



How Does Collaborative Reflection Unfold in Online Communities? An Analysis of Two Data Sets

(E-mail: michael.prilla@tu-clausthal.de; oliver.blunk@tu-clausthal.de)(E-mail: chounta@ut.ee)Michael Prilla¹, Oliver Blunk^{1*}  & Irene-Angelica Chounta²

¹*Clausthal University of Technology, Clausthal-Zellerfeld, Germany*; ²*Tartu Ülikool, Tartu, Estonia*

Abstract. People can learn a lot through (collaborative) reflection at work: In organizations, staff debate experiences and due to issues every day, thus reflecting together and learning from each other. While this is desirable, it is often hindered by differences in time and space. Online discussions in community-like systems may provide a means to overcome this issue and enable staff to share experiences and learn from them. In this paper, we study two different data sets from two software systems to support online collaborative reflection in order to analyze what possible factors influence the occurrence of aspects of reflection in online discussions, like experiences, suggestions and learning. Our results include findings that partially approve existing models of (collaborative) reflection, but also others that add to or even challenge these models. Overall, we found that collaborative reflection may take different paths towards learning. From this we derive that facilitation mechanisms need to take up these paths, and that facilitation may be successful in different ways than anticipated from existing literature. We describe possible facilitation mechanisms and discuss their implementation.

Keywords: Collaborative reflection, Content analysis, Conversation analysis, Design for reflection, Reflection

1. Introduction

Reflection is a process of learning from experience by going back to past activities, re-assessing them in the light of current knowledge and drawing conclusions for further activities (Boud et al. 1985). Reflection has been recognized as an important skill in modern workplaces and as a mindset of modern workforce (Cressey et al. 2006; Prilla et al. 2013; Raelin 2002; Schön 1983). It has also been emphasized that reflection often takes place in social settings, in which people explicitly exchange their experiences and learn from them together (Cressey et al. 2006; Prilla et al. 2013; Scott 2010). This social or *collaborative* (Prilla et al. 2013) reflection has a lot of potential, as it may lead to solutions for problems that often go beyond solutions found by individuals (Hoyrup 2004; Mercer and Wegerif 1999). The discourse about collaborative reflection also links to central concepts of the CSCW community such as sensemaking, common ground, group decision support and collaborative problem solving (cf. Prilla et al. 2013, see below).

In this paper, we focus on collaborative reflection *at work*, which is done in *online media* and is mostly *asynchronous*: In many cases, reflection participants are geographically dispersed or lack the time to regularly come together for reflection. Support for collaborative reflection at work can help workers to improve as an individual as well, just like when doing group work (Prilla et al. 2013). This is mirrored by recent work on support for reflection in social settings (Marcu et al. 2014; Prilla et al. 2015; Slovák et al. 2017; Tang and Chen 2015). Despite this work and recent interest in research on design for corresponding reflection support (Baumer 2015), little is known about support for collaborative reflection in practice and at work (Slovák et al. 2017).

Literature provides several models explaining steps and activities in (collaborative) reflection (e.g., Moon 1999; Fleck and Fitzpatrick 2010; Krogstie et al. 2013). However, other work provides cause for doubts regarding whether collaborative reflection processes can be described by these models (Cressey et al. 2006; de Groot et al. 2013), and calls reflection ‘messy’ (Cressey et al. 2006, p. 23) rather than structured (as in the models). While it seems reasonable to doubt that collaborative reflection follows singular paths depicted in models, work available on the facilitation (Blunk and Prilla 2017a; Daudelin 1996; Fessler et al. 2017; Hoyrup 2004) show that certain aspects, phases and utterances are important for collaborative reflection and can be supported. Despite this, there is no work available that attempts to identify these phases beyond theoretic models. To close this gap, in this paper we ask the questions *how do reflective conversations unfold online in communities* as well as (based on answering the former question) *how can we support collaborative reflection*, and we investigate them based on two data sets.

In particular, we aim to shed light on the design of support for collaborative reflection by investigating the course of reflection and influences on certain aspects of reflection, like experiences, suggestions, and learning in reflective discussions in online tools and deriving insights from this. We looked at sequences of contributions to collaborative reflection and tried to find sequences or combinations that foster reflection. By this, our work focuses on understanding how collaborative reflection emerges between people, that is, how reflective discussions unfold in online communities.

For this we use an approach inspired by content analysis, and we analyze two data sets with a total of 135 (65 and 70) conversations created by 93 (48 and 45) users. More specifically, we look into individual conversations and analyze how certain types of (reflective) contributions influence the course of reflective discussions, looking at different factors that may influence reflection. This includes sequence analysis, correlation analysis and regression analysis as well as sequential pattern mining. Our results include findings that support existing literature, but also show that collaborative reflection may take several paths not to be anticipated from the literature. We also find relations between contributions to reflective discussions that challenge or add to existing models for collaborative reflection. From our results, we

derive suggestions for facilitation support of collaborative reflection and discuss how to implement them.

Our work adds to the discourse on the (IT-based) support of human practices prominent in the (E)CSCW community. Collaborative reflection is a common practice at work, from which a lot of people learn and grow. At the same time and like many other practices, it is unlikely that collaborative reflection can be modelled in a way that describes how reflection unfolds in online communities. Looking at collaborative reflection as situated in a sense that it is always ‘messy’ and cannot be supported by specific means at all is also dissatisfying. To dissolve this dichotomy, our work identifies paths along which collaborative reflection unfolds and provides an initial view into the multiplicity of these paths. In that way, it ties in with the practice lens of (E)CSCW and provides a new way to look at collaborative reflection when compared to existing literature.

2. Related work

2.1. Collaborative reflection

In line with the understanding of reflection provided in the introduction, collaborative reflection can be understood as reviewing past activities together and drawing conclusion from them together (Baumer 2015; Cressey et al. 2006; Prilla et al. 2015). Experience is one of the corner stones of reflection, as learners refer to their experience in order to learn from it (Boud et al. 1985; Schön 1983). Literature therefore describes the process of collaborative reflection as making experiences explicit (by e.g. writing), sharing and collaboratively making sense of them (Dyke 2006; Fleck and Fitzpatrick 2010; Scott 2010). This needs support for individual articulations of experiences, ideas and reflection as well as for relating to the articulations of others (Daudelin 1996; de Groot et al. 2013; Prilla et al. 2015).

Available literature provides a lot of models and characterizations that aim to explain reflection (e.g., Boud et al. 1985; Dewey 1933; Kolb 1984; Moon 1999). One of the most common models is the reflective learning cycle by Boud (see Figure 1). It describes reflection as learning from experiences and re-evaluating them with current knowledge in order to learn for the future (Boud et al. 1985). Returning to experiences and re-evaluating them is described as an iterative process that eventually should have outcomes such as new perspectives on experiences or changes in behavior. Learning is understood in a bandwidth that reaches from deriving new insights on past experiences to changing behavior. While this model explains the cognitive processes of reflection, it only applies to individual reflection and does not support the understanding of collaborative reflection.

Other models like the CSRL model (Figure 2, left) describe reflection with a more formal approach and define inputs and outputs of different stages, e.g. gathering data to initiate the reflection, or creating a frame for the reflection session to conduct it (Krogstie et al. 2013). The model also includes aspects of collaborative reflection

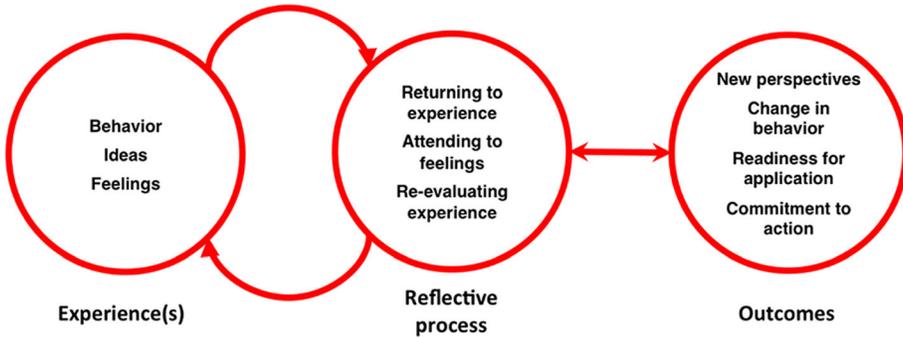


Figure 1.. The Boud et al. (1985) model of reflection.

such as articulating meaning created in reflection sessions but makes no assumptions on how collaborative reflection takes place in detail.

For reflection to become a collaborative process, there is a need to articulate and share experiences (Daudelin 1996; Scott 2010) as well as suggestions on how issues can be solved or what can be learned from reflection processes (Cressey et al. 2006; Prilla et al. 2013). This process, as articulated in the Prilla (2015) model of collaborative reflection (Figure 2, right), works as an interplay of individual (cognitive) and collaborative (explicit) reflection activities, in which the articulation of experiences and ideas is crucial.

Despite the amount and value of models and characterizations explaining (collaborative) reflection, it has also been argued that collaborative reflection is a ‘complex, multifaceted and messy process that is tamed and domesticated at the risk of destroying what it can offer’ (Cressey et al. 2006, p. 23). This suggests that the process of collaborative reflection cannot (and should not) be structured (see also de Groot et al. 2013 on this). In fact, much of the support and many studies available from the literature provide insights on supporting early phases of reflection such as

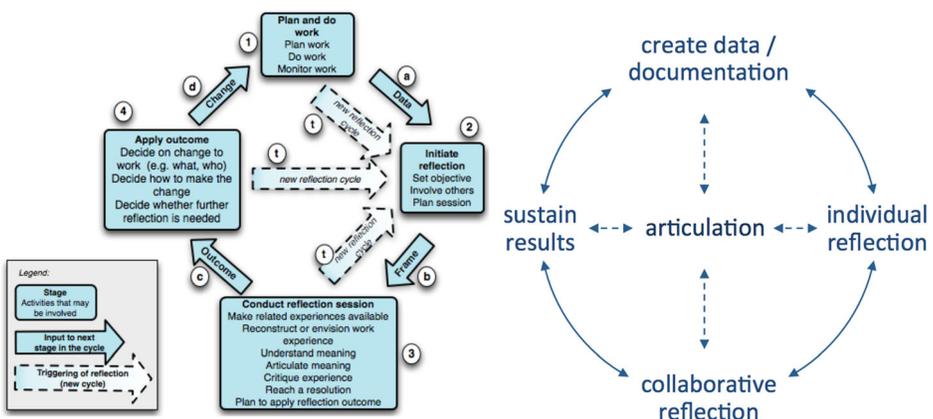


Figure 2.. The Krogstie et al. (2013) CSRL model (left) and the Prilla (2015) collaborative reflection model (right).

data collection and exchange, but leave the process of collaborative reflection to its participants with very little structure or intervention (e.g., Fleck and Fitzpatrick 2009; Marcu et al. 2014; Scott 2010).

2.2. Collaborative reflection and CSCW

Collaborative reflection and its support are important topics for HCI and, more specifically, CSCW. This is reflected in the past discourse on support for collaborative reflection (Fleck and Fitzpatrick 2009; Slovák et al. 2017), the value of collaborative reflection in methods (Bjørn and Boulus 2011), the adaptation of social groups (Convertino et al. 2007) and collaborative reflection as a means to (re-) design socio-technical systems (Prilla et al. 2013; Tang and Chen 2015).

Collaborative reflection is closely linked to established concepts of CSCW and related communities such as sensemaking (Weick 1995), grounding (Clark and Brennan 1991), decision support (Gray et al. 2011) and collaborative problem solving (Roschelle and Teasley 1995), which has been discussed in Prilla et al. (2013). As an excerpt from that discussion that is relevant for the work presented here, we note that collaborative reflection goes beyond concepts such as sensemaking and grounding: As described above, after establishing a common understanding of a topic, reflection is targeted towards learning from this understanding. In addition, collaborative reflection can be seen as a specific means for collaborative problem solving that relies on the perception of all reflection participants and aims to find a solution that stems from their experiences.

Moreover, concepts from CSCW and related disciplines can be used to analyze and support collaborative reflection (see also below). For example, articulation work (Schmidt and Bannon 1992; Suchman 1996) is an important concept for collaborative reflection, which affords explicit statements of experiences and ideas (Prilla et al. 2012). Complementing articulation, reciprocity (Robertson 2002) is key to successful collaborative reflection, as reflection participants need to refer to each other and link to others' statements to arrive at common results (Prilla et al. 2015). Articulation and reciprocity are key concepts for the work presented here, which looks at how reflective discussions in online communities unfold, that is, how people articulate reflective statements and how they relate to each other.

2.3. Collaborative reflection and learning

In our work, we understand learning from or by collaborative reflection as informal learning (Eraut 2004) that happens as part of practice (Schön 1983). This understanding differs from reflection as part of a (formal) education process, in which reflection is a primary process and for which there are well-designed procedures for reflection. In contrast to that, in (work) practice reflection often is a secondary process (Prilla et al. 2015), which is conducted in a rather messy instead of a structured or systematic way (Cressey et al. 2006). As a result, collaborative

reflection at work is a process that is done whenever there is time, and therefore needs opportunities for asynchronously sharing and relating to experiences and ideas.

Learning from collaborative reflection happens in processes in which individuals articulate their experiences and ideas (Järvinen and Poikela 2001), when they link their knowledge to explicated experiences and ideas (Daudelin 1996), and when people draw conclusions from explicated experiences and ideas together (Hoyrup 2004). In the sense of Stahl (Stahl 2000), collaborative reflection can be understood as a process of collaborative knowledge building that consists of individual and collaborative learning activities that complement each other.

Collaborative reflection resembles problem based learning (Barrows 1986; Hmelo-Silver 2004), as both processes rely on experiences in learning (Dewey 1933). Other than focusing solely on problems and how to solve them, (collaborative) reflection is usually triggered by a perceived discrepancy between what was assumed and what happened such as ‘contradicting information, incongruent feelings, interpersonal conflicts and other occurrences during work, leading to a state of discomfort that the individual or group wants to overcome’ (Krogstie et al. 2013, p. 155), which is also called ‘breakdown’ (Baumer 2015) or ‘puzzling’ (Schön 1983). This means that if a situation happening does not match one’s personal expectation of how it should happen, it triggers reflection during which tries to analyze how and what this ‘breakdown’ is and what to learn from it. Schön lists lawyers who reflect during or after a court session on how their strategy has played out as an example (Schön 1983). This means that the space for reflection is wider and less concrete (compared to focusing on a ‘problem’), and that reflection often has to include a phase of sensemaking on what is reflected upon. In addition, other than focusing on ‘solving’ a problem (in which learning is mostly considered a ‘by-product’ (Eraut 2004, p. 250), reflection aims at creating ‘new perspectives’ and a ‘change in behavior’ (Boud et al. 1985, see above) as well as questioning assumptions that led to earlier behavior (Argyris and Schön 1978).

2.4. Support for collaborative reflection: state of the art

Despite the availability of models for reflection as presented above, *supporting* collaborative reflection – that is, according to the models above, outcomes such as new perspectives, changes in behavior and learning that can be applied in practice – is not fully explored yet in terms of what to support and how to support it. This increases the complexity of designing support for such reflection. Some approaches include the use of pictures as memory aids and reflection triggers (Fleck and Fitzpatrick 2010) and generic tools such as shared whiteboards, on which groups may exchange their experience with the help of facilitators (Kim and Lee 2002).

Despite the value of these approaches, they provide triggers or spaces for collaborative reflection, but leave the reflection process to users. For the support of such processes, scholars of collaborative learning have emphasized the importance of understanding discourses in learning and *facilitating* them properly, proposing

support such as guidance (Suthers 2000) or scaffolds (Pea 2004) as well as smooth transitions between them (Dillenbourg et al. 2009). With regard to support for collaborative reflection by facilitation, this is echoed by work on supporting face-to-face reflection groups. Among this work, Daudelin (1996) emphasizes a need for the facilitation of collaborative reflection to structure the process and get the most out of it (see also Cressey et al. 2006) as well as suggesting questions to be asked in face-to-face meetings. Asking the right questions in face-to-face reflection has been reported to help people to articulate experiences to be discussed (Bjørn and Boulus 2011) and to refer to others in reflection (de Groot et al. 2013).

With regard to supporting facilitation of collaborative reflection in online tools, Davis (2000) suggests prompts as means to provoke reflective discussions. Such prompts may be used to periodically remind people of things to reflect on (Isaacs et al. 2013), to question their own thinking (Lin and Lehman 1999), to increase quality and quantity of contributions (Renner et al. 2016), and to structure reflective interaction (Davis 2000). However, despite the potential these approaches may have for reflection, the question of when, how and for what to prompt users of tools for collaborative reflection remains (see Thillmann et al. 2009 for a similar discussion). The work presented in this paper – besides other goals – aims at providing answers to these questions.

2.5. The key to support? Activities and phases of collaborative reflection

General models of reflection as presented above help to understand what reflection is. However, they are not sufficient for the implementation of support of collaborative reflection such as facilitation or guidance, as they do not describe in detail which factors need to be present in collaborative reflection and what communicative activities lead to its success. While, for example, the CSRL model (see Figure 2, left) by Krogstie et al. (2013) contains phases and input factors that have to be met in order to be able to continue the cycle, it falls short in providing directions for the facilitation of reflective conversations. To create such support, there is a need for more knowledge about how collaborative reflection goes on, which specific factors influence collaborative reflection and how the flow of collaborative reflection in online discussions looks like.

Corresponding work on reflection has focused on phases and activities in reflection that may lead to learning outcomes. Among this work, Fleck and Fitzpatrick (2010) describe six core activities of reflection, including returning to experiences, sharing thoughts and offering alternative interpretations. Moon proposes nine stages of reflection including the expression of experiences, the clarification of issues in the experience, reviewing experiences and transforming ideas into actions (Moon 1999). Van Woerkom and Croon (2008), analyzing discussions in face-to-face group settings, identify activities such as objecting to accepted ideas in a group or asking for feedback as decisive for the occurrence of reflective discourse.

In an attempt to analyze online collaborative reflection, de Groot et al. (2013) analyzed types and levels of reflection in different communities, finding that ‘critical opinion sharing’ is crucial for fruitful reflective conversations. Finding that such sharing does not occur often and that there are different factors that influence whether outcomes can be derived from reflective conversations, they state that ‘critically reflective dialogues is not a single concept’ (de Groot et al. 2013, p. 17). This – as indicated in section 2.1 – indicates that there may be different ways in which reflective conversations unfold in communities as opposed to phases or activities that need to happen in certain sequences, but stays on a rather vague level of describing these levels without providing further details.

Our prior work builds on the work described in this section and aims to provide these details, looking at aspects of reflective conversations that co-occur in successful and less successful threads composed from initial statements with follow-up (reflective) answers (Prilla et al. 2015). While this shows that, for example, the exchange of experiences often co-occurs with suggestions based on experiences, it uses complete threads as units of the analysis and therefore does not shed light on how contributions to a discussion thread influence each other, and how sequences of contributions lead to the occurrence of reflection in a thread. Therefore and because of the differences in the literature on this topic laid out above, another aim of the work in this paper is to look at this question and find out how reflective conversations unfold, and if this unfolding is in line with common models of reflection.

3. Analysis of online collaborative reflection content

3.1. Research questions

From the state of the art as described earlier, we can see that despite the amount of valuable research and models on (collaborative) reflection, there is a gap in available knowledge, which is about understanding how reflective conversations unfold in online communities. This gap leads to difficulties in implementing support for the facilitation of collaborative reflection in online tools. Literature offers several models describing phases of collaborative reflection as described above. Using these models, we may assume that collaborative reflection develops along a certain path of articulating experiences, sharing them, discussing them and collaboratively drawing conclusions from this process. However, there is hardly any evidence whether online reflective conversations follow these models. Rather, the diversity inherent to human communication and work by de Groot et al. (2013) question this. AS such, the first set of questions guiding our work focuses on the nature of online collaborative reflection.

RQ 1:How can the nature of online collaborative reflection be described?

RQ 2:(How) Do reflective conversations unfold in communities along the elements mentioned in common models of reflection?

Answering these questions is directed towards two different goals. First these questions aim at putting existing (phase based) models to the test, asking if they can be used to create facilitating support for collaborative reflection. Second, the questions aim at discovering the flow of reflection in practice to inform this support.

In a second step, our work aims to identify the elements of reflective conversations that lead to outcomes from reflection such as new insights and learning (see the goals of reflection as described above). In particular, we are interested in what leads to reflection outcomes such as suggestions and learnings, and what may hinder this. The next set of research question is concerned with this:

RQ 1: Which are the elements of reflection that lead to successful outcomes of reflection?

RQ 2: Which are the elements that diminish reflection in the conversations?

Answering these questions may inform support for the facilitation in that it may enable designers to build features such as prompts (see section 2.2) to foster the occurrence of elements that lead to successful reflection and thereby supporting reflectivity in conversations. In addition, it provides insights on which elements to avoid, thus influencing the design of such features further.

3.2. Two data sets and the corresponding software platforms

Data set M contains four smaller data sets, collected using a tool called “TalkReflection App” that supports individual and collaborative reflection in workplaces. The tool was used in four cases, which belong to the domains of care homes and public administrations. The tool was designed to support collaborative reflection amongst its members (for the following also see (Prilla and Renner 2014)). In order to achieve this, it allowed users to write down their experiences and share them with others (Figure 3, right). While providing initial experiences, users also had the opportunity to provide a first initial reflection (Figure 3, (2)). Others could then comment on the discussion (Figure 3, (3)). Using these features, users could share their experiences (Figure 3, left) and discuss amongst themselves. Additionally, users could share their post with individual users exclusively in case they did not want to share something e.g. with their superior. The tool was not integrated into other software at the workplace of the people reflecting together. Workshops were held to introduce the tools to the users.

Each case had small teams of between 9 and 18 users, and in three of the four cases, the platform was used alongside regular face-to-face meetings. Users mostly knew each other. The tool was used over a time span of 42 to 80 days. We described the data set M in detail in previous work (Prilla et al. 2015; Prilla and Renner 2014).

Data set E was collected in a platform aimed at collaborative reflection and learning (Figure 4). The users of the platform in project E were employees in a public administration organization. Starting with 18 users in the initial launch workshop, over 200 users had registered to the platform after one year. Of these,

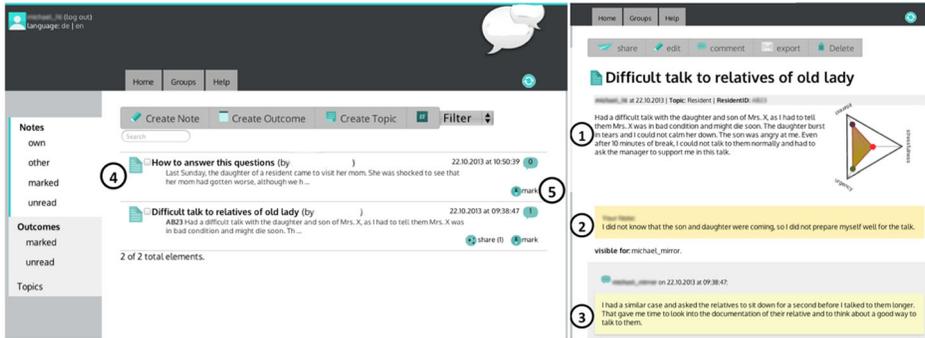


Figure 3.. The TalkReflection App used during the creation of data set M.

data set E contains utterances by the 45 users who contributed content to the platform. The platform was intended to connect a community of workers throughout a small European country and therefore most of the participants did not have face-to-face meetings. Additionally, most users did not know each other before joining the community.

The platform was based on WordPress with plugins for forums (bbPress) and social profiles (buddyPress). Users owned user profiles and could create and discuss in ‘Groups’, which had different privacy levels. Groups were used to organize topics thematically, e.g. by grouping all topics concerning discussions about a specific type of customer. Users could write their own topics (Figure 4, left) within the groups. Users could then normally discuss those topics and share their insights (Figure 4, right).

There are similarities and differences between the users engaged in creating the data sets that we analyze in this paper. Most importantly, data set M was created by small sets of users, in which people usually knew and specifically addressed each other. Data set E stems from the interaction in a larger community, in which some participants did not know each other and the (initial) sharing of experiences was more of a broadcast than directed to certain participants. In both case, users used the respective tools to share experiences on and to solve current issues, but in the participants associated to data set M these issues were more concrete. Two of the groups were dealing with administrative changes in their organizations, while the

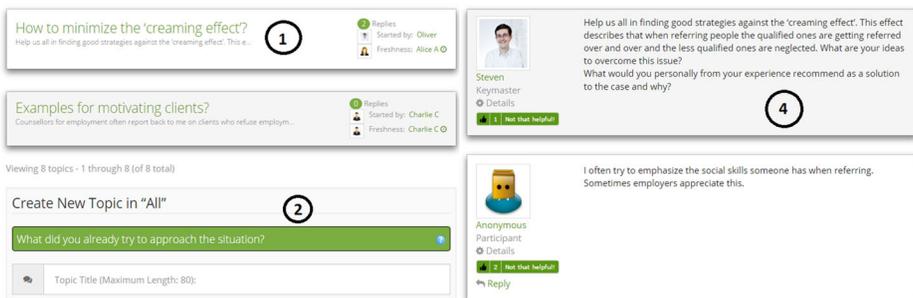


Figure 4.. The community tool used during the creation of data set E.

Table 1. Differences between the users who created the data sets.

Criterion	Data set M	Data set E
Sample size	Four groups, 9–18 users each	45 users active, >200 total
Purpose	Difficult interactions with patients (two groups), organizational change (two groups)	Difficult interactions with clients
Domain	Public administration, medicine/care	Public administration
Time of usage	42–80 days	One year (max, some participants joined during the year)

other two reflected on interactions with relatives of their patients. Users used the platforms to pose questions regarding situations they encounter and how to deal with them. Other colleagues tried to help with adding their experiences or possible solutions (based on their experiences or other knowledge). Users used the tools overwhelmingly for important work related issues and only a tiny fraction of the posts could be considered off topic (e.g. where to meet for after-hour drinks). While the participants associated with data set E share the working domain with two of the groups in data set M (public administration) and the overall topic with the other two groups (difficult social interactions, e.g. with clients), their topic was much wider and therefore less concrete. Moreover, the data set M groups used the tool for a shorter timespan than the users associated to data set E. These differences are summarized in Table 1 and Table 2 will be used for the interpretation of results below.

As the groups stem from different organizations and countries, there are differences between their organizational cultures. In addition, the used tools are similar, but differ slightly. Because of these differences, we do not combine the data sets. We described the domain and factors influencing the adoption of the tool of data set E in earlier work (Blunk and Prilla 2017b).

In our analysis, we use threads as a unit of analysis. In our case threads are discussions consisting of an initial topic and one or more answers. Of course, the author of the initial topic can also post replies within the discussion threads on the community systems. Both data sets include posts (both the initial topic and replies) as their main content as well as meta-data such as who posted what and when. Discussions are plain as there are no hierarchies and they do not contain nested

Table 2. Descriptive statistics for both data sets.

Category	Data Set M	Data Set E
# Users	48	45 (who actively contributed)
# Threads	65	32
# Replies/Thread	2.43	5.2 (only threads with replies)
Longest Thread	6	15 (number of posts)

replies. Although a feature for nested replies was present in the tools, the users did not use it. Both data sets were subject to data cleaning. In all cases, threads which contained non-work-related content were removed from the data set (e.g. agreeing where to meet for after-hour drinks).

3.3. Content analysis

For the analysis of the threads in the data sets described above, we relied on content analysis. Content analysis (De Wever et al. 2006) provide means to gain an understanding of (reflective) learning discourses by providing a ‘theoretical base and the operational translation’ (De Wever et al. 2006, p. 6) of them, and by relating conversation elements to practices of learning (Gee 2004). Content analysis uses two or more coders, who apply a coding scheme that depicts certain elements to be looked at in the analysis. It is a common method of understanding group communication in CSCW (e.g. Newman et al. 1995; Prinz and Zaman 2005) and regarded as the preferred method of analyzing communication and interaction if the amount of material permits manual coding (Introne and Drescher 2013). Using content analysis, we aim at understanding how reflective conversations unfold in online communities, and how tools may be used to facilitate this process.

For the work presented here, we used a content scheme created in earlier work (Prilla et al. 2015), as it was successfully applied for the analysis of collaborative reflection. Besides, the authors are not aware of other schemes for content analysis of collaborative reflection. Tools for automatic content analysis such as LIWC (Tausczik and Pennebaker 2010) and Empath (Fast et al. 2016) do not contain elements specific enough for a detailed analysis of reflection.

The scheme we used comprises the three important steps of sharing experiences, reflecting on them and deriving insights from it as described above. Without going into details, in Table 3 we describe the most relevant elements of this scheme. Please refer to Prilla et al. (2015) and Appendix A for details.

A major downside to content analysis is that it can only capture what has explicitly been stated or written down. From previous work (de Groot et al. 2013; Prilla et al. 2015), we know that this may provide a problem for the analysis of reflection with respect to learning outcomes. In the previous work mentioned above, users reported learning and change as outcomes of reflection support, but these learning outcomes were not explicitly found in the documented discussions.

3.4. Coding procedure

The schemes used for coding data set M and E differed slightly: In data set M we did not differentiate between suggestions based on experience or knowledge as described above, but regarded suggestions based on knowledge as advice (see Table 3). We introduced this differentiation for the analysis of data set E, which was done some time after the analysis of data set M, in order to differentiate plain advice (without

Table 3. Elements of the coding scheme used in our analysis (cf. Prilla et al. 2015).

Code	Description
Emotions (EMO_*)	Emotions help users to understand experiences of others and to relate to them (Boud et al. 1985; Fleck and Fitzpatrick 2010; Moon 1999; Tigelaar et al. 2008). We differentiate between emotions in an experience (provided by the author of a contribution) by using code EMO_OWN and emotions of others towards this experience by using code EMO_OTH .
Interpretation of own experiences (INT)	Some authors provide their initial interpretations while writing the contribution, which may help others to make sense of them (Raelin 2002). We use code INT for these utterances.
Sharing Experiences (EXP) vs. Sharing Knowledge (KNO)	For collaborative reflection, providing experiences needs to be separated from contributions stemming from knowledge, as only the former constitutes a reflection process (Zhu 1996). Therefore, for contributions mentioning experiences code EXP is used, and for contributions relying on knowledge or rules (not explicitly based on personal experiences) we used code KNO .
(Dis) Agreement (AGR*)	Collaborative reflection benefits from challenging or supporting experiences, ideas and suggestions (Raelin 2002; Tigelaar et al. 2008; Fleck and Fitzpatrick 2010). To differentiate agreement from disagreement we used codes AGR and DISAGR .
Providing Advice (ADV) vs. Suggestions (SUG*)	For reflection to be more than ruminating, it is vital to differentiate suggestions without explanation from suggestions based on experiences (Hatton and Smith 1995). There is also a need to differentiate suggestions made based on experiences from those based on knowledge. In the coding scheme, advice articulated without further explanation is marked by code ADV , solutions based on experiences brought are marked by SUG_EXP , and solutions related to knowledge or not related explicitly on experiences are coded with SUG_KNO .
Learning (*_LOOP) and Change (CHANGE)	The goal of reflection is learning from previous experiences, which might include 'a new way of doing something, the clarification of an issue, the development of a skill or the resolution of a problem' (Boud et al. 1985, p. 34). This includes learning from a problem (single loop learning, code S_LOOP) or learning on a more general level (double-loop learning, code D_LOOP) (cf. Argyris and Schön 1978). In addition, it may lead to changes in practices and the work environment, which we code by CHANGE (Daudelin 1996; Moon 1999).

(continued on next page)

Table 3. (continued)

Code	Description
Questions (Q_*)	Questions are important facilitators in collaborative reflection (Woerkom and Croon, 2008; Zhu 1996). Questions posed to request further information from others (e.g., to get a better understanding of the situation) are supposed to increase activity in reflection and are coded with Q_INF . Questions asking for interpretations or viewpoints of others are supposed to increase the quality of outcomes in reflection and are marked by Q_INT .

justification) from suggestions based on knowledge. Therefore, data set E includes all information and codes produced for data set M (the way code **SUG** was used for data set M is the same that **SUG_KNO** was used for data set E), data set E is coded slightly more specific regarding suggestions. However, this does not affect this analysis, as we focus on solutions based on experiences for reflection.

Two researchers used the content coding scheme to code the first data set (data set M) which contained threaded online discussion data in English. Codes were assigned to each contribution, and coders were asked to mark parts of a sentence, full sentences or multiple sentences that led them to applying the code. This means that multiple codes could be assigned to the same unit of coding (the contribution), for example when someone talked about own experiences (**EXP**) while including own emotions (**EMO_OWN**). This was done for two reasons. First, the complex nature of reflection makes it hard for coders to code exactly the same parts of sentences: For example, whether a phrase or a part of a sentence belongs to an experience or not is too subjective to be formalized (see the examples from the data as provided in this paper). This is also not core to our analysis, as we were interested in how contributions of users may influence collaborative reflection rather than in the analysis of concrete statements. Second, there is no need for such formalization or stricter rules for the unit of analysis, as this would mean quantifying reflective utterances based on the number of text fragments coded. This would be arbitrary and not backed by literature.

To reduce subjective interpretations in the coding, the coders employed a strict rule which states that pieces of text (e.g. parts of a sentence) can only be coded with a specific code if the coder can point at the words that lead to a code. This way, for example, only text pieces in which someone explicitly mentioned that they made a specific experience were coded (code **EXP**, referring to own experiences) and other phrases which *sound* like an experience were discarded as explicit references to the experience were missing. Using this rule, phrases like ‘After initially struggling with taking calls I sat down with my manager and talked through the different various different calls we would get and how best to deal with them.’ (data set M) were coded

as an experience (EXP) because of their explicit reference to past activities ('I sat down (...) and talked'). Likewise, phrases like 'I think keeping 'to do lists' for each day is effective then just work through them throughout the day. I would also suggest staying late and working harder.' (data set M) were not coded as related to own experiences (although sounding like it) because there was no explicit mentioning of past activities. This way, we made sure that personal interpretation of possible thoughts and intentions in a piece of text was kept to a minimum.

We used Krippendorff's alpha (Hayes and Krippendorff 2007) to assess intercoder reliability and to ensure a high degree of agreement between both coders. After coding parts of the content, the researchers discussed differences between their codings in order to learn about their respective understanding of the content coding scheme and communicate their perspective. This was repeated after all content had been coded. After discussing, both researchers reviewed their coding again, and changed codes that did not fit the new understanding after the discussions. This is common procedure to ensure intersubjectivity and thus reliability in content coding (e.g., Johri and Yang 2017). For the final coding, Krippendorff's alpha was on average .91 across all codes, with each code being above the threshold of .66, which is proposed by Krippendorff as the minimum acceptable agreement for further analysis (Hayes and Krippendorff 2007).

As data set E contains text written in a foreign language not spoken by the researchers fluently¹, two student researchers, who are native speakers for this language and fluent in English, conducted the content coding based on the content in their native language. This was important to ensure that meanings or statements are not lost in translation. They were already familiar with the process of content coding and the researchers who coded data set M trained them in applying the same content coding scheme used for set M, using samples from that set to ensure the coding process was done in the same fashion. After an extensive explanation of the coding scheme the students coded sample data from set M to train applying the coding scheme. Afterwards their coding was compared to the coding for data set M as described above and differences were discussed. This was repeated until the coding of the students was consistent with the initial coding. In doing so, it was ensured that the understanding of the other coders did not deviate from the understanding of the researchers, and that data sets M and E were coded in a comparable way. Afterwards the students coded data set E following the same process as described for data set M, resulting in Krippendorff's alpha (Hayes and Krippendorff 2007) of 0.67.

It should be noted that we used Krippendorff's alpha (Hayes and Krippendorff 2007) as a measurement of agreement to ensure a high data quality only. For the analysis, we used a data set that included only those codes which were assigned by

¹ See above for reasons of non-disclosing the language.

both coders. On the one hand, this led to excluding data entries but on the other hand to a higher reliability of the final data sets.

4. Results

To be open for all possible ways of how collaborative reflection could unfold in our data, we applied several methods to analyze the data. This included a frequency analysis to calculate how often a given code was present at different stages of a thread, that is, how often a certain code was present in the first, second and other posts in a thread. This was supposed to inform us on the role of specific utterances in earlier or later stages of reflection as predicted by the reflection models presented above. In addition, we looked at the relationship between posts and their predecessors. Taking into account that its immediate predecessor may not only influence a post and to consider different reading styles, we calculated correlations between the occurrences of codes in a post and the codes in all predecessors. Based on the results of these analyses, we looked for causal relationships by using sequential pattern mining and regression models (overview in Appendix B).

4.1. Code frequency: looking for conversation flows

Common reflection models suggest a sequential or at least iterative flow of reflection. Therefore, we should see a distribution of certain codes among the phases that reflect this flow (e.g., experiences should be provided before suggestions based on experiences set in). We therefore started with a frequency analysis, computing how often a given code was present at which position within a post.

In data set M, shown in Table 4, we can see that people are more likely to refer to experiences (EXP) rather than linking to knowledge (KNO). The difference is small, and the numbers fluctuate. Critical discussions (discussions containing disagreements (DISAGR)) are rare as well, and disagreement only tends to happen more often in later stages of discussions, which may be a result of disagreeing with suggestions made. Agreement (AGR) starts slowly in data set M but is building up as a thread progresses. We can see that the amount of agreement stated rises slightly with the amount of answers in a thread. This may be attributed to the small groups responsible for the data: Agreement on experiences, suggestions and other contributions may have led some discussion to become longer, whereas lack of interest or even disagreement may have led people to leave the discussion and thus stop it, given that there were only a few potential participants anyway.

For solutions, we see that advice (ADV) is much more common than suggestions accompanied by reasons (SUG), which is different from data set E (see below). This may be attributed to the fact that in three of four groups of data set M superiors were active in the groups and provided advice. Questions asking for more information are less common than questions asking for opinions or interpretation.

Table 4. Frequency analysis of data set M in percentages. It shows how often a given code was applied in the respective answer number within a thread. Post 1 is the initial topic which starts a thread. In this data set, we did not differentiate between suggestions based on experience or knowledge. The second column shows how often a thread reached a certain length. The last row shows how often the corresponding code was assigned in total. For example, this table shows that the first reply contained experience reports in 22% of the cases. Bold face highlights depict figures characterizing the data or separating it from data set E.

Post (#)	EMO_ OWN in %	EMO_ OTH in %	INT in %	EXP in %	KNO in %	DISAGR in %	AGR in %	ADV in %	SUG in %	S_LOOP in %	D_LOOP in %	CHANGE in %	Q_INF in %	Q_INT in %
1	65	17	18	14	2			2	2	5	5	3	5	22
2	65		9	11	22		8	25	12	2	5	3	11	8
3	50	4	2	14	14		16	24	14	2	6		4	6
4	26			8	12		23	15	12			8	8	23
5	11			9	18		27	27		9	18	9	9	9
6	5			20			20		20		20	20		20
Σ	17		19	24	27	1	23	36	20	6	12	8	15	30

In data set M, double loop learning (D_LOOP) and change are reported more often in data set E, and (like in data set E) occur more often in later stages of the thread as the discussion progresses.

For data set E, results of this analysis are shown in Table 5. Threads are longer than in data set M, which is due to the number of users per group. We can see that (except for the sixth answer (post #7)), experiences are mentioned in roughly 40% of all replies. The numbers do not fluctuate as much as in data M, and the difference between the amounts of occurrences of the codes is bigger. Contrasting this to linking knowledge as a possible solution to someone's issue, the code for linking knowledge was used less often (in around 20% of all replies). We can therefore note that in the discussions in data set E, referencing experiences (EXP) was more common than linking knowledge (KNO).

As agreement (AGR) and disagreement (DISAGR) in data set E are the only codes in the coding scheme to mark interaction within a topic, we can see that there were not many critical discussions: The amount of code for disagreement (DISAGR) is quite low, though increasing as the discussion evolves. Agreement (AGR) appeared more often, indicating that the users engaged in discussion, that the contributions were perceived as helpful, and that users could relate to the discussion. As in data set M, the amount of questions asking for opinions or interpretation (Q_INT) is higher than for questions on further information (Q_INF), and these questions are asked throughout the threads. This suggests that there was interest in continuing the discussion and that these questions may have affected it.

When it comes to solution orientation, users provided little advice (ADV), but rather tried to provide suggestions with reasons (SUG). The high presence of experience-based contributions (EXP) is also visible in the suggestions, as the overall frequency of experience-based suggestions (SUG_EXP) is higher than the occurrence of knowledge-based suggestions (SUG_KNO). What is striking is that the data shows that after the third reply, suggestions based on knowledge vanished, which underpins the focus on experience exchange in data set E. Only a few threads reached a state in which someone indicated that they learned something, and the overall occurrence of questions asking for more information is rather low. However, the codes indicating learning only start to appear after the second answer given (post #3), which indicates that they build on what was exchanged before. In addition, the low number of codes concerning learning does not mean that learning did not happen: The users of the system that we met during the study told us they had taken away various learnings from their interaction with the others in the system, but had not documented it.

Besides enabling a comparison between the data sets, this analysis shows that the elements suggested by reflection literature were present in our data, and that therefore we can call the corresponding discussions reflective.

Table 5. Frequency analysis on data set E shows how often a given code was applied in the respective answer number within a thread for threads, which received at least one reply. All frequencies are percentages. Post 1 is the initial topic, which starts a thread. The second column shows how often a thread reached a certain length. The last row shows how often the corresponding code was assigned in total. The code EMO_OWN is not shown (3% of the third reply contained own emotion). For example, this table shows that the first reply in discussion contained in 43% of the cases codes for experiences and that 35 threads reached this length. Bold face highlights depict figures characterizing the data or separating it from data set M.

Post (#)	EMO_OTH in %	INT in %	EXP in %	KNO in %	DISAGR in %	AGR in %	ADV in %	SUG_EXP in %	SUG_KNO in %	S_LOOP in %	D_LOOP in %	CHANGE in %	Q_INF in %	Q_INT in %
1	3	14	46	23			6	17	3		3		3	60
2		11	43	17	3	26	9	26	14			6	3	14
3	10	17	45	21		28	10	34	14			3		7
4		13	38	29		29		21	13	8	4		4	4
5	5	5	45	14	5	27	18	23		5	5			5
6	5	5	42	11	11	16		32		5	5			11
7			19	13		19	13	19			17	6		19
8	8	8	33	33		33		25					8	8
9		11	44	33	11	33	11	33						11
10	7	14	43	14		14		14		14				14
11	4	25	50			25		25						
Σ	20	26	94	43	5	45	15	54	14	5	6	4	7	46

4.2. Cluster analysis: focusing on types of threads

While the frequency analysis did not show a clear singular flow of collaborative reflection, we found that there was an inner differentiation of threads: For example, we found ‘experience-heavy’ threads, which contained a higher amount of posts including experiences (code EXP from Table 7) than others, and ‘knowledge-heavy’ threads that included more posts based on knowledge (KNO). As described above, reflection literature suggests that the provision of experiences created more reflection than the provision of knowledge, that advice is inferior to suggestions based on experiences when reflection is the goals, and that learning is the positive outcome of successful reflection. Using the relative frequency of codes in a discussion thread, we created clusters in the data for these distinctions to analyze whether reflection went differently in e.g. threads heavily relying on experiences as opposed to threads relying on knowledge provided.

To create the clusters, we computed the relative frequency of a code in a thread by dividing the number of occurrences of the code by the number of its overall occurrences in the data set. We then characterized a thread as ‘heavy’ for a code if the code appeared 1.5 times (using this factor as a heuristic) more often than on average. Note that this way the same thread is assigned to a cluster in each clustering. Based on the different elements present in reflection, we created the following clusterings:

- **Experience-Knowledge Clustering:** Depending on whether the thread was dominated by *Experience* or *Knowledge* the threads were assigned to the corresponding cluster. If neither the codes for experience (EXP) nor knowledge (KNO) were dominant to the extent described above, threads were assigned to an *Undecided* cluster.
- **Advice-Suggestion Clustering:** Threads were put into an *Advice* cluster or in a *Suggestion* cluster (uniting codes SUG_EXP and SUG_KNO in order to compare data sets E and M), and again if neither was overwhelmingly represented, threads were assigned to an *Undecided* cluster.
- **Learning Clustering:** We created two clusters. The first cluster contained all threads that included *Learning* or the willingness to change (uniting codes S_LOOP, D_LOOP, or CHANGE) and another cluster containing all threads not reaching this status (*No-Learning*). Here, we did not use the weighting of code frequency described above, as given the overall numbers of learning related codes found in the data, we assumed the binary distinction whether a thread contains learning to be sufficient to create the clusters.

Within these clusters, we performed a correlation analysis to derive possible indicators of which two codes might have a relationship within the analyzed conversations. An overview of the clusters is shown in Table 6. As can be seen in Table 6, the clusters differed in size. As a result, we excluded the *Advice*, *Knowledge* and *Undecided (Experience-Knowledge)* cluster for data set E in the analysis, as the

Table 6.. Cluster sizes for the different data sets. The table shows the number of threads assigned to each cluster (#T) and the total number of posts in each thread (including the initial post; #P). The clusters Advice, Knowledge, and Undecided (Experience-Knowledge) have been removed due to having a too small size.

Clustering	Cluster	Data set M		Data set E	
		#T	#P	#T	#P
Experience-Knowledge	Experience	12	49	19	133
	Knowledge	10	37	(6)	31
	Undecided	43	137	(7)	38
Advice-Suggestion	Advice	23	74	(3)	23
	Suggestion	18	71	16	118
	Undecided	24	78	13	61
Learning	Learning	18	72	12	109
	No-Learning	47	151	20	93

number of posts assigned to those clusters were too small. Other clusters like the *Knowledge* cluster in data set M are borderline acceptable.

From these numbers in Table 6, we can already derive that in data set E experience exchange dominated whereas the discussions in data set M were not heavily focused on either experience or knowledge.

In terms of whether advice (ADV) or suggestions (SUG) dominated, we can see that the threads in data set M are very balanced, and that in data set E most threads were focused on suggestions. Concerning learning, we can note that roughly half of the posts in data set E belonged to threads that reached a state of learning, whereas in data set M this number is clearly below half.

We performed independent t-tests to compute whether the thread-length varies from cluster to cluster in each set of clusters. Although the averages of thread lengths differ from cluster to cluster, we found that there is no significant difference in terms of thread-length among the clusters in all sets.

4.3. Influencing variables within a thread

We computed correlations between codes in a current post and codes in previous posts to evaluate to what extent certain codes are related to each other. To reflect different reading behaviors, we computed one set of correlations based on codes assigned to a post and *all* previous posts in the same thread, and in another set, we computed only the correlation between the codes in a post in relation to the codes in the *immediate* previous post. This reflects two different assumptions on how people would contribute to a reflective online discussion: the first style encompasses people who read an entire thread before phrasing an answer and the second style describes users who focus on the last post before phrasing an answer. In this section we report correlations between two codes found in *both* analyses, including regression models.

4.3.1. *Procedure of analysis*

Due to the very high number of codes in the data sets, we were very likely to observe correlations in the data sample. We therefore only looked at correlations with an effect size of 0.2 and above. In both data sets this eliminated roughly half of all correlations found. In addition, we removed all correlations of codes without a reasonable explanation of how the respective utterances would influence each other as well. For example, we found a relationship between double loop learning (D_LOOP) and reports about other people's emotions (EMO_OTH). This and other correlations were likely to have occurred by chance and some low numbers of code assignments within the **data** sets (cf. Table 4 and Table 5).

In the following paragraphs, we report on correlations found between codes which were present in multiple clusters, indicating that their correlation might hold true across various circumstances within a thread. This includes correlations from both calculations (all predecessors and only the immediate predecessor). For the sake of simplicity, we report both the weakest and the highest effect size of the correlation in question. Additionally, we conducted linear regression analysis for the strongest correlation of two codes, which is reported together with the correlation analysis. We also removed all correlations which had a low number of occurrences in the data sets. For example, we found a high correlation between challenging suggestions (DISAGR) and agreement (AGR), but discarded it as it occurred only once. For a comprehensive overview of all kept findings refer to the appendix. Note that when we report on the correlation of code A and code B, we mean that B often occurred after A.

4.3.2. *Data set M*

We observed three correlations in multiple clusters showing that mentioning experience (EXP; min $r = 0.247$, $p < 0.001$; max $r = 0.435$, $p = 0.002$; max $r^2 = 0.189$, $F = 10.978$, $p = 0.002$), knowledge (KNO; min $r = 0.211$, $p = 0.001$; max $r = 0.329$, $p = 0.004$; max $r^2 = 0.183$, $F = 10.551$, $p = 0.002$) and double loop learning (D_LOOP; min $r = 0.309$, $p = 0.008$; max $r = 0.557$, $p < 0.001$; max $r^2 = 0.310$, $F = 32.367$, $p < 0.001$) correlate with agreement (AGR). This shows that users engaged with content contributed by others, which is a prerequisite for collaborative reflection to happen in the models presented above. This is supported by the moderate to good (D_LOOP) explanatory power of the regression models we computed. The following sequence from data set M² exemplifies this by showing how one user contributed an experience and the other refers to it with agreement, thus re-assuring the original contributor in their action:

‘(...) I was soo terrified to tell my manager I needed days off for my manager. However, when I told her she was so shocked and happy for me. She spread the

² All examples from data set M are direct quotes containing original orthography.

news to almost the entire office and didnt ask me to make up the time. I guess our managers are cool :)’ (Code EXP)

‘Well done, its good that you told him at least it’ll give him time to prepare and a a less shock lol’ (Code AGR)

Our data indicates that questions may also have an influence on how a discussion thread unfolds. Questions for opinions or interpretation (Q_INT) correlate with the provision of experience, and explain up to 27% of the variance of code for experience in the data (EXP; min $r = 0.233$, $p = 0.04$; max $r = 0.442$, $p < 0.001$; max $r^2 = 0.266$, $F = 12.697$, $p = 0.001$). From this, we may interpret that users answer those questions with a viewpoint based on their own experience. We observed this correlation in the clusters *Learning* and *Undecided* for both the Experience-Knowledge clustering as well as the Advice-Suggestion clustering. This may indicate that answering questions for interpretation with own experiences may be beneficial for thread with regard to learning as an outcomes. A typical example is shown below:

‘How did you find the situation and did it help’ (Code Q_INT)

‘He is always saying that he can cope but he has to realise that he’s not as fit as he were, he will end up having a fall’ (Code EXP)

There are no codes which directly correlate with single codes for learning (S_LOOP, D_LOOP or CHANGE). This may be attributed to the low number of learning codes overall, which (as mentioned in sections 3.3 and 4.1) was also observed in previous studies. Therefore, we decided to focus on analyzing on any kind of learning outcome, taking into account that single and double learning as well as change are all desirable outcomes of collaborative reflection and cannot be preferred over each other. For the corresponding analysis, we computed a new variable as the unification of the three learning variables and called it LEARN. In a correlation analysis, we then found correlations of own emotions (EMO_OWN; $r = 0.414$, $p < 0.01$), experience (EXP; $r = 0.564$, $p < 0.01$) and disagreement (DISAGR; $r = 0.417$, $p < 0.01$) with this variable. A linear regression analysis shows that the occurrence of the code for mentioning own experience (EXP) explains 32% of the variance of LEARN ($r^2 = 0.318$, $F = 102.829$, $p < 0.01$) and that the model gets better when adding the code indicating own emotion (EMO_OWN) and the code for challenging existing suggestions or opinions (DISAGR; $r^2 = 0.412$, $F = 51.065$, $p < 0.01$). This suggests that sharing experiences has a positive impact on learning documented as a result in threads, and that the articulation of own emotions and disagreement amplify this effect.

Interestingly, single loop learning (S_LOOP) and change (CHANGE) correlate with suggestions (SUG; min $r = 0.265$, $p = 0.002$; max $r = 0.360$, $p = 0.002$; max $r^2 = 0.315$, $F = 16.081$, $p < 0.001$) and experience reports (EXP min $r = 0.256$, $p = 0.028$; max $r =$

0.456, $p < 0.001$; max $r^2 = 0.307$, $F = 33.615$, $p < 0.001$) respectively. Those correlations were observed in the clusters based on the immediate predecessor calculation in the clusters for *Learning*, *Suggestion* and *Knowledge*. The corresponding regression models shown that both the code for single loop learning (S_LOOP) and the code for intended or planned change (CHANGE) explain one third of the variance in the occurrence of the code for experience reports (EXP), which supports this interpretation. This may suggest that adding own experiences or knowledge to documented learning helps people to make sense of this learning and potentially adopt it.

4.3.3. Data set E

We observed that advice (ADV; min $r = 0.236$, $p = 0.006$; max $r = 0.552$, $p < 0.001$; max $r^2 = 0.103$, $F = 6.740$, $p = 0.012$) correlates with suggestions based on own knowledge (SUG_KNO), suggesting that users might not be content with plain advice and thus are trying to provide more reasoning. In addition, suggestions following on advice are not based on experience (SUG_EXP) but on knowledge (SUG_KNO), which may indicate that advice leads a thread into a direction not desirable for collaborative reflection. Those correlations were observed mainly in the *Undecided* cluster of the Advice-Suggestion clustering, as well as the *Learning* cluster. It should be noted, however, that corresponding linear regression models stayed at low explanatory power, and that therefore further work is necessary to investigate this correlation. The example³ below illustrates it:

‘Hi. Your prior work is very useful because it makes it easier to make sense of the laws and the actual application of the law. You should of course read everything. Much success and kind regards.’ (Code ADV)

‘Hey there, (...) When it comes to the written part, it would be good to go over the laws at least once, mark the important chapters so that you know exactly where everything is because you will have 4 open-ended questions and will have to describe/write a certain article (...)’ (Code SUG_KNO)

Suggestions based on experiences (SUG_EXP; min $r = 0.220$, $p = 0.011$; max $r = 0.326$, $p = 0.001$; max $r^2 = 0.106$, $F = 12.716$, $p = 0.001$) correlate with consecutive suggestions based on experience in the *Learning* and *Experience* cluster, showing that some threads revolve around multiple suggestions from personal experience. This means that making such suggestions may be a trigger for additional suggestions, leading to a conversation in which people exchange their practices and suggest them to others (see the example below for a typical conversation). The regression models

³ As content from data set E was created in a different language, examples given for data set E are translations of original contributions created for this publication. As such, some details may be missing or changed because of the translation.

we computed, however, remain at low explanatory power, and so further work needs to be done to look at this relation.

'I send the latest information and events at the Employment Service to my clients approximately once a month. Clients welcome this type of information provision and are very satisfied. As recently as this week, a client told me that if she had not received such information, she would have missed out on a lot (...)' (Code SUG_EXP)

'I have created a mailing list using [system A], and [system B] has (finally) added the clients' e-mail addresses. [a number] have subscribed for this mailing list and I now send them e-mails that are formulated better. (...) Every Thursday, I check who has been added to my record anew (...), I add them to the list and notify them of this (and send them the latest news).'

 (Code SUG_EXP)

At the same time, suggestions based on experiences (SUG_EXP; min $r = -0.232$, $p = 0.001$; max $r = -0.477$, $p < 0.001$; max $r^2 = 0.227$, $F = 34.125$, $p < 0.001$) negatively correlate with questions for opinions or interpretations (Q_INT), hinting that those suggestions already provide enough explanation so that asking an additional question as well as further discussion is not needed. Those correlations were observed in the *Learning*, *Suggestion* and *Experience* cluster.

4.4. The Influence of immediate predecessors in threads

The previous analysis covered correlations present in the calculations for all predecessors and for the immediate predecessor, showing what possible influences exist throughout the discussion. Now we focus on specific results for immediate predecessors, which bring forward interesting insights.

4.4.1. Data set M

As we saw earlier, conversations in the online discussions do not stop when someone mentions something which can be coded as learning (S_LOOP, D_LOOP or CHANGE). When referring to only immediate predecessors we can observe that double loop learning (D_LOOP) may have influenced the subsequent post, as in those posts experience reports (EXP; min $r = 0.250$, $p = 0.034$; max $r = 0.296$, $p = 0.039$; max $r^2 = 0.194$, $F = 17.382$, $p < 0.001$), links to knowledge (KNO; min $r = 0.254$, $p < 0.001$; max $r = 0.301$, $p = 0.01$; max $r^2 = 0.091$, $F = 6.985$, $p = 0.01$) and even change (CHANGE; min $r = 0.205$, $p = 0.002$; max $r = 0.388$, $p = 0.001$; max $r^2 = 0.151$, $F = 12.781$, $p = 0.001$) can be identified often. Those correlations were found in the unclustered correlation analysis as well as in the *Learning* and *Advice* cluster. While the relation between learning and change is not very surprising but desirable in reflection, the other two correlations may indicate an interesting pattern, as they suggest that relating a statement about learning to own experiences or knowledge may

ease the adoption of this learning for oneself. The moderate explanatory power of the linear regression models for the code for experience reports (EXP) backs this interpretation up. This finding resembles our findings for the codes for single loop learning (S_LOOP) and planned respectively intended change (CHANGE) in data set M as reported above. The following example from data set M underpins this:

‘It is hard to deal with phone calls. Sometimes it makes me anxious, but the best way I have found out to deal with them is having a list of FAQs and a list of contacts to whom I can transfer the tricky calls... (...)’ (Code D_LOOP)

‘The list of FAQ’s/useful contacts is actually really useful - I can’t think of the amount of times I have picked up a call, then wasted time trying to find the correct procedure - will definitely be using that.’ (Code EXP; Code CHANGE)

4.4.2. Data set E

We can observe feedback sequences of agreement (AGR; min $r = 0.204$, $p = 0.05$; max $r = 0.381$, $p = 0.002$; max $r^2 = 0.145$, $F = 9.998$, $p = 0.002$) correlating with agreement. This shows that in data set E there was a lot of positive reinforcement throughout the discussions. These correlations appeared in the *No-Learning* and both *Undecided* cluster.

‘[name] your advice was appropriate. It is important for her body language to be corresponding – i.e. that the things she is saying are also demonstrated by her posture + open posture, eye contact, just as you advised her. (...)’ (Code AGR)

‘(...) just as we talked the last time, I agree with [name], namely that your advice was appropriate – that her posture is open, eye contact and being at ease (situation permitting). (...)’ (Code AGR)

We also observed one possible influence on learning: Agreement (AGR; min $r = 0.201$, $p = 0.004$; max $r = 0.286$, $p = 0.003$; max $r^2 = 0.154$, $F = 10.747$, $p = 0.002$) correlates with single loop learning (S_LOOP). This happened in the *Learning*, but also in the *Undecided* (Advice-Suggestion) cluster, and especially the explanatory power of the linear regression model for agreement (AGR) suggests that – as trivial as it seems – positive reinforcement may foster learning from collaborative reflection.

4.5. The influence of all predecessors in threads

There was only one correlation to be found only for all predecessor in a thread. In data set E we found that questions for information (Q_INF) negatively correlated with the provision of experience (min $r = -0.284$, $p = 0.003$; max $r = -0.302$, $p = 0.018$; max $r^2 = 0.091$, $F = 5.916$, $p = 0.018$). This suggests that asking for further information hinders

experiences to be articulated, which sounds reasonable: Usually, such further information may clarify issues (probably including experiences), but is usually not made up by experiences. This relationship, however, is only weakly supported by the regression models we computed, and therefore it needs further investigation.

4.6. The influence on answers provided to a thread

In order to evaluate what a topic needs to provide to receive replies, we compared the mean values for codes in the topics that received a reply and the topics that did *not* receive a reply. We found a significant difference in the occurrence of advice (ADV) between the topics with replies ($M = 0.25$, $SD = 0.44$) and those topics without replies ($M = 0.52$, $SD = 0.51$) in an independent-samples t-test ($t(47.575) = 2.105$, $p = 0.041$). This suggests that the more users are including plain advice in the initial topic, the less likely the user is to receive a reply to the discussion. This suggests that providing advice initially is not desirable for online discussions in general and for online collaborative reflection specifically. The independent-sample t-test did not show significant differences for other combinations of codes.

4.7. Pattern mining

In order to look deeper into sequences of code and what they may mean for the flow of collaborative reflection, we performed an analysis with sequential pattern mining algorithms in order to uncover common sequences in users' posting activity that may indicate reflective learning. Using sequential pattern mining – in particular the PrefixSpan (Pei et al. 2004) and SPADE (Zaki 2001) algorithms – we attempted to identify threads consisting of the same patterns, which means that they include causal relationships of types of statements, which could serve as indications of successful reflective learning. However, we did not find any meaningful patterns other than the ones confirming the correlation analysis (presented in section 5). One reason for this may be the level of detail of the coding scheme (many different codes) in combination with the size of the dataset. That is, we may need more data in order to be able to detect meaningful coding sequences. We plan to further investigate on this research line in future work.

5. Discussion

Our analysis of discussion threads in different cases provides initial insights into how reflective discussions unfold in online communities in workplace settings and which elements of discussions foster or hinder collaborative reflection. We are aware of the fact that given the sizes of our data sets and scarce causal relationships, our results cannot be generalized, and that there is a need to further investigate them. However, the results show that there is a diversity in online collaborative reflection that is not predicted by common models, and that reflection is not 'messy' but follows certain paths, which have not been identified and described fully, but characterize

collaborative reflection. Therefore, our work brings forward new insights into how reflection unfolds in online communities. Rather than deriving rules for the flow and support for collaborative reflection from our results, we derive hypotheses and suggestions from our findings, which need to be evaluated in subsequent studies.

5.1. Comparing results from the data sets: the multiple paths of collaborative reflection

When comparing the correlations for both all and immediately preceding comments in the two data sets M and E, they seem to differ from each other very much. Both contain reflective content, but the flow of the threads seems to be different. We attribute this to the user groups and tools having slightly different purposes, different group sizes, slightly different domains, different time spans of usage and possibly also different cultures in the work places (see Table 2). Reflective discussions may unfold differently in different contexts, and potentially support needs to be adapted accordingly. For example, in the smaller groups reflecting together in data set M, people might have put more emphasis in documenting their learnings, as they might have felt they had closer relationships to each other. As another example, the culture of the organization data set E was created in was rather little fault-tolerant, which may have led to less people sharing experiences.

Data set M contains more codes that point to the documentation of learning. In addition, we found relations between the codes related to learning (S_LOOP, D_LOOP, and CHANGE) and follow-up experience- and knowledge exchange. This may be attributed to the fact that in the dataset's four groups, in which people know each other, the participants worked together closely and therefore discussed issues in-depth, even continuing after initial learning success was documented.

In comparison to data set M, data set E contains a lot of suggestions – half of the clusters created are suggestion-heavy, and most correlations found include suggestions based on experiences or knowledge. There was no indication of what may have caused suggestions to appear often in the discussions. This focus on suggestions may be attributed to the size of the group, which was considerably larger than in data set M. More participants mean more potential ideas for dealing with certain challenges. At the same time, in contrast to smaller groups, there is an increased likelihood that an issue shared with 200 (and more) people is received by some people who have already made similar experiences and can provide their solutions. Thus, rather than engaging in sensemaking of the experience, these people may have provided solutions that worked for them right away.

These differences show the diversity that reflection processes may demonstrate in practice, and how group size and setting may influence the overall way of reflecting together. This may be regarded as one of the central findings our work points to: Rather than being messy or strictly adhering to paths shown in models, *collaborative reflection may follow different paths*, which are influenced by factors such as group size, familiarity among purpose of the reflection, cultural context, and others. While

this may not seem surprising on first sight, with the exception of de Groot et al. (2013), it has rarely been discussed in research on collaborative reflection. Our work resonates with de Groot et al. (2013) and adds on it: To the knowledge of the authors, this is the first investigation of online collaborative reflection content that shows that there are multiple paths and presents certain sequences of reflection that describe (parts of) these paths. While our work cannot (and does not aim to) present an exhaustive list of paths that collaborative reflection may take, it shows the need to investigate these paths in future work for the implementation of proper support for collaborative reflection.

5.2. Elements of reflection and their relations

In this section we discuss our findings in more detail to analyze and interpret the different sequences and paths we found.

5.2.1. *Exchange of experience*

In both data sets, we observed that the discussion threads often contain experience reports (and more experience reports than links to knowledge). This is in line with reflection literature (Boud et al. 1985; Schön 1983) and highlights that the threads show experience exchange as part of reflective interaction. We also observed that sets reporting on experiences (EXP) correlate with others agreeing with the perspective articulated in this (AGR) for both data sets, which shows the reciprocity of collaborative reflection as discussed in section 2.2. This suggests that people often found experiences similar to their own in this exchange, and that they related to these. It also indicates that people discussed in a ‘healthy’ environment in which colleagues are supporting each other while talking about issues and ideas. In general, having other people present is helpful to discuss issues and to receive feedback on ideas (Fleck and Fitzpatrick 2010; Raelin 2002).

We found at least two specific roles the exchange of experiences took in the data sets, among which one is in line with what could be expected from the literature, of which one represents an interesting and rather surprising new insight. First, we found that the occurrence of code for experience reports (EXP) explains 32% of the variance of learning outcomes documented in a thread for data set M, with additional codes for personal emotion (EMO_OWN) and disagreement (DISAGR) raising the explanatory power to 41%. This is a good example for the need for articulation of experiences in collaborative reflection, which, in line with reflection literature as described above, suggests that the exchange of experience is a strong factor for the success of collaborative reflection and needs to be supported explicitly. Second, and more surprisingly, we found codes for experience (and knowledge) to occur after learning had been documented in a thread, which is not mentioned in literature on (collaborative) reflection. Looking closer at the respective correlation, it makes sense for collaborative reflection, as relating the learning and change documented by others to one’s own experiences (or knowledge) helps to make sense of the solution, to

decide on its applicability for oneself or to apply it. This could mean that facilitation mechanisms can continue to support users' reflection even after learning was documented and can encourage users to relate their experiences to the learning documented.

5.2.2. *Provision of suggestions based on experiences*

Our analysis on the occurrence of suggestions based on experiences and what may have fostered or caused them did not provide the results we expected. Instead, the interesting finding here lies in what we did not find: Looking at reflection models and reflection literature, we assumed that the provision of suggestions based on experiences (SUG_EXP) should be positively influenced by the provision of experiences (EXP), and that the provision of knowledge (KNO) should have similar effects on the provision of suggestions based on experience (SUG_KNO). However, we did not observe any of these relationships in our coding. Instead, we observed that suggestions are made from the beginning, that is, without any clear preceding type of statement. Additionally, we found that for some threads learning preceded the provision of suggestions based on experience (data set M) and for others, suggestions based on experiences followed each other. This is counterintuitive and surprising, and it suggests that reflection may go different ways than suggested in the literature.

What we may derive from this is that the provision of suggestions may be caused by different elements of a conversation and can happen anywhere in a thread, and that it may be worthwhile to encourage users to provide them from the beginning on. It should be noted, however, that such statements of suggestions based on experiences by definition include statements on own experiences as shown in the examples provided in this paper (see e.g. second example in section 4.3.3 and Appendix Table 7). Nevertheless, for the support of collaborative reflection this means that rather than fostering a process of exchanging experiences and then deriving suggestions from them (as described in almost all models), mechanisms may foster the provision of suggestions based on experiences any time and without necessary preceding elements of the conversation.

5.2.3. *The role of questions: fostering questions for interpretation*

One of the common factors in reflection models are questions to be asked in order to structure, moderate and lead reflection to success. In our coding scheme, we built on Zhu (1996), who found that with respect to their influence on reflection, there is a need to differentiate between questions for further information and questions for interpretation of experiences, as the latter stimulate reflection while the former do not. While we cannot prove this finding, our data suggests that this differentiation is important: In data set M we found correlations between the codes for questions for interpretations (Q_INT) and experience reports (EXP), and for data set E we found negative correlations between the codes for questions for information (Q_INF) and experience reports (EXP). In addition, we found a negative correlation between codes for suggestions based on experiences (SUG_EXP) and questions for more

information (Q_INF). These all suggest that questions for interpretation are helpful for the articulation of experiences, which then supports collaborative reflection. For the support of online collaborative reflection, this may mean that mechanisms such as prompts (see above) may (point users to) ask these questions in order to stimulate collaborative reflection.

5.2.4. *Learning*

We did not observe any codes correlating with individual codes that document learning (S_LOOP, D_LOOP, and CHANGE), which can be attributed to the fact that in our data sets these single learning outcomes were not represented to a large extent (low number of explicit statements on learning as described above). As mentioned above, this has also been observed in other studies (de Groot et al. 2013; Prilla et al. 2015) and does not mean that learning did not take place (see the discussion of this in the explanation of the codes in section 3.3), but that it was not explicitly documented very often. In fact, this was also the case for our groups: As mentioned in section 4.1 our participants told to us that they learned from discussions in the platform, and this indicates that learning was often not documented.

In the analysis of all explicit statements on learning, we found that there was a considerable influence of the articulation of experiences (EXP) on learning documented in the tools, which again shows the need for articulation work in collaborative reflection. On first sight, this seems to fit the reflection models described in this paper nicely, as the articulation of experience is an integral part of reflection. On a second consideration, however, this leaves out steps such as creating suggestions based on the experience exchange as described by most reflection models.

Another interesting finding is that many of the correlations and corresponding regression models we found showed highest effect sizes in the *Learning* cluster created during the analysis – for both data sets. This includes relations between questions for interpretation and the provision of experiences, between learning documented and relating experiences to it and the relation to suggestions from experience to more suggestions of this kind, and it means that these relations were strongest when learning occurred in a thread. This suggests that learning may have been mediated by the occurrence of these code combinations or vice versa. We did not find a mediating effect in the analysis, and, as such, this is another aspect to be taken on board in further work.

We also observed correlations from learning (S_LOOP, D_LOOP, and CHANGE) to codes for experience reports (EXP) or suggestions (SUG), meaning that both experience reports (EXP) and suggestions (SUG_KNO and SUG_EXP) followed the explication of learning. As mentioned earlier, this indicates that facilitation mechanisms should not stop facilitating reflective discussion once someone states that they have learned, as discussions still continue.

5.3. Research questions

The insights discussed above provide a good basis to answer the research questions guiding the work presented here. Below, we relate them to the respective question. RQ 1: *How can the nature of online collaborative reflection be described?*

We did not observe clear or singular reflection flows, the relations we found deviated from what could have been assumed based on existing literature, and different relations were found in the different data sets. As discussed above, this shows that there most likely are multiple paths of collaborative reflection, and that these are potentially influenced by different factors. This opens up a new way to look at the nature of collaborative reflection as neither model-based nor messy, but constituted by multiple paths. For these paths, we found a good amount of causal relationships between types of contributions to collaborative reflection, which may make up elements of the “nature” we looked for in RQ1. This includes both sequences well known from existing models and relationships not commonly featured in existing models (e.g., the effect of relating own experiences to learning documented by others).

Our findings also show the importance of articulation work and reciprocity in collaborative reflection. We found that asking questions and sharing experiences were correlated with documented learning outcomes in a considerable way, which underpins that articulation of issues is crucial for reflecting together and needs to be supported. We also found reinforcement (agreeing when others agreed in data set E, agreeing on issues when experiences were shared in data set M) as an example for reciprocal interaction to be common and of potential importance for collaborative reflection. The important role of reciprocity was also found in multiple suggestions co-occurring (data set E) and statements on change following statements indicating learning had happened (data set M). All of these examples create a situation in which the participants of a collaborative reflection process relate to each other and reflect together rather than sharing thoughts and taking these as input for parallel individual reflections.

Additionally, we have to take into consideration that in online collaborative reflection certain steps may be done by the users of a tool supporting collaborative reflection outside the system, that is, cognitively without being articulated in the tool or in face-to-face interaction. This may explain why the articulation of experiences had a direct influence on learning, why suggestions based on experiences came up without experiences articulated in a thread before, and why some learning outcomes were not documented. If this is the case, then reflection support may even benefit from this, as there would be no need to emphasize or force certain steps but mechanisms could rely on users often fulfilling these steps on their own. In any case, we must regard collaborative reflection in online media as an *online and offline* process.

RQ 1: *(How) Do reflective conversations unfold along the elements mentioned in common models of reflection?*

While the answer to RQ1 is largely based on our insights about the multiple paths that collaborative reflection may take in practice, the answer to RQ2 needs to describe these paths. The multiplicity of sequences we found means that from our work we cannot (and should not) derive a single description of how reflection threads unfold in online communities. Comparing our findings to existing reflection models, the threads we analyzed seem to “jump” certain steps in the models, and other steps occur without the triggers that would be expected.

Our findings include insights on the positive influence of experiences exchange, questions and the provision of suggestions based on experience on collaborative reflection. There are also new insights that add additional facets on collaborative reflection. One of these is that collaborative reflection often continued in our groups after learning had been documented. This may show that, as assumed in previous work (Krogstie et al. 2013), collaborative reflection does not terminate but is an iterative process. As an example of activities most likely performed offline, we found that the provision of experiences correlated with the documentation of learning, implicating that the step of deriving learning insights was done cognitively or face-to-face.

An important aspect to take away from the answers to RQ1 and RQ2 is that reflective conversations may go different ways than suggested by most reflection models, and that reflection may create valuable outcomes on all of these different ways. RQ 1: *Which are the elements of reflection that lead to successful outcomes of reflection?*

Our work suggests that providing experiences influences reflection positively and lead to documented outcomes and suggestions based on experiences. We also found that suggestions based on experiences were followed by more suggestions based on experiences, and that questions for interpretation were helpful for the creation of outcomes. These relations are in line with the literature, but (as discussed above) represent only some findings amongst others.

Besides relations between codes and corresponding utterances that we assumed to be present, we also found surprising relations in the data. Among these, we found that relating own experiences and knowledge to learning documented by others may lead to understanding and adoption of this learning for oneself. This is not prominently featured in reflection models and adds to these models. In addition, this finding also may point to a way in which the provision of knowledge, which is usually referred to as not being helpful for collaborative reflection, may foster learning from collaborative reflection.

We also observed that agreement has various influences on outcomes of collaborative reflection, as among others it is related to suggestions based on knowledge and single loop learning. While this may sound trivial initially, it points towards a culture to be established in online collaborative reflection tools, in which users engage with what others share and reassure them that these experiences have been made by others as well. This culture cannot be taken for granted in many organizations, which is also supported by results we gathered when we applied the tool

resulting in data set E in practice (Blunk and Prilla 2017b). In terms of support, facilitation mechanisms should pick this up and encourage users to agree and disagree with each other to continue the discussion.

Our observation that many of the correlations and models we found showed their highest effect sizes in the *Learning* cluster of threads in the data sets supports the notion that these combinations of codes following each other in threads are helpful to lead to outcomes from collaborative reflection.

RQ 1: *Which are the elements that diminish reflection in the conversations?*

We found minor relations between codes that provide insights on elements or behavior to avoid in collaborative reflection. Among these, we found a weak relationship in which providing advice led to suggestions based on knowledge, which are not favorable in collaborative reflection (as opposed to suggestions based on experience). We also found occasions in which advice or suggestions based on knowledge provided by users led to (single loop) learning, which is also not what is desired in reflection support. It should be noted that both effects may still include valuable insights for users, but that the scope of pure analysis was on fostering learning from reflection.

Moreover, as we observed various correlations occurring in the No-Learning clusters (e.g. multiple consecutive questions for more information (Q_INF)), we may hypothesize an effect of these kinds of questions on threads turning into a direction that does not lead to learning. This is supported by the negative correlation we found between this type of question and the provision of experiences. Facilitation support may therefore try to encourage other types of questions as described below.

5.4. Impacts on modeling reflection

Although reflection literature often implies a sequence of things that need to happen before reflective learning can take place (see Schön 1983; Boud et al. 1985; Krogstie et al. 2013), our data does not support these models fully but shows much more variety and even unexpected relations. For example, from the frequency analysis, we found we can see that both solution proposals as well as learning occurred right from the beginning. Our results therefore suggest that there is not one single way to describe online collaborative reflection (RQ 1 and 2), and that there were both commonly assumed and surprising elements that fostered (and hindered) reflection in the discussions we looked at (RQ 3 and 4).

This affects the way models of collaborative reflection should be built and used. Rather than showing or implicating sequences, such models should instead focus on the collaborative process. This may include how people related to each other (interactivity, reciprocity) and how online collaborative reflection is not only a process of individual and collaborative activity as described in Prilla (Prilla 2015), but also a process that happens online and offline and therefore misses some traces and trajectories in the online medium. Dealing with these gaps in the visibility and

availability of reflection aspects for the group reflecting together means allowing more flexible ways for reflection to unfold (rather than prescribing paths implicated by models) and the need to emphasize the benefit created by sharing these aspects with the other participants in order to allow them to relate to these aspects.

Our findings also suggest that we need to carefully use existing models when analyzing and designing for collaborative reflection. Using these models to explain online collaborative models may automatically result in an incomplete view of what happens in reflective conversations and the question arises whether models can capture collaborative reflection after all, given the diversity we found. This may be the true meaning of reflection being ‘tamed and domesticated at the risk of destroying what it can offer’ as stated by Cressey et al. (2006, p. 23). On the other hand, we must not stop at the notion that reflection is messy, as we have shown that there are multiple paths that it unfolds along. In any case, our work points towards the need to at least partially reconsider these models, especially linking our new findings to them. Further work in this direction, as stated above many times, needs further investigation of our findings and additional data to examine.

5.5. Designing for (collaborative) reflection: implications for facilitation support

In our analysis we gained several insights into how collaborative reflection in online discussion unfolds and which factors might influence others. One implication from our work is that facilitation mechanisms, which have been presented as key to the support for collaborative reflection above, may not have to strictly adhere to specific steps or prerequisites in order to support reflection, but provide freedom for collaborative reflection to unfold along different paths and directly ask users to provide certain contributions. Based on our analysis we derived several suggestions to deal with the variety of ways we found for collaborative reflection:

- Facilitation mechanisms should support users in articulating and sharing their experiences, in relating their statements to each other (reciprocity), and in providing solution proposals based on experience from the beginning of a thread on. Both describe paths that led to outcomes in the threads we analyzed. One way of providing this support could be in prompting users to articulate corresponding contributions. Our initial work on prompts for reflection supports this (Blunk and Prilla 2017a; Renner et al. 2016), but further work is needed to build and evaluate this support.
- Facilitation mechanisms should encourage users to continue reflective discussion even after learning has been documented by referring to their experiences and knowledge to the documented learning outcomes. For the implementation of support this means two things: First, mechanisms need to help users to make documented learning visible to others, as this may help others to benefit from it. Second, promoting mechanisms should encourage users to articulate related experiences in order to help them to learn as well. However, support for collaborative reflection also needs to take into account its nature as an online and

offline process, and therefore we should not force users to explicate everything but allow them to carry out phases in the ways they can reflect best.

- Facilitation mechanisms should foster a culture in which users engage with other users' contributions and explicitly agree or disagree with it in their contributions. Both ways of referring to other contributions were found to be useful in our analysis. This could be done by encouraging and showcasing respective openness and engagement. In addition, prompts may make users aware of the positive aspects of engaging with other content.
- Facilitation mechanisms should encourage and help users to ask questions for interpretation of content shared with them, as in our analysis this showed to support reflection. One way to gently direct online discussions in directions that foster collaborative reflection is to provide users with blueprints for such questions, which they can adopt and integrate into their contributions. We have implemented a prototype for this, and at the time of writing this paper, we are evaluating it (see Blunk and Prilla 2017b for examples and very early results).
- Facilitation mechanisms should adapt to the specific context of the collaborative reflection and to the user(s). Enabling users to take multiple paths in collaborative reflection needs context-dependent mechanisms that provide support for the respective reflection path a thread follows or should follow according to its context or users. It may also afford the personalization of facilitation, that is, providing the right kind of support for a certain user. However, whether and how the success of certain facilitation support can be related to the situation or the person receiving the support is subject to further work. Our work suggests that this is a path worthwhile following.

We also noticed that the data sets differed in the influences we found on collaborative reflection. We attributed this to the difference in the contexts the data stems from. For the design of reflection support, this means that we need to understand the context (e.g., small vs. large groups, short or long-term exposure etc.), what it means for the flow of reflection, and tailor support to it. Following up on our discussion above, in smaller groups we may have to prompt for activity that includes sensemaking such as the provision of additional experiences, while in larger groups we may directly ask for solution proposals. However, while our work points at differences in the context, it can only provide initial insights and further work is needed to explore this influence.

For the implementation of context-dependent support, there is a need to gain an understanding of what is happening in a collaborative reflection thread. The manual content coding approach that we employed to analyze the data, while being well-suited for this analysis, does not work for that. Instead, there is a need for on-the-spot detection of the situation in thread. Numerous tools for automated content analysis such as LIWC (Tausczik and Pennebaker 2010) and EMPATH (Fast et al. 2016) exist. However, as stated above, these tools do not include elements needed to analyze collaborative reflection. Adding these elements based on the work presented

here and other work as well as using their existing elements to analyze, for example, the topics reflected upon could enable us to automatically assess and categorize the content. In this case, facilitation features could be selected based on what may be most helpful in a thread. This, however, needs additional work, as none of the existing tools contains the categories included in the coding scheme we used.

5.6. Limitations

Our work brings forward interesting insights on how collaborative reflection unfolds in online communities, and how to potentially support it. However, we are aware of the fact that our findings cannot be generalized. We discuss limitations of our work below.

Our findings are based on sequence analyses and correlations backed up by corresponding models from linear regression analysis. While this is sufficient to draw the tentative conclusions we draw in the paper (e.g., multiple paths of reflection that are potentially influenced by contextual factors), we did not find particular patterns, and the explanatory power of some models is low. We mentioned this explicitly in our results, and we marked those relations between contributions to collaborative reflection that need further investigation. However, despite this limitation, our work clearly shows the plurality of paths that collaborative reflection may take, and it indicates that we may need to support some of these different paths rather than following certain models.

In addition, given the amount of data, we were almost guaranteed to observe several correlations. Thus, we focused on correlations appearing multiple times (that is, in the overall data sets and in the clusters we created) as well as on correlations that we found acceptable regression models for. Despite this, we emphasize that further work is needed to approve and strengthen our findings, while we also emphasize that these findings provide new insights on their own.

We also applied the content analysis to content written in a language the authors do not speak. The coding scheme by Prilla et al. (2015) used was developed and evaluated for the analysis of German and English texts, and therefore it is not guaranteed that it picks up all language and culture related subtleties of this language. In addition, coding needed to be done by different coders for the data sets due to this language barrier. We accounted for these issues by training the coders for data set E with the coding applied to data set M by the researchers, making sure there was a higher interrater reliability between the coders and the researchers in order to make sure that the coders' understanding of the coding scheme was the same than that of the researchers. This way we could make sure that both groups of coders had a comparable understanding of the coding scheme.

6. Conclusion

This paper provides an analysis of reflective conversations from two online tools gathered in different workplace settings. The analysis is based on a manual content coding scheme developed for the investigation of collaborative reflection. We

employed different methods of analysis to provide answers to the four research questions guiding our work. From this analysis, we derived a number of insights into the structures of those discussions, among which some support existing knowledge on collaborative reflection, while many others question this knowledge and add to it. Most importantly, we found that collaborative reflection unfolds along multiple paths, which are likely to be influenced by the context reflection is conducted in. While this may seem intuitive or even trivial, there is hardly any work that points into this direction and no work that identifies and describes the paths, which is done in the work presented here. In addition, our results indicate that experiences and suggestions based on experiences lead to outcomes of reflection, and that engaging with contributions of others and asking the right questions fosters this process as well. While this could have been assumed from the literature, we also found that reflection may be successful on many more different paths than discussed in the literature. Two examples for this are that providing experiences after learning has been documented may be important for learning from collaborative reflection, and that suggestions based on own experiences can be provided without prerequisites or other content present. Based on these insights we provide several suggestions on how reflection could be supported by facilitation mechanisms, including a discussion how to apply these mechanisms and further work needed for this.

While we are aware of the limitations of our work, our results provide interesting and novel insights into the course of collaborative reflection, and the insights from our analysis can be used to inspire the design of means to facilitate and support reflection.

Acknowledgments

Open Access funding provided by Projekt DEAL. This work was part of the EmployID (<http://employid.eu>) project on “Scalable & cost-effective facilitation of professional identity transformation in public employment services” supported by the EC in FP 7 (project no. 619619). We thank all colleagues and associated partners for their cooperation and our fruitful discussions. Special thanks to our student researchers Nika and Neža.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

1. Appendix

1.1. Appendix A: Content coding scheme

Table 7. Content coding scheme used in the analysis (adapted from (Prilla et al. 2015)). To improve readability we substituted the original code numbers with labels depicting the meaning of the codes.

Code	Description (adapted from (Prilla et al. 2015))
EMO_OWN	Mentioning own emotions in an experience (e.g. ‘Was not fun man’ or ‘this made me angry’) with an explicit emotion-keyword being present (e.g. ‘happy’, ‘sad’, ‘frustrated’, ‘angry’).
EMO_OTH	Mentioning emotions of others in an experience (e.g. ‘[resident] said he is unhappy living here’) with an explicit emotion-keyword being present (e.g. ‘happy’, ‘sad’, ‘frustrated’, ‘angry’).
INT	Interpretation or justification of actions and situations (e.g. ‘As far as I am aware I had done nothing to deserve this’). This code is assigned when the thread author adds interpretation or justification. This includes assessments of the situation (e.g., explanations why the situation is problematic or relevant for work going beyond only describing the problem), as well as hypotheses for problems / success.
EXP	Linking an experience to other experiences (e.g. ‘I made a similar experience’, or ‘X told me he was through this before ...’). Explicit reference to past experience needed, otherwise a different code is used (e.g. KNO).
KNO	Linking an experience to knowledge, rules or values not explicitly linked to a particular experience (e.g. ‘I was reading an article (guardian newspaper) on do’s and don’ts in an interview. One of a don’t was, not to keep looking in the interviewees eyes but look at their face’).
DISAGR	Responding to interpretation of the action by challenging existing interpretation(s)/ suggestions or adding perspectives (e.g. ‘Hmmm. Is this really different from [...]’).
AGR	Responding to interpretation of the action by supporting interpretation(s) and suggestions (e.g. ‘Agreed!’). This includes also authors repeating previous suggestions.
ADV	Giving advice without a reason or reference to experiences (e.g. ‘Never accept blame for another’s mistake’).
SUG ^a	Giving possible solutions with a reason given (see examples for codes SUG_EXP, SUG_KNO).
SUG_EXP ^b	Giving possible solutions with a reason given based on experience (e.g. ‘from my experience a list of FAQs is useful’). When this code is assigned, code EXP is assigned to the experience part of the solution, as the author also makes a reference to her personal experiences.
SUG_KNO2	Giving possible solutions with a reason given based on knowledge (e.g. ‘It is good to do X, because it helps to...’). When this code is assigned, code KNO is assigned to the knowledge part of the solution proposal, as the author also makes a reference to her knowledge.

(continued on next page)

Table 7. (continued)

Code	Description (adapted from (Prilla et al. 2015))
S_LOOP	Insights from reflection as single-loop learning: Different or better understanding of experiences . Expressed by reporting insights (e.g. ‘I realised that I shouldn’t have been so worried about this [...]’ or ‘It is good to know that I personally haven’t done something wrong’). This code is only assigned if the learning resulted from the current reflection process within the discussion, and statements about prior learning are not coded.
D_LOOP	Insights from reflection as double-loop learning: Generalizing from experiences . Expressed by patterns or roots of a problem (e.g. ‘The best way I found to deal with this is [...]’). This code is only assigned if the learning resulted from the current reflection process within the discussion, and statements about prior learning are not coded.
CHANGE	Drawing conclusions and implications from reflection by suggesting to apply new practice or reporting on changes <i>done</i> or <i>planned</i> , like how to implement a change (e.g. ‘Will definitely try and do [...] in the future’).
Q_INF	Questions for information on the experience , if not all information is given or additional information is supposed to be available (e.g. ‘what do you mean by [...]’ or ‘what happened afterwards’).
Q_INT	Questions triggering discussions , asking for an <i>interpretation</i> of a situation, for opinions or proposals (e.g. ‘what do people think about [...]’). This code is also assigned for questions aimed to start discussions or keep them running (e.g. ‘what do people think about [...]’ or ‘how can we solve this problem?’)

^aCode SUG was used in data set M only

^bCodes SUG_EXP and SUG_KNO were used in data set E only

1.2. Appendix B: Data from the analysis

Table 8. Overview of all reported findings across both data sets. Each row shows one a possible influence of the code in all of the or the immediate predecessor post and the subsequent post together with the highest observed correlation and the result of the regression analysis for that specific pair. The regression analysis has been performed on the cluster containing the highest correlation.

Predecessor	Current code	Correlation	Regression
Data Set M – All and Immediate Predecessors			
CHANGE	EXP	$r = 0.554, p < 0.001$	$r^2 = 0.307, F = 33.615, p < 0.001$
Q_INT	EXP	$r = 0.516, p = 0.001$	$r^2 = 0.266, F = 12.697, p = 0.001$
EXP	AGR	$r = 0.435, p = 0.002$	$r^2 = 0.189, F = 10.978, p = 0.002$
KNO	AGR	$r = 0.428, p = 0.002$	$r^2 = 0.183, F = 10.551, p = 0.002$
D_LOOP	AGR	$r = 0.557, p < 0.001$	$r^2 = 0.310, F = 32.367, p < 0.001$
S_LOOP	SUG	$r = 0.561, p < 0.001$	$r^2 = 0.315, F = 16.081, p < 0.001$
Q_INF	Q_INF	$r = 0.703, p < 0.001$	$r^2 = 0.083, F = 6.910, p = 0.01$
Data Set E – All and Immediate Predecessors			

(continued on next page)

Table 8. (continued)

Predecessor	Current code	Correlation	Regression
SUG_EXP	SUG_EXP	$r = 0.326, p = 0.001$	$r^2 = 0.106, F = 12.716, p = 0.001$
ADV	SUG_KNO	$r = 0.552, p < 0.001$	$r^2 = 0.103, F = 6.740, p = 0.012$
SUG_EXP	- Q_INT	$r = -0.477, p < 0.001$	$r^2 = 0.227, F = 34.125, p < 0.001$
Data Set M – Only Immediate Predecessors			
ADV	EXP	$r = 0.561, p < 0.001$	$r^2 = 0.315, F = 33.081, p < 0.001$
D_LOOP	EXP	$r = 0.441, p < 0.001$	$r^2 = 0.194, F = 17.382, p < 0.001$
D_LOOP	KNO	$r = 0.301, p = 0.01$	$r^2 = 0.091, F = 6.985, p = 0.010$
D_LOOP	CHANGE	$r = 0.388, p = 0.001$	$r^2 = 0.151, F = 12.781, p = 0.001$
Data Set E – Only Immediate Predecessors			
AGR	AGR	$r = 0.381, p = 0.002$	$r^2 = 0.145, F = 9.998, p = 0.002$
AGR	S_LOOP	$r = 0.393, p = 0.002$	$r^2 = 0.154, F = 10.747, p = 0.002$
S_LOOP	D_LOOP	$r = 0.338, p < 0.001$	$r^2 = 0.114, F = 14.981, p < 0.001$
Data Set E – All Predecessors			
Q_INF	- EXP	$r = -0.302, p = 0.018$	$r^2 = 0.091, F = 5.916, p = 0.018$

References

- Argyris, Chris; and Donald A. Schön (1978). *Organizational learning: A Theory of Action Perspective*. Reading, Mass.: Addison-Wesley Pub. Co.
- Barrows, Howard S. (1986). A taxonomy of problem-based learning methods. *Medical Education*, vol. 20, no. 6, November 1986, pp. 481–486.
- Baumer, Eric P. S. (2015). Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. *CHI'15. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, 18 - 23 April 2015*. New York, NY, USA: Association for Computing Machinery, pp. 585–594.
- Bjørn, Pemille; and Nina Boulus (2011). Dissenting in reflective conversations: Critical components of doing action research. *Action Research*, vol. 9, no. 3, September 2011, pp. 282–302.
- Blunk, Oliver; and Michael Prilla (2017a). Prompting to Support Reflection: A Workplace Study. In É. Lavoué, H. Drachler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (eds.): *ECTEL'17. Data Driven Approaches in Digital Education, Tallinn, Estonia, 12 - 15 September 2017*. Cham: Springer, pp. 367–372.
- Blunk, Oliver; and Michael Prilla (2017b). Supporting Communities of Practice in Public Administrations: Factors Influencing Adoption and Readiness. *C&T'17. Proceedings of the 8th International Conference on Communities and Technologies, Troyes, France, 26 - 30 June 2017*. New York, NY, USA: ACM, pp. 36–45.
- Boud, David; Rosemary Keogh; and David Walker (1985). Promoting reflection in learning: A model. In D. Boud, R. Keogh, & D. Walker (eds.): *Reflection: Turning Experience Into Learning*. London, New York: RoutledgeFalmer, Taylor & Francis Group, pp. 18–40.
- Clark, Herbert; and Susan E. Brennan (1991). Grounding in Communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (eds.): *Perspectives on Socially Shared Cognition*. Washington, DC: American Psychological Association, pp. 127–149.
- Convertino, Gregorio; Thomas P. Moran; and Barton A. Smith (2007). Studying activity patterns in CSCW. *CHI'07. Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, USA, 28 April - 3 May 2007*. New York, NY, USA: Association for Computing Machinery, pp. 2339–2344.

- Cressey, Peter; David Boud; and Peter Docherty (2006). The emergence of productive reflection. In D. Boud, P. Cressey, & P. Docherty (eds.): *Productive reflection at work: Learning for changing organizations*. London: Routledge, pp. 11–26.
- Daudelin, Marilyn Wood (1996). Learning from experience through reflection. *Organizational Dynamics*, vol. 24, no. 3, Winter 1996, pp. 36–48.
- Davis, Elizabeth A. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education*, vol. 22, no. 8, August 2000, pp. 819–837.
- de Groot, Esther; Maaikje D. Endedijk; A. Debbie C. Jaarsma; P. Robert-Jan Simons; and Peter van Beukelen (2013). Critically reflective dialogues in learning communities of professionals. *Studies in Continuing Education*, vol. 36, no. 1, May 2013, pp. 15–37.
- De Wever, Bram; Tammy Schellens; Martin Valcke; and Hilde Van Keer (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, vol. 46, no. 1, January 2006, pp. 6–28.
- Dewey, John (1933). *How we think: A restatement of the relation of reflective thinking in the educative process*. Chicago: Henry Regnery.
- Dillenbourg, Pierre; Sanna Järvelä; and Frank Fischer (2009). The Evolution of Research on Computer-Supported Collaborative Learning. In D. N. Balacheff, D. S. Ludvigsen, D. T. de Jong, D. A. Lazonder, & D. S. Barnes (eds.): *Technology-Enhanced Learning*. Dordrecht: Springer Netherlands, pp. 3–19.
- Dyke, Martin (2006). The role of the 'Other' in reflection, knowledge formation and action in a late modernity. *International Journal of Lifelong Education*, vol. 25, no. 2, March 2006, pp. 105–123.
- Eraut, Michael (2004). Informal learning in the workplace. *Studies in Continuing Education*, vol. 26, no. 2, July 2004, pp. 247–273.
- Fast, Ethan; Binbin Chen; and Michael S. Bernstein (2016). Empath: Understanding Topic Signals in Large-Scale Text. *CHI'16. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7 - 12 Mai 2016*. New York, NY, USA: Association for Computing Machinery, pp. 4647–4657.
- Fessl, Angela; Oliver Blunk; Michael Prilla; and Viktoria Pammer (2017). The known Universe of Reflection Guidance: A Literature Review. *International Journal of Technology Enhanced Learning*, vol. 9, no. 2–3, January 2017, pp. 103–125.
- Fleck, Rowanne; and Geraldine Fitzpatrick (2009). Teachers' and tutors' social reflection around SenseCam images. *International Journal of Human-Computer Studies*, vol. 67, no. 12, December 2009, pp. 1024–1036.
- Fleck, Rowanne; and Geraldine Fitzpatrick (2010). Reflecting on reflection: Framing a design landscape. *OZCHI'10. Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction, Brisbane, Australia, 22 - 26 November 2010*. New York, NY, USA: Association for Computing Machinery, pp. 216–223.
- Gee, James Paul (2004). Discourse Analysis: What Makes It Critical? In R. Rogers (ed.): *An Introduction to Critical Discourse Analysis in Education*. New York, NY, USA: Routledge, pp. 23–46.
- Gray, Paul; Bob Johansen; Jay Nunamaker; Jeff Rodman; and Gerald R. Wagner (2011). GDSS Past, Present, and Future. In D. Schuff, D. Paradice, F. Burstein, D. J. Power, & R. Sharda (eds.): *Decision Support: An Examination of the DSS Discipline*. New York, NY: Springer, pp. 1–24.
- Hatton, Neville; and David Smith (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, vol. 11, no. 1, January 1995, pp. 33–49.
- Hayes, Andrew F.; and Klaus Krippendorff (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, vol. 1, no. 1, April 2007, pp. 77–89.
- Hmelo-Silver, Cindy E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, vol. 16, no. 3, September 2004, pp. 235–266.

- Hostrup, Steen (2004). Reflection as a core process in organisational learning. *Journal of Workplace Learning*, vol. 16, no. 8, December 2004, pp. 442–454.
- Introne, Joshua E.; and Marcus Drescher (2013). Analyzing the flow of knowledge in computer mediated teams. *CSCW'13. Proceedings of the 2013 conference on Computer supported cooperative work, San Antonio, Texas, USA, 23 - 27 February 2013*. New York, NY, USA: Association for Computing Machinery, pp. 341–356.
- Isaacs, Ellen; Artie Konrad; Alan Walendowski; Thomas Lennig; Victoria Hollis; and Steve Whittaker (2013). Echoes from the Past: How Technology Mediated Reflection Improves Well-being. *CHI'13. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April - 2 May 2013*. New York, NY, USA: ACM, pp. 1071–1080.
- Järvinen, Annikki; and Esa Poikela (2001). Modelling reflective and contextual learning at work. *Journal of Workplace Learning*, vol. 13, no. 7–8, October 2001, pp. 282–289.
- Johri, Aditya; and Seungwon Yang (2017). Scaffolded Help for Learning: How Experts Collaboratively Support Newcomer Participation in Online Communities. *C&T'17. Proceedings of the 8th International Conference on Communities and Technologies, Troyes, France, 26 - 30 June 2017*. New York, NY, USA: ACM, pp. 149–158.
- Kim, Dongsik; and Seunghee Lee (2002). Designing Collaborative Reflection Supporting Tools in e-Project-Based Learning Environments. *Journal of Interactive Learning Research*, vol. 13, no. 4, 2002, pp. 375–392.
- Kolb, David A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Krogstie, Birgit; Michael Prilla; and Viktoria Pammer (2013). Understanding and Supporting Reflective Learning Processes in the Workplace: The CSRL Model. In D. Hernández-Leo, T. Ley, R. Klamma, & A. Harrer (eds.): *ECTEL'13, Scaling up Learning for Sustained Impact, Paphos, Cyprus, Greece, 17 - 21 September 2013*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 151–164.
- Lin, Xiaodong; and James D. Lehman (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, vol. 36, no. 7, August 1999, pp. 837–858.
- Marcu, Gabriela; Anind K. Dey; and Sara Kiesler (2014). Designing for collaborative reflection. *PervasiveHealth '14. Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare, Oldenburg, Germany, 20 - 23 May 2014*. Brussels, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), pp. 9–16.
- Mercer, Neil; and Rupert Wegerif (1999). Is “exploratory talk” productive talk? In K. Littleton & P. Light (eds.): *Learning with computers: Analysing productive interaction*. London, New York: Routledge, pp. 79–101.
- Moon, Jennifer A. (1999). *Reflection in learning and professional development: Theory and practice*. London: Kogan Page.
- Newman, David R.; Brian Webb; and Clive Cochrane (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology*, vol. 3, no. 2, April 1995, pp. 56–77.
- Pea, Roy D. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *Journal of the Learning Sciences*, vol. 13, no. 3, July 2004, pp. 423–451.
- Pei, Jian; Jiawei Han; B. Mortazavi-Asl; Jianyong Wang; H. Pinto; Qiming Chen; U. Dayal; and Mei-Chun Hsu (2004). Mining sequential patterns by pattern-growth: The PrefixSpan approach. *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 11, November 2004, pp. 1424–1440.

- Prilla, Michael (2015). Supporting Collaborative Reflection at Work: A Socio-Technical Analysis. *AIS Transactions on Human-Computer Interaction*, vol. 7, no. 1, January 2015, pp. 1–17.
- Prilla, Michael; and Bettina Renner (2014). Supporting Collaborative Reflection at Work: A Comparative Case Analysis. *GROUP'14. Proceedings of the 18th International Conference on Supporting Group Work, Sanibel Island, Florida, USA, 9 - 12 November 2014*. New York, NY, USA: Association for Computing Machinery, pp. 182–193.
- Prilla, Michael; Martin Degeling; and Thomas Herrmann (2012). Collaborative Reflection at Work: Supporting Informal Learning at a Healthcare Workplace. *GROUP'12. Proceedings of the 17th ACM international conference on Supporting group work, Sanibel Island, Florida, USA, 27 - 31 October 2012*. New York, NY, USA: Association for Computing Machinery, pp. 55–64.
- Prilla, Michael; Viktoria Pammer; and Birgit Krogstie (2013). Fostering Collaborative Redesign of Work Practice: Challenges for Tools Supporting Reflection at Work. In O. W. Bertelsen, L. Ciolfi, M. A. Grasso, & G. A. Papadopoulos (eds.): *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, Paphos, Cyprus, Greece, 21 - 25 September 2013*. London: Springer, pp. 249–268.
- Prilla, Michael; Alexander Nolte; Oliver Blunk; Dennis Liedtke; and Bettina Renner (2015). Analyzing Collaborative Reflection Support: A Content Analysis Approach. In N. Boulus-Rødje, G. Ellingsen, T. Bratteteig, M. Aanestad, & P. Bjørn (eds.): *ECSCW '15. Proceedings of the 14th European Conference on Computer Supported Cooperative Work, Oslo, Norway, 19 - 23 September 2015*. Cham: Springer International Publishing, pp. 123–142.
- Prinz, Wolfgang; and Baber Zaman (2005). Proactive support for the organization of shared workspaces using activity patterns and content analysis. *GROUP'05. Proceedings of the 2005 international ACM SIGGROUP conference on supporting group work, Sanibel Island, Florida, USA, 6 - 9 November 2005*. Sanibel Island, Florida, USA: Association for Computing Machinery, pp. 246–255.
- Raelin, Joseph A. (2002). “I Don’t Have Time to Think!” (vs. The Art of Reflective Practice). *Reflections*, vol. 4, no. 1, Fall 2002, pp. 66–79.
- Renner, Bettina; Michael Prilla; Ulrike Cress; and Joachim Kimmerle (2016). Effects of Prompting in Reflective Learning Tools: Findings from Experimental Field, Lab, and Online Studies. *Frontiers in Psychology*, vol. 7, no. 820, May 2016, pp. 1–9.
- Robertson, Toni (2002). The Public Availability of Actions and Artefacts. *Computer Supported Cooperative Work (CSCW)*, vol. 11, no. 3, September 2002, pp. 299–316.
- Roschelle, Jeremy; and Stephanie D. Teasley (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In C. O’Malley (ed.): *Proceedings from the NATO Advanced Research Workshop on Computer Supported Collaborative Learning, Acquafredda di Maratea, Italy, 24 - 28 September 1989*. Springer Berlin Heidelberg, pp. 69–97.
- Schmidt, Kjeld; and Liam Bannon (1992). Taking CSCW Seriously: Supporting Articulation Work. *Computer Supported Cooperative Work (CSCW)*, vol. 1, no. 1–2, March 1992, pp. 7–40.
- Schön, Donald A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- Scott, Susanne G. (2010). Enhancing Reflection Skills Through Learning Portfolios: An Empirical Test. *Journal of Management Education*, vol. 34, no. 3, June 2010, pp. 430–457.
- Slovák, Petr; Christopher Frauenberger; and Geraldine Fitzpatrick (2017). Reflective Practicum: A Framework of Sensitising Concepts to Design for Transformative Reflection. *CHI'17. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6 - 11 May 2017*. New York, NY, USA: ACM, pp. 2696–2707.
- Stahl, Gerry (2000). A model of collaborative knowledge-building. In B. J. Fishman & S. F. O’Connor-Divelbiss (eds.): *ICLS 2000. Proceedings of the Fourth International Conference of the Learning Sciences, Ann Arbor, Michigan, USA, 14 - 17 June 2000*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, pp. 70–77.

- Suchman, Lucy (1996). Supporting articulation work. In R. Kling (ed.): *Computerization and controversy: Value conflicts and social choices*. San Diego: Academic Press, 2nd ed, pp. 407–423.
- Suthers, Daniel D. (2000). Initial evidence for representational guidance of learning discourse. *ICCE'02. Proceedings of International Conference on Computers in Education, Taipei, Taiwan, 21 - 24 November 2000*. Taipei, Taiwan, pp. 21–24.
- Tang, Charlotte; and Yunan Chen (2015). Technology Readiness through Collaborative Reflection: A Qualitative Case Study of an Under-resourced Nurse-Managed Clinic. In T. X. Bui & R. H. Sprague (eds.): *HICCS 2015. Proceedings of the 48th Annual Hawaii International Conference on System Sciences, Kauai, HI, USA, 5 - 8 January 2015*. Los Alamitos, California ; Washington ; Tokyo: IEEE Computer Society, pp. 3013–3022.
- Tausczik, Yla R.; and James W. Pennebaker (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, vol. 29, no. 1, March 2010, pp. 24–54.
- Thillmann, Hubertina; Josef Künsting; Joachim Wirth; and Detlev Leutner (2009). Is it Merely a Question of “What” to Prompt or Also “When” to Prompt?: The Role of Point of Presentation Time of Prompts in Self-Regulated Learning. *Zeitschrift Für Pädagogische Psychologie*, vol. 23, no. 2, January 2009, pp. 105–115.
- Tigelaar, Dineke E. H.; Diana H. J. M. Dolmans; Paulien C. Meijer; Willem S. de Grave; and Cees P. M. van der Vleuten (2008). Teachers’ Interactions and their Collaborative Reflection Processes during Peer Meetings. *Advances in Health Sciences Education*, vol. 13, no. 3, August 2008, pp. 289–308.
- Woerkom, Marianne van; and Marcel Croon (2008). Operationalising critically reflective work behaviour. *Personnel Review*, vol. 37, no. 3, April 2008, pp. 317–331.
- Weick, Karl E. (1995). *Sensemaking in organizations*. Thousand Oaks: Sage Publications.
- Zaki, Mohammed J. (2001). SPADE: An Efficient Algorithm for Mining Frequent Sequences. *Machine Learning*, vol. 42, no. 1–2, January 2001, pp. 31–60.
- Zhu, Erping (1996). 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*, pp. 821–844.