How do you feel, how much were you able to collaborate with your peers?”. Both the First and Second Laboratory Studies could identify team usability problems in teams that rated collaboration low. The number or type of problems did not differ between teams that rated collaboration as high or low, therefore the team-level usability testing method is able to explore usability problems regardless of the rating of collaboration.
Thesis point 5.

Using the team-level usability test, it was discovered that teams with certain communication-patterns can be characterized by certain types of usability problems: the overwriting team usability problem could be avoided by teams that discoursed about awareness, while not by others.

During the development of the method, I examined the communication patterns of the teams participating in the First and Second Laboratory Studies with the help of sequence analysis. I considered this analysis necessary because the results of the laboratory studies showed that each team experienced different types of team usability problems. The study of team communication is a significant topic in the psychological literature, however, it is less researched in software usability. In case of team research, the goal is to identify intervention points by exploring the aforementioned breakdowns and problems, thus improving team communication to increase efficiency (Behfar et al., 2008; Marlow, Lacerenza, Paoletti, Burke & Salas, 2018). Therefore, I considered it important to examine team communication patterns in the context of software usability as well, and to explore the occurrence of usability problems in teams with different dynamics.

The analysis of the communication patterns related to all usability problems is beyond the scope of the dissertation, therefore I chose the overwriting problem for the analysis. My choice is justified by the fact that it is the most serious team usability problem, as in this case, one participant accidentally (due to inadequate support of collaborative functions) overwrites or deletes the work of another participant in the common workspace. Certain types of teams could avoid overwriting, while others could not, therefore I examined the nature of the differences in the communication patterns of teams in which overwriting occurred comparing to teams in which it did not.

Overwriting do not occur in teams, where there is a discourse about awareness. Thus, when someone communicates or asks about exactly what’s going on in the workspace, team members respond with that type of information. Additionally, overwriting can be avoided for teams whose members help each other effectively (if a team member asks for help, he/she gets help), or closely organize and plan of collaborative work (if a team member shares information about coordinating collaborative work, he/she gets this type of information in response).

The results are in line with previous research, which demonstrate that by identifying communication patterns, we have the opportunity to explore the communication dynamics of
each group and thus identify difficulties, breakdowns, problems (Hámornik, 2013; Juhász, 2015; Soós, 2012). The novelty of the results is that examination happened it in the context of software usability. As certain conflicts and problems occur in teams that communicate differently, it is also true in a usability test situation: certain usability problems occur in certain teams, teams with different communication characteristics will encounter different problems. It can also affect the interpretation of the results of the usability test, as the communication strategy of the teams can compensate or cover serious usability problems. These results support the statement that teams with certain communication patterns can be characterized by certain types of usability problems.

**Thesis point 6.**

| The coding system developed for the sequential analysis of the team communication process is capable of exploring the communication patterns of teams related to team usability problems and is therefore a component of the team-level usability test. |

Several studies focus on exploring team communication patterns, and several reliable coding systems are available in this topic (Hámornik, 2013; Marlow, Lacerenza, Paoletti, Burke & Salas, 2018; Soós, 2012). The most common coding system for examining communication is Marks's (2001) team processes framework, which the authors believe to be suitable for examining teams working in any context (Marks et al., 2001). It has been used in the past in areas such as healthcare, sales teams and grocery store employees (Killumets, D’Innocenzo, Maynard & Mathieu, 2015; Rapp & Mathieu, 2007; Valentine, Nembhard & Edmondson, 2015).

The study of team communication is a less researched topic in the context of software usability, therefore in my doctoral research I developed a suitable coding system by combining the theory of team processes and the mechanics of collaboration theory (Marks et al., 2001; Pinelle et al., 2003). The team process theory is a commonly used theory for analyzing team communication, while the mechanics of collaboration theory has not been used for such purposes. The two theories are similar in that they both define relatively basic actions, only while one is related to team processes, the other is related to the basic actions of software usability. By merging the two coding systems, I developed my own coding system. The reliability of the final coding
system is excellent (Cohen kappa = 0.825, p <0.001), thus it is suitable for exploring the communication patterns of teams in connection with team usability problems. The results are in line with the fact that team process theory and the mechanics of collaboration theory consist of analytical categories that can be used in practice (Marks et al., 2001; Pinelle et al., 2003). The results show that the coding system developed based on the two theories can reliably identify team communication patterns in the context of usability, hence I consider it part of the method developed in my work.

**Publications related to the thesis points:**

**Thesis point 1**

   [http://acta.uniobuda.hu/Geszten_Komlodi_Hercegfi_Hamornik_Young_Koles_Lutters_84.pdf](http://acta.uniobuda.hu/Geszten_Komlodi_Hercegfi_Hamornik_Young_Koles_Lutters_84.pdf)


**Thesis point 2**

*accepted, under publication*

**Thesis point 1 and 3**


**Further scientific publications related to the preparation of my doctoral research**


Further scientific publications


Abstract publications

References


plan, design, and conduct effective tests (2nd ed.). Indianapolis, IN: Wiley Pub. https://doi.org/10.1007/s13398-014-0173-7.2


