This is an Accepted Manuscript of an article published by Taylor & Francis in Scandinavian Journal of Educational Research on 25th Aug 2015, available online: http://www.tandfonline.com/10.1080/00313831.2015.1066440.

Monitoring makes a difference – quality and temporal variation in teacher education students' collaborative learning

Abstract

The aim of this process-oriented video-observation study is to explore how groups that perform differently differ in terms of the number, quality and temporal variation of their content-level (knowledge co-construction) and meta-level (monitoring) activities. Five groups of teacher education students (N = 22) were observed throughout a three-month course. Video recordings (33 hours) of face-to-face group interaction (N = 12,931 speech turns) and pre- and post-tests of students' knowledge were collected. The results show that the well-performing group was more engaged in high-level knowledge co-construction and monitoring activities. The well-performing group was also capable of maintaining a higher level throughout the tasks, whereas the lower performing groups' knowledge co-construction and monitoring activities was reduced during the course.

Keywords: knowledge co-construction, monitoring, process-orientation, teacher education, video data

Monitoring makes a difference – quality and temporal variation in teacher education students' collaborative learning

Collaborative skills and an ability to take an active role in one's own learning and in the learning of others are highlighted as success factors in modern society (e.g. Griffin, McGaw, and Care, 2012). To develop these skills, students need their own experience of inquiry, along with collaboration-based instructional approaches to engage in resolving authentic, ill-structured and complex problems (Hmelo-Silver, 2004; Hmelo-Silver and Barrows, 2008).

In general, collaborative learning offers opportunities for the construction of knowledge by sharing, questioning and justifying ideas and understanding (Chi, 2009; Dillenbourg, 1999; Roschelle, 1992). These types of interaction and attending to other people's understanding have been defined as a co-construction of knowledge (Hogan, Nastasi, and Pressley, 1999). The theoretical ideas behind knowledge co-construction follow Roschelle's (1992) notion of convergence; group members construct shared meanings by constantly monitoring the degree to which they understand each other's thinking, extend other people's ideas, acknowledge divergent interpretations and resolve inconsistencies between the ideas that have been proposed. The premise underlying such learning relates to a process of explicating one's own ideas and engaging in the ideas of others (Webb, Franke, Ing, Wong, Fernandez, Shin, and Turrou, 2014). In sum, prior research has shown that when learners externalize and explain their still-developing understanding to each other, they learn more effectively (Baker, Hansen, Joiner, and Traum, 1999).

However, students' adaptation to collaborative learning situations – sharing knowledge and maintaining coordinated activity (Mäkitalo-Siegl, 2008; Roschelle and Teasley, 1995) – requires strategies, self-regulation, and socially shared regulation of learning, which are different from and often more challenging than the strategies required for individual learning (Barron, 2003; Hämäläinen and Arvaja, 2009; Kirschner, Sweller, and Clark, 2006). Effective collaborative learning requires groups to set goals and standards together, and to jointly monitor and evaluate their progress against these standards (Järvelä and Hadwin, 2013).

Despite extensive research on collaborative learning over the last decade, there is relatively little observational and process-oriented research that combines the analysis of content-level and meta-level interactions (Volet, Summers, and Thurman, 2009; Volet and Vauras, 2013). These interactions include how student groups engage each other's thinking by asking content-related questions, providing explanations, and particularly by monitoring their own and others' developing understanding, interests and task awareness, and group progress. The current study contributes to the growing corpus of process-oriented research (Volet and Vauras, 2013) by using a video study to evaluate and follow teacher education students' collaborative learning processes during a course of several collaborative learning tasks. More specifically, it develops the previous findings on collaborative learning by combining an analysis of content-level interaction (i.e. the number and level of questions and answers), and meta-level interaction (i.e. the number and focus of monitoring activities). It also makes use of the temporal variation between groups that are performing well, averagely and weakly across learning situations in order to determine what makes a collaborative group more or less successful.

Content-level activities: asking thought-provoking questions and elaborating answers

Previous studies have shown that high-order thinking processes, such as asking complex questions and elaborating on answers, enhance individuals' learning outcomes (Veenman, Denessen, van den Akker, and van der Rijt, 2005). Complex explanations are especially beneficial for learning (i.e., further evidence is provided and multiple concepts are integrated into the explanation) (Chinn, O'Donnell and Jinks, 2000; Roscoe and Chi, 2008). Several reasons have been suggested as to why such interactions may benefit participants during these interactions. First, explicating one's own ideas to others requires students to monitor their own thinking; they must transform their own knowledge into a communicative form that is relevant, coherent and complete so that others may attempt to understand it (Baker, 2002). During this process, students may recognize their own incompleteness or any misconceptions in their ideas more easily than they would when learning individually (Miyake, 1986; Shirouzu, Miyake, and Masukawa, 2002). Second, listening to others' ideas offers the students opportunities to monitor their own thinking and recognize possible gaps in their understanding. Third, having one's own ideas challenged and being asked for justifications may encourage students to seek new information, develop new ideas and build new connections between pieces of information (Chi, 2009; Wittrock, 1989).

Elaborating on one's own thinking and engaging with other people's ideas at a high level are central to many researchers' perspectives on productive small group interaction (e.g., Barron, 2003; Volet, Summers, and Thurman, 2009; Webb, Troper, and Fall, 1995). Barron's (2003) study explored the interaction processes of successful and less successful groups in detail. Differences in performance between the groups were found in relation to the ways in which students proposed ideas, acknowledged each other's ideas and elaborated on ideas proposed by others. Successful groups were more open to the contributions of all the group members, whereas less successful groups rejected the proposed solutions at critical moments (Barron, 2003). Additionally, in our previous studies we have explored how students' approach to own and other people's ideas contribute to successful collaborative interaction (Näykki and Järvelä, 2008) and we have also identified unsuccessful collaborative interactions, where students ignore, overrule or undermine others' contributions and the negative effects of these interactions on the group's learning process and outcomes (Näykki, Järvelä, Kirschner and Järvenoja, 2014).

Meta-level activities: monitoring evolving understanding

Järvelä and Hadwin (2013) highlighted that research into successful collaborative learning needs to be extended to target meta-level regulation processes within group interactions. To succeed in collaborative learning, students are required to engage in shared meta-level activities; that is, to go beyond an individual awareness of the knowledge gaps and to engage in monitoring each other's evolving content understanding (Lee, O'Donnell, and Rogat, 2014; Leinonen, Järvelä, and Häkkinen, 2005).

The role of meta-level activities in small groups is to coordinate, structure and regulate the cognitive and motivational processes of the group members (Baker et al., 1999; Hogan, 2001; Winne, Hadwin, and Perry, 2013). These mutual regulation activities that take place in groups have recently been referred to as a socially shared regulation of learning (Hadwin, Järvelä, and Miller, 2011). In other words, in successful groups, learners use a repertoire of appropriate cognitive activities to attain their learning goals, and use metacognitive activities to control and monitor their learning (e.g. Goos, Galbraith, and Renshaw, 2002; DiDonato, 2013; Khosa and Volet, 2014; Rogat and Linnenbrink-Garcia, 2011). Monitoring involves evaluating and judging one's understanding and progress during the task. As follows, students may check their understanding of the content and skills required for successful engagement in a group activity. This type of regulating the group's content understanding during a shared task refers to coordinating or negotiating each other's cognitive and metacognitive processes linked to content (Salonen, Vauras, & Efklides, 2005). Students can also monitor their progress toward task goals, and the assigned time for the task. Finally, once the task has been completed, effective regulators also reflect back and evaluate on their content understanding and task performance. Thus, group members plan the group's activities, monitor its actions and evaluate and reflect on its learning.

Previous studies have shown that in collaborative problem-solving activities, groups where members monitor their own and their peers' learning and thinking processes seem to have an advantage over groups where members do not (Goos et al., 2002; Hurme, Palonen, and Järvelä, 2006; Iiskala, Vauras, Lehtinen, and Salonen, 2011; Lee et al., 2014). For example, Roscoe and Chi (2008) evaluated events where explaining one's own understanding by using metacognitive statements such as "I didn't understand this before" was useful for making new connections and building understanding at the group level. In other words, meta-level statements activated longer episodes of active knowledge co-construction. In a similar manner, Volet, Summers, and Thurman (2009) explored what activated and sustained high-level regulation episodes within collaborative learning. Their findings indicated processes such as asking questions and tentativeness in explaining as meaningful activities for the group learning process. Furthermore, recent findings from Lee and colleagues' study (2014) indicate that the cognitive regulatory sub-process that enacted most frequently within collaborative interaction was content monitoring and that content monitoring can play a key role in the high quality joint activity.

Overall, previous findings indicate the value of locating meta-level events within group interaction as well as evaluating the transactional and intertwined role of such activities at the cognitive, motivational and social level (Rogat and Linnenbrink-Garcia, 2011; Volet and Summers, 2013). In other words, what kinds of effects do meta-level activities have on ongoing group interactions (see also Iiskala et al., 2011; Lee et al., 2014; Volet, Vauras, Khosa, and Iiskala, 2013).

Variations in time and quality in monitoring activities during collaborative learning

The small group as a learning context is dynamic and constantly evolving; every new contribution is a result of a previous discussion or decision, which affords the group members new possibilities in terms of remaining involved, becoming involved or withdrawing from the group and its activities (Mercer, 2008; Näykki et al., 2014). This kind of process-oriented focus in collaborative learning implies the need for the analysis of temporal differences within group activities (Molenaar and Järvelä, 2014). Reimann (2009) states that group learning unfolds over time, it is dependent on the groups' previous actions, it is cumulative in that current knowledge influences future knowledge, and that it has an anticipated future based on the interpretation of past experiences. Thus, tracking the temporal nature of content- and meta-level (monitoring) activities as a part of the groups' developmental process are essential for understanding group learning processes (Järvelä et al., 2013).

The temporal perspective can highlight differences at different points in time during group interaction (Reimann, 2009). In practice, the temporal perspective implies two points of view. First, some activities are needed and used more often at certain points in a group's learning process (e.g. some activities are precursors of others, while some activities follow naturally from others). Second, when learning proceeds, groups become more capable of undertaking some activities more often (e.g. as the group learns certain processes that were difficult before, these new processes become easier and/or even necessary) (Arrow, Poole, Henry, Wheelan, and Moreland, 2004). Previous studies have suggested the importance of the early stages of group activities. For example Kapur, Voiklis, and Kinzer (2008) examined the temporal patterns of problem-solving groups and emphasized the fact that there are high levels of sensitivity in early exchanges within groups. They noticed that group performance could be predicted based on the earlier phases of group discussion. In addition, Fransen, Weinberger, and Kirschner (2013) found that building a shared understanding of the characteristics of the tasks and the group's abilities in the early stages of group work were conditional factors required for group success. Thus previous studies highlight the importance of characterizing how collaboration evolves over time and across multiple learning situations.

Aim

The aim of this study is to investigate content-level and meta-level monitoring activities within differently performing groups. The specific research questions are: (1) How do differently performing groups engage in knowledge co-construction activities during collaborative learning? (1.1) How do knowledge co-construction activities vary temporally in differently performing groups? (2) How do differently performing groups engage in monitoring activities during collaborative learning? and (2.1) How do monitoring activities vary temporally in differently performing groups?

Methods

Participants and the research setting

The participants were 22 adult students on an educational science master's program at a Finnish university (17 women and 5 men; M_{age} = 39 years; SD = 11.2, range = 23–55). The prevalence of women reflects the gender ratio of education students at the university. All of the students had previously obtained a bachelor's degree in education, and had teaching and/or administrative work experience (for example as a kindergarten teacher, kindergarten principal, primary school teacher or adult educator). The master's program that the students were enrolled in employs a broad range of different kinds of group learning activities. Therefore, it can be assumed that all of the participating students had prior experience of group working. In this research setting, this existing knowledge and experience was not measured. However, the students' existing content knowledge was measured with a pre-test, and it was used also to form homogenous groups. Groups were formed so that the gender ratio (one male in each group) and the level of existing knowledge (based on the pre-test outcomes) were equally distributed between the groups.

The task

The students participated in a compulsory twelve-week course entitled 'Future scenarios and technologies in learning'. The overall aim of the course was to engage students in learning activities that would inspire them to implement new teaching practices in their own teaching work. The contents of the course included theoretical and pedagogical orientation to the technology-enhanced learning. The participants worked in groups of four to five students on assigned tasks. Following the approach of progressive inquiry (Hakkarainen, 2003), the learning activities in the course were organized in a cyclical model that emphasized shared expertise and collaborative work for knowledge building and inquiry. This was done by setting the context, and using questions, explanations, theories and scientific information in the cycle of deepening inquiry. The course structure included recurrent classroom, solo and collaborative phases mediated by the use of social media services (Figure 1).

Figure 1. The design

The course consisted of six lectures, and the main ideas of each lecture were discussed in collaborative learning sessions (which were the focus of this study's data collection). The face-to-face collaborative sessions always followed the same structure. In practice, groups were first asked to formulate the main ideas of the lecture and decide on one topic to serve as the basis for further group work. In the next phase, groups were advised to set their own learning objectives and to undertake an experiment with pictorial knowledge representations using smartphones (Näykki and Järvelä, 2008), weblogs and wikis (Laru, Näykki, and Järvelä, 2012).

The different phases of the pedagogical design are listed and explained below.

1. Lecture. In the first phase of the course design, the students attended a series of introductory lectures on the topics in question. The aim of the lectures was to introduce the different course topics (six themes) to the students and to support their conceptual grounding and theoretical understanding. The themes were: learning infrastructures, learning communities, metacognition, self-regulated learning, learning design, and the social web as a learning environment.
2. Face-to-face group work. After each introductory lecture, the students worked in face-to-face groups (six meetings). The aim of this was to discuss the main approaches of the lecture and to formulate a working problem for the group, to be continued in subsequent phases. The groups were advised to set their own learning objectives and to undertake an experiment with pictorial knowledge representations using smartphones, blogs and wikis.
3. Individual blog work (after every face-to-face session, with six topics). Students were asked to evaluate their everyday environment to find real-life examples and case scenarios to illustrate their ideas related to their group's problem. Within this phase, the idea of pictorial knowledge representation was implemented; the students used their individual blog learning environment to share their pictorial knowledge representations of and theoretical ideas on the topic.
4. Face-to-face group work (two meetings). The aim of the group work (in the middle and at the end of the course) was to evaluate the previous work phases (three topics at a time) and to share ideas from their individual blog work. Students were asked to negotiate and choose the best examples from their pictorial knowledge representations to illustrate their group's shared ideas, which were then used as the basis for their shared wiki work in the subsequent phase.
5. Group wiki work (after the first blog work). The aim of this last phase of the design was to co-construct a shared wiki for the group. In this phase, the students were asked to use their material (ideas and pictorial knowledge representations) from the previous phases and to generate new ideas and understanding when creating the wiki. The group wikis functioned as shared group products, which were presented to the whole class at the end of the course. The assignment required students to not only learn and apply content knowledge, but also to generate their own learning objectives and determine what information to include in their final presentations.

Data collection and analysis

The data for this study is composed of video-recorded face-to-face collaborative sessions and pre- and post-tests of student knowledge. The video recording captured the six collaborative learning sessions from each group (1980 minutes of video data). The duration of the sessions was determined by the group; the average duration of each session was 44 minutes.

To assess individual knowledge, the students completed identical paper-and-pencil pre- and post-tests. The tests consisted of six constructed-response questions based on the key concepts of the course. Students were asked to write answers to the following questions: 1. What is learning infrastructure? 2. What are learning communities? 3. What is meant by metacognition? 4. What is meant by self-regulated learning? 5. What is learning design? 6. How can social media be used as a learning environment? This meant that each question was also connected to the specific lectures and its associated collaborative task, and thus it was used to measure the students' learning outcomes for a particular task. The outcomes of the tests were used to characterize groups as performing well, averagely or weakly.

Phase 1. Analysis of the knowledge test: Defining differently performing students

The purpose of the first phase of the data analysis was to identify students' performance based on their individual knowledge tests. Three independent researchers (including the first author) developed the criteria and marked students' answers to all six questions in the knowledge tests (minimum score = 0, maximum score = 3 per question, resulting in a maximum of 18 points; Table 1). Initially, the researchers marked the tests independently, then they compared the results and calculated the proportional agreement (%) as reliability indices. The pre-test coding agreement between all three of the coders was 73.5% and the post-test agreement was 65.9%. Possible differences were negotiated until a consensus was reached. In general, the outcomes of the individual knowledge tests show that students scored higher on the post-test (M = 7.95) than on the pre-test (M = 3.95) ($t(22) = 8.33$, $p < .01$). The effect size (Cohen's d) was 1.69. This indicates that the students' understanding of the main concepts increased during the course.

Table 1. Rating criteria and examples of marking the learning test answers

The groups' learning gain was defined by first calculating a learning gain for each student (range: 0-8, gain score = ∑ post-test scores – ∑ pre-test scores). Next, the mean of all the gain scores was calculated ($M_{Gain} = 4.23$). Based on the M_{Gain}, the students were divided into good (gain score > 5), average (3 ≤ gain score ≤ 5) and weak performers (gain score < 3). The difference between the learning gain scores was further weighted by multiplying the good performers' gain by three, the average performers' gain by two, and the weak performers' gain by one. Finally, the groups' learning gain was calculated (group learning gain = ∑ group members' weighted learning gain/number of the group members) to indicate the groups' performance (Table 2). Based on this analysis, Group E ($M_{gain} = 14.25$) is defined as a well-performing group, groups C ($M_{gain} = 11.25$) and D ($M_{gain} = 10.5$) as averagely-performing groups, and groups A ($M_{gain} = 8.4$) and B ($M_{gain} = 8.2$) as weakly-performing groups.

Table 2. Groups' performance

Phase 2. Group-level analysis: Exploring knowledge co-construction activities

The second analysis phase was based on identifying the level of knowledge co-construction within each group; that is, analyzing the groups' processes for asking questions and providing answers. This was done to evaluate whether there were process differences between groups in addition to the above-described individual (and calculated group-level) outcome differences. The transcribed video data of all of the groups' face-to-face learning situations (33 hours, 12,931 speech turns) were analyzed through qualitative content analysis (Chi, 1997; Krippendorf, 2004) to capture what every member of each group was doing at a particular time.

First, all of the content-related questions and answers were detected. These included only those interactions in which the group members were sharing or questioning content information.

In practice, all of the speech turns where group members were asking questions or providing answers related to the task progress or off-task activities were excluded from the analysis. Second, the levels of questions and answers were evaluated and speech turns were divided into either high-, average- or low-level categories. Coding rules and analysis examples are presented in Table 3.

Table 3. Coding rules and analysis examples of content-related questions and answers

Two researchers developed criteria and marked the levels of questions and answers. The first researcher independently coded the speech turns, then the other researcher coded for 10% of the data. Reliability analysis was used to refine the coding scheme and ensure reliability of the analysis. Both researchers responsible for the reliability coding participated in the refinement of the coding system and, while coding, were blind to the performance of the students (individuals and groups). The intercoder reliability was.700 (Cohen’s kappa) for the question category and.729 (Cohen’s kappa) for the answer category.

Phase 3. Group-level analysis: Exploring monitoring activities

The third phase in the analysis explored what kinds of monitoring activities different groups use. The data-driven coding categories and overall coding protocol were developed in three phases. First, prior to viewing the videotapes, a list of preliminary areas of interest was developed according to the stated research questions. Second, the coding protocol was developed and elaborated on further after viewing the videotapes and reading the transcribed group discussions. Third, the final coding categories were formulated and tested several times. This involved the reorganization and renaming of categories, as well as the specification of sub-codes and providing examples of the specified categories. The final version of the data-driven coding protocol included the following four categories: monitoring task understanding, monitoring task progress, monitoring content understanding, and monitoring task interests. The coding rules and examples of the analysis are presented in Table 4. Within this analysis phase, the qualitatively analyzed interaction data was quantified to detect the possible differences between groups (Chi, 1997; Strijbos, Martens, Prins, and Jochems, 2006).

Table 4. Coding rules and examples of the data analysis for monitoring activities

The reliability of the coding was ensured by selecting 10% of the speech turns to be classified by the independent coder. The speech turns were randomly selected so that all categories (different types of monitoring activities) were represented in the selected data. Both researchers responsible for the reliability coding participated in the refinement of the coding system and, while coding, were blind to the performance of the students (individuals and groups). Reliability analysis was used to refine the coding scheme and the analysis. The overall intercoder reliability was.741 (Cohen’s kappa).

Results

The 30 face-to-face situations were analyzed in order to explore the differently performing groups’ number, quality and temporal variations in knowledge co-construction activities (content-related questions and answers) and monitoring activities (monitoring content understanding, monitoring task understanding, monitoring task interests and monitoring task progress).

How do differently performing groups engage in knowledge co-construction activities during collaborative learning? (RQ 1)

The results from the group-level content analysis show that the groups differ in terms of the quality of their knowledge co-construction. In practice, the quality of each groups’ knowledge co-construction was determined by evaluating the number and the level of content-related questions and answers (high, average, and low). The number of content-related questions and answers for each group is summarized in Table 5.

Table 5. The number of content-related questions and answers for each group

A Kruskall-Wallis H-test showed that there was a significant difference between the groups in *the number of high-level questions* ($H_{(4)}$ = 13.89, p <.01), with a mean rank of 23.17 for Group E (the well-performing group), 16.83 for Group A (a weakly-performing group), 12.50 for groups B (a weakly-performing group), C (an averagely-performing group) and D (an averagely-performing group). A pairwise comparison (post-hoc test) showed significant differences between groups E (a well-performing group) B, C, and D (*p* <.01). The effect size for all three significant pairwise differences (calculated based on the Mann-Whitney U test z value) is *r* = -0.66. A Kruskall-Wallis H-test also showed a significant difference between the groups in *the number of average level questions* ($H_{(4)}$ = 13.87, *p* <.01), with a mean rank of 24.25 for Group E (a well-performing group), 19.88 for Group A (a weakly-performing group), 13.50 for Group B (a weakly-performing group), 11.50 for Group D (an averagely-performing group), and 8.50 for Group C (an averagely-performing group). A pairwise comparison (post-hoc test) showed a significant difference between Group E and Group C (*p* <.01). The effect size for the significant pairwise difference is *r* = -0.86. Finally, a Kruskall-Wallis H-test also showed a difference between the groups in the number of low-level questions ($H_{(4)}$ = 12.72, *p* <.01), with a mean rank of 21.92 for Group A (a weakly-performing group), 21.67 for Group E (a well-performing group), 15.75 for Group B (a weakly-performing group), 9.92 for Group C (an averagely-performing group), and 8.25 for Group D (an averagely-performing group). However, a pairwise comparison did not reveal any significant differences between the groups.

A Kruskall-Wallis H-test showed that there was a significant difference between the groups in *the number of high-level answers* ($H_{(4)}$ = 14.56, p <.01), with a mean rank of 23.50 for Group E (a well-performing group), 18.50 for Group A (a weakly-performing group), 16.50 for Group B (a weakly-performing group), and 9.50 for groups C and D (an averagely-performing groups). A pairwise comparison (post-hoc test) showed significant differences between groups E (a well-performing group), C, and D (p <.01). The effect size for both significant pairwise differences (calculated based on the Mann-Whitney U test z value) is r = -0.77. A Kruskall-Wallis H-test also showed a significant difference between the groups in *the number of average-level answers* ($H_{(4)}$ = 18.45, p <.01), with a mean rank of 27.33 for Group E (a well-performing group), 19.17 for Group A (a weakly-performing group), 11.33 for Group B (a weakly-performing group), 9.75 for Group C (an averagely-performing group), and 9.92 for Group D (an averagely-performing group). A pairwise comparison (post-hoc test) showed significant differences between Group E and Group B (p <.01), Group E and Group C (p <.01), and Group E and Group D (p <.01). The effect size for all three significant pairwise differences is r = -0.83. There were no significant differences in the number of low-level answers between the groups.

How do knowledge co-construction activities vary temporally in differently performing groups? (RQ 1.1)

The different collaborative tasks were evaluated in more detail in order to evaluate the temporal variation in knowledge co-construction activities (content-related questions and answers) in well-, averagely- and weakly-performing groups. This revealed situational differences in the activities. The results show that all the groups had a variation between the tasks in terms of the number and the level of questions (high, average, low). However, throughout the course, well-performing Group E operated at a higher level than the other groups. In particular, the difference between the groups was visible at the beginning and at the end of the course. In terms of asking content-related questions, well-performing Group E was the most active during the tasks (tasks 1, 4, 5, and 6). Weakly-performing Group B's and averagely-performing Group C's and D's number of questions for the first (Task 1) and last tasks (tasks 5 and 6) are the lowest among the groups. However, during Task 2 the difference was not significant and, actually, in Task 3, weakly-performing Group B was more active than well-performing Group E in terms of asking questions. However, after the third task, the number of content-related questions Group B asked dropped, and during the final tasks, there were only a few content questions in their interactions. Contrary to this, well-performing Group E was the only group that managed to keep their level of on-task working high to the end and asked content-related questions more often than the other groups did (Figure 2).

Figure 2. Temporal variation of content-related questions

The analysis of content-related answers from different tasks follows the same trend that was visible in the analysis of the content-related questions above, in that all of the groups showed variance between the tasks in terms of quality as well as the number of answers given. However, there were apparent differences between well-performing Group E and the other groups. The differences between the groups is particularly visible at the beginning and at the end of the course. In terms of providing content-related answers, well-performing Group E was the most active during tasks 1, 2, 4, 5, and 6. This is a similar finding to the number of questions. In the first task, the well-performing group provided answers more often (f = 129) than the weakly-performing group did (f = 21). In the second task, the difference was not so great and, actually, in the third task, weakly-performing Group B was more active (f = 118) than the well-performing group (f = 84). In the final tasks, the well-performing group was again more able to keep their content-related discussion at a high level by providing answers more often, whereas the weakly- and averagely-performing groups' number of content-related answers dropped.

Figure 3. Temporal variation of content-related answers

How do differently performing groups engage in monitoring activities during collaborative learning? (RQ 2)

The results from the group-level content analysis show that all of the groups monitor their group processes by monitoring their content understanding, task understanding, task progress and task interests. The number of monitoring activities for each group is summarized in Table 6. A Kruskall-Wallis H-test shows that there was a significant difference between the groups in *the number of monitoring content understanding* ($H_{(4)} = 11.15$, $p < .05$), with a mean rank of 23.58 for Group E (a well-performing group), 19.50 for Group A (a weakly-performing group), 13.58 for Group C (an averagely-performing group), 12.08 for Group D (an averagely-performing group), and 8.75 for Group B (a weakly-performing group). A pairwise comparison (post-hoc test) showed significant differences between Group E (a well-performing group) and Group B ($p < .01$). The effect size for the significant pairwise difference (calculated based on the Mann-Whitney U test z value) is $r = -.71$. There were no significant differences between the number of events for monitoring task progress, monitoring task understanding and monitoring task interests between the groups.

Table 6. The number of monitoring events for each group

How do monitoring activities vary temporally in differently performing groups? (RQ 2.1)

To evaluate the temporal variation in monitoring content understanding in well-, averagely- and weakly-performing groups, the different collaborative tasks were evaluated in more detail, revealing situational differences in the activities (Figure 4). The results show that all groups had a variation between the tasks in terms of monitoring understanding. However, throughout the course, well-performing Group E operated at a higher and more stable level than the other groups. In particular, the difference between the groups is visible at the beginning of the tasks.

At the start of the tasks (Task 1), the most active group in monitoring understanding was Group A. However, its monitoring activities dropped during the next tasks (tasks 2 and 3), whereas Group E continued to monitor their group members' content understanding throughout the tasks (tasks 2, 3, 4, and 6). Weakly-performing Group B was one of the most active at monitoring content understanding during Task 3, but was inactive in all the other tasks. Overall, weakly-performing Group B used the lowest number of content monitoring activities, and significantly, this group did not start its content monitoring activities until the third task, whereas well-performing Group E was one of the most active in monitoring their group members' content understanding from the beginning to the end of the group work.

Figure 4. Temporal variation of monitoring content understanding

Discussion

This study explored the numerical, quality and temporal variations in knowledge co-construction and monitoring activities in differently performing groups. To summarize, the results for the knowledge co-construction processes as well as their temporal variation indicate differences between groups. The results confirm that Group E was well-performing in terms of quality and the number of knowledge co-construction activities. That is, even though all the groups showed variations in their collaborative processes, well-performing Group E was more capable of maintaining a high level of content processing throughout the tasks. Furthermore, this well-performing group was able to monitor their content understanding throughout their group activities more effectively than the other groups did. One interesting finding in this study was that the weak performing group B was exceptionally active in their content-level interaction and also their meta-level monitoring activities during the third task. However, to ponder about the possible reasons for their active mode of interaction is above the focus of this article, but it certainly requires further studies to find out possible situational explanations for this group's change of activity.

In general, these results are in line with previous findings which suggest that high-level cognitive processing, such as asking complex questions, elaborating on answers and monitoring an evolving understanding, is beneficial for collaborative learning (e.g., Chinn et al., 2000; Roscoe and Chi, 2008; Veenman et al., 2005). However, this study also extends previous findings by pointing to the possible connection between meta-level monitoring activities and content-level activities and learning outcomes by emphasizing the different focus of monitoring activities within differently performing groups and explicating temporal differences within groups' monitoring.

Monitoring content understanding has previously been recognized as a necessity for the effective development of a shared understanding during collaborative learning (e.g., Goos et al., 2002; Mäkitalo-Siegl, 2008). Additionally, the results of this study show that during collaborative learning, groups in which participants monitor their own and their peers' content understanding throughout the tasks have an advantage over groups who do not. The results also point out how groups differed in their focus on monitoring activities. It can be concluded that in addition to monitoring the development of their individual content understanding, in successful collaborative learning group members are actively involved in monitoring each other's task understanding, task progress and task interests.

Earlier studies have often used self-reporting and other “static” measurements, resulting in a lack of solid empirical evidence of what actually happens within a group (Järvelä et al., 2013); and how interaction unfolds over time and affords different kinds of opportunities for group members to engage in group activities (Barron, 2003; Näykki, et al., 2014). The use of video observations and analyzing real-time interactions in this study provides an extension by targeting the content-level and meta-level interaction and following it across several tasks to make effective collaborative learning processes visible. In this study, analyses of the video recordings enriched the understanding of group interaction as a moment-by-moment process (Barron et al., 2013). The indicators of effective collaborative learning were detected based on the group’s interaction and calculated group-level outcome measures. The results indicate possible reasons for success, but generalizable information and causal relations are not assured. This type of process-oriented methodological approach can, however, fill a gap in the research on content-level and meta-level processes in collaborative learning and provide an explanation of how learning situations develop in interactions between group members (Järvelä et al., 2013; Khosa and Volet, 2014; Salonen, Vauras, and Efklides, 2005).

We can conclude that the advantages of the in-depth interaction analysis can be seen in evaluating how interaction emerge and how group members’ contributions are contingent on others in the group. However, more analysis is needed to understand what constitutes success in collaborative learning and to comprehend the qualitative differences of monitoring activities. For example, further studies could follow Rogat and her colleagues (e.g. 2011) approach to differentiate high quality and low quality content monitoring activities. According to their study (Rogat and Linnenbrink-Garcia, 2011) a key element of high level content monitoring was the focus on ensuring group member’s understanding of the content. Accordingly, high quality content monitoring often occurred when group members jointly examined their task responses to evaluate content contributions and if their understandings could be further elaborated, corrected, or improved. Low quality content monitoring was observed when members of the group were primarily focused on monitoring for the right answer, rather than ensuring conceptual understanding of the group members. Furthermore, a second pattern of low quality content monitoring was evident in that not all group members were considered equal contributors to the process of group monitoring. For instance, certain individual’s monitoring efforts were ignored by the group or a particular group members’ incomplete or inaccurate understanding was not scaffolded. Another pattern of low quality content monitoring was when monitoring was especially critical or harsh. The harsh nature of the monitoring made it challenging for the group member to be responsive and negotiate a shared understanding of the content. These examples highlight the further need to look more closely at the differences between group members in terms of the focus and trigger for monitoring; that is the ways monitoring is enacted but also when and how monitoring is followed by controlling activities. In addition to observational data also group members’ own explanations of why and how they engage in monitoring activities can offer important input for understanding group processes in details (Winne et al., 2013).

According to Son and Schwartz (2002) merely monitoring one’s learning does not have significant effect on learning performance. The difference is made when both regulation functions are involved; monitoring is followed with appropriate controlling. It is known from earlier studies on collaborative learning that how group members listen, share perspectives and build on others ideas can make the difference in learning success (Barron, 2003; Chinn et al., 2000; Hogan et al., 1999; Näykki and Järvelä, 2008; Miyake, 1986; Webb et al., 1995). In a similar manner meta-level monitoring and controlling activities, and particularly their situational specific functionalities can make a difference in groups' performance (Lee et al., in press; Khosa and Volet, 2014; Rogat and Linnenbrink-Garcia, 2011; Volet et al., 2009).

The temporal difference between groups' monitoring activities over time, particularly at the beginning and the end of the course, requires further exploration. One of the key idea in exploring successful collaborative learning is to identify how groups progress during their tasks and study courses. There is an apparent lack of studies evaluating the regulation processes in collaborative learning situations that would characterize for example a progress or development of skills to regulate learning together as a group.

Furthermore, the findings of this study have practical implications for the design of student group learning in higher education. As monitoring makes a difference for successful collaborative interaction, it needs to be supported and trained. A better quality of collaborative learning processes and outcomes could be reached by learning environment that provide instructional scaffolding for monitoring activities and offer opportunities for students to practice and receive feedback about their learning.

Acknowledgement

This work was supported by the Research Council of Cultural and Social Science Research, Academy of Finland Grant no. [107667] and [201759].

References

Arrow, H., Poole, M.S., Wheelan, S. and Moreland, R. (2004). Time, change and development: The temporal perspective on groups. *Small Group Research, 35*(1), 73-105.

Baker, M. (2002). Forms of cooperation in dyadic problem-solving. In P. Salienber and H. Benchekroun (Eds.), *Cooperation and complexity* (pp. 587-620). Paris: Hermés.

Baker, M., Hansen, T., Joiner, R. and Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 31-64). Amsterdam: Pergamon.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences, 12*(3), 307-359.

Barron, B.J., Pea, R., and Engle, R.E. (2013). Advancing understanding of collaborative learning with data derived from video records. In C.E. Hmelo-Silver, C.A. Chinn, C.K.K. Chan and A. O'Donnell (Eds.), *The International Handbook of Collaborative Learning* (pp. 203-219). New York: Routledge.

Chi, M.T.H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences, 6*(3), 271-315.

Chi, M.T.H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science, 1*(1), 73-105.

Chinn, C.A., O'Donnell, A.M., and Jinks, T.S. (2000). The structure of discourse in collaborative learning. *Journal of Experimental Education, 69*(1), 77-97.

Dillenbourg, P. (1999). Introduction: What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-19). Oxford: Elsevier.

DiDonato, N.C. (2013). Effective self- and co-regulation in collaborative learning groups: An analysis of how students regulate problem solving of authentic interdisciplinary tasks. *Instructional Science*, *41*(1), 25-47.

Fransen, J., Weinberger, A. and Kirschner. P.A. (2013). Team effectiveness and team development in CSCL. *Educational Psychologist, 48*(1), 9-24.

Goos, M., Galbraith, P., and Renshaw, P. (2002). Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics, 49*(2), 193-223.

Griffin, P., McGaw, B., and Care, E. (2012). Assessment and Teaching of 21st Century Skills. Dordrecht: Springer.

Hadwin, A., Järvelä, S. and Miller, M. (2011). Self-regulated, co-regulated and socially shared regulation of learning. In B. Zimmerman, and D. Schunk (Eds.), *Handbook of Self-Regulation of Learning and Performance* (pp. 65-84). New York: Routledge.

Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of Research in Science Teaching, 40*(10), 1072–1088.

Hmelo-Silver, C. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*, 235-266.

Hmelo-Silver, C. and Barrows, H. S. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction, 26*(1), 48-94.

Hogan, K. (2001). Collective metacognition: the interplay of individual, social, and cultural meanings in small groups' reflective thinking. In F. Columbus (Ed.), *Advances in Psychology Research*, Vol. 7 (pp. 199-239). Hauppauge, NY: Nova Science.

Hogan, K., Nastasi, B.K., and Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*(4), 379-432.

Hurme, T.-R., Palonen, T., and Järvelä, S. (2006). Metacognition in joint discussions: an analysis of the patterns of interaction and the metacognitive content of the networked discussions in mathematics. *Metacognition and Learning, 1*(2), 181-200.

Hämäläinen, R. and Arvaja, M. (2009). Scripted Collaboration and Group-Based Variations in a Higher Education CSCL Context. *Journal of Scandinavian Research, 53*(1), 1-19.

Iiskala, T., Vauras, M., Lehtinen, E., and Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction, 21*, 379-393.

Järvelä, S. and Hadwin, A. (2013). New Frontiers: Regulating Learning in CSCL. *Educational Psychologist, 48*(1), 25-39.

Järvelä, S., Järvenoja, H., and Näykki, P. (2013). Analyzing regulation of motivation as an individual and social process – A situated approach. In S. Volet and M. Vauras (Eds.), *Interpersonal Regulation of Learning and Motivation: Methodological Advances* (pp. 170-187). New York: Routledge.

Kapur, M., Voiklis, J. and Kinzer, C. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers and Education, 51*(1), 54-66.

Kirschner, P.A., Sweller, J. and Clark, R.E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.

Khosa, D. and Volet, S. (2014). Productive group engagement in cognitive activity and metacognitive regulation during collaborative learning: Can it explain differences in students' conceptual understanding. *Metacognition and Learning, 9*(3), 287-307.

Krippendorff, K. (2004). Content Analysis: An Introduction to its Methodology. Thousand Oaks: Sage.

Laru, J., Näykki, P. and Järvelä, S. (2012). Supporting small-group learning using multiple Web 2.0 tools: A case study in the higher education context. *The Internet and Higher Education*, *15*(1), pp. 29–38.

Lee, A., O'Donnell, A. M., and Rogat, T. K. (2014). Exploration of the cognitive regulatory sub-processes employed by groups characterized by socially shared and other regulation in a CSCL context. *Computers in Human Behavior* (in press).

Leinonen, P., Järvelä, S. and Häkkinen. P. (2005). Conceptualizing awareness of collaboration: A qualitative study of a global virtual team. *Computer Supported Cooperative Work, 14*(4), 301-322.

Mercer, N. (2008). The seeds of time: why classroom dialogue needs a temporal analysis. *Journal of the Learning Sciences, 17*(1), 33-59.

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, 10, 151-177.

Molenaar, I. and Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning, 9*(2), 75-85.

Mäkitalo-Siegl, K. (2008). From Multiple Perspectives to Shared Understanding: A small group in an online learning environment. *Scandinavian Journal of Educational Research*, *52*(1), 77-95.

Näykki, P. and Järvelä, S. (2008). How Pictorial Knowledge Representations Mediate a Collaborative Knowledge Construction in Groups. In C. Angeli (Ed.), Special Issue: Distributed Cognition Approaches on Technology in Teacher Education. *Journal of Research on Technology in Education*, *40*(3), pp. 359-387.

Näykki, P., Järvelä, S., Kirschner, P., and Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning – A process-oriented case study in a higher education context. *International Journal of Educational Research*, *68*, 1-14.

Reimann, P. (2009). Time is precious: Variable and event-centered approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 239-257.

Rogat, T.K. and Linnenbrink-Garcia, L. (2011). Socially shared regulation in collaborative groups: An analysis of the interplay between quality of social regulation and group processes. *Cognition and Instruction*, *29*(4), 375–415.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences, 2*(3), 235-276.

Roschelle, J., and Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C.E. O'Malley (Ed.), *Computer-Supported Collaborative Learning* (pp. 69-97). Heidelberg: Springer-Verlag.

Roscoe, R.D. and Chi, M.T.H. (2008). Tutor learning: the role of explaining and responding to questions. *Instructional Science, 36*(4), 321-350.

Salonen, P., Vauras, M., and Efklides, A. (2005). Social interaction: What can it tell us about metacognition and co-regulation in learning? *European Psychologist 10*(3), 199-205.

Shirouzu, H., Miyake, N., and Masukawa, H. (2002). Cognitively active externalization for shared reflection. *Cognitive Science*, 26, 469-501.

Son, L. K. and Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. J. Perfect and B. L. Schwartz (Eds.), *Applied Metacognition*. Cambridge, UK: Cambridge University Press, pp.15–38.

Strijbos, J.W., Martens, R.L., Prins, F.J., and Jochems, W.M.G. (2006). Content analysis: What are they talking about? *Computers and Education*, 46, 29-48.

Veenman, S., Denessen, E., van den Akker, A., and van der Rijt, J. (2005). Effects of a cooperative learning program on the elaborations of students during help seeking and help giving. *American Educational Research Journal*, *42*(1), 115-151.

Volet, S. and Summers, M. (2013). Interpersonal regulation in collaborative learning activities: Reflections on emerging research methodologies. In S. Volet and M. Vauras (Eds.), *Interpersonal regulation of learning and motivation: Methodological advances* (pp. 204-220). New York: Routledge.

Volet, S., Summers, M., and Thurman, J. (2009). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? *Learning and Instruction*, *19*(2), 128-143.

Volet, S. and Vauras, M. (2013). Interpersonal Regulation of Learning and Motivation: Methodological Advances. New York: Routledge.

Volet, S., Vauras, M., Khosa, D., and Iiskala, T. (2013). Metacognitive regulation in collaborative learning: Conceptual developments and methodological contextualizations. In S. Volet and M. Vauras (Eds.), *Interpersonal regulation of learning and motivation: Methodological advances* (pp. 67-101). New York: Routledge.

Webb, N.M., Franke, M.L., Ing, M., Wong, J., Fernandez, C.H., Shin, N., and Turrou, A.C. (2014). Engaging with others' mathematical ideas: interrelationships among student participation, teachers' instructional practices and learning. *International Journal of Educational Research*, *63*, 79-93.

Webb, N.M., Troper, J.D., and Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology*, *87*(3), 406-423.

Winne, P.H., Hadwin, A.F., and Perry, N.E. (2013). Metacognition and computer-supported collaborative learning. In C. Hmelo-Silver, A. O'Donnell, C. Chan and C. Chinn (Eds.), *International Handbook of Collaborative Learning* (pp. 462-479). New York: Taylor and Francis.

Wittrock, M.C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345-376.