International Journal of Computer-Supported Collaborative Learning
https://doi.org/10.1007/s11412-019-09296-0

Improving the quality of vocational students' collaboration and knowledge acquisition through instruction and joint reflection

Elise H. Eshuis[1] · Judith ter Vrugte[1] · Anjo Anjewierden[1] · Lars Bollen[2] · Jakob Sikken[1] · Ton de Jong[1]

Received: 17 August 2018 / Accepted: 7 February 2019

Abstract
New societal demands call for schools to train students' collaboration skills. However, research thus far has focused mainly on promoting collaboration to facilitate knowledge acquisition and has rarely provided insight into how to train students' collaboration skills. This study demonstrates the positive effects on the quality of students' collaboration and their knowledge acquisition of an instructional approach that consists of conventional instruction and an online tool that fosters students' joint reflection on their collaborative behavior by employing self- and peer assessment and goal setting. Both the instruction and the collaboration reflection tool were designed to promote students' awareness of effective collaboration characteristics (the RIDE rules) and their own collaborative behavior. First-year technical vocational students ($N = 198$, $M_{age} = 17.7$ years) worked in heterogeneous triads in a computer-supported collaborative learning environment (CSCL) on topics concerning electricity. They received either 1) conventional instruction about collaboration and the online collaboration reflection tool, 2) collaboration instruction only, or 3) no collaboration instruction and no tool. Analysis of chat data ($n = 92$) and knowledge tests ($n = 87$) showed that students from the instruction with tool condition outperformed the other students as far as their collaborative behavior and their domain knowledge gains.

Keywords Collaborative learning · Collaboration instruction · Collaboration reflection tool · Inquiry learning

Introduction

With the progressive embedding of technology in society, professionals in technical vocations (e.g., car mechanics, electrical engineers) are increasingly required to function in multidisciplinary

✉ Elise H. Eshuis
e.h.eshuis@utwente.nl

Extended author information available on the last page of the article

teams and to work on complex multifaceted problems in which collaboration is essential for successful problem-solving. Hence, such technicians are expected to be not only experts in their field, but also efficient and effective collaborators (Christoffels and Baay 2016). However, schools struggle with explicitly teaching most interdisciplinary skills such as collaboration (Onderwijsraad 2014) and thus fail to prepare their students to meet these workplace requirements. Aside from being a problem for the transition from student to employee, this is also a missed opportunity for education, because collaboration, if done effectively, can contribute to students' knowledge acquisition (e.g., Gijlers et al. 2009; ter Vrugte and de Jong 2017).

Currently, technical vocational education does encourage collaboration. The most common integration of collaboration seems to be through projects in which students work in teams on a joint product. However, schools do not typically provide instruction on effective collaboration, nor do they assess whether students can demonstrate the desired skills, aside from occasional self-reflection reports. Although this format allows students to practice collaboration in a relevant context, which is essential for development of skills such as collaboration (Hattie and Donoghue 2016), it is unlikely to be the most effective, productive, and efficient way to improve students' collaborative behavior. Research has shown that merely placing students together does not automatically result in the desired collaborative behavior (Johnson et al. 2007; Mercer 1996) and that without support, collaborating students often fail to reach the desired goal within the set timeframe or fail at the task completely (Järvelä et al. 2016; Rummel and Spada 2005; Anderson et al. 1997). Moreover, "Inappropriate use of teams can undermine the educational process so badly that learning does not take place, students learn how not to learn, and students build an attitude of contempt for the learning process" (Jones 1996, p. 80).

The popularity of collaboration in educational settings is also demonstrated by the considerable amount of research directed at collaboration. This research, however, mainly focuses on the use of collaboration to optimize knowledge acquisition (e.g., Dehler et al. 2011; Kollar et al. 2006; Noroozi et al. 2013; Wecker and Fischer 2014). As a logical consequence, the majority of the studies have focused their interventions and analyses on knowledge gain; far less emphasis has been put on how to train the collaboration itself and how instructional approaches affect students' actual collaborative behavior. Therefore, the current study investigates whether instructional support related to collaboration in the form of a combination of conventional instruction and prompted joint reflection (incorporating principles of self- and peer-assessment and goal setting) can steer technical vocational students towards behavior that is desirable for effective collaboration. As a frame of reference for identifying desired collaborative behavior, the outcomes of the analyses on essential characteristics of collaboration by Saab et al. (2007) were used.

Theoretical framework

Instruction

Research has shown that instructing students about characteristics of effective collaboration fosters their collaborative behavior (Chen et al. 2018). A frequently used method for such instruction is scripting, a form of guided instruction in which students are instructed about how they should interact and collaborate (Dillenbourg 2002). A considerable amount of research has been conducted in which a variety of scripts were implemented (Kollar 2006; Vogel 2017). Although students may learn from the information introduced by scripts, there is a risk that these scripts spoil the (natural) collaboration process (Dillenbourg 2002).

Other more conventional forms of instruction that precede students' collaborative activities and therefore do not interfere with ongoing collaborative process have also proven to be effective. Rummel and Spada (2005) and Rummel et al. (2009) showed that observing a worked-out example of a model collaboration can positively affect students' subsequent collaborative behavior. This is in line with more general findings that students can learn from observing other students' dialogue (Stenning et al. 1999).

More evidence for the effectiveness of preceding students' collaboration with instruction on characteristics of effective collaboration comes from Saab et al. (2007), who performed a study in which students received computerized instruction on characteristics of effective collaboration combined with examples of good and poor employment before collaborating in an inquiry learning environment. The outcomes demonstrated that students who received the instruction, collaborated more constructively compared to students who did not receive the instruction. The content of the instruction was based on their literature review, in which they identified the following behaviors as essential for effective collaboration:

> …allow all participants to have a chance to join the communication process; share relevant information and consider ideas brought up by every participant thoroughly; provide each other with elaborated help and explanations; strive for joint agreement by, for example, asking verification questions; discuss alternatives before a group decision is taken or action is undertaken; take responsibility for the decisions and action taken; ask each other clear and elaborated questions until help is given; encourage each other; and provide each other with evaluative feedback. (p. 75).

Saab et al. (2007) summarized these essential behaviors in four rules: Respect, Intelligent collaboration, Deciding together, and Encouraging (the RIDE rules). These rules, which reflect communicative activities that are seen as essential for effective collaboration, formed the basis for the definition of effective collaboration in the current study.

Although instruction and examples are already effective, stimulating students to connect desired behavior to their actual behavior, and providing students with the opportunity to close gaps between desired and actual behavior might be necessary to create substantial change (Sadler 1989). Reflection (i.e., looking back on past behavior in order to optimize future behavior) fosters this connection.

Joint reflection

Reflection is "a mental process that incorporates critical thought about an experience and demonstrates learning that can be taken forward" (Quinton and Smallbone 2010, p.126). It creates awareness of processes that are normally experienced as self-evident and is considered an essential element of the learning process (Chi et al. 1989). Although the majority of the research focusing on reflection in collaborative settings has related reflection to learning outcomes (Gabelica et al. 2012), there is also some research that has related reflection to collaboration skills (e.g., Phielix et al. 2010, 2011; Prinsen et al. 2008; Rummel et al. 2009). However, the outcomes of these studies were ambiguous. Though this ambiguity is possibly partially due to the fact that reflection on collaboration was implemented differently in the different studies and collaboration was measured in different ways, it could also be explained by the reliability of the information students reflected upon.

Reflection requires students to assess their own performance, which ideally involves comparing their performance to the goal performance, identifying gaps in their performance, and working towards fixing these gaps (see, Quinton and Smallbone 2010; Sadler 1989; Sedrakyan et al. 2018). Hence, the starting point for reflection is often students' self-assessment. However, research has demonstrated that in general, students tend to overestimate their own skills and performance (Dunning et al. 2004). With regard to the specific focus of the current study, Phielix et al. (2010) and Phielix et al. (2011) found that students hold unrealistic positive self-perceptions regarding their collaboration skills. Hence, even when reflecting, students' unrealistic self-perceptions can cause them to fail to identify gaps in their skills, which leads them to exert less effort than needed to optimize their behavior. Gabelica et al. (2012) stated that they consider feedback from an external agent to be a necessary precondition for reflection. Complementing students' self-assessment with additional data (e.g., peer-assessments) is likely to help optimize the effectiveness of their reflection. Providing students with multiple views can complement the results of their self-perception, which, in turn, provides a better basis for reflection (Dochy et al. 1999; Johnston and Miles 2004).

Based on a review of 109 studies, Topping (1998) concluded that the effects of peer assessment are "as good as or better than the effects of teacher assessment" (p. 249); furthermore, research has demonstrated that peer assessment is beneficial for both the assessors as well as the assessees (e.g., Li et al. 2010; Topping 1998; van Popta et al. 2017). In peer assessment, social processes might stimulate students to increase their effort when comparing their scores to those of their peers or to set higher standards for themselves (Phielix 2012; Janssen et al. 2007). The studies from Phielix and colleagues (Phielix et al. 2010, 2011) in which they used a peer feedback and a reflection tool to enhance group performance in a computer-supported collaborative learning environment provided examples of this. They found positive effects of peer feedback and reflection on perceived group-process satisfaction and social performance.

After identifying their performance, reflection requires students to compare their performance to the goal performance, identify gaps, and make a plan to work towards fixing those gaps (see, Quinton and Smallbone 2010; Sadler 1989; Sedrakyan et al. 2018). Though all of these steps are essential for successful reflection, the importance of the last step has been emphasized by research demonstrating that reflection should not only be about students' current behavior, but should also include their future functioning (Gabelica et al. 2012; Phielix et al. 2011; Hattie and Timperley 2007; Quinton and Smallbone 2010). More specifically, it seems important that students are stimulated to set goals for further improvement of their behavior. Gabelica et al. (2012) compared students who received feedback with students who received feedback but additionally were also prompted to reflect on this feedback, identify gaps, and explain how their behavior could be improved. They termed this second step "reflexivity", and found that this proactive analysis is essential to the effectiveness of feedback.

Current study

From the above, it can be concluded that schools are in need of effective tools for teaching their students how to collaborate. Though research shows that simply having students collaborate – a common practice in technical vocational education – is not enough to improve students' collaboration skills, it is as yet unclear how instructional approaches can foster students' collaborative behavior or affect their collaboration skills. There is limited evidence showing that instructing students on characteristics of effective collaboration is beneficial (Rummel and Spada 2005; Rummel et al. 2009; Saab et al. 2007). More research is necessary to substantiate these findings. In addition, it is likely that the inclusion of reflection could increase the effectiveness of such instruction. Considering reflection, joint reflection seems preferable over independent reflection (Renner 2016), and principles of self- and peer assessment and goal setting can be employed to optimize the effectiveness of this joint reflection. The rare studies that have coupled joint reflection (through principles of self- and peer assessment and goal setting) and collaboration have demonstrated promising results (Phielix et al. 2010, 2011). However, the focus of these studies was on perceived collaborative behavior. Results of the joint reflection for students' actual behavior therefore remains unclear.

The current study extends knowledge in the field of collaborative learning by investigating the effect of instructional support (conventional instruction together with joint reflection using principles of self- and peer-assessment and goal setting) not only on students' knowledge acquisition, but also on their actual collaborative behavior, while working in a computer-supported collaborative learning (CSCL) environment. It employs approaches used in prior studies and unites them in a unique way.

More specifically, the instructional support was designed to inform students about the RIDE rules and stimulate students to use these rules during their collaboration. The RIDE rules (i.e., Respect, Intelligent collaboration, Deciding together, and Encouraging) are communication rules based on essential characteristics of collaboration and have been tested as a support for synchronous distance communication (Saab et al. 2007; Gijlers et al. 2009). The instructional support took the form of face-to-face conventional instruction and an online tool to prompt students' joint reflection, supported by studies by Renner et al. (2014, 2016) showing evidence of the effectiveness of online prompts for joint reflection. To optimize students' reflection, the tool incorporated self- and peer-assessment (students assessed their own and each other's collaborative behavior) to provide a more reliable information source for reflection, and collaborative goal setting (students collaboratively planned how to optimize their future collaboration) to stimulate students to connect their reflection to their future behavior.

The following research questions were addressed:

1. What is the effect of instruction about characteristics of effective collaboration on the quality of students' collaboration?
2. What is the effect of a combination of instruction about characteristics of effective collaboration and joint reflection on the quality of students' collaboration?
3. What is the effect of instruction about characteristics of effective collaboration on students' knowledge acquisition?
4. What is the effect of a combination of instruction about characteristics of effective collaboration and joint reflection on students' knowledge acquisition?

To answer these questions, three conditions were compared in the current study. In one condition, students received instruction in which they were taught about the RIDE rules before entering into collaboration. In a second condition, students received similar instruction complemented with an online tool (also based on the RIDE rules) that required them to reflect on their collaborative behavior. In a third condition (a control condition), students received no instruction and no tool.

Based on the above-mentioned literature, it was expected that instruction on the RIDE rules would foster chat activities that contribute to effective collaborative behavior and would therefore positively affect the quality of students' collaboration (measured in terms of desired communication activities) and that the students' use of the tool would further strengthen this effect. In addition, as several studies have shown that effective collaboration positively affects students' knowledge construction (Lou et al. 2001; van der Linden et al. 2000; Johnson et al. 2007), it was expected that the improved collaborative behavior would positively affect knowledge acquisition.

Method

Design

This study utilized a pretest - intervention - posttest design. During the intervention, students worked collaboratively in heterogeneous teams of three students in a computer-supported collaborative learning (CSCL) environment. To ensure heterogeneous grouping, the following procedure was followed: per class, pretest scores were ranked from high to low and divided into three equal parts (i.e., high, average, and low pretest scores). Students who did not complete the pretest, for assignment purposes only they received the class average as their pretest score. Within each class, triads were then composed by grouping three students from each part. The first triad would consist of the three students who scored highest within their part, and the last triad was made up of the three students who scored lowest within their part.

Each triad was assigned to one of the following conditions: the instruction with tool condition, the instruction only condition, or the control condition. To ensure that triads' average pretest scores were equally distributed among the three conditions, triads within each class were ranked on their average pretest score and alternately assigned to the different conditions.

The conditions were identical in terms of learning material (i.e., the CSCL environment, in which an information icon with information about the RIDE rules was integrated that was accessible for all conditions), but differed in whether or not students received the collaboration instruction and had to use the collaboration reflection tool.

Participants

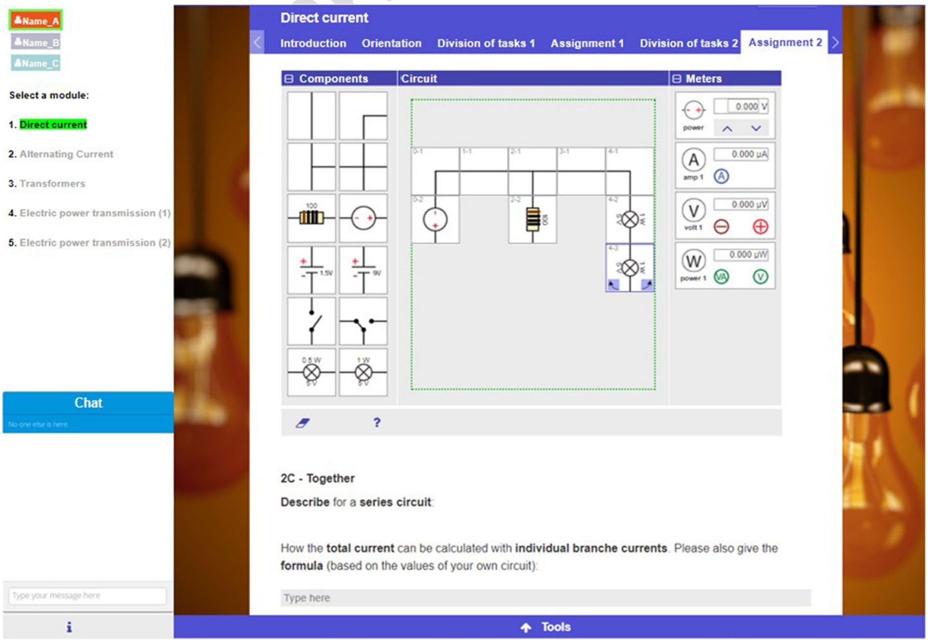
A total of 198 secondary vocational education students (192 males, 6 females), with a mean age of 17.67 years ($SD = 1.25$) - based on the exact age by using the birthday - participated in this study. Participants were first-year students from nine classes divided over four schools for secondary vocational education (in Dutch: MBO) in the Netherlands. Within this sector, students are prepared for their role as a vocational professional. Courses are offered at four levels of education and in two learning pathways (i.e., school-based and work-based). Participants in the current study were enrolled in a technical training program at the fourth level (i.e., specialist training) that includes electrical engineering as a fundamental part of their curriculum and has a total duration of four years.

As teamwork is part of their curriculum, all participants had experience with collaborative activities to a greater or lesser extent. Although these students were familiar with digital learning materials and software, they had no learning experiences in similar CSCL environments with supportive tools.

CSCL environment

Research has shown that it is important that skills such as collaboration (part of what are termed twenty-first century skills) are learned when embedded in contexts that resemble the students' future professional practice (Hattie and Donoghue 2016). Therefore, in this study, students were trained in a CSCL environment. This enabled students to practice collaboration not only with relevant topics (i.e., while working on problems they might encounter during their profession) but also within a relevant setting (i.e., one where collaboration is not face-to-face and communication is mainly digital). To ensure that the content of the CSCL environment would be similar to actual school tasks and would include situations considered important for students' future profession, the environment was co-designed with teachers.

The CSCL environment contained a series of assignments, two online labs, instructive multimedia material, and a chat facility (see Fig. 1). It was designed with the Go-Lab Authoring Platform (de Jong et al. 2014) and covered three topics (i.e., direct/alternating current, transformers, and electric power transmission) divided over five modules (i.e., direct current, alternating current, transformers, electric power transmission (1), and electric power transmission (2)). All modules were similarly structured by means of tabs at the top of the



Q5

Fig. 1 Screenshot of the CSCL environment (translated from Dutch) with the current module with the Electricity Lab; the names of the team members, the chat, and an overview of the modules are on the left screen. The first tab opened an introduction in which the purpose and use of the module was briefly explained. The introduction was followed by an orientation tab where the topic of that particular module was introduced, either through an introductory video or through an overview of relevant concepts. The remaining tabs opened assignments that were connected to one of the online labs. Completion of a module was a prerequisite for starting the next one.

Two *online labs* were integrated within the CSCL environment. In the Electricity Lab (see Fig. 1), students could create electrical circuits based on direct or alternating current, perform measurements on them, and view measurement outcomes. In the Electric Power Transmission Lab (see Fig. 2) students could design a transmission network by choosing different power plants and cities, and by varying different components within the network (e.g., properties of the power line, number of power pylons, and the voltage). Depending on the assignment, the labs were used either individually (i.e., student actions were not synchronized) or together (i.e., student actions were synchronized).

All modules contained similar types of assignments that required collaboration for completion (e.g., information sharing and shared decision making). The assignments stimulated this collaboration through individual accountability and positive interdependence, two elements that aim to trigger interactions within teams (Johnson et al. 2007). More specifically, every module contained one or more assignments that students had to complete individually, after which they had to share their information in order to complete a joint assignment. For example, when students had to create an electrical circuit in the Electrical Circuit Lab, they were each assigned a different task (i.e., specific components that only they could manipulate). Similarly, in the Electrical Power Transmission Lab the ultimate joint goal was to create an optimal power transmission network, while each individual student pursued a unique goal (i.e., highest efficiency, lowest costs, and highest safety/sustainability). In this way, students could only reach their joint goal if the individual tasks were completed (i.e., positive interdependence) and all students could be held responsible for the team's success (i.e., individual accountability).

Collaboration instruction

The goal of the classroom collaboration instruction was to inform students about important characteristics of effective collaboration. The instruction was based on the RIDE rules (see introduction), which were slightly adapted in order to make them more suitable for this target group and learning environment (see Table 1).

The instruction, delivered by the researcher, followed a teacher-centered approach and was structured in accordance with the first principles of instruction (Merrill 2002): activation of prior experience, demonstration of skills, application of skills, and integration of these skills. The instruction started with a short introduction during which the learning goals (i.e., knowing what are important characteristics of effective collaboration and improving collaboration skills) were explained. Thereafter, the relevance of being able to collaborate effectively was emphasized by stressing the importance of collaboration skills in their future jobs. Prior experiences were activated by recalling situations in which students had to work together in teams during school projects. After this, each RIDE rule was explained by defining its relevance and by introducing the sub-rules, which were each illustrated by a good and a poor example. The instruction continued with an interactive portion in which chat excerpts showing good and

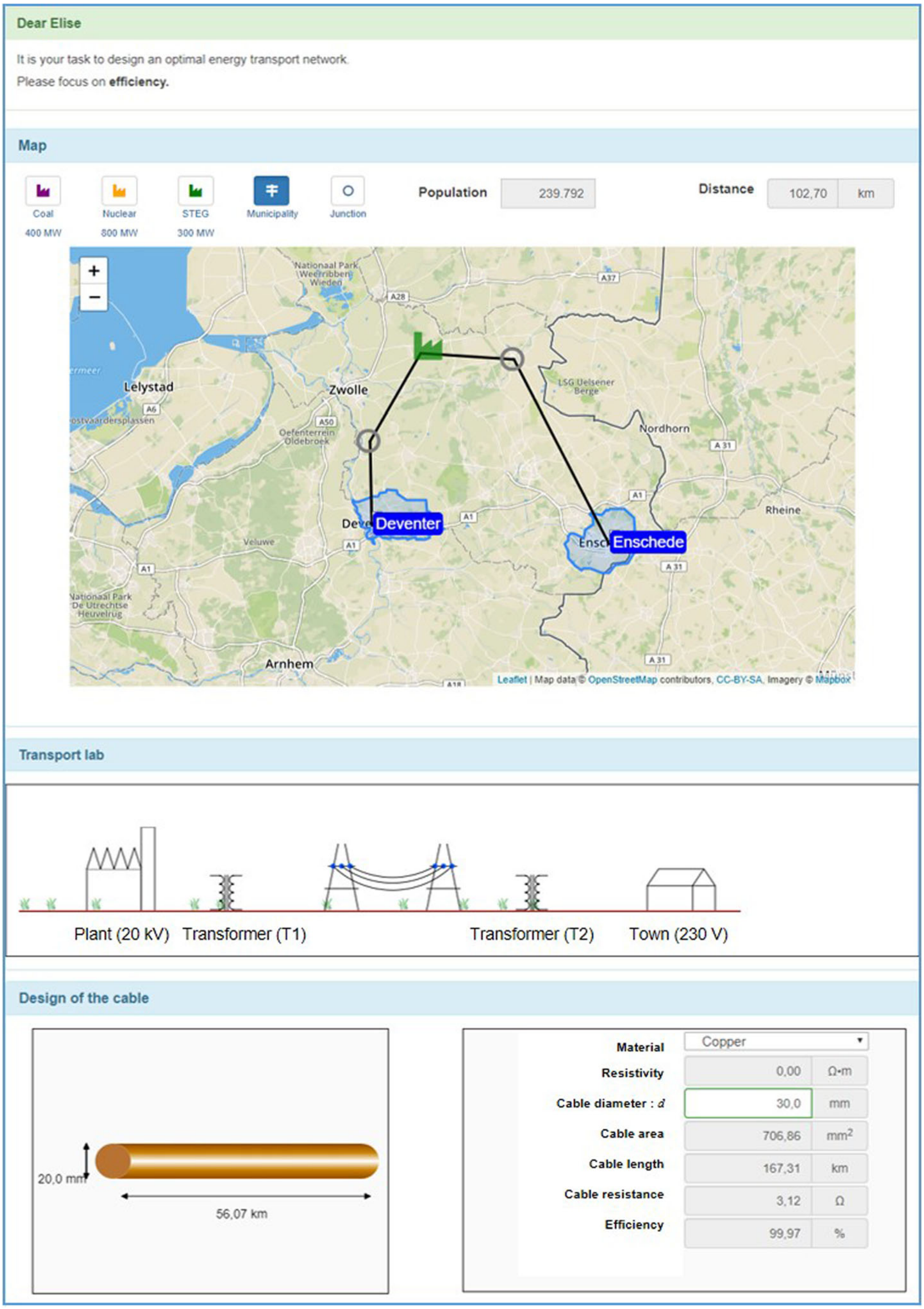


Fig. 2 The electric power transmission lab (translated from Dutch), with efficiency as the main goal

poor examples were evaluated. In this way, students were encouraged to apply what had been demonstrated and explained during the instruction in another context. Finally, students were asked to think about how they could apply the RIDE rules when working together in the CSCL environment with the online labs.

Table 1 RIDE rules with their sub-rules, adapted from Saab et al. (2007)

RIDE rule	Sub-rules
Respect	Give everyone a chance to talk
	Consider other students' input
	Don't judge students personally after they make mistakes
Intelligent collaboration	Share all relevant information and ideas
	Clarify the information/answers given
	Ask for explanations if they have not been given or when something is unclear
	Give constructive criticism of other students' ideas (not of the person himself or herself)
Deciding together	Check if everyone agrees before taking actions or giving answers
	Contribute to the decision-making process if others want to make a decision
Encouraging	Encourage others to participate actively
	Give compliments when others make a useful contribution

Collaboration reflection tool

The collaboration reflection tool was designed to prompt and scaffold students' joint reflection on collaborative behavior. It incorporated principles of self- and peer-assessment and goal setting.

As described, the content of the tool addressed the four RIDE rules, while the structure of the tool was based on a study of peer feedback and reflection by Phielix et al. (2011). In line with the model for successful feedback proposed by Hattie and Timperley (2007), the tool consisted of three phases that were completed either individually (1. feed up and 2. feed back) or in collaboration with the other team members (3. feed forward). This feedback model is also aligned with the major steps in reflection, which require students to assess their performance (feed up), compare their performance to the goal performance (feed back), identify gaps in their performance and work towards fixing these gaps (feed forward) (see, Quinton and Smallbone 2010; Sadler 1989; Sedrakyan et al. 2018).

In the *feed up phase*, students had to rate their own and each other's collaborative behavior on a ten-point scale for each RIDE rule (see Fig. 3a). Students could access information about the rules by clicking on the corresponding information icons. This information was similar to the information given in the collaboration instruction. Once students had completed the assessment, they indicated this by pressing the 'finished' button, after which they entered the feed back phase (see Fig. 3b).

In the *feed back phase*, students received a graphical representation of the team's and their own evaluated collaborative behavior. The initial graph showed the average team score for each RIDE rule. By clicking on one of the bars, self- and peer assessment scores for each student were shown (see Fig. 3b). Once all students indicated they had finished viewing the feedback, they entered the feed forward phase (see Fig. 3c).

The *feed forward phase* was designed to stimulate reflection and goal setting. Students received a set of questions that structured their reflection and supported them in constructing goals for improvement. This phase was complemented by the graphical representation from the feed back phase. Similar to the information given in the feed up phase, the RIDE rules were explained briefly. For each RIDE rule, students had to discuss what went well and what could be improved, after which they had to write down their joint goals (i.e., what will they keep on doing? and what are they going to improve?). If the tool had been completed before,

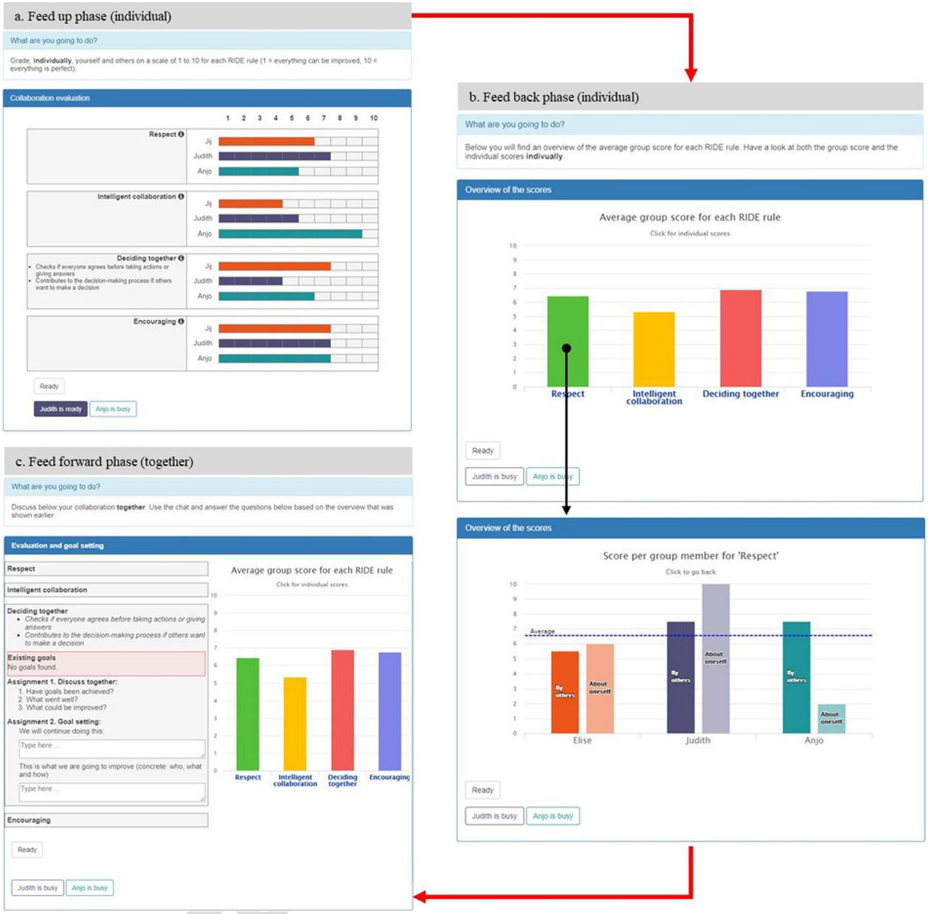


Fig. 3 The three phases of the collaboration reflection tool: feed up, feed back, and feed forward (translated from Dutch)

previously formulated goals were shown and students were asked to discuss whether or not these goals had already been reached. Once the current goals for each RIDE rule had been formulated, students could click on the 'finished' button. The tool closed when all students did so. Completion of the collaboration reflection tool takes up to ten minutes.

Domain knowledge tests

Two parallel paper-and-pencil tests were used to measure students' domain knowledge about the three topics in the CSCL environment, both before and after the intervention. The parallel items differed from each other in context or formulation. Counterbalancing was used to prevent order effects. That is, approximately 50% of the students of each condition received version A as a pretest and version B as a posttest, while the rest of the students received version B as a pretest and version A as a posttest. Reliability analysis revealed a Cronbach's alpha of .87 on the pretest and .85 on the posttest.

The test consisted of 12 open-ended questions, which were tailored to the content of the CSCL environment: four questions per topic, of which two assessed knowledge at the conceptual level and two assessed knowledge at the application level. A rubric was used to score the tests. For each test, a maximum of 25 points could be earned. The total maximum score for the conceptual knowledge questions was 9 points. The total maximum score for the application questions was 16 points. A second coder coded 37% of the tests independently, which resulted in an interrater reliability (Cohen's Kappa) of .76 for the pretest and .82 for the posttest.

Procedure

The experiment took place in a real school setting during regular school hours. It comprised five sessions, the first and the last of which took 60 min each, while the second, third, and fourth took 90 min each. The first session started with a short introduction during which students were informed about the upcoming lessons and learning goals. The ultimate goal of improving their collaboration was emphasized. Subsequently, students were given the domain knowledge pretest. They were told that during the learning sessions, content-related questions would not be answered by either the teacher or the researcher. The first session ended with the introduction of the CSCL environment and the labs. At the start of the second session, students in the instruction with tool condition and the instruction only condition received the collaboration instruction, which took 20 min. Students in the control condition received no collaboration instruction and had to wait in another room. After this instruction, students in the instruction only condition were asked to go to the room in which the students from the control condition were seated. Students in the instruction with tool condition then received brief instruction about how the collaboration reflection tool worked. After the instructions were given, all students were gathered and assigned a seat. To discourage face-to-face communication, students within the same triad were seated apart from each other. Thereafter, students received the URL of the CSCL environment and a login code and started working in the environment that was specific for the condition they were assigned to. Students continued their work in the CSCL environment during the third and fourth sessions or until they had completed all the modules. For the students in the instruction with tool condition, the collaboration reflection tool was offered four times at set points in the CSCL environment. To give students a chance to get used to the environment, before reflecting on their collaboration, there was no tool after module one, but only at the end of modules two, three, four, and five. In the fifth session students completed the domain knowledge posttest.

Data analysis

The inclusion criteria for students' data were based on students' grouping, attendance, and progress: Students who had to work individually because they could not be grouped in a triad due to the total number of students in a class not being a multiple of three were excluded from the final dataset. Moreover, not all students attended all three intervention sessions (i.e., sessions 2–4). If only one team member did not attend all intervention sessions, the data for the team (which was a dyad in one or more sessions and a triad in the remaining sessions) were still analyzed, but only the data from the members who attended all intervention sessions were included in the final dataset. If more than one team member did not attend all intervention sessions, the team was removed from the analysis. Also, students who did not finish the fourth (and thus also fifth) module, which means that they could not have obtained all of the necessary knowledge (as measured by the test) and did not complete at least three iterations of the tool, were excluded from the final dataset. As a result, the final sample included 92 students (27 in the instruction with tool condition, 28 in the instruction only condition, and 37 students in the control condition). Five of these students were excluded from the sample for the analysis of the domain knowledge tests, because they missed either the pretest or the posttest or both; results for these students were included in the log file analyses.

Students' chat logs were used to assess their collaborative behavior. Chat activities were derived from the log files and were coded based on their content. Students' communications were evaluated as being on-task or off-task, responsive, and respectful. Other evaluations included whether students kept each other posted, helped each other when necessary (i.e., whether they shared information, asked questions, and were critical regarding others' input), and whether they took responsibility for their actions and decisions.

At the start of the coding, the chat data per student were segmented into utterances. An utterance is a coherent entry by a student that was submitted to the chat by pressing Enter. If two consecutive chat entries by a single student contained an exact repetition or spelling correction, both entries were combined and considered as one utterance. A coding scheme, based on the coding scheme developed by Saab et al. (2007) was developed to code each utterance (see Fig. 4). The coding scheme of the current study shows communalities with the framework developed by Weinberger and Fischer (2006) and the coding scheme used by Gijlers et al. (2009).

A total of 17,412 utterances were coded (see Table 2 for the number of utterances per condition and the average number of utterances per team). Each utterance was coded on two levels. On the first level an utterance was coded as *off-task* (utterance related to neither the task nor the domain), or on-task. Off-task utterances received no further codes. On-task utterances were further specified as *domain* (domain-related utterance), *coordination* (task-, but not

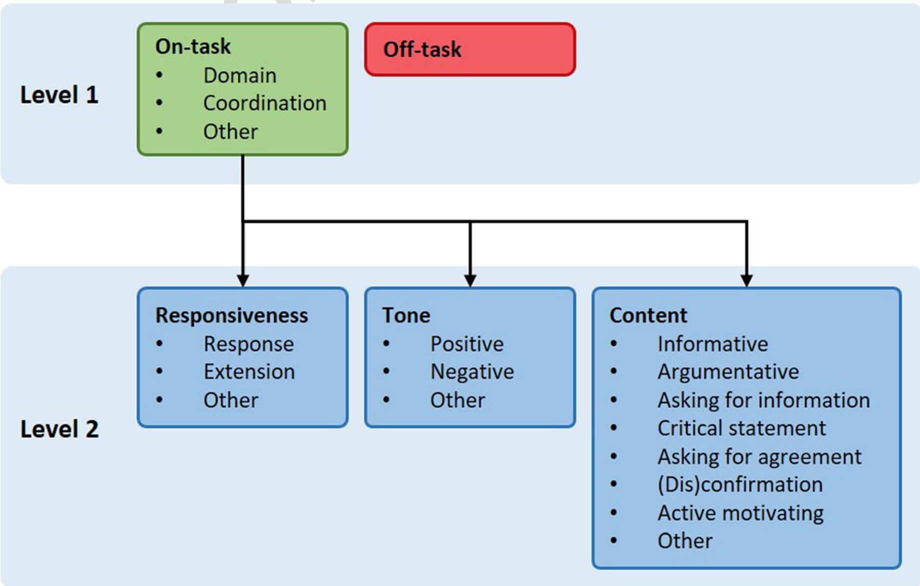


Fig. 4 Overview of codes at the utterance level

domain-related utterance), or *other* (utterance related to actions within the lab, (dis)functioning of the lab or the CSCL environment, or that could not be coded as either domain or coordination).

Only on-task utterances received three additional codes at the second level. First, the responsiveness of an utterance was determined. It was decided whether an utterance was a *response* (a reaction to one of the previous utterances within the chat). If an utterance was not a response, it was coded as either an *extension* (extension of one's own previous utterance) or *other* (an utterance that could not be coded as either response or extension). Second, utterances were assessed on their tone, which could be *positive* (towards a person, a task or in general), *negative* (towards a person, a task or in general), or *other* (neutral or undeterminable). Third, the content of an utterances was determined using one of the following codes: *informative* (informative utterance), *argumentative* (argumentative utterance aiming at clarifying, reasoning, interpreting, stating conditions or drawing conclusions), *asking for information* (asking for understanding, explanations or clarification), *critical statement* (asking a critical question or making a critical statement), *asking for agreement* (asking for agreement or action), *(dis)confirmation* (explicitly (dis)agreeing or (not) giving consent), *active motivating* (encouraging team member(s) to participate or to take action), *other* (utterances that could not be coded in one of the previous categories, i.e., 'haha', 'hmm'). See Table 3 for examples of each utterance.

A second coder coded 12.6% of the chats independently (level 2), which was a selection of 21 chat excerpts, randomly divided across students and modules. This resulted in an interrater reliability (Cohen's Kappa) of .74 (with 85.5% agreement) for responsiveness, .76 (with 94.5% agreement) for tone, and .76 (with 84.3% agreement) for content. These codes were used to create codes in terms of percentage (the proportion of a specific type of utterance out of the total number of utterances) that could serve as indicators for the quality of students' collaboration. These codes in terms of percentage are used as communication variables for further analysis. The extensive data set that results from this, allows for conducting quantitative analyses to objectively identify differences in communication activities between conditions and, therefore, gives insight in the possible effect of the intervention on the quality of students' collaboration.

One of the communication activities (i.e., give everyone a chance to talk) cannot be expressed in terms of frequencies of particular utterances, but is related to equal participation. Therefore, a group-level measure was created in addition to the measurements at the student level, to gain insight into how equal team members' participation was (i.e., the extent to which all students contributed to the dialogue). When all students in a group are given the chance to talk and therefore a chance to contribute to the group process, this would become manifest in equal participation (Janssen et al. 2007), whereas highly unequal participation can be an indicator of social loafing (Weinberger and Fischer 2006). Therefore, the total number of utterances (minus the utterances coded as extensions) per student within a team was used to

Table 2 Total number of utterances per condition, average number of utterances per team.

Condition	n	Total	M	SD	Min	Max	Range
Instruction with tool	11	5114	464.91	159.26	162	716	554
Instruction only	11	5242	476.55	133.97	308	723	415
Control	13	7056	542.77	342.09	191	1310	1119
Total	35	17,412	494.74	211.77	220.33	916.33	696

t3.1 **Table 3** Overview of codes with two corresponding example chat excerpts (translated from Dutch)

t3.2	**Codes**	**Example**
t3.3	**Level 1**	
t3.4	Domain	What is the formula of the effective value? That's true, because here the 0 is the middle of the sine wave.
t3.5	Coordination	We should do this assignment together. I only see 2 rows of text.
t3.6	Other	Nflsdfkjsd Oooh...
t3.7	Off-task	Time for beer! I'm going to watch Netflix
t3.8	**Level 2**	
t3.9	*Responsiveness*	
t3.10	Response	Student A: [name of student B], that's incorrect. Student B: What should it be? *(response)* Student A: Must also be calculated with Ohm's law *(response)* (…) Student B: What do they mean by efficiency? Student A: Loss, 100% efficiency means no loss. *(response)*
t3.11	Extension	Student A: In a parallel circuit, it's V1 = V2 Student A: etc. *(extension)* (…) Student A: yes, Student A: think you're right *(extension)*
t3.12	Other	What's happening? Next assignment?
t3.13	*Tone*	
t3.14	Positive	Yes, saw it. Well done! Great teamwork, nice!
t3.15	Negative	You're so stupid [name of student].. Pff.. stop harping
t3.16	Other	Go ahead I really don't know
t3.17	*Content*	
t3.18	Informative	27 indeed. I'm done
t3.19	Argumentative	Best analogy? I think blood circulation, because this must be a full circulation, that's not the case with water and transportation. That's because you need I1, I2 and I3, so 3 × 200 mA
t3.20	Asking for information	Which factors do influence the cable resistance? Do you know what we should do here?
t3.21	Critical statement	[Name of student], I think you didn't get it. It should not be truck, but the whole transportation cycle as the analogy. Shouldn't that be 15 V?
t3.22	Asking for agreement	Do you agree? Shall we use just 1 power plant?
t3.23	(Dis)confirmation	Yes, I agree. No, I'm not ready yet
t3.24	Active motivating	[Name of student], what do you think? Come on, we at least can try
t3.25	Other	….. Hmmm

calculate the Gini coefficient. This coefficient is a group-level measure and is often used to measure (in)equality of participation (e.g., Janssen et al. 2007). For each team, it sums the deviation of team members from equal participation. This sum is divided by the maximum possible value of this deviation. The coefficient ranges between 0 and 1, with 0 indicating perfect equality and 1 perfect inequality. In this case, perfect equality would mean a perfectly equal distribution of utterances between team members, while perfect inequality would mean that one member was responsible for all of the utterances within a team. A team with a high number of utterances can have the same Gini coefficient as a team with a low number of utterances as long as the distribution of utterances is similar within each team.

Results

Chats

To investigate whether there were differences in individual students' communication activities between conditions, multivariate analysis of variance (MANOVA) was conducted, with condition as the independent variable and the communication variables (i.e., domain-related talk, coordination-related talk, off-task talk, responsive talk, positive talk, negative talk, informative talk, argumentative talk, questions asking for information, critical statements, questions asking for agreement, (dis)confirmations, and motivating talk) as dependent. Results show an overall effect of condition on the communication variables, Wilk's $\Lambda = .446$, $F(28, 152) = 2.70$, $p < .001$, $\eta_p^2 = .332$. Table 4 presents mean scores with standard deviations and summarizes the results of the subsequent univariate tests per communication variable.

Post-hoc analyses (Bonferroni) of the significant effects were performed to identify which differences between conditions were significant. Results revealed that students in the *instruction with tool condition when compared to those in the control condition* scored *significantly higher* on the variables domain-related talk ($p < .001$), responsive talk ($p = .024$), critical statements ($p = .015$), questions asking for agreement ($p = .010$), and (dis)confirmations ($p < .001$), and *significantly lower* on the variables off-task talk ($p = .024$), negative talk ($p = .002$), and informative talk ($p < .001$).

Table 4 Mean scores (%) and standard deviations per communication variable and results of univariate tests by condition

	Instruction with tool ($n = 27$)		Instruction only ($n = 28$)		Control ($n = 37$)		Univariate tests		
	M	SD	M	SD	M	SD	$F(2,89)$	p	η_p^2
Domain-related talk	31.91	10.33	21.68	9.33	20.97	7.20	13.79	< .001	.237
Coordination	40.47	10.82	42.80	8.79	43.52	26.42	0.22	.801	.005
Off-task talk	12.03	14.15	19.42	14.12	21.29	12.54	3.89	.024	.080
Responsive talk	24.74	10.93	18.54	7.58	18.28	9.52	4.32	.016	.089
Positive talk	3.68	2.81	3.47	2.03	2.73	2.54	1.33	.269	.029
Negative talk	2.21	2.32	6.39	6.66	7.03	6.08	6.67	.002	.130
Informative talk	54.43	7.32	60.59	7.67	62.44	8.46	8.36	< .001	.158
Argumentative talk	4.98	3.15	3.49	2.37	3.46	2.64	2.09	.060	.061
Questions asking for information	14.21	6.41	14.37	5.17	13.22	5.70	0.39	.678	.009
Critical statements	3.20	2.91	1.36	1.52	2.00	1.64	7.81	.001	.149
Questions asking for agreement	6.69	4.41	4.88	3.16	3.87	3.55	4.54	.013	.093
(Dis)confirmations	7.80	5.88	4.30	3.20	2.96	2.58	11.77	< .001	.209
Motivating talk	2.34	2.71	2.39	2.60	2.20	2.31	0.05	.948	.001

In addition, students in the *instruction with tool condition in comparison to those in the instruction only condition* scored *significantly higher* on the variables domain-related talk ($p < .001$), critical statements ($p = .001$), and (dis)confirmations ($p = .005$), and *significantly lower* on the variables informative talk ($p < .015$) and negative talk ($p = .017$).

To gain insight into the equality of team members' *participation*, Gini coefficients between conditions were compared. These coefficients turned out to be relatively close to zero in all conditions (instruction with tool: $n = 11$, $M = .11$, $SD = .08$; instruction: $n = 11$, $M = .10$, $SD = .04$; control: $n = 13$, $M = .13$, $SD = .08$), indicating that the distribution of participation in all conditions was fairly equal. A Kruskal-Wallis test revealed no significant differences between conditions ($H(2) = .160$, $p = .923$).

Domain knowledge test

Table 5 presents an overview of the mean pretest scores, posttest scores, and learning gains (posttest scores minus pretest scores) for every condition. Univariate analyses of variance (ANOVA) indicated no significant difference on pretest scores between conditions, $F(2, 84) = 0.52$, $p = .595$, $\eta_p^2 = .012$, which indicates that students in all conditions were comparable in terms of their prior knowledge. To calculate whether students' domain knowledge improved after the intervention, a paired samples t-test comparing pre- and posttest scores was performed for each condition. Results show that posttest scores were significantly higher than pretest scores in all three conditions (instruction with tool condition: $t(24) = 5.93$, $p < .001$, $d = 1.19$; instruction only condition: $t(26) = 5.74$, $p < .001$, $d = 1.10$; control condition: $t(34) = 3.38$, $p = .002$, $d = 0.57$), which indicates that, on average, students in all conditions did learn. In addition, an ANOVA with learning gains as dependent variable and condition as independent variable revealed that these learning gains differed significantly between conditions, $F(2, 84) = 3.39$, $p = .038$, $\eta_p^2 = .075$. Post-hoc comparisons (Bonferroni) showed that learning gains were significantly higher in the instruction with tool condition compared to the control condition ($p = .038$).

Discussion and conclusion

The unique contribution of this study is that the focus is not simply on the use of collaboration to optimize knowledge acquisition, but on how to affect students' collaborative behavior, and whether this behavior impacts students' knowledge acquisition. In the current study, two interventions to improve students' collaborative behavior and knowledge acquisition were compared to each other and to a control condition: instruction with a tool that facilitated joint

Table 5 Mean pretest scores, posttest scores and learning gains (max = 25)

Condition	n	Pretest				Posttest				Gain			
		M	SD	Min	Max	M	SD	Min	Max	M	SD	Min	Max
Instruction with tool	25	4.96	3.35	.00	12.66	7.58	3.73	2.08	15.16	2.63	2.21	−1.58	9.83
Instruction only	27	4.02	3.07	.00	10.08	6.09	3.28	1.25	11.99	2.08	1.88	−1.25	6.00
Control	35	4.55	3.52	.00	14.58	5.78	3.38	.50	13.33	1.23	2.16	−2.00	6.58
Total	87	4.50	3.32	.00	14.58	6.39	3.50	.50	15.16	1.89	2.15	−2.00	9.83

reflection through principles of self- and peer-assessment and goal setting (instruction with tool condition), instruction only (instruction only condition), and a condition where students received no instruction on collaboration and no tool (control condition). Both the instruction with tool and the instruction only conditions were designed to raise students' awareness of important factors that influence the quality of collaboration, but the former also included support that was designed to stimulate students to connect these characteristics to their own current behavior, identify gaps, and adjust their behavior accordingly (i.e., through goal setting or feed forward).

In general, the findings of this study show that only providing students with instruction on essential characteristics of collaboration benefits neither their collaborative behavior nor their knowledge acquisition, compared to providing no instruction (i.e., research question 1 and 3). However, instruction combined with instructional support (i.e., the collaboration reflection tool) that prompts students to connect their experience with the instructed characteristics (through self- and peer assessment) and to collaboratively set goals to improve their collaborative behavior, does affect students' collaborative behavior and knowledge acquisition (i.e., research question 2 and 4). Based on the finding that collaborative behavior was affected and knowledge acquisition was improved we carefully conjecture that the designed tool can advance the quality of students' collaboration.

Effect on students' collaborative behavior

In the current study, in order to evaluate the effectiveness of the instruction and tool, the focus was not on whether students knew how to behave—many people may know how to behave without putting it into practice (the knowing vs doing gap)—but whether students actually engaged in the desired behavior.

From the results of the chat analyses, it can be deduced that students who received both the instruction and the tool demonstrated relatively more behaviors that are related to higher quality collaboration in comparison to the control condition. These students dedicated a higher proportion (i.e., the proportion of a specific type of utterance out of the total number of utterances) of their communication to on-task activities (e.g., more domain related utterances, less off-task utterances), employed better social hygiene (e.g., less negative utterances, more often making responsive utterances, asked more frequently for agreement), and adopted a more critical attitude (e.g., more critical statements, shared more opinions about other students' decisions). Specifically this last category, which can also be described as 'conflict- and integration-oriented consensus building' (i.e., critiquing another students' ideas or contributions, or actively adapting one's conceptions to include the ideas of peers) has been identified as predicting students' knowledge acquisition (Gijlers et al. 2009; Weinberger and Fischer 2006). However, it must be noted that in the current study, in line with findings by Gijlers et al. (2009), students' critical utterances were rare.

It is noteworthy that students who received the instruction only, did not demonstrate any of the desired collaborative behaviors more often in comparison to students in the control condition. This is not in line with our expectations. Based on earlier studies, it was expected that students who received instruction would become more aware of effective collaborative behavior and adjust their behavior accordingly (Stenning et al. 1999; Rummel and Spada 2005; Rummel et al. 2009; Gijlers et al. 2009). More specifically, Saab et al. (2007) and Gijlers et al. (2009) found that instruction on the RIDE rules led to more constructive communication in comparison to a control condition. However, the main difference between those studies and the current study is that the collaboration in those studies was directed at dyads (not triads) and that they practiced the RIDE rules in the learning environment before proceeding to the task. Although practice was incorporated in the instruction on the RIDE rules in the current study, this practice was via questions that required students to provide examples from their own experience, and assessment of good and poor examples (chat excerpts). We can, however, argue that the current study did provide an extensive intervention (3 lessons of 90 min each) that should have provided sufficient time for students to get acquainted with the environment. Therefore, the influence that lacking that kind of practice with the rules might have on the results of the current study remains unclear. Future research could focus on the influence of group size on the effectiveness of this kind of instructional support. Another direction could be to look into the effect of repeated instruction in comparison to the repeated use of the tool and to compare students' collaborative behavior each time after the intervention is offered, as the quality of their collaboration might improve when their collaboration is repeatedly supported, due to internalization of desirable behavior (Vogel et al. 2017).

Effects on students' knowledge acquisition

Although the average posttest score of 6.39 out of 25 might be interpreted as relatively low, on average, students in all conditions in the current study learned from the CSCL environment (as demonstrated by a significant knowledge gain in each condition). When considering that typical lessons in vocational education offer the learning material repeatedly over longer periods of time, the learning results achieved within this short time frame can be considered to be quite satisfactory. Nonetheless, it might be worthwhile to consider the effect on learning gains of how students were teamed for collaboration and whether other team compositions might generate different results. For the current study, heterogeneous triads were composed based on their prior knowledge. This might have affected the observed learning gain, as research on ability grouping has shown that contributions within heterogeneous groups are often less equal than in homogeneous groups. This implies that the observed knowledge gain might be a result not only of collaboration, but also of a teacher-learner relationship that might have emerged between students with differing prior knowledge (Lou et al. 1996; Saleh et al. 2005). Although the Gini coefficients being close to zero indicates that the contribution within the teams was in fact rather equal, which makes this interpretation somewhat unlikely, future studies could control for this possible effect.

Students who received both the instruction and the tool were most successful in terms of knowledge acquisition, as they showed significantly higher gains on the knowledge test than students who received no instruction and no tool. We carefully conjecture that this is due to the fact that these students engaged in higher-quality collaboration than the other students did. This would be in line with other findings that (good) collaboration has a positive impact on learning (Dillenbourg et al. 1995; Barron 2003; Weinberger and Fischer 2006). As discussed above, several of the presented affected behaviors have been related to knowledge gain. However, less direct behavior or qualities of communication such as 'tone' (e.g., being polite) and supportive behavior are known to affect people's responsiveness, receptivity, and (depth of) processing (Kirschner et al. 2015). We are therefore carefully arguing that it is a set of behaviors, rather than a specific activity, that is responsible for the observed effect on knowledge acquisition. In the current study, only students' own chat activities were considered; therefore, it is unclear how much of the other communications students were aware of and how communication between other team members might have affected their involvement and learning. A future study might include eye-tracking measures to gain more insight into students' processing of the chat communications within a team (i.e., whether or not certain communications activities are observed by the different team members).

Considerations and conclusion

The outcomes of this study show the effectiveness of the described tool in combination with conventional instruction. However, as yet, it is unclear whether all of the included working mechanisms of the tool are equally essential for creating the described effect on behavior. The tool included two working mechanisms to improve students' reflections: peer assessment and goal setting. It would be beneficial to understand which working mechanism generated what effect and how. For instance, peer assessment was employed to foster reflection because it can help overcome problems that arise from self-overestimation. In addition, whether there is reflection or not, it has been shown that students can gain both knowledge and skills by performing peer assessment (Gonzalez de Sande and Godino Llorente 2014; Strijbos et al. 2009). More specifically, research has shown that in peer assessment, what fosters learning is not receiving feedback, but the activity of providing feedback (Lundstrom and Baker 2009). This was demonstrated in a study by Li et al. (2010), in which they found a relation only between the quality of provided feedback and learning, but not between the quality of received feedback and learning. For the current study, this could mean that it matters less whether students received correct assessments of their behavior than whether they provided high quality assessments of their peers' behavior. In that situation, the activity of assessing drives the students to think critically about their peers' behavior, which stimulates awareness and could also trigger self-reflection. Future studies focusing on data about students' assessment process and its quality can add to our understanding. A comparison between the current tool and versions without either peer assessment or goal setting can demonstrate whether both mechanisms are essential for the currently established effectiveness, and might also give further insight into whether students need additional support for these processes. Casual observations, for example, suggested that students were not very specific in their goal setting (i.e., they either stated that there was no need for improvement, or they suggested improving a particular activity, without specifying how and by whom this should be done). This might indicate that the effectiveness of the tool can be further improved. More support tailored to goal setting (which requires additional meta-cognitive skills) or specific instruction on how to assess might optimize the tool's effectiveness.

It is important for the interpretation of the results that the quality of collaboration in this study was measured by the relative frequency of communicative activities that are known to contribute to effective collaboration. However, although these activities represent characteristics related to the quality of collaboration as defined by Saab et al. (2007), these characteristics can also be observed in other collaborative behavior aside from communication. For instance, turn-taking behavior could demonstrate respect. Nevertheless, as the chat was permanently available, a likely assumption is that decisions regarding a task would have been discussed within the chat. As a direction for future research, though, it would be interesting to understand what behaviors students took into account when they assessed each other and how the tool affected behaviors that did not become manifest in students' communication, such as the contribution individual team members made to a task, their approach to tackling a task, and turn-taking behavior.

In conclusion, the current study demonstrated the effectiveness of an intervention that used instruction and joint reflection by means of self- and peer assessment and goal setting to improve the quality of students' collaboration and their knowledge acquisition. From the results, it can be concluded that this combination is successful. This study makes a contribution to the research in this area, as studies measuring the direct effect of interventions on the quality of students' collaboration are scarce. In addition, the tool in this study provides a practical, effective, and time-efficient solution for schools that can help support students' collaboration skills. An advantage here is that the tool was designed for generic application. More concretely, it can be used in other CSCL environments, and also as a supplement to face-to-face collaboration. An interesting focus of future studies could be to investigate whether the observed improvement in the quality of collaboration appears to be sustainable (i.e., whether students have internalized the desired behavior), either in a similar CSCL environment or in a face-to-face or workplace setting.

Acknowledgements The authors wish to thank Emily Fox for her contribution to the article. This work, with project number 409-15-209, was financed by the Netherlands Organisation for Scientific Research (NWO), TechYourFuture, and Thales Nederland BV.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 932–945. https://doi.org/10.1037/0278-7393.23.4.932.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12*(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1.

Chen, J., Wang, M., Kirschner, P. A., & Tsai, C.-C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research, 88*(6), 799–843. https://doi.org/10.3102/0034654318791584.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145–182. https://doi.org/10.1016/0364-0213(89)90002-5.

Christoffels, I., & Baay, P. (2016). 21ste-eeuwse vaardigheden in het mbo. 'Vaardig' voor de toekomst. (21st century skills in middle vocational Education 'Skillful' for the future). (pp. 1–8). 's Hertogenbosch (the Netherlands): Expertisecentrum Beroepsonderwijs.

de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in stem education: The go-lab federation of online labs. *Smart Learning Environments, 1*(1), 3. https://doi.org/10.1186/s40561-014-0003-6.

Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior, 27*(3), 1068–1078. https://doi.org/10.1016/j.chb.2010.05.018.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61–91): Heerlen: Open Universiteit Nederland.

Dillenbourg, P., Baker, M. J., Blaye, A., & O'Malley, C. (1995). The evolution of research on collaborative learning. In P. Reiman, & H. Spada (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning* science (pp. 189–211). Oxford (UK): Elsevier.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education, 24*(3), 331–350. https://doi.org/10.1080/03075079912331379935.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x.

Gabelica, C., van den Bossche, P., Segers, M., & Gijselaers, W. (2012). Feedback, a powerful lever in teams: A review. *Educational Research Review, 7*(2), 123–144. https://doi.org/10.1016/j.edurev.2011.11.003.

Gijlers, H., Saab, N., van Joolingen, W. R., de Jong, T., & van Hout-Wolters, B. H. A. M. (2009). Interaction between tool and talk: How instruction and tools support consensus building in collaborative inquiry-learning environments. *Journal of Computer Assisted Learning, 25*(3), 252–267. https://doi.org/10.1111/j.1365-2729.2008.00302.x.

Gonzalez de Sande, J. C., & Godino Llorente, J. I. (2014). Peer assessment and self-assessment: Effective learning tools in higher education. *International Journal of Engineering Education, 30*(3), 711–721.

Hattie, J., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. npj. *Science of Learning, 1*, 16013. https://doi.org/10.1038/npjscilearn.2016.13.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*(4), 1037–1065. https://doi.org/10.1016/j.compedu.2006.01.004.

Järvelä, S., Kirschner, P. A., Hadwin, A. F., Järvenoja, H., Malmberg, J., Miller, M., et al. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning, 11*(3), 263–280. https://doi.org/10.1007/s11412-016-9238-2.

Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review, 19*(1), 15–29. https://doi.org/10.1007/s10648-006-9038-8.

Johnston, L., & Miles, L. (2004). Assessing contributions to group assignments. *Assessment & Evaluation in Higher Education, 29*(6), 751–768. https://doi.org/10.1080/0260293042000227272.

Jones, D. W. (1996). Empowered teams in the classroom can work. *The Journal for Quality and Participation, 19*(1), 80–86.

Kirschner, P. A., Kreijns, K., Phielix, C., & Fransen, J. (2015). Awareness of cognitive and social behaviour in a CSCL environment. *Journal of Computer Assisted Learning, 31*(1), 59–77. https://doi.org/10.1111/jcal.12084.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts–a conceptual analysis. *Educational Psychology Review, 18*(2), 159–185. https://doi.org/10.1007/s10648-006-9007-2.

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology, 41*(3), 525–536. https://doi.org/10.1111/j.1467-8535.2009.00968.x.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(4), 423–458. https://doi.org/10.3102/00346543066004423.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449–521. https://doi.org/10.3102/00346543071003449.

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing, 18*(1), 30–43. https://doi.org/10.1016/j.jslw.2008.06.002.

Mercer, N. (1996). Cooperation and social context in adult-child interactionthe quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*(4), 359–377. https://doi.org/10.1016/S0959-4752(96)00021-7.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43–59. https://doi.org/10.1007/bf02505024.

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers & Education, 61*, 59–76. https://doi.org/10.1016/j.compedu.2012.08.013.

Onderwijsraad. (2014). *Een eigentijds curriculum*. Den Haag: Onderwijsraad.

Phielix, C. (2012). *Enhancing collaboration through assessment & reflection* (Unpublished doctoral dissertation). Utrecht University, Utrecht, the Netherlands.

Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL -environment: Effects of peer feedback and reflection. *Computers in Human Behavior, 26*(2), 151–161. https://doi.org/10.1016/j.chb.2009.10.011.

Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior, 27*(3), 1087–1102. https://doi.org/10.1016/j.chb.2010.06.024.

Prinsen, F., Terwel, J., Volman, M., & Fakkert, M. (2008). Feedback and reflection to promote student participation in computer supported collaborative learning: A multiple case study. In R. M. Gillies, A. F. Ashman, & J. Terwel (Eds.), *The teacher's role in implementing cooperative learning in the classroom* (pp. 132–162). Boston, MA: Springer US.

Quinton, S., & Smallbone, T. (2010). Feeding forward: Using feedback to promote student reflection and learning – A teaching model. *Innovations in Education and Teaching International, 47*(1), 125–135. https://doi.org/10.1080/14703290903525911.

Renner, B., Kimmerle, J., Cavael, D., Ziegler, V., Reinmann, L., & Cress, U. (2014). Web-based apps for reflection: A longitudinal study with hospital staff. *Journal of Medical Internet Research, 16*(3), e85. https://doi.org/10.2196/jmir.3040.

Renner, B., Prilla, M., Cress, U., & Kimmerle, J. (2016). Effects of prompting in reflective learning tools: Findings from experimental field, lab, and online studies. *Frontiers in Psychology, 7*(820). https://doi.org/10.3389/fpsyg.2016.00820.

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14*(2), 201–241. https://doi.org/10.1207/s15327809jls1402_2.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning, 4*(1), 69–92. https://doi.org/10.1007/s11412-008-9054-4.

Saab, N., van Joolingen, W. R., & van Hout-Wolters, B. H. A. M. (2007). Supporting communication in a collaborative discovery learning environment: The effect of instruction. *Instructional Science, 35*(1), 73–98. https://doi.org/10.1007/s11251-006-9003-4.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/bf00117714.

Saleh, M., Lazonder, A. W., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science, 33*(2), 105–119. https://doi.org/10.1007/s11251-004-6405-z.

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2018). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. Computers in Human Behavior https://doi.org/10.1016/j.chb.2018.05.004

Stenning, K., McKendree, J., Lee, J., Cox, R., Dineen, F., & Mayes, T. (1999). Vicarious learning from educational dialogue. *Proceedings of the Computer Support for Collaborative Learning (CSCL) Conference*, 1999. In C. M. Hoadley, & J. Roschelle (Eds.), *International society of the learning sciences* (pp. 341–347). Palo Alto, California: Stanford University.

Strijbos, J., Ochoa, T. A., Sluijsmans, D. M. A., Segers, M. S. R., & Tillema, H. H. (2009). Fostering interactivity through formative peer assessment in (web-based) collaborative learning environments. In C. Mourlas, N. Tsianos, & P. Germanakos (Eds.), *Cognitive and emotional processes in web-based education: Integrating human factors and personalization* (pp. 375–395). Hershey, PA: IGI Global.

ter Vrugte, J., & de Jong, T. (2017). Self-explanations in game-based learning: From tacit to transferable knowledge. In P. Wouters & H. van Oostendorp (Eds.), *Instructional techniques to facilitate learning and motivation of serious games* (pp. 141–159). Cham: Springer International Publishing.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249–276. https://doi.org/10.2307/1170598.

van der Linden, J., Erkens, G., Schmidt, H., & Renshaw, P. (2000). Collaborative learning. In P. R. J. Simons, J. van der Linden, & T. Duffy (Eds.), *New learning* (pp. 37–54). Dordrecht: Springer.

van Popta, E., Kral, M., Camp, G., Martens, R. L., & Simons, P. R. J. (2017). Exploring the value of peer feedback in online learning for the provider. *Educational Research Review, 20*, 24–34. https://doi.org/10.1016/j.edurev.2016.10.003.

Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: A meta-analysis. *Educational Psychology Review, 29*(3), 477–511. https://doi.org/10.1007/s10648-016-9361-7.

Wecker, C., & Fischer, F. (2014). Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning. *Computers & Education, 75*, 218–228. https://doi.org/10.1016/j.compedu.2014.02.016.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education, 46*(1), 71–95. https://doi.org/10.1016/j.compedu.2005.04.003.

Affiliations

Elise H. Eshuis[1] • Judith ter Vrugte[1] • Anjo Anjewierden[1] • Lars Bollen[2] • Jakob Sikken[1] • Ton de Jong[1]

Judith ter Vrugte
j.tervrugte@utwente.nl

Anjo Anjewierden
a.a.anjewierden@utwente.nl

Lars Bollen
larsbollen@gmail.com

Jakob Sikken
j.sikken@utwente.nl

Ton de Jong
a.j.m.dejong@utwente.nl

[1] Department of Instructional Technology, Faculty of Behavioral, Management, and Social Sciences, University of Twente, Enschede, The Netherlands Q1

[2] DB Schenker, Essen, Germany