JrnlID 11412_ArtID 9307_Proof# 1 - 22/10/2019

International Journal of Computer-Supported Collaborative Learning
https://doi.org/10.1007/s11412-019-09307-0

It is not either or: An Initial Investigation into combining collaborative and individual learning using an ITS

Jennifer K. Olsen[1] • Nikol Rummel[2,3] • Vincent Aleven[2]

Received: 18 January 2019 / Accepted: 16 September 2019

Abstract
Research on Computer-Supported Collaborative Learning (CSCL) has provided significant insights into why collaborative learning is effective and how we can effectively provide support for it. Building on this knowledge, we can investigate when collaboration is beneficial to support learning. Specifically, collaborative and individual learning are often combined in the classroom, and it is important for the CSCL community to understand when a combination is beneficial compared to individual or collaborative learning alone. Before investing significant work into discovering these details, an initial investigation is needed to determine if there may be any value in a combination. In this study, we compared a combined condition to individual or collaborative-only learning conditions using an intelligent tutoring system for fractions. The study was conducted with 382 4th and 5th grade students. Students across all three conditions had significant learning gains, but the combined condition had higher learning gains than the other conditions. However, this difference was restricted to the 4th grade students. By analyzing the hints and errors of students over time from process data, we found that students in the combined condition tended to make fewer errors both when working collaboratively and individually, and asked for fewer hints than the students in the other conditions. Students who collaborated (collaborative and combined conditions) also reported having higher situational interest in the activity. By finding support for the effectiveness of combining collaborative and individual learning, this paper opens a broader line of inquiry into how they can effectively be combined to support learning.

Keywords Collaboration · Problem-solving · Erroneous examples

✉ Jennifer K. Olsen
jennifer.olsen@epfl.ch

Extended author information available on the last page of the article

Introduction

Research on Computer-Supported Collaborative Learning (CSCL) has provided significant insights into why collaborative learning is effective (Chi and Wylie 2014) and how we can effectively provide support for it (Vogel et al. 2017). Building on this knowledge, we can continue the investigation into when collaborative learning is beneficial for supporting learning. In the recent conceptual paper by Wise and Schwarz (2017), they engage in a conversation around the question whether it would be beneficial for the CSCL community to research "if, when and for what ends" collaboration is beneficial (Wise and Schwarz 2017, p. 433). We argue that as part of the when question, it is important not just to understand when collaborative learning can be beneficial by itself but also why and how it may be productively combined with individual learning. In this paper, we present an initial investigation into the combination of collaborative and individual learning. This first, and necessary, step provides a basis for the value that further research on a combination may provide before investing significant work into the why, when, and how a combination can be beneficial.

When investigating the value of a combination between individual and collaborative learning, it is important to consider the role that each task plays within the lesson as a whole. As Wise and Schwarz (2017) point out, research in CSCL has repeatedly demonstrated the benefits of collaborative learning (Chen et al. 2018; Jeong et al. 2019; Lou et al. 2001; Slavin 1996). The research has primarily focused on understanding the processes that students engage in while collaborating (Chi and Wylie 2014) as well as how we can better support collaboration for improved student learning (Fischer et al. 2013a; Lou et al. 1996, 2001; Magnisalis et al. 2011; Rummel et al. 2008). However, collaboration is often not used in isolation in the classroom. For example, many well-known scripts, such as Jigsaw and ArgueGraph (Aronson 1978; Dillenbourg 2002), combine collaborative learning with an individual phase. Integrative scripts are collaboration scripts that incorporate multiple social levels (e.g., individual, group, whole class) to support student learning (Dillenbourg 2004; Dillenbourg and Tchounikine 2007). Some scripts use an individual phase (or phases) to prepare students for a productive collaboration phase (Dillenbourg 2002; Diziol et al. 2007). Other designs, such as productive failure, allow students to work collaboratively or individually on complex problems to prime them for whole class direct instruction (Kapur 2010; Kapur 2014). In these cases, it is important to extend the frameworks of collaboration support (Rummel 2018) to include a dimension that discusses combining social levels to be able to fully capture the benefits and uses of collaboration within these integrative scripts. In this paper, we aim to provide an initial investigation into a combination of collaborative and individual learning and if it shows potential benefits compared to either social level alone.

Related work

It is difficult to find studies where a combination of collaborative and individual learning was used in the literature because conditions with combinations are often referred to as "collaborative conditions" without distinguishing that students also have a chance to work individually. However, there are some examples where a combined condition has been explicitly explored. For instance, Celepkolu et al. (2017) compared students working on paired programming to students who had individual time to assess the problem before working collaboratively. Although they found that the combined condition was better than just the paired programming, the conditions did not have the same number of phases, which may have impacted the results. Additionally, Wang et al. (2011) found that a combined condition around brainstorming led to outcomes greater than those working individually only and less than those working collaboratively only. In this study, the students in the combined condition were doing the same activity both collaboratively and individually with the individual portion being a short initial brainstorm before mainly engaging in collaborative brainstorming. The combination was not intended to target different knowledge and learning processes but instead to approach the same task in different ways. However, the alignment of the learning method and target knowledge may be important for a combination to be successful (Mullins et al. 2011).

In other words, a combination of collaborative and individual learning may be beneficial for learning as the social levels may support different types of knowledge acquisition. The Knowledge Learning and Instruction (KLI) framework (Koedinger et al. 2012) proposes that different types of skills have different levels of complexity and that there may be an alignment between the complexity of instructional methods and those of skills in terms of supporting knowledge efficiently. For example, collaborative learning supports students in giving and receiving of explanations and co-constructing knowledge (Hausmann et al. 2004), which may help students develop a deeper conceptual understanding (Teasley 1995) around rules and principles. On the other hand, when students are working individually, they are able to optimize the pace of the work to develop fluency and memory necessary for certain skills (Frank and Gibson 2011; Koedinger et al. 2012) since students are not sharing tasks with a partner and do not necessarily have to pause to explain their actions. This alignment of instructional design complexity and skill complexity may help to explain why collaborative learning has not always been found to foster greater learning gains compared to individual learning (Lou et al. 2001).

However, previous work has found conflicting evidence regarding the hypothesized complementary strengths of collaborative and individual learning when aligned with conceptual and procedural tasks respectively (Mullins et al. 2011; Olsen et al. 2014a; Olsen et al. 2016). While some studies have found that students working on conceptually oriented tasks perform better when collaborating compared to working individually and the opposite on procedurally oriented tasks (Mullins et al. 2011), other studies have not found a significant difference in learning performance between those working collaboratively or individually for either type of task (Olsen et al. 2014b; Olsen et al. 2016). These differences in findings may be due to differences between participants in the studies. If the students did not have the necessary prior knowledge needed to engage in the intended knowledge acquisition, they may have engaged in additional learning processes to gain the needed prior knowledge as they solved the problem. For example, if the students did not enter with prior conceptual knowledge, they may have spent time focused on gaining this knowledge even when working on the procedurally oriented problems. Rittle-Johnson et al. (2001) claim that both conceptual and procedural knowledge are important for learning and may interrelate. In other words, with the hypothesized alignment, the students collaborating would be more successful in gaining the conceptual knowledge than the students working individually (Olsen et al. 2017) and, hence, may be overall more successful even when engaged in a procedural task. In this case, we may not observe the hypothesized alignment between the collaborative and individual learning and the knowledge acquisition. In such a situation, students may benefit from getting to practice both types of knowledge.

Outside of the hypothesized benefits from the alignment of skills and instructional methods, the combination of collaborative and individual learning has both benefits and drawbacks that may impact its effectiveness. When students are working in both contexts, they may benefit just from getting to work both individually and collaboratively. Previous research has found that when students spend a medium amount of time in groups, compared to high or low, there is a trend towards greater learning effects (Springer et al. 1999). Thus, students may benefit from the variation provided by both the collaborative and individual learning as suggested by variation theory (Ling and Marton 2012). On the other hand, it is also possible that switching between individual and collaborative learning adds overhead to the learning process, which could have a negative impact on the student performance that outweighs the benefits of a combination. In this case, the transition between the collaborative and individual learning takes time that could be spent on instruction. As the instruction time is decreased, student learning may be negatively impacted (Fraser et al. 1987).

In the study reported in this paper, we conducted an initial investigation into whether a combination of collaborative and individual learning is more efficient for student learning than either alone. We had 4th and 5th grade students work on both conceptual and procedural knowledge through erroneous example problems and procedural problem sets respectively for fractions using a collaborative intelligent tutoring system (CITS). This study, along with previous work, provides a foundation in which to begin exploring the details around when and how a combination of collaborative and individual learning can be effective.

Learning context and hypotheses

To design a combined collaborative and individual condition, we created learning activities that, based on theoretical grounds such as those discussed above in the KLI framework (Koedinger et al. 2012), may align with the strengths of individual and collaborative learning. Specifically, we used erroneous example problems for collaborative learning and tutored procedural problem-solving for individual learning. For the erroneous example problems, the students were asked to not only study the problem, as is typical with worked examples, but to engage in the problem-solving process by identifying, fixing, and writing how to prevent the error. We chose to have the students work on erroneous example problems collaboratively to align with research on example-based learning that has shown that learning from both correct and erroneous worked examples is successful for supporting learning (McLaren et al. 2012; Renkl 2005; Tsovaltzi et al. 2010; Van Gog et al. 2019). Specifically, examples, compared to problem-solving, allow students to focus on underlying rules and principles compared to memorizing facts and procedures (Atkinson et al. 2000). In other words, example problems support the acquisition of the higher complexity skills as defined in the KLI framework (Koedinger et al. 2012). The use of erroneous examples specifically can help to foster reflection and more fruitful explanations compared to standard problem-solving (Isotani et al. 2011; Siegler 1995; Tsovaltzi et al. 2009). Given that erroneous examples foster the acquisition of higher complexity skills, within the KLI framework, we would hypothesize that collaboration would be beneficial for supporting this knowledge acquisition (Koedinger et al. 2012). We find evidence for this alignment in that prior research has shown that when students study worked examples collaboratively, they tend to avoid shallow processing, ask for fewer hints, and spend more time on explanations than when working individually (Hausmann et al. 2009; Hausmann et al. 2008a; Hausmann et al. 2008b). When students are able to collaborate around erroneous example problems, the sense-making that they engage in through their collaborative interactions (Chi and Wylie 2014) may be beneficial given that the erroneous example problems bring focus to the rules and principles used within the problems (Koedinger et al. 2012).

On the other hand, we chose to have the students work individually on tutored problem-solving to align the fact and rule memorization and fluency that tutor problem-solving supports with the memory and fluency building that working individually may foster as hypothesized by the KLI framework (Koedinger et al. 2012). Tutored problem-solving supports student learning of procedures through step-by-step support (VanLehn 2006) that focuses the attention of students on the facts and rules that form the procedure. Given the fact-based nature of these problems, tutored problem-solving can often be used to support students in building their memory and fluency of the problem-solving steps (Mullins et al. 2011), which within the KLI framework would be beneficially supported through less complex learning designs, such as individual learning (Koedinger et al. 2012). Students working individually may support memory and fluency learning due to the ability for them to work at their own pace. When working individually, students do not have to divide tasks with another student or stop often to discuss a problem step, which likely allows each student to get more practice. This alignment of tutored problem-solving and individual learning may foster students to take advantage of the fluency support implicit within the problem (Koedinger et al. 2012) to develop procedural knowledge (Anderson 1983).

Hypotheses

For our study, we wanted to investigate both the overall learning gains between conditions through the use of pretest and posttests as well as understand these findings by analyzing variables collected during the learning process and analyzing the students' interest in the task. Our main hypothesis centered on the effectiveness of combining students working collaboratively on erroneous example problems and individually on procedurally oriented problems compared to students either working collaboratively or individually on both problem types. We hypothesized that the students who have a combination of collaborative and individual learning (i.e., combined condition) will have higher learning gains than students who carry out the same set of activities while working only collaboratively or only individually (H1). This hypothesis is based in the reasoning outlined in the previous sections.

To help explain any overall difference found between conditions, we additionally investigated process variables including errors students made and hints they received from the system. Past research has found that collaborating students tend to make fewer errors and ask for fewer hints than students working individually (Hausmann et al. 2008a; Hausmann et al. 2008b). Within a collaboration, the students can discuss the problem before submitting an answer allowing them to engage in a sense-making process (Chi and Wylie 2014) leading to needing less system support. While working with the fractions CITS, we hypothesized that students in the combined condition would make fewer errors (H2a) and request fewer hints (H3a) with a greater decrease in errors and hints over time than those working only individually even when the students in the combined condition are working individually because the students may benefit from the previous collaboration. We also hypothesize that the students in the combined condition will not make more errors (H2b) or request more hints (H3b) with the same decreased rate of errors and hints over time compared to those in the collaborative condition, even when working individually because of possible learning during the collaborative phase that is carried over to the individual. Together, these process analyses could provide insights into how the different students performed while working with the tutor.

Finally, we investigated the interest that the students had in the task and how this may have differed between conditions. Discussions that happen during collaboration can potentially support the students' social goals (e.g., responsibility goals, popularity goals) and make them feel more connected to their group members, which can increase their motivation for the activity (Rogat et al. 2013) and increase the desire to continue working on the task. Specifically, situational interest in the task, which is interest that arises due to a response to the factors in the environment (Linnenbrink-Garcia et al. 2010), can increase when a task involves collaboration. For the situational interest in the fractions CITS, we hypothesized that students who have a chance to work collaboratively (i.e., combined and collaborative conditions) will have more situational interest in the activity than students that only work individually (H4) (Linnenbrink-Garcia et al. 2010) due to the opportunity and anticipation of getting to work with a peer.

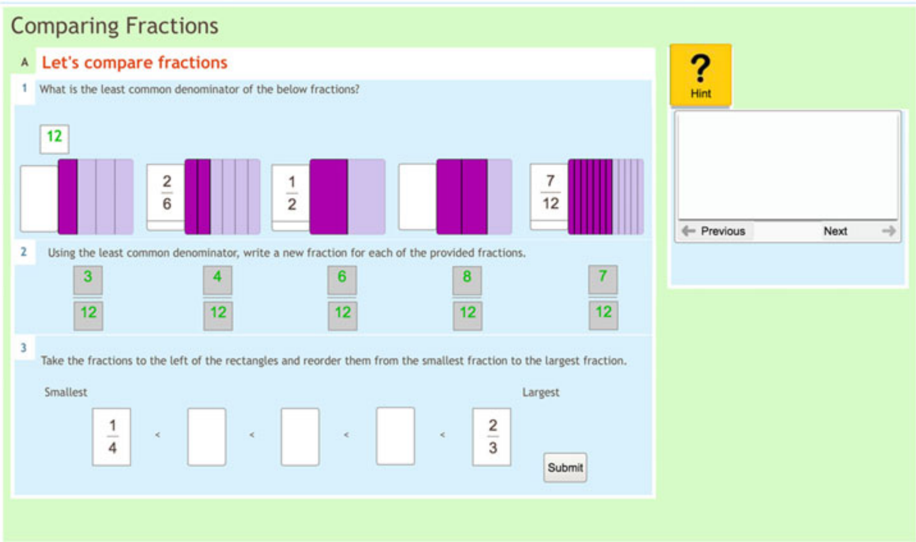
Tutor design

In our experiment, we supported students through fractions intelligent tutoring systems (ITSs) as platforms for our research. We chose to use ITSs for two primary reasons: best practices for individual learning and to prevent students going in the wrong direction when collaborating. ITSs have been shown to be beneficial for student learning (Kulik and Fletcher 2015; Ma et al. 2014) and are effective by providing cognitive support for students as they work through problem-solving activities. This cognitive support comes in the form of step-level guidance, namely, an interface that makes all steps visible, error feedback, and on-demand hints (VanLehn 2006). ITSs may be successful through their ability to create an individualized learning environment for each student where they can work at their own pace. Within previous research, ITSs have been found to improve learning by as much as one standard deviation (Anderson et al. 1990), indicating their effectiveness at supporting individual learning and an appropriate choice for supporting problem-solving tasks.

Although the majority of ITSs have been developed for individual use, the integration of collaboration within an ITS, in prior studies, has effectively supported learning (Baghaei and Mitrovic 2005; Diziol et al. 2010; Olsen et al. 2016). Support for collaboration can be directly embedded into the tutor to support the students both cognitively and socially. The cognitive supported provided through the standard ITS features prevent the students from spending too much time working in the incorrect direction when collaborating. Additionally, it provides students with correctional feedback and common grounding in which to focus their discussions (Olsen et al. 2018b). However, because collaboration does not occur spontaneously, the following section provides additional information on how we designed the collaborative condition to align with state-of-the-art support for collaborative learning.

Informed by prior work on fractions tutors (Olsen et al. 2014a; Olsen et al. 2016), we developed a new ITS for three fractions units: equivalent fractions, least common denominator, and comparing fractions. The ITS versions were built with the Cognitive Tutoring Authoring Tools (CTAT), extended to support collaborative tutors (Aleven et al. 2015; Olsen et al. 2014a). For each of the three units, we created both procedurally oriented activities (see Fig. 1) and erroneous examples (see Fig. 2). Further, we created both individual and collaborative versions of both types of activities, for use in different conditions. For each unit, there were eight problems of the same type.

At the beginning of the experiment, the students were asked to complete a tutorial that consisted of six problems. These problems introduced the concept of the unit for each of the three representations (i.e., pie chart, rectangle, and number line). By going through the tutorial, the students are able to learn what the different interaction types are that they will have with the



Q2 **Fig. 1** An example of a procedurally-oriented tutor. The students are guided in finding a common denominator for comparing fractions

interface and how the interface will provide feedback. In addition, when the students are collaborating, the tutorial allows them to understand how their interactions are shared within the interface with their partner.

The procedurally oriented problems were designed to provide students with practice completing the steps needed to solve the problem type within the unit. Procedural knowledge is the ability to perform steps and actions in sequence to solve a problem (Rittle-Johnson et al. 2001). For example, Fig. 1 shows a problem asking students to compare fractions. The

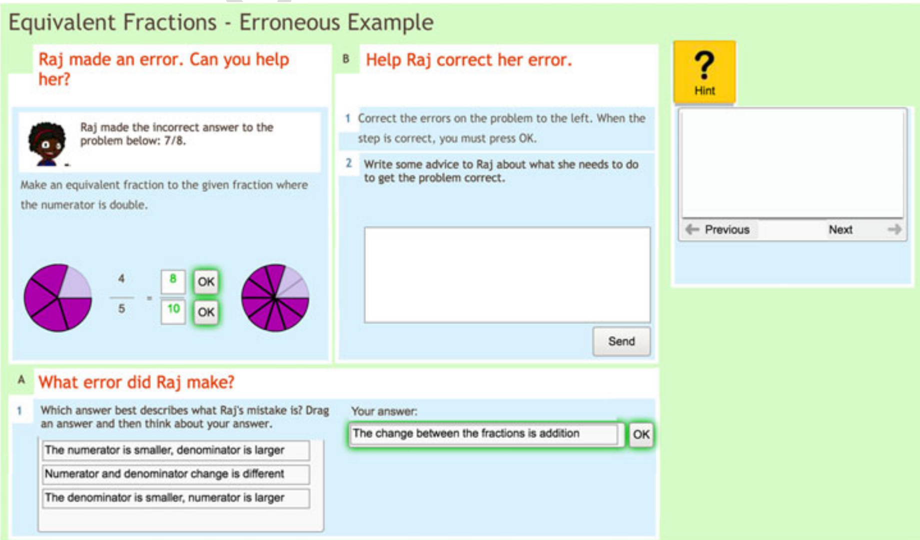


Fig. 2 An example if an erroneous example problem. The students were instructed to find and correct the error and provide advice for solving the problem

students are first asked to find the least common denominator for all of the fractions. They are then asked to convert the fractions into equivalent fractions using the least common denominator. After this step, since the fractions can now be more easily compared, the students are asked to order the fractions from smallest to largest.

The second problem type developed for each unit was erroneous example problems. The erroneous example problems were designed to address the errors that often arise within the procedural problems. To find these errors, we analyzed log data collected during previous experiments. For each unit, we found the common errors that students were making across problems and developed problems to directly address the errors. For the erroneous example problems, each problem had a fictitious student that had made an error when solving the problem (see Fig. 2). By providing a student in each problem, the students solving the problem could feel more connected and invested in helping the student (Lester et al. 1997). When beginning the problem, the students were first asked to identify the error that the fictitious student had made when solving the problem. After identifying the error, the students were asked to correct it (i.e., provide the correct answer, with feedback from the tutor) and write to the fictitious student to provide them with advice on what they could do better the next time.

Collaboration support

For each of the units and problem types covered in the fractions CITS, problem sets were designed for both individual use and collaborative (dyadic) use. The individual and collaborative problem types were designed to have the same format and to go through the same set of steps. The students also had the same access to error messages and on demand hints for the tutor. The individual and collaborative tutor types did differ in the social support that was provided to the students in the collaborative tutor through an embedded collaboration script and the sharing of information across tutor interfaces between partners. Controlling for the differences between the collaborative and individual tutors allowed us to make comparisons between the social levels rather than different outcomes being due to task differences.

However, because we know from the CSCL literature that for collaborative learning to be effective, support needs to be designed for the collaborative process (Kollar et al. 2006), we could not use identical tutoring systems. Instead, based off of best CSCL practices, we designed an embedded collaboration script to support the students in their collaborative interactions. Following the framework proposed by Rummel (2018), we designed our collaboration script with the goal of supporting the students' interactions in a way that supports the acquisition of the domain knowledge. To meet this goal, we provided the support during the collaboration with a fixed implementation delivered through the ITS. We chose a fixed implementation given the amount of time that the students would be engaged in any single problem set providing little time to adapt. Because there was already cognitive support provided in the ITS, the CSCL script focused on the social support at the step level (to align with the provided cognitive support). Based upon the desired dimensions of collaboration support, we chose collaborative script features that could be applied at the step level and were proven ways of supporting the social aspects of the collaboration. These features are outlined in more detail below.

The collaboration tutor used synchronized, networked collaboration. Each student sat at their own computer and had a shared, but differentiated view of the problem. The students were able to see their partner's actions before being checked by the tutor, which allowed them to have a discussion around the answer. However, because the students also each had their own screen, each student was able to receive different information or take different actions on the problem, which allowed us to implement the collaboration script delivered through the system. For example, for making equivalent fractions, one student could be put in charge of the numerators while the other in charge of the denominators (see Fig. 3). To be able to make a full fraction, each student would have to interact with the problem. Although all of the collaboration was designed for students to be at separate computers, the actual features of the collaboration script were designed to correspond with the learning objectives of the individual problems (Kollar et al. 2006). Within the fractions CITS, we used two main collaboration features to support learning: cognitive group awareness and group accountability through the use of separate information and actions. By using these features together, the tutor could better engage each member of the dyad in the problem-solving to avoid free riding.

The first form of collaborative support that was implemented in the fractions CITS was producing group accountability through the use of separate information and actions. Individual accountability has been argued to be essential for group work to be successful (Slavin 1989). Individual accountability within a group provides each student with a sense of responsibility for the task completion. Within the fractions CITS, we support individual accountability by giving each student within a dyad separate information and actions in a way that they have to understand what their partner is doing to finish their part. In this way, we encourage social interactions and cognitive exchanges between the students. Within the procedurally oriented

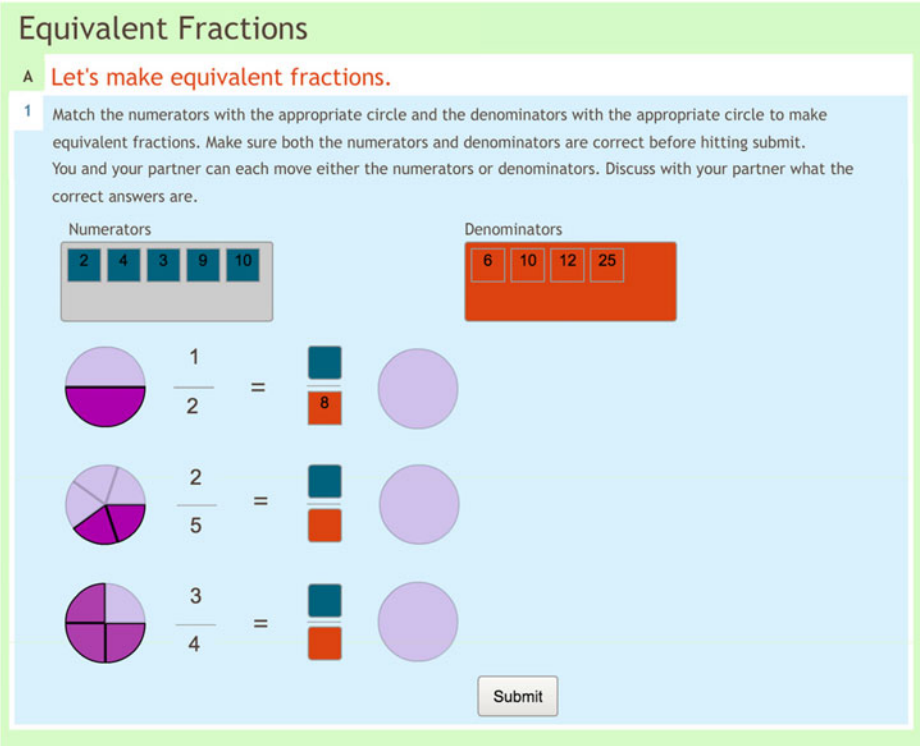


Fig. 3 An example of division of responsibilities. One student in a collaborating dyad is responsible for selecting numerators, the other for selecting denominators. Each student can see the numerators and denominators but can only interact with one set (i.e., drag them into the open slots)

problems, on some steps within the problem, students would only be able to interact with half of the available answer choices. For example, within the procedural equivalent fractions, one student would be able to move the numerators while the other student could only move the denominators (see Fig. 3). The correct choice for the numerator depends upon the choice of the denominator and vice versa.

The second form of collaborative support that was used within the fractions CITS was cognitive group awareness (Dehler et al. 2011; Janssen and Bodemer 2013). Cognitive group awareness can be defined as providing information to the group members about the other group members' knowledge, information, or opinions. Within collaborative learning, providing cognitive group awareness tools, which explicitly display a student's knowledge to the group, to students has been found to be effective in supporting their learning (Janssen and Bodemer 2013). When students are more aware of their group members' expertise, they are better able to make use of their partners' knowledge and to coordinate the task. Additionally, by making the knowledge and opinions of the different group members more salient, the students can be more aware of when they have differing answers, which can lead to more discussion. We chose to use cognitive group awareness to make disagreements on the tutoring steps more explicit leading to discussions between the students instead of quickly passing by the question. This disagreement and discussion leads to the students each updating their mental models and strengthening correction connections through explanation (Schwarz et al. 2000). Within the fractions CITS, cognitive group awareness was supported by giving the students an opportunity to answer a step individually before working on the step as a group (see Fig. 4). After each student enters an answer, the individual answers are shared with the whole group so that each group member can see what their partner answered. The group is then asked to choose a group answer. Only on the group answer does the system provide correctness feedback. By supporting cognitive group awareness through this method, students are provided with an equal opportunity to express their opinion on the answer before getting feedback from the system, which can lead to more conversations between the students, especially when their answers are different.

Fig. 4 To support cognitive group awareness, the students are first asked to answer the question individually (top) before answering as a group (bottom)

Like previous CITS, the two different collaboration features are embedded directly into the system to provide support for the social dynamics of the students as they work through the different problem sets. However, unlike the collaboration support provided in previous CITS, our focus was on supporting a balanced collaborative dynamic rather than peer tutoring (Walker et al. 2011). Given the differences in dynamics between the students, the collaboration support also needed to differ. In the peer tutoring paradigm, there is no concern of one student taking over and doing all of the work because the collaborating students are not equal. In this case, we provided the students with support to provide accountability with both parties. Additionally, the support could be given equally to both students since, overall, they were in the same role. In peer tutoring, the support provided must be different between the tutor and tutee as they engage in very different tasks in the learning process. With our support primarily focused on the social support, the students were able to step in to provide more of the cognitive support (with the ITS features providing support when needed).

Methods

Research Design

To test our hypotheses, we conducted a study with a quasi-experimental, between subject design where condition was randomly assigned to the classroom with variables measured at the individual or dyad level. At the class level, students were randomly assigned to one of three conditions: combined, collaborative, or individual. In the combined condition, the students worked collaboratively on the erroneous example problems and individually on the procedural problem-solving activities. In the other conditions, students either worked collaboratively on both types of problems or individually on both types of problems. For the tutor, we controlled for time on task, giving all students the same amount of time to complete a problem set.

Participants

The quasi-experimental study was conducted in a classroom setting with 382 4th and 5th grade students from 18 classrooms (7 fourth grade and 11 fifth grade), 12 math teachers, and five school districts. Seven classes were assigned to the combined condition, 6 classes to the collaborative only condition, and 5 classes to the individual only condition. As the study was conducted at the end of the school year, both 4th and 5th grade students had experience with fraction concepts but only the 5th grade students had learned the concepts covered within the units of our fractions tutor.

Experimental procedure

The study took place during the students' regular class periods. All students worked with the fractions CITS described above. In all three conditions, the erroneous example problems for a unit came before the procedural problems to allow the students to address errors before getting more instruction through the procedural problems sets (Renkl and Atkinson 2003). Students in all conditions completed one unit each day; they switched from the erroneous example problems to the procedural problem-solving activities half way through class. Within each class, all of the students were instructed to switch problem sets at the same time. Because the time-on-task was constant for all conditions within each unit, each student finished a different number of problems.

The study ran across five class periods of 45 min each. On the first day, the students took the pretest individually. At the beginning of the second day, the students took a short tutorial either individually or in groups, depending upon how they would work for the erroneous example problems, that gave some instruction on how to interact with the tutor. The students then worked with the tutor for the next three days in their condition. On the fifth day, the students took a posttest individually and answered a short survey to gauge their situational interest when working with the tutor.

Within each class, teachers paired their students based on who would work well together and had similar math abilities to avoid extreme differences that could hinder collaboration. Students worked with the same partner as much as possible and only changed partners due to absenteeism. If a student's partner was absent in the collaborative conditions, the student would be paired with another student working in the same condition for the remainder of the study. When students were collaborating, they each sat at their own computer. The students within each collaborating dyad were instructed to sit next to each other and were able to communicate through speech. This speech was recorded for each student individually using a tablet.

Dependent measures

In this study, we collected pretest and posttest measures, tutor log data, and situational interest measures. For the pretest and posttest measures, we assessed students' fractions knowledge at two different time points using two equivalent test forms in counterbalanced fashion. The tests targeted isomorphic problems for both the erroneous and procedurally oriented problem types and were administered on the computer. The tests also had transfer problems for naming, making, adding, and subtracting fractions as these units were not covered within the instruction. Each test had 15 questions, namely, seven erroneous examples, six problem-solving items, and two fractions explanations questions. For each question on the test, the students were able to get a point for each step completed correctly. On the tests there were 81 possible points for the 13 erroneous example and procedural knowledge questions. Within the results, all test scores are reported as a percentage of the total possible points.

During the tutoring session, we also collected log data from the students. The log data contained information around the students' transactions with the tutor, including attempts at solving steps, errors, and hint requests. Because some students were changing social levels between the different problem types, we compared the log data variables within the problem types rather than across all problems. In other words, from the log data we computed the number of errors and hint requests separately for both the erroneous example problems and the procedural problem-solving. For each student, we calculated the number of errors made and hint requests per problem. For errors and hints, there was no limit to the number of errors that could be made or the number of times a student could request a hint (although there were only three distinct hints for each step). Because the students encountered a different unit each day of different difficulties, we could not compare the number of errors and hints between days.

To assess the students' situational interest in the tutoring activity, we had the students answer a brief survey before completing the posttest. The questions were adapted from the Linnenbrink-Garcia et al. (2010) situational interest scale. The scale consists of three separate factors: trigger, maintained feeling, and maintained value. Situational interest can consist of both the attentional as well as the affective reaction to a situation (Mitchell 1993) and can then be divided into two forms: triggered and maintained. The triggered situational interest refers to the initiated interest that is associated with the environment (Linnenbrink-Garcia et al. 2010). On the other hand, the maintained situational interest is the connection that the students make with the material or domain and the realization of its importance. The learning environment can impact the maintained situational interest by allowing the students to make a connection with the knowledge presented (Mitchell 1993). The maintained situational interest provides the link between the triggered situational interest and personal interest, which is interest in a topic than endures over time (Hidi and Renninger 2006; Schraw and Lehman 2001). Maintained situational interest can take a form that is similar to individual interest with both feeling and value components (Linnenbrink-Garcia et al. 2010). The maintained feeling focuses on the enjoyment that the student has had while the value focuses on the perceived meaningfulness of the topic. The situational interest survey consisted of 12 questions, four within each factor. We adapted the questions from asking about the math teacher and math classroom to asking about the time spent on the fractions CITS. Each question was presented to the student on a Likert scale that ranged from one to seven, yielding a score for each factor in the range from 4 to 28. We report the percentage of the maximum score (28) for each of the three factors.

Analysis

To analyze the outcomes of the pretest and posttest measures as well as the log data, we used a multilevel approach to take into account differences between school districts. We used a hierarchical linear model (HLM) with student/dyad at the first level and school district at the second level. For the situational interest measures, we conducted a MANOVA analysis to take into account the dependence between the dependent variables. For all comparisons, the p value was set to .05, and we measured the effect size with Pearson's correlation coefficient (r) where 0.1 is considered a small effect size, 0.3 a medium effect size, and 0.5 a large effect size.

To assess the student process, we analyzed the hints and errors that students made over time. Using the problem number as an indicator of the passage of time within a session, we investigated the temporal change in errors and hints within a problem set. In other words, we analyzed the learning curve, the change of student learning over time, of the students at the problem level. We chose the problem level rather than the step level, which is typical in learning curve analysis, as there were no repeated skills within a problem so all of the steps in a problem were at the same opportunity level. The analysis for the errors and hints were done separately for the erroneous example problems and procedural problems so that dyads could be compared to students working individually. We compared hints and errors across conditions, grades, and problem number using an HLM to account for the nested nature of the data. However, given that there were differences in the number of problems completed between the conditions (for the erroneous example problems the individual completed more problems than the combined, $t(96.35) = 1.89$, $p = .06$, $r = .19$, and the combined more than the collaborative, $t(118.82) = -2.13$, $p < .05$, $r = .19$ and for the procedural problems there was no significant difference between the individual and combined, $t(193) = -0.35$, $p = .73$, and the combined condition completed more problems than the collaborative condition, $t(193) = -4.47$, $p < .05$, $r = .31$), each progressive problem number has fewer student data points.

To test our hypotheses of equivalence, we tested for statistical equivalency using the confidence interval approach (Rogers et al. 1993). Based upon prior studies and the examination of related literature, we used an equivalence interval of ±0.5 for both the errors and hints per question. The equivalence interval indicates the difference between the means that would indicate a meaningful difference. For this study, we calculated a 90% confidence interval. If the confidence interval lied within the equivalence interval, equivalence was concluded.

Results

Out of the 382 students who participated in the study, 75 students were excluded from the analyses because of absenteeism during parts of the study, thus leaving us with a final set of 307 students. Out of the 307 students, 104 were in the collaborative only condition, 83 in the individual only condition, and 120 in the combined condition. There was no significant difference between conditions with respect to the number of students excluded, $F(379,2) = 0.59$, $p = .56$. There was, however, a significant difference in the pretest scores across conditions, $F(2, 304) = 9.4$, $p < .05$. In post hoc analysis using a Bonferroni correction, we found that the collaborative condition was significantly lower than the other two conditions.

Hypothesis H1: learning gains

To investigate whether students learned using our tutor and if there was a difference in learning between the students in the different conditions (H1), we conducted an HLM. At level 1, we modeled the pretest and posttest scores along with the student's grade (4th or 5th) and condition, and at level 2, we accounted for differences that could be attributed to the school district. For the different variables, we chose pretest for the test baseline, combined condition for the condition baseline, and 4th grade for the grade baseline. For each variable, the model includes a term for each comparison between the baseline and other levels of the variable. We did not include dyads as a level because of the added complexity of some students working with no partner (i.e., individuals), some students having one partner, and some students having two partners because of absenteeism. We are aware of non-independence issues such as common fate and reciprocal influence within dyads that may have impacted our results (Cress 2008).

For the learning gains analysis (see Table 1), there was a significant increase in test scores between pretest and posttest across all conditions, $t(301) = 12.56$, $p < .05$, $r = .59$ (see Fig. 5). For the main effects of condition, the combined condition had higher test scores compared to the collaborative condition, $t(297) = -3.12$, $p < .05$, $r = .18$, and marginally higher scores compared to the individual condition, $t(297) = -1.83$, $p = .07$, $r = .11$. Furthermore, the learning gain (pretest to posttest) was higher for the combined condition compared to the collaborative condition, $t(301) = -2.78$, $p < .05$, $r = .16$, and individual condition, $t(301) = -3.56$, $p < .05$, $r = .20$, supporting our hypothesis that the combined condition would be more effective for learning.

For the student's grade level (i.e., 4th v. 5th grade), the 5th grade students had higher test scores compared to the 4th grade students, $t(297) = 2.93$, $p < .05$, $r = .17$ (see Fig. 6, left). Surprisingly, the 4th grade students had higher learning gains than the 5th grade students, $t(301) = -5.53$, $p < .05$, $r = .30$. There was not a significant interaction between grades for either the combined and individual conditions, $t(297) = 0.90$, $p = .37$, or the combined and collaborative conditions, $t(297) = 0.80$, $p = .42$, (see Fig. 6, right) as these differences were captured in the higher order interaction.

Table 1 Percent Correct: means (SD) for test items at pretest/posttest

Grade	4th	5th	Condition Mean
Pretest			
Collaborative	0.22 (0.12)	0.36 (0.26)	0.30 (0.23)
Individual	0.28 (0.15)	0.55 (0.24)	0.42 (0.24)
Combined	0.31 (0.17)	0.53 (0.23)	0.43 (0.23)
Grade Mean	0.27 (0.16)	0.47 (0.26)	
Posttest			
Collaborative	0.46 (0.23)	0.66 (0.24)	0.59 (0.26)
Individual	0.49 (0.21)	0.74 (0.20)	0.62 (0.24)
Combined	0.69 (0.17)	0.68 (0.24)	0.69 (0.21)
Grade Mean	0.57 (0.22)	0.69 (0.23)	

For the three-way interaction between grade, condition, and test time, the slope differences were confined to the 4th grade students between the combined and collaborative conditions, $t(301) = 4.57$, $p < .05$, $r = .25$, and combined and individual conditions, $t(301) = 3.19$, $p < .05$, $r = .18$ (see Fig. 7). These interactions indicated that the combined condition, compared to the other conditions, was more beneficial in terms of learning gains of 4th grade students than those of 5th grade students.

Finally, to investigate if the difference between conditions was different for the 4th graders than the 5th graders because of the initial lower pretest scores allowing the 4th graders to have more room to grow, we ran an HLM using normalized gain scores. Using normalized learning gains also allowed us to account for the differences found in the pretest scores between conditions. The gain scores were calculated as the posttest minus the pretest over one minus the pretest (both the posttest and pretest scores are reported as percentages). Our results confirmed the earlier findings with the combined condition having a higher learning gain than the collaborative, $t(296.55) = -3.05$, $p < .05$, $r = .17$, or individual conditions, $t(297.16) = -3.25$, $p < .05$, $r = .19$. Additionally, the 4th grade students had higher gain scores than the 5th grade students, $t(281.62) = -2.79$, $p < .05$, $r = .16$. Finally, the gain score differences were more pronounced between the combined and collaborative conditions in the 4th grade students than the 5th, $t(276.83) = 3.04$, $p < .05$, $r = .18$, but no significant difference between the individual and collaborative conditions and grade, $t(200.76) = 0.75$, $p = .46$ (Fig. 8). Q3

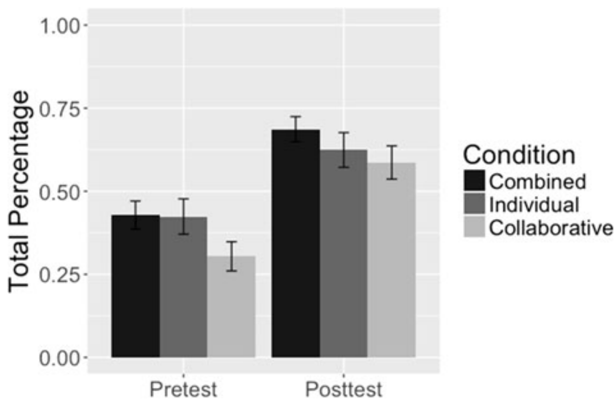


Fig. 5 Test score percentage at pretest and posttest by condition

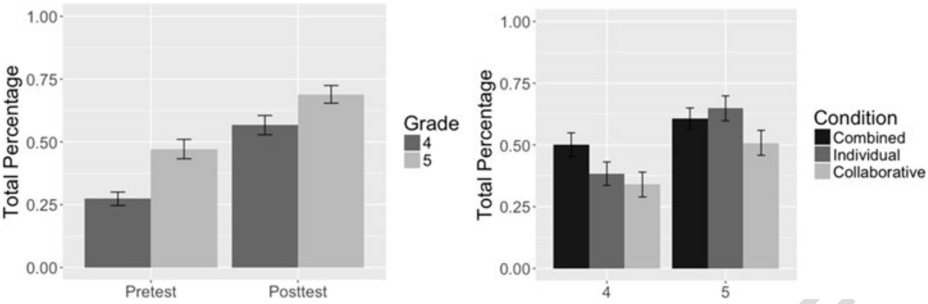


Fig. 6 (Left) Test score percentage for pretest and posttest by grade and (Right) test score percentage for grade by condition

In summary, we found that the students learned across all three conditions. However, not surprisingly, the 5th grade students had higher test scores than the 4th grade students. Additionally, confirming our hypothesis (H1), we found across both the learning slopes and the normalized learning gains that the learning gains were higher for the combined condition compared to the individual or collaborative conditions but that the differences may have been confined to the 4th grade students.

Hypotheses H2a and H2b: error analysis

To investigate the hypothesis that students in the combined condition make fewer errors than those in the individual condition (H2a) and not more errors than those in the collaborative condition (H2b) and how these errors may change over time (see Table 2), we ran two HLMs for the erroneous example problem types and the procedural problem types. For the erroneous problem type, the number of errors decreased over time (problem number), $t(1218) = -3.54$, $p < .05$, $r = .10$. Furthermore, the combined condition made fewer errors per problem compared to the individual condition, $t(218) = 2.78$, $p < .05$, $r = .19$, and the collaborative

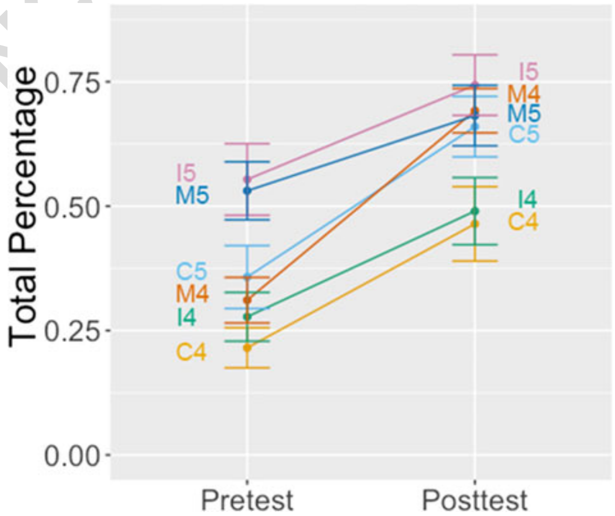


Fig. 7 Pre- and post-test scores for students working either collaboratively and individually (M), only collaboratively (C), or only individually (I), separated by grade level

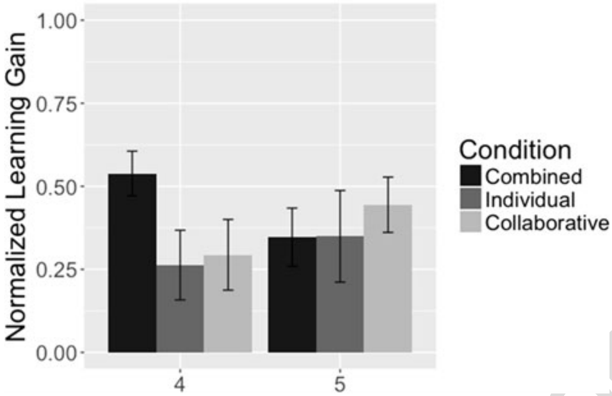


Fig. 8 Normalized learning gains by condition and grade level

condition, $t(218) = 3.04$, $p < .05$, $r = .20$ (see Fig. 9). There was no significant main effect of grade, $t(218) = -0.27$, $p = .79$. For the interactions, more in the 4th than the 5th grade, students' errors increased from the combined condition to the individual condition, $t(218) = -2.15$, $p < .05$, $r = .14$, and the collaborative condition, $t(218) = -1.97$, $p < .05$, $r = .13$. Additionally, there was not a difference in errors over time by grade, $t(1218) = 0.15$, $p = .88$, nor between the combined and collaborative conditions, $t(1218) = -0.17$, $p = .86$. However, the error rate did decrease faster in the individual condition compared to the combined, $t(1218) = -2.47$, $p < .05$, $r = .07$. Finally, there was not a significant interaction between problem number, grade, and condition, $t(1218) = 1.45$, $p = .15$ (combined and individual) and $t(1218) = -0.35$, $p = .72$ (combined and collaborative).

For the procedural problem types, the number of errors also decreased over time, $t(1545) = -2.88$, $p < .05$, $r = .07$. The combined condition made fewer errors per problem compared to the individual condition, $t(253) = 3.38$, $p < .05$, $r = .21$, and the collaborative condition, $t(253) = 9.61$, $p < .05$, $r = .52$ (see Fig. 9). The 5th grade students made fewer errors than the 4th grade students, $t(253) = 2.41$, $p < .05$, $r = .15$. For the interactions, in the 4th but not the 5th grade, students' errors increased from the combined condition to the individual condition, $t(253) = -2.69$, $p < .05$, $r = .17$, and the collaborative condition, $t(253) = -6.17$, $p < .05$, $r = .36$. There was not a difference in errors over time by grade, $t(1545) = -0.62$, $p = .53$, nor between the combined and individual conditions, $t(1545) = -0.99$, $p = .32$, but the error rate did decrease faster in the collaborative condition compared to the combined, $t(1545) = -6.21$, $p < .05$, $r = .16$. Finally, the 4th grade students made marginally fewer errors over time in the individual condition, $t(1545) = 1.78$, $p = .07$, $r = .04$, and significantly fewer errors over time in the collaborative condition, $t(1545) = 4.23$, $p < .05$, $r = .10$, compared to the combined condition whereas these differences were less pronounced with the 5th grade students.

Table 2 Mean errors per problem (SD) for all conditions

Condition	Err. Example Problems		Procedural Problems	
Grade	4th	5th	4th	5th
Collaborative	7.72 (4.14)	4.57 (2.19)	19.94 (23.69)	7.29 (8.69)
Individual	5.51 (2.79)	3.46 (1.93)	9.75 (6.69)	5.95 (6.16)
Combined	4.48 (2.04)	4.54 (1.99)	3.57 (4.36)	5.49 (7.35)

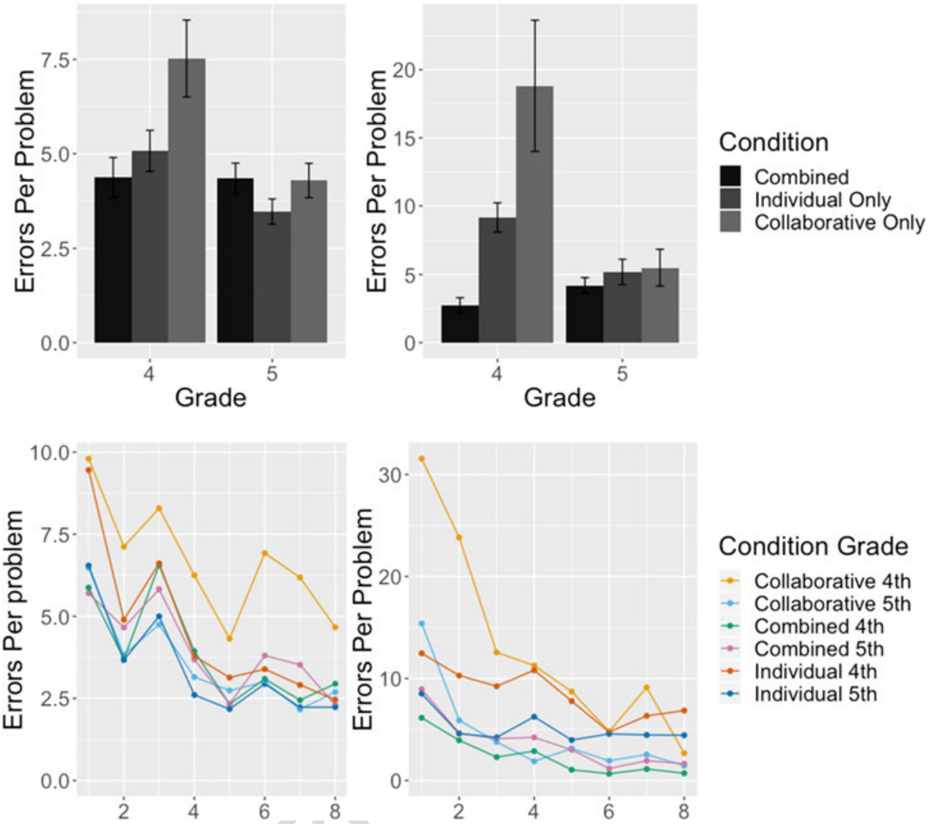


Fig. 9 Errors per problem made for (Left) erroneous example problems and (Right) procedural problems for grade by condition (Top) and over time (Bottom)

To test for significant equivalence of the combined and collaborative conditions (H2b), we used the confidence interval approach. For the errors made per problem, we did not find a statistically significant equivalence for the erroneous example problems or the procedural problems (see Table 3).

In summary, we found that students made significantly fewer errors over time in both problem types. In support of our hypothesis H2a, we found that students made fewer errors in the combined than the individual condition across both problem types, but the students in the individual condition had more of a change in errors (decrease) over time. In contrast to our hypothesis H2b, we also found that students made fewer errors in the combined compared to the collaborative condition, which we hypothesized would be equivalent. However, the students in the collaborative condition had more of a change in errors (decrease) over time.

Table 3 90% confidence interval for mean differences between the combined and collaborative conditions. The equivalence interval is set to ±0.5

	Lower Bound	Upper Bound
Err. Example Problems	−1.8	−0.59
Procedural Problems	−10.3	−4.5

As with the main results, there was a difference between grades and conditions with the 4th grade students in the combined condition having fewer errors than the other 4th grade students but this result less pronounced with the 5th grade students.

Hypotheses H3a and H3b: Hint Analysis

In addition to analyzing student performance through error rates, we also analyzed the request for hints. To investigate the hypothesis that students who work collaboratively will request fewer hints (H3a) than students working individually and will not request more hints (H3b) than those working collaboratively (see Table 4), we ran two HLMs for the erroneous example problem types and the procedural problem types. For the erroneous problem type, there was not a significant difference in the number of hints requested over time, $t(1218) = -1.15$, $p = .25$. However, the combined condition requested fewer hints per problem than the individual condition, $t(218) = 3.21$, $p < .05$, $r = .21$, and the collaborative condition, $t(218) = 4.10$, $p < .05$, $r = .27$ (see Fig. 10). The 5th grade students requested marginally fewer hints than the 4th grade students, $t(218) = 1.72$, $p = .08$, $r = .12$. For the interactions, students' hints increased from the combined condition to the individual condition, $t(218) = -3.77$, $p < .05$, $r = .22$, and collaborative condition, $t(218) = -3.49$, $p < .05$, $r = .23$, in the 4th grade but not in the 5th grade. There was not a significant difference in hints over time by grade, $t(1218) = -1.23$, $p = .22$, nor between the combined and individual conditions, $t(1218) = -1.08$, $p = .28$, but the collaborative condition did request fewer hints than the combined group over time, $t(1218) = -3.77$, $p < .05$, $r = .11$. Finally, there was not a significant interaction between problem number, grade, and the combined and individual conditions, $t(1218) = 1.18$, $p = .24$, but the 4th grade students in the collaborative conditions requested fewer hints over time compared to the combined condition, $t(1218) = 3.07$, $p < .05$, $r = .09$, whereas these differences were less pronounced with the 5th grade students.

For the procedural problem types, there was not a significant difference in the number of hints requested over time, $t(1545) = -1.18$, $p = .24$. The combined condition requested fewer hints per problem than the individual, $t(253) = 4.74$, $p < .05$, $r = .29$, and the collaborative, $t(253) = 4.82$, $p < .05$, $r = .29$ (see Fig. 10). There was not a significant main effect of grade, $t(253) = 1.38$, $p = .17$. For the interactions, students' hints increased from the combined condition to the individual, $t(253) = -4.01$, $p < .05$, $r = .24$, and the collaborative, $t(253) = -3.82$, $p < .05$, $r = .23$, in the 4th grade but less in the 5th grade. There was not a significant difference in hints over time by grade, $t(1545) = -0.13$, $p = .90$, but the individual, $t(1545) = -1.71$, $p = .08$, $r = .04$, and collaborative conditions, $t(1545) = -3.09$, $p < .05$, $r = .08$, requested significantly fewer hints over time than the combined condition. Finally, the students in the individual condition requested marginally fewer hints, $t(1545) = 1.91$, $p = .06$, $r = .05$, and the collaborative condition requested significantly fewer hints, $t(1545) = 2.47$, $p < .05$, $r = .06$,

Table 4 Mean hints per problem requested (SD) for all conditions

Condition	Err. Example Problems		Procedural Problems	
Grade	4th	5th	4th	5th
Collaborative	6.10 (7.07)	2.32 (3.72)	3.22 (3.91)	0.39 (0.64)
Individual	4.30 (4.25)	1.17 (1.96)	6.19 (6.35)	1.74 (4.09)
Combined	2.23 (4.39)	2.52 (3.49)	1.20 (2.56)	1.47 (3.49)

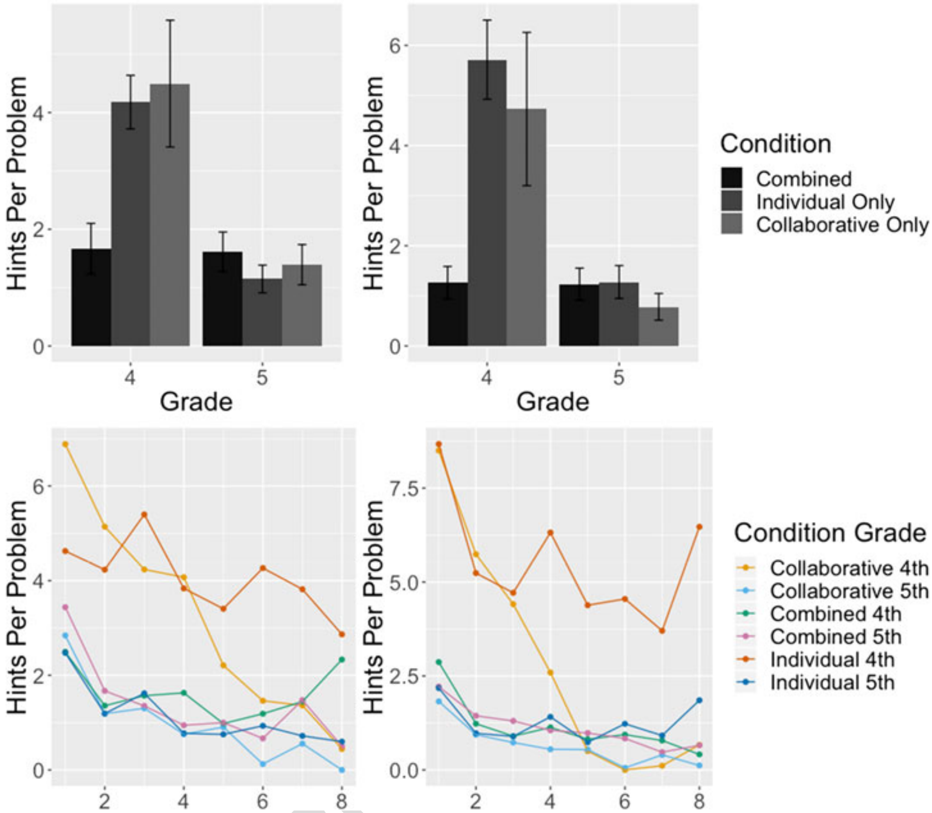


Fig. 10 Hints per problem made for (Left) erroneous example problems and (Right) procedural problems for grade by condition (Top) and over time (Bottom)

over time compared with the 4th grade students, whereas these differences were less pronounced with the 5th grade students.

To test for significant equivalence of the combined and collaborative conditions (H3b), we again used the confidence interval approach. For the hints requested per problem, we did not find a statistically significant equivalence for the erroneous example problems or the procedural problems (see Table 5).

In summary, we did not find a significant main effect for a change in hint requests over time across either problem type. However, like the errors, we found support for our hypothesis H3a in that the students in the combined condition requested fewer hints than those in the individual condition. Also like with the errors, we found that the students in the combined condition requested fewer hints than the students in the collaborative condition instead of being

Table 5 90% confidence interval for mean differences between the combined and collaborative conditions. The equivalence interval is set to ±0.5

	Lower Bound	Upper Bound
Err. Example Problems	−2.4	−0.28
Procedural Problems	−0.74	0.53

equivalent across both problem types, which does not support our hypothesis H3b. However, for change in hints across problems, we found the slopes to decrease at a faster rate for both the individual and collaborative conditions compared to the combined. Finally, as with the main results, there was a difference between grades and conditions with the 4th grade students in the combined condition requesting fewer hints than the other 4th grade students but this pattern being less pronounced with the 5th grade students.

Hypothesis H4: situational interest

To investigate the impact that working with a partner may have had on the students' situational interest in the tutoring activity (H4), we conducted a MANOVA with the trigger, maintained feeling, and maintained value as dependent variables and condition and grade as independent variables (see Table 6). There was a significant effect of condition on the three situational interest factors, $F(6, 600) = 7.69$, $p < .05$. There was not a significant main effect of grade on the three situation interest factors, $F(3, 299) = 0.89$, $p = .45$, but there was a significant interaction between grade and condition for the three situational interest factors, $F(6, 600) = 7.69$, $p < .05$.

Given the significance of the MANOVA analysis, we conducted a follow-up analysis using three HLMs, one for each dependent measure, with student at the first level and school district at the second level. At level 1, we modeled the situational interest scores, grade, and condition, and at level 2, we accounted for random differences that could be attributed to the school district. For trigger situational interest, the combined condition had a higher interest score than the individual condition, $t(229.99) = -2.64$, $p < .05$, $r = .17$, but there was not a significant difference between the combined and collaborative conditions, $t(225.02) = -1.58$, $p = .12$. There was no main effect for grade, $t(151.38) = -0.96$, $p = .34$, or any interactions between grade and conditions, $t(292.32) = -0.30$, $p = .77$ (individual/combined) and $t(126.27) = 1.62$, $p = .11$ (collaborative/combined).

For the maintained feeling situational interest factor, we found the students in the combined condition had a higher maintained feeling than the students in the individual condition, $t(186.92) = -2.07$, $p < .05$, $r = .15$ (see Table 6). As with the trigger situational interest, there was no significant main effect between combined and collaborative conditions, $t(180.84) = -1.36$, $p = .18$, or grade, $t(104.61) = -0.60$, $p = .55$. There was also no significant interaction between the conditions and grade, $t(276.65) = -0.80$, $p = .43$ (combined/individual) and $t(91.20) = 1.19$, $p = .24$ (combined/collaborative).

The maintained value situational interest measure did not follow the same pattern of results at the other factors (see Table 6). For the maintained value situational interest, the students in the combined condition reporting a higher maintained value than the students in the individual

Table 6 The situational interest mean scores (SD) for trigger, maintained feeling, maintained value for Collaborative (C), Individual (I), and Combined (M)

	Trigger		Maintained Feeling		Maintained Value	
	4th Grade	5th Grade	4th Grade	5th Grade	4th Grade	5th Grade
C	0.65 (0.22)	0.75 (0.20)	0.65 (0.26)	0.73 (0.21)	0.74 (0.21)	0.82 (0.19)
I	0.61 (0.19)	0.50 (0.20)	0.62 (0.22)	0.49 (0.22)	0.75 (0.17)	0.62 (0.23)
M	0.75 (0.17)	0.69 (0.20)	0.74 (0.21)	0.70 (0.24)	0.86 (0.14)	0.79 (0.18)

or collaborative conditions, $t(167.01) = -2.87$, $p < .05$, $r = .22$ and $t(160.75) = -2.85$, $p < .05$, $r = .22$ respectively. The 4th grade students reported marginally higher maintained value than the 5th grade students, $t(87.56) = -1.71$, $p = .09$, $r = .18$. For the interactions, there was not a significant interaction between grade and the combined and individual conditions, $t(264.38) = -0.41$, $p = .68$, but 4th grade students in the combined condition had significantly higher maintained value scores than those in the collaborative condition while this same effect was not found with 5th grade students, $t(75.35) = 2.45$, $p < .05$, $r = .27$.

In summary, these results indicate that the students who had an opportunity to work with a partner found the fractions CITS more immediately interesting than students only working individually confirming our hypothesis H4. This interest may have been extended to the domain as well as indicated by the maintained situational interest measures.

Discussion

In this paper, we investigated if a combination of collaborative and individual learning is more effective than engaging in either alone. The analysis of the pretest and posttest data confirmed our hypothesis (H1) that a combination of collaborative and individual learning can be more beneficial than either alone. Specifically, our result was confined to the 4th grade students. These results resemble those from other research where the age of the students had an impact on the effectiveness of the learning intervention (Mazziotti et al. 2015). This difference in grade may indicate that the given combination of individual and collaborative learning is particularly effective early in the learning process when students may need more support targeted at the skills they are trying to acquire. The 5th grade students may have already learned correct knowledge for the targeted fractions skills, so the support from a partner would not be as beneficial. It may also be that the 5th grade students had higher pretest scores so could not have similarly high learning gains. However, at posttest, the students were still not at ceiling and when comparing the normalized learning gains, there was still the impact of condition and grade. Below, we explore why the combined condition had higher learning gains for the 4th grade students than the 5th grade students based on the results from the process analyses. In addition, the 5th grade students in the collaborative condition had higher learning gains than the other 5th grade conditions. However, the difference may be an effect of differences at pretest where the 5th grade students in the collaborative condition performed substantially lower than the other 5th grade students. The 5th grade collaborative condition did not have significantly different posttest scores than the other 5th grade students.

The differences in learning gains between conditions may have been due to the way that the students engaged with the learning process. To explore this question, we analyzed indicators of student process while working with the tutor. Previous research has shown that there is a negative correlation between frequency of errors and hint requests with posttest scores (Aleven and Koedinger 2001). Q4 Students who do not attempt to game the system and request hints when it is helpful, may learn more because they are able to struggle and work through the problem. The combined condition may have been more effective if they were able to apply good habits around errors and hints learned from working with a partner to their individual sessions where they had fewer interruptions.

From the analysis of the errors and hints, we found similar trends to those seen in the learning gain analysis with students in the combined condition engaging in more productive learning processes from the beginning of the sessions. Based upon previous research that found students working collaboratively asked for fewer hints and made fewer errors than students working individually (Hausmann et al. 2009; Hausmann et al. 2008a; Hausmann et al. 2008b), we hypothesized that students in the combined condition would ask for fewer hints and making fewer errors than students working individually and ask for the same number of hints and make the same number of errors as those in the collaborative condition when working on the erroneous example problems (H2a,b, H3a,b). We found that the students in the combined condition tended to make fewer errors and request fewer hints than the other conditions with an interaction with grade level, such that 4th grade students not in the combined condition tended to make more errors and request more hints than 4th graders in the combined condition or 5th graders, only partially supporting our hypotheses.

For the procedural problems, we again had hypothesized that students working collaboratively would need the same amount of assistance as those in the combined condition and that the combined condition would need less assistance than those in the individual condition because they apply the good practices that they learned when collaborating to working individually. Again, our findings did not support our hypothesis in terms of the collaborative condition but did for the individual condition. Like with the erroneous example problems, the students in the combined condition had fewer errors and hints than the other conditions and an interaction between grades. Again, the students in the 4th grade combined condition had results much closer to those in 5th grade while the 4th grade students in the other conditions were much higher.

When looking at the changes in the hints and errors over time, we found that the students made fewer errors over time but there was not a main effect for a decrease in hints. These changes may indicate that although the students still needed support in solving the problems later in the problem sets (request for hints), they were able to apply the support more efficiently and made fewer errors per problem. Surprisingly, we found that the students in the combined condition had shallower error and hint slopes over time than the students in the individual or collaborative conditions. This may have been due to their starting point. The 4th grade students in the combined condition began the problem sets with a much lower error and hint rate than the other 4th grade students so had less of a distance to change until reaching floor (having no errors and requesting no hints on a problem). In other words, the students in the combined condition had fewer hints and errors than the other conditions perhaps not because they were learning at a faster rate but because they began with better habits from the beginning.

The actions that the students take while working with the ITS may help to explain the differences in the learning outcomes. From the log data, we see that the 4th grade students perform significantly worse than the 5th grade students when they are not in the combined condition, but the 4th grade combined students have similar results to the 5th grade students, which echoes the learning gain results. This finding again might be explained by the fact that the 5th grade students may have already been familiar with the concepts and procedures associated with the units covered. While working with the tutor, we would then expect to not see them have as many hints and errors, which is what we found. This is also true for the 5th grade students in the collaborative condition despite the fact that they had significantly lower pretest scores indicating that they also may have already known the domain material before entering the study. On the other hand, the combined condition may have been able to appropriately support the 4th grade students where needed by having a partner available when more sense making was necessary, such as with the erroneous example problems. Students could then take this knowledge and apply it to the procedural problems without having to negotiate and share steps with a partner.

Finally, the students may have learned more in the combined condition because they found the task more engaging. When students are more interested in a task, they are willing to put more time and effort into completing that task (Rogat et al. 2013). From the literature and as we had hypothesized (H4), the students who have a chance to collaborate (collaborative and combined) would have higher interest in the task. Our results support this hypothesis. Students in both the collaborative and combined conditions expressed higher interest in the immediate task and their feelings towards the domain (i.e., maintained feeling) than the students working individually. These results show, interestingly, that even when students are not working collaboratively the whole time, the collaboration can still be motivating but do not fully explain why the students in the combined condition (among 4th graders) may have had higher learning gains (since the students in the collaborative condition also had higher interest).

Being in the combined condition was not only motivational for the students in the moment, but also influenced their perceived value of the domain. We found that the students in the combined condition had a higher reported situational interest on the maintained value factor. This finding indicates that the combined condition may impact how the students value fractions in the short term, which can lead to maintained personal interest (Hidi and Renninger 2006; Schraw and Lehman 2001). However, because the interest measure was only administered at the end of the experiment, we cannot rule out that the students in the combined condition already had a greater interest in the domain, which influenced their learning. Also, although the higher value was only in the combined condition, we did not find any differences between the grades. Allowing students to collaborate on tasks thus might be one way to both motivate students and to create a beneficial learning environment that could lead to a personal interest in the domain, but the interest in the task does not help to explain the differences between the grades in the combined condition.

Through both our analysis of the learning gains and process analysis, we found that the combined condition was more effective than either social level alone, especially for 4th grade students. Having a combined condition may be more important for students that are less familiar with the material being taught. The combined condition can then provide students with an environment where they make fewer errors, request fewer hints, and report being engaged, which may lead to the higher learning gains. However, it is still unclear what about the combined condition leads to these effects. For future work, it would be beneficial to analyze the dialogues between the students to see how the support from the partners was different between conditions and how the support may have impacted the effectiveness the conditions.

This study contributes to the understanding within CSCL of when collaborative learning can be beneficial with our result indicating that there is promise in further investigating the combination of collaborative learning with other social levels. Although we found positive results for a combination of collaborative and individual learning, these findings are in contrast to the results from Wang et al. (2011), in which they found the combined condition to have gains less than those working collaboratively and more than those working individually only. Taking these studies together, there is some indication that it is not enough to just combine collaborative and individual learning as variation theory may suggest (Ling and Marton 2012), but we must begin to explore how this combination is done and, as our results show, when a combination should be used.

Although our comparison supported the alignment of the learning activities and knowledge acquisition as proposed by the KLI framework (Koedinger et al. 2012), from the analysis of the hints and errors, we found that there may be more benefit than can just be explained by the alignment due to students in the combined condition making fewer errors and requesting fewer hints even in comparison to when the separate conditions (individual or collaborative only) would have been well aligned. In this case, as researchers in CSCL further explore what an effective combination entails and when, it will be important to consider how working in the different social levels may influence the learning process. For example, Celepkolu et al. (2017) had the students work individually and then collaboratively in their combined condition to have the students prepare for the collaborative discussion. In contrast, in our study, the students worked collaboratively to first address misconceptions before working individually on the fluency of the procedures. In these cases, the orderings of the social levels were different, but both studies found a positive impact of the combinations. When considering when a combination of collaborative and individual learning may be useful, it may be important to not only consider the alignment learning support and skills for the individual activities, but how working on one activity may positively influence the next, which is integral to many CSCL integrative scripts (Dillenbourg and Tchounikine 2007) and may contribute to explaining the positive impact that the combination has on the learning processes.

Conclusion

This paper opens up a broader line of inquiry in CSCL that focuses on the question of how collaborative and individual learning can most effectively be combined. In our study, we supported student learning through the use of erroneous example problems and procedurally oriented problems. We chose these activity types because the strengths of collaborative and individual learning theoretically aligned with the knowledge targets being acquired in each of the learning activities. Specifically, this combination may have been effective because it allowed the students to address misconceptions with a partner and thus develop a deeper understanding. After addressing misconceptions, the students then had an opportunity to build fluency with individual problem-solving. This alignment of the learning activities with the hypothesized strengths of the individual and collaborative learning may have enhanced the support to the students more than either could provide alone.

Although our results support that this combination of collaborative and individual learning with the learning tasks was more effective than either alone, our study is still only an initial step into understanding the combination of collaborative and individual learning carried out in a very specific ITS context that may have influenced our findings. However, it provides an indicator that combining collaborative learning with other social levels may be a promising direction. Our results taken along with previous research indicate that it is not just a combination that is important, but to understand what combinations of collaborative and individual learning can be effective for learning and when, additional research is needed. One direction for this future research is to investigate how our findings may transfer to other domains and technologies, such as those used in Wang et al. (2011) and Celepkolu et al. (2017). Additionally, it is important to explore how the order and combination of the individual and collaborative learning activities influence student learning and the learning process as to contribute to the understanding of what learning mechanisms may be at work within a successful combination. This research contributes to the CSCL literature by opening the investigation into why integrative scripts that combine collaborative learning with other social levels are impactful for learning.

Furthermore, as we have seen with previous CSCL technology, it may not be enough to only explore these fixed types of support and combinations (Fischer et al. 2013b). To continue the exploration of the combination of collaborative and individual learning into more personalized and adaptable areas, it is important to consider when these transitions between social levels would be most beneficial for individual students. For example, in our study, all transitions occurred at the same set time. It may also be beneficial for students to transition between social levels adaptively based on student characteristics, such as repeated errors on a skill when working individually. In this case, one of the major hurdles to this task is to support the teacher orchestration that is needed for these transitions to occur in the classroom (Olsen et al. 2018a). Only once we have the technological support needed for the orchestration of these more complex designs can we begin to develop adaptive combinations that can be feasibly used, and, therefore, empirically tested, without making the learning design inconsequential for student learning because the orchestration load is too high for teachers.

The results of our study are notable because of the complexity in supporting both collaborative and individual learning in the classroom and providing real-time support. This study adds to the CSCL literature by exploring when collaborative learning may be effective by comparing a combination of collaborative and individual learning to both alone, which is so far uncommon. By finding support for the effectiveness of combining collaborative and individual learning, this paper has opened a broader line of inquiry into how collaborative and individual learning can most effectively be combined to support learning. Within this space, we can begin to evaluate integrative scripts (Dillenbourg 2004) to better understand what aspects of the scripts are proving to be effective for student learning.

References

Aleven, V., & Koedinger, K. R. (2013). Knowledge Component (KC) Approaches to Learner Modeling. *Design Recommendations for Intelligent Tutoring Systems, 1*, 165–182.

Aleven, V., Sewall, J., Popescu, O., van Velsen, M., Demi, S., & Leber, B. (2015). Reflecting on twelve years of ITS authoring tools research with CTAT. *Design Recommendations for Adaptive Intelligent Tutoring Systems, 3*, 263–283.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 261–295.

Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence, 42*, 7–49.

Aronson, E. (1978). *The jigsaw classroom*. Sage.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*(2), 181–214.

Baghaei, N., & Mitrovic, A. (2005). COLLECT-UML: supporting individual and collaborative learning of UML class diagrams in a constraint-based intelligent tutoring system. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* (pp. 458–464). Berlin Heidelberg: Springer.

Celepkolu, M., Wiggins, J.B., Boyer, K.E., McMullen, K. (2017). Think First: Fostering Substantive Contributions in Collaborative Problem-Solving Dialogues. In *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning.* (pp. 295-302).

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.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243.

Cress, U. (2008). The need for considering multilevel analysis in CSCL research–An appeal for the use of more advanced statistical methods. *International Journal of Computer-Supported Collaborative Learning, 3*(1), 69–84.

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.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *Three worlds of CSCL. Can we support CSCL?* 61-91.

Dillenbourg, P. (2004). Framework for integrated learning. Kaleidoscope Network of Excellence Deliverable D23.5.1.

Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of Computer Assisted Learning, 23*(1), 1–13.

Diziol, D., Rummel, N., Spada, H., & McLaren, B.M. (2007). Promoting learning in mathematics: Script support for collaborative problem solving with the Cognitive Tutor Algebra. In C.A. Chinn, G. Erkens & S. Puntambekar (Eds.) *Mice, minds and society. Proceedings of the Computer Supported Collaborative Learning Conference* (pp. 39-41). International Society of the Learning Sciences, Inc.

Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2010). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review, 22*(1), 89–102.

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013b). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist, 48*(1), 56–66.

Fischer, F., Kollar, I., Stegmann, K., Wecker, C., Zottmann, J., & Weinberger, A. (2013a). Collaboration scripts in computer-supported collaborative learning. *The International Handbook of Collaborative Learning*, 403-419.

Frank, M. C., & Gibson, E. (2011). Overcoming memory limitations in rule learning. *Language, Learning, & Development, 7*, 130–148.

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research, 11*(2), 147–252.

Hausmann, R. G., Chi, M. T., & Roy, M. (2004) Learning from collaborative problem solving: An analysis of three hypothesized mechanisms. In *26nd annual conference of the Cognitive Science society*, 547-552.

Hausmann, R. G., Nokes, T. J., VanLehn, K., & van de Sande, B. (2009). Collaborative Dialog While Studying Worked-out Examples. In *International Conference on Artificial Intelligence in Education.* (pp. 596-598).

Hausmann, R. G. M., Van de Sande, B., & VanLehn, K. (2008a). Are self-explaining and coached problem solving more effective when done by pairs of students than alone? In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2369–2374). Austin, TX: Cognitive Science Society.

Hausmann, R. G., van de Sande, B., & VanLehn, K. (2008b). Trialog: How Peer Collaboration Helps Remediate Errors in an ITS. In *Proceedings of the 21st International Florida Artificial Intelligence Research Society Conference* (pp. 415–420). Menlo Park, CA: AAAI Press.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Isotani, S., Adams, D., Mayer, R. E., Durkin, K., Rittle-Johnson, B., & McLaren, B. M. (2011). Can erroneous examples help middle-school students learn decimals? In *Towards Ubiquitous Learning* (pp. 181–195). Berlin Heidelberg: Springer.

Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist, 48*(1), 40–55.

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten Years of Computer-Supported Collaborative Learning: A meta-analysis of CSCL in STEM education during 2005-2014. *Educational Research Review, 100284*.

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523–550.

Kapur, M. (2014). Comparing learning from productive failure and vicarious failure. *Journal of the Learning Sciences, 23*(4), 651–677.

Koedinger, K. R., Corbett, A. C., & Perfetti, C. (2012). The Knowledge-Learning-Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36*(5), 757–798.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts–a conceptual analysis. *Educational Psychology Review, 18*(2), 159–185.

Kulik, J. A., & Fletcher, J. D. (2015). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *Review of Educational Research*. https://doi.org/10.3102/0034654315581420.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: affective impact of animated pedagogical agents. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (pp. 359-366). ACM.

Ling, L. M., & Marton, F. (2012). Towards a science of the art of teaching. *International journal for lesson and learning studies*.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*.

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.

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.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology, 106*(4), 901.

Magnisalis, I., Demetriadis, S., & Karakostas, A. (2011). Adaptive and intelligent systems for collaborative learning support: A review of the field. *IEEE Transactions on Learning Technologies, 4*(1), 5–20.

Mazziotti, C., Loibl, K., & Rummel, N. (2015). Collaborative or Individual Learning within Productive Failure: Does the Social Form of Learning Make a Difference? In O. Lindwall et al. (Eds.), *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning,* (570-575). Gothenberg, Sweden.

McLaren, B. M., Adams, D., Durkin, K., Goguadze, G., Mayer, R. E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S., & Van Velsen, M. (2012). To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In *21st Century Learning for 21st Century Skills* (pp. 222–235). Berlin Heidelberg: Springer.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology, 85*, 424–436.

Mullins, D., Rummel, N., & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer-Supported Collaborative Learning, 6*(3), 421–443.

Olsen, J. K., Aleven, V., & Rummel, N. (2017). Exploring Dual Eye Tracking as a Tool to Assess Collaboration. In *Innovative Assessment of Collaboration* (pp. 157–172). Cham: Springer.

Olsen, J. K., Belenky, D. M., Aleven, A., & Rummel, N. (2014b). Using an intelligent tutoring system to support collaborative as well as individual learning. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Proceedings of the 12th International Conference on Intelligent Tutoring Systems* (pp. 134–143). Berlin, Heidelberg: Springer.

Olsen, J. K., Belenky, D. M., Aleven, A., Rummel, N., Sewall, J., & Ringenberg, M. (2014a). Authoring tools for collaborative intelligent tutoring system environments. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Proceedings of the 12th International Conference on Intelligent Tutoring Systems* (pp. 523–528). Berlin, Heidelberg: Springer.

Olsen, J.K., Rummel, N., & Aleven, V. (2016). Investigating effects of embedding collaboration in an intelligent tutoring system for elementary school students. In *the International Conference of the Learning Sciences*.

Olsen, J. K., Rummel, N., & Aleven, V. (2018a). Co-designing orchestration support for social plane transitions with teachers: Balancing automation and teacher autonomy. In J. Kay, R. Luckin, & R. (Eds.), *Proceedings of the 13th International Conference of the Learning Sciences* (pp. 1541–1542). London: International Society of the Learning Sciences.

Olsen, J. K., Sharma, K., Aleven, V., & Rummel, N. (2018b). Combining gaze, dialogue, and action from a collaborative intelligent tutoring system to inform student learning processes. In J. Kay, R. Luckin, & R. (Eds.), *Proceedings of the 13th International Conference of the Learning Sciences* (pp. 689–696). London: International Society of the Learning Sciences.

Renkl, A. (2005). The worked-out example principle in multimedia learning. In R. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning* (pp. 229–246). Cambridge: Cambridge University Press.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*(1), 15–22.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346–362.

Rogat, T. K., Linnenbrink-Garcia, L., & DiDonato, N. (2013). Motivation in collaborative groups. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 250–267). New York: Routledge.

Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin, 113*(3), 553.

Rummel, N. (2018). One framework to rule them all? Carrying forward the conversation started by Wise and Schwarz. *International Journal of Computer-Supported Collaborative Learning, 13*(1), 123–129.

Rummel, N., Weinberger, A., Wecker, C., Fischer, F., Meier, A., Voyiatzaki, E., … & Koedinger, K. R. (2008). New challenges in CSCL: Towards adaptive script support. In *Proceedings of the 8th International Conference for the Learning Sciences, 3,* (pp. 338-345). International Society of the Learning Sciences.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review, 13*(1), 23–52.

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right... If they argue together! *Cognition and Instruction, 18*(4), 461–494.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28*(3), 225–273.

Slavin, R. E. (1989). Cooperative learning and student achievement: Six theoretical perspectives. In *Advances in Motivation and Achievement, 6, 161-177*. Greenwich: JAI Press, Inc..

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43–69.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*(1), 21–51.

Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology, 31*(2), 207–220.

Tsovaltzi, D., Melis, E., Mclaren, B. M., Dietrich, M., Goguadze, G., & Meyer, A. K. (2009). Erroneous examples: A preliminary investigation into learning benefits. In *Learning in the Synergy of Multiple Disciplines* (pp. 688–693). Berlin Heidelberg: Springer.

Tsovaltzi, D., Melis, E., McLaren, B. M., Meyer, A.-K., Dietrich, M., & Goguadze, G. (2010). Learning from erroneous examples: When and how do students benefit from them? In M. Wolpers, P. A. Kirschner, M. Scheffel, S. Lindstaedt, & V. Dimitrova (Eds.), *Proceedings of the 5th European Conference on Technology Enhanced Learning, Sustaining TEL: From Innovation to Learning and Practice, Barcelona, Spain* (pp. 357–373). Berlin Heidelberg: Springer-Verlag.

Van Gog, T., Rummel, N., & Renkl, A. (2019). Learning how to solve problems by studying examples. In J. Dunlosky & K. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education*. New York: Cambridge University Press.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education, 16*(3), 227–265.

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.

Walker, E., Rummel, N., & Koedinger, K. R. (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer-Supported Collaborative Learning, 6*(2), 279–306.

Wang, H. C., Rosé, C. P., & Chang, C. Y. (2011). Agent-based dynamic support for learning from collaborative brainstorming in scientific inquiry. *International Journal of Computer-Supported Collaborative Learning, 6*(3), 371.

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 423–467.

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

Affiliations

Jennifer K. Olsen[1] • **Nikol Rummel**[2,3] • **Vincent Aleven**[2]

Nikol Rummel
nikol.rummel@rub.de

Vincent Aleven
aleven@cs.cmu.edu

[1] Ecole Polytechnique Fédérale de Lausanne, EPFL IC IINFCOM CHILI, RLC D1 740 (Rolex Learning Center), Station 20, CH-1015 Lausanne, Switzerland

[2] Carnegie Mellon University, Pittsburgh, PA, USA

[3] Institute of Educational Research, Ruhr-Universität Bochum, Bochum, Germany