

Supporting Collaborative Reflection at Work: A Comparative Case Analysis

Michael Prilla

IAW, Ruhr University of Bochum
Universitätsstr. 150
44780 Bochum
michael.prilla@rub.de

Bettina Renner

Knowledge Media Research Center (KMRC)
Schleichstr. 6
72076 Tübingen
b.renner@iwm-kmrc.de

ABSTRACT

Reflection is a common activity at work. Collaborative reflection is an activity in which multiple participants add to reflection by sharing experiences, perspectives and insights together, thus transcending individual capabilities. Despite its potential for change at workplaces, there are little insights on how to support collaborative reflection with technology. To close this gap, this paper analyses four cases in which a tool to support collaborative reflection has been used at different workplaces. It uses qualitative data from app usage, an analysis of content from the tool and feedback gathered from participants to shed light on how people use support tools for collaborative reflection. The results of this work include factors supporting and constraining reflection in tools as well as implications for tool design.

Categories and Subject Descriptors

K.3.1 [Computer Uses in Education]: Collaborative learning;
K.4.3 [Organizational Impacts]: Computer-supported collaborative work.

General Terms

Measurement, Performance, Design, Human Factors

Keywords

Reflection, Collaborative Reflection, Workplace, Learning

1. INTRODUCTION: SUPPORTING COLLABORATIVE REFLECTION

Reflection is a common activity performed by people in different work places every day: Workers ask themselves whether they can do better in certain situations, teams analyze their performance and think about improving their cooperation, and during work people question their doings and try to improve. Conceptually, reflection can be understood as returning to experiences, re-assessing them in the light of current knowledge and drawing conclusion for future work from this [3]. This refers to what Schön [23] describes as reflection-on-action in contrast to

reflection-in-action which happens right while conducting a task. This process of drawing conclusions from the insights gained by reflecting on past experiences needs attention to work being done and the rationales for doing it the way it was done [32] as well as a mindset of being open to learn from current or past experiences, which needs to be established and spread to make reflection effective in organizations [30].

Despite the benefit of reflection and the awareness of people for this benefit reflection may not always be possible for workers – there might be no time and space to do this, or there may be other constraints. If it is then shifted to a certain time after the experience, fading or incomplete memories may hinder adequate reflection. In addition, relevant information to reflect on certain experiences may not be available for individuals, but only from a group of people. These and other constraints of reflection in daily practice may be diminished by tool support for reflection that helps people to keep memories of experiences, gather different perspectives on these experiences, reflect on them individually or in groups and share insights derived from reflection (see e.g., [13,15,26]).

It has been found that collaborative reflection may create insights beyond reflection outcomes that individuals can produce by combining different people's perspectives and knowledge [13,20], and that this needs specific support for the collection, coordination and combination of different contributions to the reflection process. Unfortunately, research on reflection support mainly covers individual reflection and education settings, in which constraints such as time and space occurring in many workplaces often do not play a decisive role, as reflection is part of the curriculum. Consequently, there are only a few insights on how to provide collaborative reflection support with tools (e.g., [16,26,33]). Given the potential of collaborative reflection at work there is a need to fill the resulting research gap.

2. COLLABORATIVE REFLECTION: RELATED WORK

2.1 Understanding Collaborative Reflection

Collaborative reflection differs from individual reflection, and therefore imposes additional requirements for tools: If people want to reflect together, they have to make experiences explicit, share and compare them, collaboratively gain insights and create ideas for change in future work [6,33]. This needs communication support e.g. for the exchange of similar experiences and to collaboratively understand them [4,10,33]. In work on face-to-face collaborative reflection it has been shown that structuring the process can be helpful [4], that there is need to support the articulation of issues to reflection upon [2] and the possibility to refer to each other [11].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, or post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

GROUP '14, November 06 – 09, 2014, Sanibel Island, FL, USA
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3043-5/14/11...\$15.00.
<http://dx.doi.org/10.1145/2660398.2660400>

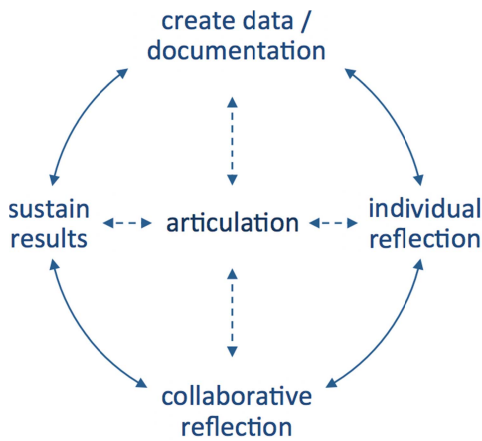


Figure 1. Collaborative reflection support blueprint [25,28].

Collaborative reflection overlaps with well-known concepts from CSCW and CSCL such as sensemaking, problem solving or group decision support (e.g., [5,24,31,36]), and this may imply that concepts from these areas may be easily transferred to collaborative reflection support. However, despite these overlaps, the discriminating aspects co-occurring in reflection – focus on past work, understanding the experience *and* planning for the future – are not present in such combination in other concepts [26]. Therefore research on collaborative reflection may draw from such concepts, but needs to explore how collaborative reflection works in practice and how people use technology to support it at their workplaces (see [26] for a detailed discussion).

Based on empirical and design work feeding into the studies presented in this paper [25,26,27] we created a model for collaborative reflection, which we use as a blueprint for the development of tools [25,28]. As shown in Figure 1 it suggests supporting collaborative reflection as interplay of reflecting in the group and leaving opportunities for individual reflection. Support includes (usually at the beginning of the process) the documentation of experiences and, if results are created from the process, possibilities to sustain these results.

2.2 Tool Support for Collaborative Reflection

An analysis on existing collaborative reflection support as part of this work revealed that this area has not been researched intensively, and that existing approaches are mainly focused on individual reflection or education settings [25,26]. This has recently also been recognized by several authors [1,19,23].

Among existing approaches on individual reflection support, a notable approach has been described by Isaacs et al [15]. In their tool, users can describe experiences, rate them according to their feelings about them and are re-prompted regularly afterwards to re-assess these rating in order to continuously reflect on them. Despite the value of this contribution, it is solely designed for individual reflection purposes. For collaborative reflection support, besides generic tools such as whiteboards also proposed for reflection purposes (e.g., [16]), Fleck and Fitzpatrick show how a series of pictures representing daily activities can trigger reflection in a social group [9], and Scott shows how learning portfolios can support collaborative reflection in education [33].

Looking at this state of the art, we can see that there is no sufficient amount of insights on tools support for collaborative work available. Our work [25,26,27] and that of a few others [19,23] aims to derive such insights and support the design of

collaborative reflection support. A decisive part of this work was the development of tools specifically tailored to collaborative reflection (see section 3.1), as there were no such apps available when we started the study.

2.3 Analyzing Reflection Support

Literature on analyzing tools to support conversations with a certain purpose among people points to the analysis of usage figures and content analysis. While the analysis of usage figures offers insights on tool adoption, activity and engagement in conversations (e.g., [8,18,35,37]), content analysis is seen as a means to look into the quality of conversations [14,37], which in our case means the quality of collaborative reflection.

Coding the content of conversations has been shown to be a key to analyzing the way these tools support people and to which extent they fulfill their goals [14,34,37]. Unfortunately, for collaborative reflection at work supported by tools there is no scheme for analyzing such content available. Existing approaches as proposed by Fleck and Fitzpatrick [9], Zhu [40] or de Groot et al. [11] stem from more general conversation analysis contexts or are related to reflection in educational settings, which differs from reflection at work. In particular, schemes from educational settings do not regard the articulation of experiences and issues in them, as the content to be reflected about is often given in such settings instead of being created by a reflection participant. In addition, these schemes are often too coarse to better understand the way people reflect. Complementing existing schemes and supporting their extension, several authors describe aspects of (collaborative) reflection that need to be considered when reflection is analyzed [3,17,21,22,29,35,39]. Among these aspects, recurring topics are

- **Returning to experiences**, including emotions and rationales for actions [3,9,21,35]
- **Probing, challenging or supporting** other contributions [11,17,29]
- **Linking** between experiences and other sources (experiences, knowledge, data) [17,22,35,40]
- **Repeating and commenting** on other contributions [22,40]
- **Sharing** perspectives and opinions [9,35,39]
- **Drawing** from experiences [3,9,11,21]
- **Transforming** insights into practice [3,21]

In combination with existing schemes for conversation analysis in related areas, these topics can be used as a foundation to develop a coding scheme for collaborative reflection.

However, such analysis may not be enough: Instead of relying solely on content coding, several authors argue for a combined approach of analyzing communication in electronic media. For example, [18] argue for using descriptive data such as communication thread length to complement content analysis. [8] propose to measure the responsiveness to others and the amount of people interacting with each other as additional information on a conversation group, and [37] describe different aspects of social network analysis to understand group activity better.

3. STUDY: COLLABORATIVE REFLECTION SUPPORT IN PRACTICE

To gain insights into how people use collaborative reflection support at work, we analyzed four groups from different workplaces and contexts using a tool we developed to support reflection on social interaction at work. The results from this analysis are based on a combination of log data analysis, social network analysis, content analysis, and qualitative data gathered from the participants of the study. Combining this data provides a

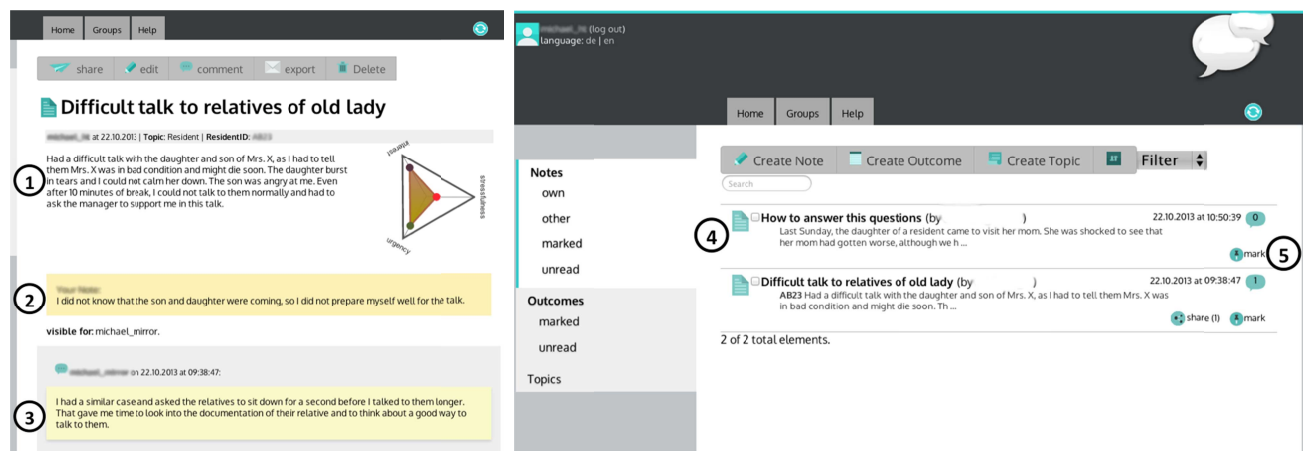


Figure 2. The TalkReflection App with an experience report commented on (left) and the list of reports created by users (right).

holistic insight into the practice of using collaborative reflection support and allows to draw conclusions both for assessing technology supported collaborative reflection groups and for designing collaborative reflection tools. The analysis of the data described in the paper was driven by three research questions:

1. How do (different groups of) people use tools supporting collaborative reflection at work?
2. Which are the factors supporting or constraining the usage of tools supporting collaborative reflection at work?
3. Which aspects of tools can support collaborative reflection at work?

This paper describes a study aiming to answer these questions, It was undertaken in four different cases, in which participants used the same tool (the ‘TalkReflection App’) to support collaborative reflection.

3.1 The TalkReflection App for Collaborative Reflection Support

For the analysis of collaborative reflection tool support in practice we conducted a study of different cases with different workforce, using a tool we had created from earlier studies of collaborative reflection at work [reference omitted for blind review]. The TalkReflection App supports collaborative reflection on social interaction such as conversations at work. To make sure that the study covered different situations and settings of reflection at work such as spontaneous reflection, reflection in meetings and others we did not restrict the use of the tool at the workplaces participating in the study but left usage to participants.

Based on earlier empirical work [references omitted for blind review], the TalkReflection App was built to support collaborative reflection on interactions among staff and clients, service providers and colleagues. This is a relevant topic for many organizations: In the workplaces covered by the study as examples, in medical and care domains workers have stressful interaction with patients of hospitals, residents of care homes and relatives, and in public administration people are facing difficult interactions with members of the public.

The TalkReflection App was developed to support such interactions by collaborative reflection among workers, clarifying issues in a peer group. For this, it supports the documentation of interactions and collaborative reflection on them by sharing resulting reports and commenting on them. This, in line with

related work [4,13,33,35,38], helps workers to explicate, share and reflect on experiences from conversations by supporting the steps of reflection described in the blueprint shown in Figure 1:

- **Creating experience reports:** The app supports users in documenting experiences by providing an opportunity to write them down. This includes a description of the experience and personal reflections on the report. Figure 2 (left) shows a report (no. 1) with a personal reflection comment (2). Writing down experiences can be a trigger for initial reflection [33] and provides a basis for later reflection.
- **Sharing experience reports:** Experience reports remain private to users initially, but can be shared with others. Once they are shared, other users can find them as shown in Figure 2 (right, no. 4). Sharing experiences with others can be regarded as asking them for feedback or opinions (cf. [39]).
- **Commenting on experience reports:** To engage in reflection on shared experience reports users can create comments on these reports as shown in Figure 2 (left, 3). Comments may contain similar experiences of a user, suggestions for acting in the situation described or other reflective content. Comments are core to collaborative reflection, as they enable the necessary communication for the exchange of perspectives and insights [4,10,33].

The app provides a certain structure for its users, assuming that the conduction of tasks such as creating reports of experiences, commenting on them and thinking about corresponding changes to work can create a flow of collaborative reflection.

3.2 Four Cases of Collaborative Reflection Support

The TalkReflection App was used with four different groups in three organizations, with each of the groups providing a unique context for using the app. Cases 1 and 2 were run in a public administration organization in the UK. Case 1 included the interns working in different departments of the organization, who were supposed to use the tool to learn how to interact with colleagues and members of the public professionally. Case 2 was done with participants of two departments conducting similar services, which were to be merged into one department. The aim of using the tool was to share and reflect on practices of the respective other department to support the merging process. Case 3 was done with physicians of a German hospital, who wanted to improve their abilities to talk to relatives, which is a task physicians are not prepared for well in medical school. Case 4 was conducted in a

British care home for people suffering from dementia. The aim of this study was to support care staff in improving their skills in conversations with residents, relatives and third parties.

The cases were conducted within different timeframes and with different amounts of users – the study configuration was based on the goals and resources of the different organizations. Table 1 gives an overview of time and participants for each case – it should be noted, however, that the numbers for the participants given in the table refer to the *total* amount of participants *in the beginning of the study*. In each study participants dropped after a short period of time, reducing the actual amount of users. This was especially relevant for case 1, in which six participants stopped using the tool after a couple of days of activity. As participants dropped out at similar times (some after initial usage, some after a couple of days) and given the sample size we cannot differentiate groups of users and will refer to the full number of participants for each study in this if not explicitly stated differently. Using the data of all participants is necessary because the content they put in the app during the first days cannot be removed from the rest of the content, as for example other participants commented on each the notes created by the dropouts.

Table 1. Overview for cases of using the TalkReflection App.

	Case 1	Case 2	Case 3	Case 4
Domain	Interns, public ad.	Public admin.	Hospital	Care Home
Participants	18	12	9	9
Duration (days)	51	80	42	50
Time	Sep-Oct 2013	Aug-Oct 2013	Jul-Aug 2013	Aug-Sep 2013
Dominant user	no	yes	yes	no
Process integr.	no	yes	no	no
Co-location	no	partly	yes	yes

In each case, the TalkReflection App was introduced to the participants in on-site workshops including an introduction of the app and a session of initially using of the app, asking some users to provide real experience reports and others to comment on them. It was also discussed how the app could be used in the respective workplace, including how to use it in meetings or during the day. No further instructions were given, and the groups were allowed to use the tool in a way that suited the group best.

The resulting usage of the tool, although voluntary in each case, differed in terms of organization and conduction (see Table 1): In case 1 and 4 participants used the app in a *self-directed* way to reflect in the team of workers, meaning that the organization and course of reflection was left to the participants. In contrast to that, cases 2 and 3 had *dominant users*, who took control over the reflection process: in case 2 the manager responsible for the departments took the role of driving the usage and in case 3 the head physician played this role. In cases 1, 3 and 4 the app was used in parallel to other work, while in case 2 it was integrated into the process of merging the departments, including meetings in which the participants discussed issues reported in the app. In addition, the opportunities for face-to-face communication differed, as in cases 3 and 4 users worked on the same floor, while in case 1 the interns worked at different workplaces in different buildings and in case 2 the two departments were located in different buildings as well. Table 1 summarizes this information.

3.3 Measures for Collaborative Reflection

Our work aimed at understanding the use of the TalkReflection App as a particular collaborative reflection support tool in different groups and workplaces by analyzing each group but also by comparing group behavior and output. We took into consideration specifics of each case and derived insights on the role of tools to support collaborative reflection at work.

The analysis was done with a mixture of tools to ensure a holistic view on how the participants in the studies use the TalkReflection App. For this we combined usage and content analysis with qualitative feedback from participants to also capture the subjective impact created by using app:

- **Usage analysis** was done by the amount of content created over time and analyzing the conversation structure.
- **Social network** analysis was applied to analyze the communication between participants.
- The content created by the participants was analyzed with a specially developed **content coding** scheme.
- The results from the analysis mentioned above were compared to feedback given in **interviews** and **focus groups**.

We are aware that in our analysis we are facing an observer problem: While we have detailed data on what happened in the app, we have comparably little data on reflection outside the app, that is, reflection possibly started in the app and continued in face-to-face situations without leaving traces in the app. Therefore we cannot make general assumptions on the collaborative reflection taking place in the cases from our analysis, but we can derive insights on how users reflected in the app and the impact resulting directly from using the tool for collaborative reflection. This reflects the perspectives of users reflecting in the community connected by the tool (which is not necessarily co-located as for example in cases 1 and 2), new users entering the app or users who cannot be present in certain face-to-face encounters.

3.3.1 Usage analysis: Measures

Besides content analysis, analyzing and comparing group behavior needs data describing group and individual behavior to set group activity and outcomes into context. Using a theoretical and empirical base showing that collaborative reflection relies on the sharing of experiences, articulating perspectives and opinions on such shared experiences and engaging in collaborative sensemaking and inference, we found that we also need means to describe and analyze user and group activity. Concerning descriptive data on user and group activity we use

- the **amount of experiences reports and comments** made on them (using the terminology from the description of the TalkReflection App) as information about general *activity* necessary to make reflection work in a group (creating, sharing, commenting, see [6,13,33,35]). To compare these figures among cases, they need to be *normalized*, that is set into relation to the timeframe data was collected in and the amount of users participating in the case.
- the **average length of communication threads** as proposed by [18] to provide insights on the *engagement* of users into conversations. The average length was calculated from the number of comments, not counting the experience report being the root of each thread.
- the **answer ratio** to experience reports as proposed by [8] as a measure for responsiveness of the group. This measure is also an indicator for attention to others' contributions, which is crucial for collaborative reflection [4,9,11,17]. The answer ratio was calculated as the ratio between the number of

experience reports commented on by at least one user compared to the total number of experience reports.

Adding to such descriptive data on user interaction, we use basic social network analysis as proposed by [8,37], including

- the **density of the graph** resulting from conversations as proposed by [8]. This graph is composed by directed edges between users, and an edge is added from user A to user B if user A has added a comment in a thread originating from user B's experience report. The density of the graph is then calculated as the ratio between the number of edges in the graph and the possible maximum of edges). It shows how many *different* pairs of people communicated in the group.
- the **ratio of unique edges** in the graph as a measure of diversity in group communication. The number of unique edges is calculated by counting how many edges are exactly once in the graph, and it shows how many different edges are in the graph – an edge may be in the graph more than once between two users A and B if there were multiple comments of user A to user B's reports. The unique edge *ratio* is then calculated as the ratio between the number of unique edges and the total number of edges in the graph. A high ratio thus means that many different pairs of users communicated with each other, thus indicating high diversity. A low ratio means that there were fewer pairs, implicating that there were some pairs in the group that communicated more often with each other. In contrast to the density of the graph this measure is stable against users dropping out early as it uses the actual amount of edges in the graph instead of the possible maximum amount as a basis.

Based on the suggestion to apply combined approaches for the analysis of conversations, we used the abovementioned measures to analyze collaborative reflection in the cases and to complement the content analysis described below.

3.3.2 A Coding Scheme for Collaborative Reflection Content

Besides data describing the interaction between users, the content created in the TalkReflection App provides good insights into how the tool was used in the cases. Facing the lack of existing schemes or methods to analyze collaborative reflection content in tools and drawing from the work described in section 3.3.1 on (collaborative) reflection we developed a coding scheme to analyze content of reflection support tools. It distinguishes different phases of reflection as describe din in section 3.3.1, starting from the description of an experience and ending with the description of changes in own behavior made or to be made. In particular, it includes nine phases as described in Table 2.

These phases in the scheme may build on each other. In a reflection session the description of an experience may be followed by the provision of one or more interpretation, one or more proposals for action and possibly the mentioning of learning and change. However, the choice of single codes does not depend on other codes, meaning that for example we may code 7b for solutions suggestions directly after using code 1 for a description of experiences without needing any other code in between.

Besides describing different aspects of collaborative reflection, the codes can be subsumed into three basic stages of reflection described similarly by Fleck and Fitzpatrick [9] and de Groot et al. [11]. The stages were aligned to the levels of collaborative reflection differentiated by Fleck and Fitzpatrick [9], and they include (1) the description and sharing of experiences and emotions

from them, (2) trying to understand and solve issues in experiences shared, and (3) describing learning and change.

Table 2. Coding scheme for content in collaborative reflection tools, with example from the content analyzed.

Code	Phase
1	Description of an experience and mentioning of an issue in an experience report or in comments, including the course of actions for the experience, e.g. "I had a very rude person on the phone. She [...]"
2	Mentioning and describing emotions of oneself or others to complement the description of experiences, helping the authors and others to later go back to the emotions, e.g. "this really made me angry".
3	Interpreting or explaining behavior in the experience , including rationales for certain own or others' behavior as well as potential reasons for issues, e.g. "It [the experience] was made worse by that fact that [...]"
4	Linking an experience explicitly to other experiences by mentioning or describing own or others' experiences relevant to the experience described, e.g. "I had a similar situation last week [...]"
5	Linking an experience to knowledge by referring to own or collective knowledge, data and other resources relevant to the experience, including advice without further explanation, e.g. "Never do [...]" or "Perhaps send an email before approaching in person".
6a	Responding to the explanation of an experience by providing possible alternative perspectives to the original interpretation, e.g. "I might have said [...]"
6b	Responding to the explanation of an experience by challenging or supporting assumptions , opinions or attributions made, e.g. "Agreed!" or "Hmmm. Is this really different from [...]"
7a	Contributing to work on a solution by providing reasons for the issue by explicitly mentioning the background of the issue or going beyond standard solutions, e.g. "Could it be down to [...]"
7b	Contributing to work on a solution by providing solution proposals , including reasons for the proposal or experiences linked to it, but without proposing to set them in practice; e.g. "I suggest this is escalated to your manager because [...]"
8a	Showing insights or learning from reflection by describing better individual understanding of the experience and drawing conclusions for own work, e.g. "I realized I should not be worried about ..."
8b	Showing insights or learning from reflection by generalizing from reflection by finding solutions that fit beyond the context of the current experience reflected on, e.g. "The key for us guys is definitely [...]"
9	Describing or implementing change such as proposing to apply practices, discussing change implementation or describing change set into practice, e.g. "Will definitely try and do [...] in the future".

Stage 1 is concerned with the basic elements of reflection such as descriptions of experiences, emotions and initial assessments, which are created by an individual. Stage 2 comprises activities in the actual collaborative reflection process such as exchanging experiences and knowledge (codes 4 and 5) as well as critically referring to each other and contributing to solutions (codes 6 and 7). Stage 3 is focused only on outcomes as the decisive part for success of collaborative reflection. This includes the description of learning insights (codes 8a and 8b) or plans for changes (code 9). This stage needs to be separated from the reflection process conceptually, as for example in contributions tagged with codes 7a or 7b people may juggle with different thoughts and aspects, while reaching stage 3 means they have to report insights and actual change. Table 3 summarizes these stages and shows which codes belong to each stage. In our analysis we assigned a communication thread to a certain stage if at least one code of the stage applied to the session. We used these stages to analyze the success of collaborative reflection that can be perceived from using the app.

Table 3. Stages of reflection and codes related to the stages.

Stage	Description	Codes
1	Provision and description of experience , but no (explicitly) traces of reflection	1, 2
2	Reflection on experiences , including analysis and potential solutions, but no (explicit) mentioning of learning or change	3-7
3	Learning or change resulting from reflection explicitly mentioned	8, 9

The differentiation of these levels allows an analysis of the success of using the tool, that is, how far (in terms of results) users got in their reflection with the tool and how much these users could benefit from using the tool by becoming aware of shared perspectives on issues or learning and change happening.

3.4 Dataset for the Analysis

For the analysis of the four cases we used a dataset for each case that contains all content created in the cases within the respective timeframes shown in Table 1 as well as log data on usage of the app and qualitative data such as feedback from participants and observations made onsite in the cases.

For the analysis we reduced the resulting dataset to collaborative content, that is, content including a report describing an experience and at least one comment on that experience. This was done to ensure we only analyzed content created in interaction between people. The figures for “answer ratio” (the proportion of reports receiving at least one comment) in Table 4 show that for every case this was the vast majority of content. After coding the content, we removed non-reflective content from the dataset, that is, content in which there were no traces of reflection, indicated by no code being applicable to the conversation. Again, this was only the case for a few cases, thus not reducing the dataset drastically. Table 4 shows the resulting number of conversations analyzed for each case. These two steps reduced the dataset from 109 conversations and 176 comments to 74 conversations and 159 comments analyzed.

3.5 Participant Feedback

Additionally to the usage data and the content in the app we received different direct feedback from participants. This feedback was received on different ways in the different cases due to different possibilities to receive this feedback. Most important, while in cases

2, 3 and 4 we were able to hold focus group meetings of participants, this was not possible in case 1 due to the different workplaces of the interns. In cases 1, 2 and 3 we were able to conduct a debriefing meeting with the manager of the group, which was not possible in case 4. In case 3 we were able to conduct short interviews with participants. In all cases, we received occasional feedback via email. Given these differences the data cannot be compared to each other but still contains valuable insights on the cases.

4. RESULTS

4.1 App Usage and Group Behavior

To understand how the app was used in the different groups, we compiled the basic usage data as described above and set it into relation to the time and amount of users in each case. Table 4 shows the resulting figures describing the usage of TalkReflection in the cases. In our analysis, we differentiated between average usage per user and day to account for the different numbers of participants, and overall usage figures of activity not regarding the number of users (e.g., reports per day). While the former gives insights into the activity and motivation of each user, the latter is a measure for the impression of activity a user gets from the app.

It can be seen from Table 4 that in all cases a fairly high proportion of experience reports created were also commented on (answer ratio between 0.71 for case 2 and 0.88 for case 3). This indicates that the app has been used collaboratively and that users had an interest in reflecting together in all cases. The lower value for case 2 may be caused by scale effects, meaning that users had more choice to comment on (nearly twice as many reports) and thus stuck to the most interesting ones. Other figures show that collaboration differed in its intensity: In cases 2 and 3 the average user created 0.31 experience reports per day (“Reports/user, day”), while this was less in cases 1 and 4 – again we need to take into account that in case 1 several users stopped using the tool after a couple of days, thus lowering the respective average values for case 1. Likewise, we can see that in cases 2 and 3 the average user created more comments per day than in cases 1 and 4. Therefore we can conclude that users in cases 2 and 3 were about twice as active as users in the other cases.

Table 4. Usage figures for the four cases.

	Case 1	Case 2	Case 3	Case 4
Reports	24	45	25	15
Reports/day	0.47	0.56	0.57	0.48
Reports/user,day	0.03	0.05	0.06	0.05
Comments	47	65	39	25
Answer ratio	0.83	0.71	0.88	0.80
Comments/day	0.92	0.81	0.93	0.50
Comments/user,day	0.05	0.07	0.10	0.06
Avg. thread length	2.35	2.03	1.77	2.08

Concerning the intensity of communication, that is how many reports and comments were created per day, activity in the cases was similar, with the exception of the low number of comments per day for case 4. In case 1 there were on average more comments in communication threads than in the other cases, with case 3 showing the lowest value for average communication thread length. Thus we may conclude that case 1 showed more overall activity than the other cases (although per user and day this was different due to the dropouts), while the others trail in at least one aspect (answer ratio

for case 2, thread length in case 3 and comments per day in case 4). This can only be partially attributed to the number of users in case 1, as – including the early dropouts in case 1 – case 2 had a similar number of users.

Regarding the collaborative reflection activity in the app, the answer ratio was fairly high in all cases, with case 2 trailing the others and case 3 showing the highest value. Likewise, the number of new experience reports per day is similar in all cases, while the number of new comments per day is significantly lower for case 4 than in the other cases. This means we can consider cases 1 and 3 to be most active in terms of output, while in case 2 some experience reports were not regarded by users and in case 4 users created less comments.

Besides a description of the groups and how they worked, the figures in Table 4 also provide insights into the average usage for tools supporting collaborative reflection: Averages of about 0.5 documents per day, 2.6 to 5.4 comments per user and 0.5 to 1 comment per day in all cases suggest that the usage we observed may be typical for small to medium groups using reflection tools like the TalkReflection App. Feedback from participants and our observations underpins this: Participants in different cases told us that they do not encounter critical situations every day and therefore did not use the app more often. Given that in case 1 several participants stopped using the app early (see section 3.2), we may expect the numbers to be slightly higher in practice. However, while other studies run in other settings suggest that reflection tools may be used more often (e.g., [15]), when using collaborative reflection tools in real work environments we should expect usage as described by the figures in Table 1 and Table 4.

4.2 Collaboration Network

Looking for insights into the collaboration network in the app, we applied basic social network analysis as proposed by [37,40] to the log data gathered from the cases (see section 3.3)¹. The network graph analyzed includes the participants from the cases as *vertices* and the connection between a user commenting on another user’s experience report as directed *edges* from the user commenting to the user being commented on. This means that for each graph the number of edges is the number of comments and the number of vertices is the number of users in the cases as shown in Table 4.

As Table 5 shows, graph density is lowest for case 1 and highest for case 4, with cases 2 and 3 in between. This indicates that in case 4 more users were connected to each other than in the other cases; the low density for case 1 can be explained by the several dropouts in the study as mentioned above. This is underpinned by the fact that the unique edge ratio is highest for cases 1 and 4, and considerably lower for cases 2 and 3. This suggests that communication in cases 1 and 4 was more diverse (more different pairs of users communicating) than in cases 2 and 3, while the low values indicate in cases 2 and 3 there might have been fewer but closer communication relationships established.

Table 5. Analysis of social network aspects in the cases.

	Case 1	Case 2	Case 3	Case 4
Graph density	0.10	0.17	0.19	0.25
Unique edge ratio	0.62	0.18	0.21	0.60

¹ The social network analysis was done using NodeXL, a Microsoft Excel based tool (<https://nodexl.codeplex.com/>).

These differences can be attributed to the different cooperation styles established or chosen in the cases: While in cases 1 and 4 participants used the tool in a self-directed manner, in cases 2 and 3 a user dominating the others was present (see section 3.2): In case 2 the manager of the participant had decided to drive the reflection in the tool, and in case 3 the senior physician had asked his staff to use the tool and promised to comment on the issues shared by them in return. This resulted in multiple communication edges between the respective manager and different other users, and (as some users were reluctant to add a comment before the dominant user had added one) fewer other edges.

4.3 Content Analysis: Collaborative Reflection Outcomes

Content coding was done independently by two coders on the level of single contributions to conversations (reports and comments). This means that a code was assigned once or not at all to an experience report or comment. The coding resulted in 597 codes assigned by the two coders.

Concerning the agreement among coders, we calculated values for Krippendorff’s Alpha for each code. Results were mixed: While we found acceptable values (.75 and slightly below) for some codes (e.g., codes 4 and 5), we also found worse for others (e.g., codes 6a and 7a). Analyzing the differences between the coders we found that despite differences in some codes there was good agreement concerning the stage of reflection reached in each conversation (see Table 3 for the stages). When calculating the inter-rater agreement on these levels, we arrived at good agreement values (97% for stage 1, 96% for stage 2, 80% for stage 3). Although the level of details is lower for these stages compared to the coding scheme, the quality of the resulting data is better. Therefore we will describe results from the coding by the levels reached in the conversations. To further enhance the quality of the coding for stage 3, the coders discussed differences in using codes 8a, 8b and 9, resulting in a coding agreed upon the two coders. This data was used for the following analysis.

Table 6. Stages of collaborative reflection reached in the cases.

Stage	Case 1	Case 2	Case 3	Case 4
# Conversations	17	24	21	12
1	100 %	95.8 %	100 %	100 %
2	94.1 %	95.8 %	95.2 %	83.3 %
3	23.5 %	33.3 %	0 %	16.7 %

Analyzing the content coding and applying it to the three stages of collaborative reflection described in Table 3, we found that stages 1 and 2 were reached in most conversations throughout the cases, while there are differences in the proportion of conversations that led to outcomes on stage 3. The results for stage 1 are not surprising, because before the coding the obviously non-reflective conversations were eliminated from the sample. The conversations that did not reach this stage (only two thread in case 2, see Table 6) were initiated by the respective user with solution suggested for a certain problem and reflected upon afterwards, thus lacking the description of concrete experiences. The high proportion of conversations that reached stage 2 is a result better than we initially expected, especially for the cases with lower user numbers. The figures for stage 3 show major differences between the groups with case 2 reaching this stage for one third of all conversations and case 3 not reaching it in any thread. Table 6 gives an overview of stages reached.

Interpreting these figures, we can see that if reports on experiences received at least one comment (that is, the data set we coded) the conversation was very likely reach stage 2 and thus to be reflective – even in the worst case (case 4) 83% of the conversations reached this stage. This can be seen as a success of the tool, especially because we found in earlier studies that many conversations in which issues had been brought up were not systematically reflected on, and for the few that people reflected on there were no traces left of the corresponding conversation afterwards [reference omitted for blind review]. This shows how the TalkReflection App changed collaborative reflection in the cases: It created opportunities to reflect together asynchronously (that is, without the need to switch from other tasks or to step back immediately from work when being told about an experience) and made outcomes from reflection available after the reflection session and outside face-to-face group meetings.

In addition, we may conclude that cases 1 and 2 were more successful in terms of learning outcomes and change, and that in case 3 no learning took place. We are aware that concluding from the lack of explicit mentioning of this stage that neither learning or change resulted from using the tool is dangerous, as communication on this may have taken place outside the tool in face-to-face situations. In fact feedback from participants and our observations in workshops revealed that this was true for many cases. However, with a focus on the role of tools in collaborative reflection support, which may bring together reflection groups unable to reflect in daily face-to-face interaction, lower percentages of stage 3 reached for a group also mean that by using the tool a user gets less information on what can be learned and derived from experiences shared. The result for stage 3 is a pity especially in cases such as case 3, in which there was good effort in using the app but not even one result from reflection documented. In such cases if people are not part of meetings or other occasions in which communication on what to draw from the discussion in the tool takes place cannot take full benefit from collaborative reflection.

The high proportion of stage 2 conversations and the existence of some conversation reaching stage 3 also show the impact of using the tool: In almost all cases the tool enabled reflection among users and – especially in cases 1 and 2 – often led to explicit articulations of outcomes from reflection. Comparing this to situations we faced in earlier case studies [reference omitted for blind review], in which many experiences to be reflected as well as ideas and insights from reflection were soon forgotten, this makes reflection not only possible, but also sustainable and understandable for those not directly involved but reading through the content of a tool afterwards.

4.4 Participant Feedback

This section describes the feedback of the different sources described in section 2.5. In **case 1** participants and the manager told us that reflection with the tool helped them to cope with challenges they met as new employees in their respective jobs. We were also told that activity in the tool only lasted a short timeframe, as interns have short-time contracts and thus the user base soon became too small – new users were not added by the interns' manager.

In **case 2** participants told us that using the tool was valuable to reflect on own practices and changes on it in the upcoming merge of the departments, especially when reflection could not be done immediately. Their manager as the leading users of the tool added that in his impression the two departments had become closer to

each other as a result of reflection, and that he had liked to bring up topics and to ask his staff to reflect on them.

In **case 3** some participants told us that despite using the tool they did not see much value in it as they had known most experiences reported from face-to-face interaction before. Others mentioned that the exchange had helped them to know that others had similar problems, and to receive advice from the senior physicians, who had added this advice to nearly all reports. Overall, the perception of value added by the tools was low among the assistant physicians. The senior physician driving the usage stated that he felt the tool gave him the opportunity to train his staff.

In **case 4** participants told us that they were asked by their manager to use the tool only in breaks and in the small office of the care home, which they only used for documentation purposes. Those who had used the tool despite these constraints told us that they found it valuable to reflect on stressful cases, but were limited severely by time constraints in using the app, which was an extra effort for them. Some participants told us that they had asked for more time and to discuss some issues from the app in their meetings, which was refused by the manager.

The feedback of participants adds a flavor to the analysis that data and content analysis cannot create. In particular, it shows the attitude of participants towards the TalkReflection App, which would have been overshadowed e.g. by orders to use the app (case 3) or constraints hindering the usage (case 4).

5. DISCUSSION

The results of our studies give good insights into how participants used the TalkReflection App in each case, and also how the groups differed from each. Given the lack of insights in this area, this paper takes the support of collaborative reflection a step further. Striving to answer the research questions given above, this section analyzed these results.

5.1 Using Tools for Collaborative Reflection: Patterns in Group Activity

The results from the studies show that in each of the cases there was activity and engagement in collaborative reflection using the TalkReflection App. The results also show that there were differences in this activity and that the participants perceived the value of using the app differently. Comparing the cases reveals that there are some decisive aspects that accounted for success or low activity and perceived value.

For **case 1** we can see good values for overall collaborative reflection activity (reports and comments per day, answer ratio) and in participation (highest unique edge ratio and average thread length). Other values on activity per user would have been higher without the dropouts in the study. The case also shows a good amount of stage 3 reflection outcomes, but the number of reports is fairly low given the amount of users. Overall, case 1 can be regarded as a case in which the TalkReflection App has added value to the group. This is underpinned by the feedback we got from the interns. This may be attributed to the spatial situation of case 1, in which the interns worked at different departments and did not have the opportunity to reflect together face-to-face on problems they were facing. The TalkReflection App gave them this opportunity, resulting in good uptake and results.

For **case 2** we can see good values for reflection activity (good values for reports and comments per user and day), and low numbers for diversity (lowest unique edge ratio) and participation (lowest answer ratio). In addition, case 2 has the largest proportion of cases reaching stage 3 of reflection. Therefore, case

2 is an example in which average or low figures for participation and diversity do not predict little reflection success – on the contrary, the focus on the leading user that has caused low diversity in this case seems to also have fostered reflection outcomes. This is reflected by the feedback of users, who saw much value in using the tool.

For **case 3** we can see that despite good collaborative reflection activity (most reports and comments per day and user, highest answer ratio) the participation (low values for thread length and from social network analysis) and output (no traces of stage 3 in reports and comments, feedback indicating little impact) seems to be low. We attribute this to the organization of reflection and the dominant user in case 3: the assistant physicians had been asked to use the tool regularly, which resulted in high usage figures, but it had not been integrated into meetings and other face-to-face discussion opportunities, lowering the perceived value and outcomes. Moreover, the senior physician in case 3 acted as a dominant user giving advice in his comments how to deal with the respective situation rather than supporting reflection. This resulted in other participants waiting for him to answer and thus lowering the diversity in communication. Both of these effects may have hindered collaborative reflection to take place to an extent that resulted in traces of learning or change (stage 3).

For **case 4** we can see mediocre to low figures for activity (least reports and comments, least comments per day), high values for diversity and participation (highest graph density, high unique edge ratio) and at least some conversations that reached stage 3. Taking into account the feedback we got on the constraints of using the TalkReflection App in this case we can take it as an example in which people saw value in the tool but could not hold usage up. The resulting low usage is grounded in lacking process integration and co-location of the users.

As one conclusion from our comparison we can say that single dominant users in reflection groups can be good or bad for reflection practice. While in cases 2 and 3 activity and participation figures show engagement in reflection, the lead user led to comparably little diversity in the communication (see Table 4). This, however, led to different effects on collaborative reflection output: In case 2 it seems that the manager driving the activity in the group caused the group to be more reflective in the app, resulting in the highest number for occurrences of stage 3 reflection traces (Table 3). In case 3 the dominant user seems to have caused the opposite, as none of the communication threads contained traces of stage 3 reflection. This may be explained with the different roles these users played (see sections 3.2 and 4.4): In case 2 the manager had brought up topics and asked people to reflect on them, thus acting as an enabler and driver of collaborative reflection. In case 3 the physician had asked staff to add experience reports and – instead of driving reflection – has added his advice from the perspective of an experienced senior physician, which users may have just taken instead of continuing reflection – here the dominant user acted as an unintentional blocker of reflection activity. In both cases the focus on the dominant user may also be responsible for the low diversity in communication, as their comments may have discouraged others from commenting, resulting in edges from them other users may have dominated the respective social network.

Another conclusion from the comparison is that scale matters: the results of the different groups go along with group size and overall activity. Even if per user and day cases 3 and 4 in some respect performed equal or slightly better than cases 1 and 2, the latter cases contain more results on the outcomes stage 3 of the

content analysis. This may be caused by the fact that with more users a higher overall number of documents and comments is created, also resulting in a higher average communication thread length. In addition, for users (new or existing) of a collaboration support tool a higher amount of content available makes using the tool much more attractive, as they can see that people engage in collaborative reflection in the tool and that issues are reflected on intensively. This assumption is supported by the fact that the lower answer ratio in case 2 seems to have had no negative effect on the stages of reflection reached or any other figures – in contrast case 2 has most outcomes on stage 3 of reflection. This might be explained by the higher number of reports to choose from, which gives users the chance to decide for engagement in discussions on cases they are really interested in and which they can contribute to. Taking these observations together, we may conclude that collaborative reflection support tools such as the TalkReflection App are more beneficial to groups of 10-15 and more people than they are for small groups.

5.2 Factors supporting or constraining Reflection Support Usage

The comparison of the groups as described above points to measures and figures that help to explain why collaborative reflection works better or worse in different settings:

- **Stages of reflection:** Differentiating between the stages described in Table 3 is a good indicator of reflection success or failure; especially the differentiation between stage 2 and 3 seems to be crucial. Our results for stage 3 outcomes as shown in Table 6 go along well with the perceived value articulated in the feedback of participants, indicating that they are a measure of success in a group.
- **Thread length:** In our cases, the average length of threads was one indicator for good reflection outcomes as described above and for satisfaction of users. This seems natural, as longer threads implicate more engagement into communication in the tool, thus increasing the chances that threads will eventually lead to good outcomes.

In contrast to these measures, the analysis shows that other measures such as answer ratio or activity per user and day showed different values for the groups, but *did not predict reflection success* in any way, as cases 1 (low values per day and user) and 2 (lowest answer ratio) show. Although disregarding group size, *figures for total activity in terms of reporting and commenting worked much better*. The other way round, particularly low values for such figures such as for comments per day in case 4 indicate that reflection has not worked well. Together with thread length as a good indicator of reflection success (see above), this points to the assumption that *size and scale of reflection groups are more important* than average individual activity. A reflection group might reach a *threshold of activity* per day from which on it might be perceived more interesting to choose a relevant case and to comment on cases – further work should look into this as a success factor of tools support for collaborative reflection.

In addition to these aspects the cases also show that *some measures cannot give insights into collaborative reflection tool uptake and usage without additional context*. For example, social network analysis such as communication graph density and unique edge ratio cannot be considered predictors of collaborative reflection quality without other information given as context: In the cases, good output on stage 3 and participation co-occurred with low (case 2) and high (case 1) values for graph density and unique edges, and vice versa.

The aspects described above may also help the design of a socio-technical reflection setting, in which tools have to be embedded into certain circumstances and managed properly. The cases also point to decisive factors in this dimension:

- **Dominant user vs. self-directed use:** The cases show that tools support for collaborative reflection can be successful in self-directed tool usage (case 1) and directed by a dominant user (case 2) – under both circumstances it may also be less successful (cases 3 and 4). This may be a matter of scale and critical mass of users being active, as for the larger group in case 1, in which different users were active, self-directed usage worked while for the small group, in which some users did not find the time to use the app often, it did not.
- **Role of a dominant user:** Looking into cases 2 and 3 we have seen that the behavior of the dominant user was decisive for the success in these cases: A coach and driver of reflection fosters collaborative reflection activity while and advisor blocks it. We should therefore make users who might take this role aware of the impact they have on reflection. This (for some) may need a change in managers' mindset in becoming a coach rather than answering all questions, and it may need training.
- **Process integration and co-location:** The cases show that when using collaborative reflection tools in practice co-location and process integration need to be taken care of together. For the cases (partly) co-located (2, 3, 4; see Table 1) we saw that when there was no process integration (case 3 and 4) the app did not produce additional value – people used face-to-face encounters to talk about the most important issues anyway. In case 2, in which half of the users were collocated, the manager had integrated the app into meetings and made it a part of people's work. This resulted in successful reflection. We can also conclude that without process integration or good facilitation reflection support tools may provide limited value for co-located groups.

Concerning the latter finding, one may argue that this does not come at a surprise, as in co-located groups much happens in direct communication. However, we saw that with good process integration and facilitation the app provided benefit in the (partly) co-located group of case 2 and that, given different behavior of the dominant user, this might have also happened in case 3.

5.3 The Role of Tools in Collaborative Reflection

The analysis of the cases described in this paper reveals three ways in which the TalkReflection App was helpful for the participants:

- **Keeping up reflection:** For all cases our results suggest that the TalkReflection App helped participants to engage in reflection over a period of time. This is indicated by continuous usage over time, by good answer ratios to experience reports any by and overall satisfying activity in all cases. However, we can also say that keeping up reflection is not enough for reaching higher levels of reflection output and perceived value, as it was also present in the less successful cases.
- **Supporting face-to-face reflection:** Especially from case 2 we can see that reflection tools show benefit if they are used to foster reflection in face-to-face situations such as meetings. The manager in case 2 had included this usage from the beginning, which may also have motivated participants to use the tool. In other cases such as cases 3 and 4 this was not done, and consequently the app was perceived less valuable as it did

not add much to everyday communication possibilities – in case 3 participants had even asked for this support. We can therefore consider support for face-to-face reflection, for example using the content of the app in meetings like in case 2, as a way tools can become helpful for reflection.

- **Connecting reflection participants:** Case 1 and partly case 2 show that the value added by the tool lies in the connection of participants who work in different places to a (virtual) reflection group. The interns reflecting in case 1 with other interns would not have had the opportunity to reflect with peers in similar positions regularly without the tool. Therefore, given a situation in which people are not co-located or have regular meeting, connecting them to virtual reflection group is a major role tools can play in reflection support.

Besides analyzing the roles the TalkReflection App played in supporting collaborative reflection in the cases we also need to ask what improved tools for this support should be able to do. This means asking how a tool (besides using a lead user and proper processes) may support individual and groups in achieving more and better results from collaborative reflection. While our study may not provide an exhaustive list of requirements for such apps, the insights taken from the cases already point to some improvements:

- **Provoke activity:** Our analysis shows that overall activity seems to be helpful to keep up reflection and reach good results. In addition we saw that if reports received comments, the discussion was likely to become reflective. Therefore, tools supporting collaborative reflection should provoke activity among users, e.g. by reminding users of reports shared with them and prompting them to create comments.
- **Scaffold dominant user role:** We have shown that dominant users can be supportive in reflection, but that they may also drag too much attention to them and thus hinder reflection. Tools have limited power in coping with this, but may for example differentiate between user roles: If, like in case 3, it is known (and intended) in advance that more experienced users are part of the reflection group, a tool might enable them to contribute only after a certain while in order not to inhibit others from commenting from their point of view. Likewise, as tool could allow a dominant user to push reflection as a coach or driver by directly addressing others to contribute to certain discussions.
- **Point to relevant issues:** Our study indicates that more users and more cases in total may lead to better choice among experiences reports, which in turn lead to more commenting. Besides relying on such scale effects tools may support reflection even more if they recommend relevant issues to users, thus making contributions more likely.
- **Balance diversity and intensity of communication:** We have seen that diversity in communication (more links between users) and intensity (recurring links between the same users) may both be helpful to foster exchange among reflection participants and to deepen reflection. Tools may support a balance between these effects by managing the social network of reflection, e.g. showing users their existing communication partners and proposing new partners that might have similar issues at work.

Our analysis revealed differences between group activity and – from combining these differences with content analysis and qualitative feedback – insights on which differences indicate reflection success or flaws in reflection. From this we identified measures that may

predict reflection success and other measures that at least point to more or less successful reflection.

In terms of *setting up technology supported collaborative reflection as socio-technical system* we can say that dominant users in the role of coaches have a positive influence in reflection tools, that process integration is decisive in small, co-located groups while connecting people is important in dislocated groups and that overall activity positively influences outcomes. In particular, our results suggest that for groups in which users are connected mainly by a tool a self-directed, non-integrated approach may be sufficient, as the surplus of being connected and able to reflect together is enough motivation for users to participate. Case 1 is a model case for this. In contrast, for small co-located groups an approach with a user driving reflection and with good integration into work such as using content from the tool in meetings seems to work, as case 2 shows. This may help to overcome flaws such as lacking time to reflect, little perceived value and others identified in cases 3 and 4.

In a way these results resemble the old-new themes of the disparity between effort invested and benefit gained [12] and added value for the individual [7], but we apply these general principles to constraints of collaborative reflection that need to be in place to make collaborative reflection work. Continuously analyzing, for example, the length of threads in a reflection tool or the overall group's activity to ensure a certain amount of activity in the tool may help to keep reflection going in a tool. However, we regard the results presented here as a start for such analysis, and we will strive to find more of such measures to allow the design of better tools to support collaborative reflection.

Our analysis also shows that results from reflection do not emerge automatically just by giving people tools. While we could see in all cases that if at least one comment was made on a report, reflection on the level of stage 2 was likely to happen. However, reflection on the level of stage 3 did not follow in the same way. Our data suggests that this stage may be reached more likely if the activity in the tool is higher and if the tool fits better to the groups, thus adding value to the setting it is used in (e.g. being used for meetings in co-located settings). Likewise, we saw that a user taking charge of the reflection process by e.g. asking others to provide stage 3 insights can foster such outcomes of collaborative reflection. Looking deeper into these results and the implications drawn from them above as well, applying them to design and orchestration of reflection tools and refining them is necessary to support collaborative reflection sufficiently.

The TalkReflection App as presented here cannot be regarded as a blueprint for collaborative reflection support tools in general, as there neither any such tools to compare it to nor a lot of insights to characterize what such tools need to do besides the support described in this paper. Therefore the results presented above cannot be generalized for all collaborative reflection support, but they can be understood as a basis to draw from for further work and research on collaborative reflection tool support. Furthermore, we expect features of other collaborative reflection support tools to overlap with the features of the TalkReflection App. In this sense, the results presented in this paper show how collaborative reflection can be done and provide insights into this currently not sufficiently researched area that can be used to build on.

6. CONCLUSION AND OUTLOOK

In this paper we have analyzed four cases of using the same reflection support tool at different workplaces. The results of the analysis shed light onto collaborative reflection support, which is an area that had not been sufficiently researched yet. From our work we could identify measures to be used when assessing collaborative

reflection, socio-technical aspects of supporting collaborative reflection and requirements for tools implementing this support. While these results are not final or generalizable, they provide good ground to start from in further research.

Besides the insights on collaborative reflection to take away from our work, we also recognized that there is need for better measures of collaborative reflection activity. We identified some measures that may help to differentiate successful from less successful reflection groups and create intervention for the later ones while learning from the former. However, these provide a rather coarse grained classification of such groups and their tool usage, which should be refined. Our future work will also look into this aspect of research on collaborative reflection. Likewise, there is a need to combine the analysis of reflection in tools better with observations of reflection practice outside tools to gain a better overview of the impact created by a tool. Combining the work presented in this paper with approaches of ethnographic work we pursued earlier on collaborative reflection [26] will thus be one direction for future work.

Our work shows that collaborative reflection support is a field worthwhile working on, as the tool used in our cases created change in terms of opportunities for collaborative reflection in each case. We also saw that there is still work to be done in order to make technology supported collaborative reflection work at workplaces and to tap from the potential that collaborative reflection has for such workplaces.

7. ACKNOWLEDGEMENTS

This work has been supported by projects MIRROR (funded by the European commission in FP7, project number 257617) and EmployID (FP7, project number 619619). We thank all members of the projects for their support and ideas on this work.

8. REFERENCES

1. Baumer, E.P., Khovanskaya, V., Matthews, M., Reynolds, L., Sosik, V.S., and Gay, G.K. Reviewing Reflection: On the Use of Reflection in Interactive System Design. Proceedings of DIS 2014, (2014).
2. Bjørn, P. and Boulus, N. Dissenting in reflective conversations: Critical components of doing action research. *Action Research* 9, 3 (2011), 282–302.
3. Boud, D. Reflection: Turning experience into learning. Kogan Page, London, 1985.
4. Daudelin, M.W. Learning from experience through reflection. *Organizational Dynamics* 24, 3 (1996), 36–48.
5. Dennis, A.R., George, J.F., Jessup, L.M., Nunamaker Jr, J.F., and Vogel, D.R. Information technology to support electronic meetings. *MIS quarterly* 12, 4 (1988), 591–624.
6. Dyke, M. The role of the 'Other' in reflection, knowledge formation and action in a late modernity. *International Journal of Lifelong Education* 25, 2 (2006), 105–123.
7. Edgington, T., Choi, B., Henson, K., Raghu, T., and Vinze, A. Adopting ontology to facilitate knowledge sharing. *Communications of the ACM* 47, 11 (2004), 85–90.
8. Fahy, P.J., Crawford, G., and Ally, M. Patterns of interaction in a computer conference transcript. *The International Review of Research in Open and Distance Learning* 2, 1 (2001).
9. Fleck, R. and Fitzpatrick, G. Reflecting on reflection: framing a design landscape. Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of

- Australia on Computer-Human Interaction, ACM (2010), 216–223.
10. Forneris, S.G. and Peden-McAlpine, C.J. Contextual learning: A reflective learning intervention for nursing education. *International journal of nursing education scholarship* 3, 1 (2006), 1–17.
 11. De Groot, E., Endedijk, M.D., Jaarsma, A.D.C., Simons, P.R.-J., and van Beukelen, P. Critically reflective dialogues in learning communities of professionals. *Studies in Continuing Education* 0, 0 (2013), 1–23.
 12. Grudin, J. Why CSCW Applications fail: Problems in the Design and Evaluation of organizational Interfaces. *Proceedings of CSCW 1988*, (1988), 85–93.
 13. Hoyrup, S. Reflection as a core process in organisational learning. *Journal of Workplace Learning* 16, 8 (2004), 442–454.
 14. Introne, J.E. and Drescher, M. Analyzing the flow of knowledge in computer mediated teams. *Proceedings of the 2013 conference on Computer supported cooperative work*, ACM (2013), 341–356.
 15. Isaacs, E., Konrad, A., Walendowski, A., Lennig, T., Hollis, V., and Whittaker, S. Echoes from the past: how technology mediated reflection improves well-being. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 1071–1080.
 16. Lee, S. Design and analysis of reflection-supporting tools in computer-supported collaborative learning. *International Journal of Instructional Technology and Distance Learning* 2, 3 (2005), 49–56.
 17. Levina, N. Collaborating on multiparty information systems development projects: a collective reflection-in-action view. *Information Systems Research* 16, 2 (2005), 109–130.
 18. Lockhorst, D., Admiraal, W., Pilot, A., and Veen, W. Analysis of electronic communication using 5 different perspectives. (2003).
 19. Marcu, G., Dey, A.K., and Kiesler, S. Designing for Collaborative Reflection. *Proceedings of PervasiveHealth'14*, (2014).
 20. Mercer, N. and Wegerif, R. Is 'exploratory talk' productive talk? *Learning with computers: Analysing productive interaction*, (1999), 79.
 21. Moon, J.A. *Reflection in learning & professional development: theory & practice*. Routledge, 1999.
 22. Newman, D.R., Webb, B., and Cochrane, C. A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology* 3, 2 (1995), 56–77.
 23. Porges, Z., Yang, X., Desai, A., et al. Achieve: Evaluating the Impact of Progress Logging and Social Feedback on Goal Achievement. *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM (2014), 221–224.
 24. Power, D.J. and Sharda, R. *Decision support systems*. Springer handbook of automation, Springer (2009), 1539–1548.
 25. Prilla, M., Degeling, M., and Herrmann, T. Collaborative Reflection at Work: Supporting Informal Learning at a Healthcare Workplace. *Proceedings of the ACM International Conference on Supporting Group Work (GROUP 2012)*, (2012), 55–64.
 26. Prilla, M., Pammer, V., and Krogstie, B. Fostering Collaborative Redesign of Work Practice: Challenges for Tools Supporting Reflection at Work. *Proceedings of the European Conference on Computer Supported Cooperative Work (ECSCW 2013)*, (2013).
 27. Prilla, M. User and Group Behavior in Computer Support for Collaborative Reflection in Practice: An Explorative Data Analysis. *COOP 2014 - Proceedings of the 11th International Conference on the Design of Cooperative Systems*, Springer (2014).
 28. Prilla, M. Collaborative Reflection Support at Work: A Socio-Technical Design Task. *Proceedings of European Conference on Information Systems (ECIS 2014)*, (2014).
 29. Raelin, J.A. I don't have time to think!" versus the art of reflective practice. *Reflections* 4, 1 (2002), 66–79.
 30. Reynolds, M. Critical reflection and management education: rehabilitating less hierarchical approaches. *Journal of Management Education* 23, 5 (1999), 537–553.
 31. Roschelle, J. and Teasley, S. The construction of shared knowledge in collaborative problem solving. *Computer Supported Collaborative Learning*, Springer-Verlag (1995), 69–97.
 32. Schön, D.A. *The reflective practitioner*. Basic books New York, 1983.
 33. Scott, S.G. Enhancing Reflection Skills Through Learning Portfolios: An Empirical Test. *Journal of Management Education* 34, 3 (2010), 430–457.
 34. Suchman, L.A. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, 1987.
 35. Tigelaar, D., Dolmans, D., Meijer, P., de Grave, W., and van der Vleuten, C. Teachers' Interactions and their Collaborative Reflection Processes during Peer Meetings. *Advances in Health Sciences Education* 13, 3 (2008), 289–308.
 36. Weick, K.E. *Sensemaking in organizations*. Sage Publications, Inc, 1995.
 37. Wever, B.D., Schellens, T., Valcke, M., and Keer, H.V. Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education* 46, 1 (2006), 6–28.
 38. White, B.Y., Shimoda, T.A., and Frederiksen, J.R. Enabling students to construct theories of collaborative inquiry and reflective learning: Computer support for metacognitive development. *International Journal of Artificial Intelligence in Education (IJAIED)* 10, (1999), 151–182.
 39. Van Woerkom, M. and Croon, M. Operationalising critically reflective work behaviour. *Personnel Review* 37, 3 (2008), 317–331.
 40. Zhu, E. Meaning Negotiation, Knowledge Construction, and Mentoring in a Distance Learning Course. *Proceedings of Selected Research and Development Presentations at the 1996 National Conventions of the Association for Education Communications and Technology*, (1996).