Procedural and Conceptual Knowledge Acquisition in Mathematics: Where is Collaboration Helpful?

Dejana Diziol, Nikol Rummel, Hans Spada
Institute of Psychology, University of Freiburg, Germany
{diziol, rummel, spada}@psychologie.uni-freiburg.de

Abstract: While research has generally shown that collaboration may facilitate student learning in mathematics, such positive effects are not always found. We argue that the effectiveness of collaboration may depend on the type of knowledge the instruction targets: The interaction with a partner can slow down students and may thus decrease the amount of practice necessary for *procedural skill fluency.* On the other hand, collaboration could be particularly useful for *conceptual knowledge acquisition,* as here, the elaborative meaning-making activities ascribed to collaboration may facilitate learning. To evaluate the differential effects of collaborative learning, we compared four conditions: individual versus collaborative learning with procedural instruction, and individual versus collaborative learning with conceptual instruction. The study results support our hypotheses: Students who learned individually showed higher test scores in a procedural far transfer test. However, a combination test requiring both knowledge types revealed a positive impact of collaboration on students' conceptual knowledge acquisition.

Introduction

Collaborative learning environments have generally shown to be beneficial for learning in mathematics (e.g. Berg, 1994; Ellis, Klahr & Siegler, 1993). The positive effect of collaboration can be explained by particular student interactions such as giving explanations and knowledge co-construction that are positively related to learning (e.g. Hausmann, Chi, & Roy, 2004). However, this beneficial impact of collaboration on learning is not always found (e.g. Dillenbourg, Baker, Blaye & O'Malley, 1996; Lou et al., 1996; Souvignier & Kronenberger, 2007). The meta analysis on collaborative learning by Lou et al. (1996) gives a good overview of this phenomenon: Although most results were in favor of collaborative learning, about a fourth of the results showed negative effects. Also in one of our own projects, comparing individual to collaborative learning with the Cognitive Tutor Algebra, results were inconsistent (Rummel, Diziol, & Spada, 2008): We found no differences between conditions in a retention test that mainly required *procedural knowledge* (e.g. computational skills). Observations of students' problem-solving process indicated that collaboration might even have impeded the acquisition of this knowledge type as the interaction with the partner slowed the students down and thus decreased the amount of practice. On the other hand, we found indications that the interaction with a peer improved students' understanding of the underlying mathematical concepts, i.e., *their conceptual knowledge*. The advantage of collaboration was particularly found when students engaged in mutual explanations and deep discussions of the learning content (Diziol & Rummel, 2008). These results led us to the conclusion that the inconsistency in collaborative learning research may in part be explained by differential effects of collaboration on procedural and conceptual knowledge acquisition. In other words the benefits of collaboration on student learning may depend on the type of task solved collaboratively and on the type of knowledge students are expected to acquire during the interaction. The aim of the study presented in this paper is to increase our knowledge of these differential effects of collaboration on student learning in mathematics. This knowledge can help to introduce collaborative learning more selectively within the school classroom.In the following sections, we will give a short overview on the differentiation between procedural and conceptual knowledge acquisition in mathematics. Then we will discuss results on collaborative learning regarding this differentiation. We will conclude the theoretical background with a short overview on our study hypotheses.

Procedural and Conceptual Knowledge Acquisition in Mathematics

Literature on knowledge acquisition in mathematics distinguishes between two different types of knowledge: procedural and conceptual knowledge. *Procedural knowledge* refers to students' ability to execute action sequences in order to solve routine problems (e.g. Rittle-Johnson & Alibali, 1999). Students learn step-by-step solution procedures and, by repeatedly solving tasks that require these procedures, their proficiency improves. A typical example for tasks that requires procedural knowledge are manipulation problems such as solving equations for *x* (Nathan, Mertz & Ryan, 1994). If students know the relevant procedures, they can easily solve the task. However, procedural knowledge is closely tied to specific problem types and thus is not widely generalizable.

In contrast, conceptual knowledge is rather flexible and thus enables students to solve problems that are based on the same mathematical principles, but have a different problem format. The ability to transfer knowledge to new problems is considered an important step in gaining mathematical literacy as it enables students to apply their mathematical knowledge in everyday life (OEDC, 2003). Conceptual knowledge is the understanding "of the principles that govern a domain and of the interrelations between pieces of knowledge in a domain" (Rittle-Johnson & Alibali, 1999, p. 175). For instance, in the domain of algebra, particularly relevant concepts are the equation, the variable, and the constant term. These concepts can be represented in different formats, for instance, verbally in a story problem ("they earn $2 per glass sold"), graphically in a coordinate plane, algebraically in an equation ("+ 2*x*"), or in a table (cf. Brenner et al., 1997). If students show the ability to flexibly translate between the different representations, this indicates that they have developed a solid understanding of the underlying algebraic concepts (Brenner et al., 1997; Mevarech & Stern, 1997).

Often, it is not possible to clearly distinguish between procedural and conceptual knowledge (cf. Hiebert & Wearne, 1996). Rittle-Johnson, Siegler and Alibali (2001) therefore describe the relation between procedural and conceptual knowledge as a continuum with procedural and conceptual knowledge as its two ends. According to their model, procedural and conceptual knowledge acquisition influence each other in an iterative way, in other words, improvement in one knowledge type can result in improvement in the second type of knowledge. A high understanding of underlying concepts can help to monitor the appropriateness of procedures and their correct execution, thus conceptual knowledge can influence the performance of procedural tasks. For instance in the domain of algebra, a good understanding of the concept "variable" can help to prevent a student from trying to add a constant and a variable term, a procedural error that is quite typical in students' problem-solving (see Booth, Koedinger, & Siegler, 2007). Conceptual knowledge can also improve when students solve procedural tasks, as long as they engage in active learning processes and try to understand the underlying principles.

To make predictions about the effectiveness of collaboration in supporting the acquisition of these two knowledge types, we need to take a closer look at the processes that yield procedural and conceptual learning. For procedural knowledge acquisition, students first have to be introduced to the procedures that are relevant to solve a particular task type. Then, to gain procedural skill fluency, practicing the application of the procedures is most crucial. One method that has shown particularly effective in improving students' procedural knowledge is guidance by an intelligent tutoring system (Koedinger, 1998). Intelligent tutoring systems monitor the student's progress, provide immediate error feedback and give help upon request that is tailored to the student's needs. The immediacy of feedback has proven particularly conducive to student's learning since it yields substantial cognitive and motivational benefits (Koedinger, 1998). For conceptual knowledge acquisition, merely solving problems is not sufficient (Hiebert & Wearne, 1996). Rather, students have to engage in active learning processes in order to gain an understanding of the domain principles, that is, they have to elaborate on the learning content to increase their understanding. For instance, these elaboration processes are required when solving algebra story problems, that is, when students have to translate between the verbal problem description and an algebraic equation. Simple translation rules based on keywords might not always yield the correct solution (cf. Nathan, Kintsch, & Young, 1992; e.g., "the depth *increases* by 3 m per hour" might have to be translated to "-3 *x*", even though the word "increase" normally might infer a positive variable term). Instead, students have to correctly represent the problem scenario described, extract the important information, and transform this information into a different, that is, a mathematical representation format (Staub & Reusser, 1995). Through elaboration on these representation translations, students can increase their understanding of the underlying mathematical concepts.

The Influence of Collaboration on Knowledge Acquisition in Mathematics

So far, research on collaborative learning does not support conclusions on the differential influence of collaboration on procedural and conceptual knowledge acquisition. First, the two knowledge types were often confused in the instruction. For instance, Berg (1994) compared individual and collaborative learning in mathematics over the course of several weeks. In the collaborative condition, student problem-solving was supported by a collaboration script. The script prompted students to engage in mutual explanation, i.e. deep cognitive processes, during problem-solving. Post-test comparisons showed that students who learned collaboratively outperformed the individual learners. Similar results were found in a study by Ellis et al. (1993) where students learned to solve decimal fraction tasks either individually or together with a partner. As in Berg's study, the collaborative learning process was supported by a script that encouraged students to give explanations. In a first step, students were asked to compare two decimal fractions individually and to decide which number they deemed to be larger. Next, they joined with their partner, discussed their individual decisions, agreed upon one solution, and received feedback on the correctness of their solution. Again, students in the collaborative condition outperformed individual learners. In both studies, the instruction targeted both knowledge types: Students were instructed to collaboratively solve problems – thereby training their procedural skills; additionally they received instructions that prompted them to collaboratively engage in deep reasoning processes – thereby fostering their conceptual knowledge. This confusion yields the following question: Does the collaborative application of problem-solving procedures itself have an effect, or is the effect mediated by joint elaboration on the underlying mathematical background? In other words, is collaborative practice in applying *procedures* effective, or is it the collaborative elaboration of mathematical *concepts* that yields differences in learning outcome? Second, additional instructions that foster conceptual knowledge acquisition are often only given to the collaborative conditions. Also in Berg (1994) and Ellis et al. (1993), students that learned individually were not encouraged to self-explain their solutions and to elaborate on their thinking, while the instructions of the collaborative conditions encouraged students to engage in deep learning processes. Indeed, positive results of collaboration can particularly be found if collaborative conditions receive additional instructions that are not given to students learning individually (cf. meta-analysis of Lou et al., 1996). This yields the question whether the effectiveness of collaboration is due to the collaboration per se or due to the additional instruction. Finally, another area of confusion concerns the tasks used for assessing the learning effect. Either the tasks that assessed student learning required both procedural and conceptual knowledge, or information on the test material was not sufficient to judge which knowledge type was improved. Thus, it is not clear if the collaboration positively influenced students' procedural performance or their conceptual knowledge acquisition.

In this project, we aim to evaluate the differential effect of collaboration on the two knowledge types. We argue that the learning mechanisms ascribed to collaboration might be particularly beneficial for conceptual knowledge acquisition, while for procedural knowledge acquisition, collaborative learning might not be beneficial or might even have a detrimental effect when compared to individual learning. As discussed earlier, in order to gain mastery in procedural skills, students have to practice the application of procedures. However, several studies have found that collaboration often takes more time than individual problem-solving as further requirements such as coordinating the interaction are added (e.g., Lou, Abrami, & d'Apollonia, 2001; Rummel et al., 2008; Walker, Rummel, & Koedinger, 2008). If the interaction with a partner slows the student down, this might decrease the amount of practice necessary for procedural skill fluency. Additionally, students have to "share" the practice opportunities with their partner, further reducing the amount of practice available for the individual student. This might result in a negative effect of collaborative problem-solving on procedural knowledge acquisition. In contrast, in order to gain conceptual understanding, it is necessary that students elaborate on mathematical concepts and try to understand their meaning. Thus, for conceptual knowledge acquisition, the elaborative meaning-making activities ascribed to collaboration may serve to support student learning. First, in a collaborative setting, students are required to make their thinking explicit and verbalize their knowledge. In other words, they have to give explanations to their partners (cf. Hausmann et al., 2004; Webb, 1989). Often they have to reformulate and clarify their statements if their partner has difficulties in understanding their explanations. This verbalization and reformulation of knowledge demands elaboration on the learning content (O'Donnell, 1999) and thus might be beneficial for conceptual knowledge acquisition. Furthermore, research has shown that students can acquire a deeper conceptual understanding by jointly elaborating on the learning material in order to construct new knowledge. Particularly in the domain of mathematics, knowledge co-construction has been shown to yield improved student achievement (e.g. Berg, 1994). Finally, by asking for help and receiving explanations from a partner (e.g., Webb, 1989), the interaction with a partner enables the student to fill knowledge gaps and correct misconceptions. While these learning processes are less important for the acquisition of procedural skill fluency, for the acquisition of conceptual knowledge, the interaction with a learning partner can foster elaborative processes and thus improve conceptual knowledge acquisition when compared to individual learning.

To assess the differential effect of collaboration on both knowledge types, we compared four conditions: individual versus collaborative learning with procedural instruction (problem-solving practice), and individual versus collaborative learning with conceptual instruction (elaboration on mathematical concepts). We assessed the effect of the four conditions both on procedural and conceptual knowledge acquisition. First, we hypothesized that the procedural instruction would mainly yield benefits on students' procedural knowledge acquisition, and that the conceptual instruction would mainly yield benefits on students' conceptual knowledge acquisition. Second, we hypothesized that regarding procedural knowledge acquisition, individual learning with procedural instruction would outperform collaborative learning with procedural instruction, while regarding conceptual knowledge acquisition, collaborative learning with conceptual instruction would outperform individual learning with conceptual instruction.

Methods

Study Design and Procedure

As was mentioned above, we compared the following four conditions: individual versus collaborative learning with procedural instruction, and individual versus collaborative learning with conceptual instruction (see Table 1). The study was an initial small scale study to establish basic effects. The study procedure consisted of three phases: a pre-test phase, a learning phase, and a post-test phase. In the pre-test, students individually solved procedural and conceptual problems to assess their prior knowledge. The tests were conducted using paper and pencil. During the learning phase, students solved problems in a tutored learning environment on the computer according to their condition. In the collaborative conditions, two students worked together on one computer to solve the task (i.e. face-to-face interaction). While students worked on the problems, they received immediate feedback concerning the correctness of their problem-solving. The system automatically logged students' problem-solving actions, and their interaction was video-recorded. After the learning phase, students took a post-test that consisted of five problem-sets: near and far transfer items for each knowledge type and a combination problem-set. The order of the procedural and the conceptual problem sets in pre- and post-test was counterbalanced across conditions. As in the pre-test, the post-test was conducted individually using paper and pencil. Students could solve the problems in their own pace both during pre- and post-test and during instruction. The pre-test phase lasted for approximately 15 minutes, the learning phase took approximately 50 minutes, and the five problem sets of the post-test phase took approximately another 50 minutes in total.

Table 1: Study Design and Procedure

	Procedural conditions		Conceptual conditions	
	Individual	Collaborative	Individual	Collaborative
Pre-test phase	individual problem-solving (paper-pencil) order of procedural and conceptual problem-set counterbalanced across conditions			
Learning phase	individually or in dyads: procedural instruction (tutored learning environment)		individually or in dyads: conceptual instruction (tutored learning environment)	
Post-test phase	individual problem-solving (paper-pencil) procedural and conceptual near and far transfer problem-sets (order counterbalanced across conditions) combination problem set			

Material

Instruction During the Learning Phase

The task domain of the study was linear algebra. During the learning phase, students were instructed to solve problems on the computer. The learning environment was developed with the Cognitive Tutor Authoring Tool (CTAT, Aleven McLaren, Sewall, & Koedinger, in press), a software tool developed at the Carnegie Mellon University (CMU), Pittsburgh, that enables researchers and teachers to author intelligent tutoring behavior. In its functionality, it resembles the intelligent tutoring systems described in the introduction; however, it does not require the same amount of time and expertise for developing and adding new tasks. Students received immediate feedback to their actions: errors were marked in red, and correct answers were marked in green. Furthermore, when students made an error, they received a hint that prompted them to reflect on the error in order to find the correct answer (see Figure 1a). Within the procedural and the conceptual conditions, respectively, students working individually and students working collaboratively received the same hints. We wanted to ensure that individual and collaborative conditions received the same amount and type of support. The procedural and conceptual conditions differed in the following way: In the *procedural instruction conditions*, students were asked to solve linear equations (see Figure 1a). The problems had increasing difficulty, reaching from simple equations with one variable and one constant term to equations with several variable terms (e.g. $8x + 5 + 6x = 12$), negative constant and variable terms, and subtraction and multiplication brackets. The total number of problems to solve was 29. In the *conceptual instruction conditions*, students were presented with a linear equation and three story problems (see Figure 1b). For each story problem, they judged the validity of the linear equation to represent the story problem (i.e. "true" or "false"). The errors inserted in the story problems were based on several typical misconceptions of students: for instance, the constant term and the coefficient of the variable term were mixed up, the algebraic signs were incorrect, or the brackets were set in a wrong way. Also some of the correct story problems contained "decoys" that typically yield difficulties in student problem-solving such as irrelevant numeric information or additional conversions that had to be accomplished to solve the task (for instance, convert percent numbers in decimal numbers). For each type of misconception and for each type of decoy, we constructed three story problems. In total, students solved nine problems with one equation and three story problems each. The number of problems to be solved during the learning phase was based on previous test runs in order to keep the time constant between conditions. In all four conditions students could only proceed to the next problem once they had solved a given problem correctly.

We compared conditions based on one outcome and two process variables. The *error rate* measures the relative number of errors on the first attempt. An error rate of 1 indicates that students solved each step incorrectly; an error rate of 0 indicates that students solved each step correctly on the first attempt. As process variables, we extracted the average *time spent prior to an action*, and the average *time spent following an error*. These variables can serve as indicators for cognitive processes in problem-solving (cf. Rummel et al, 2008).

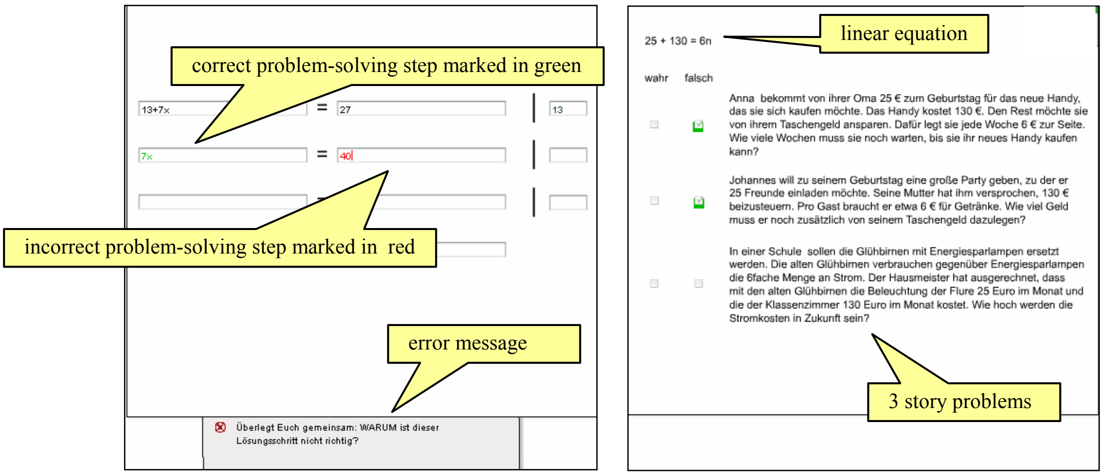

Figure 1: Screenshot of the tutorial learning environments (in German): a) procedural instruction (solving equations) b) conceptual instruction (judging the correspondence of story problems and equations).

Test Material

Both the pre-test and the post-test were solved on paper. Students worked individually without receiving feedback on their problem-solving. The *pre-test* consisted of a *procedural problem set* with eight problems and a *conceptual problem set* with two problems (i.e. two equations with three story problems each). The problems of the pre-test had a lower difficulty level than the problems of the learning phase, but were structurally equivalent.

The post-test consisted of five problem sets: procedural near transfer and procedural far transfer, conceptual near transfer and conceptual far transfer, and combination problems. The *near transfer problems* were structurally equivalent to the problems of the learning phase (12 procedural problems, 3 conceptual problems). In the *far transfer problems*, students were asked to find errors in the solution of a fictitious student and correct them (4 errors in the procedural, 3 errors in the conceptual problems). In the procedural far transfer set, the fictitious student had made several typical computational errors such as combining constant and variable terms when solving equations for the variable. In the conceptual far transfer set, the fictitious student had derived equations from story problems; some of these equations were erroneous, for instance confusing the coefficient with the constant term. Finally, the *combination problems* assessed both knowledge types: In a first step that required conceptual knowledge, students derived an equation corresponding to a story problem; in a second step that required procedural knowledge, they solved the equation for *x*. In total, students solved three combination problems. For the pre-test and the near transfer problems of the post-test, we analyzed the total number of *problems solved correctly*. For the far transfer problems, we extracted two variables: the amount of *errors detected* and the amount of *errors corrected*. Finally, in the combination problems, we evaluated both the amount of *equations that were derived correctly* (conceptual problem-solving step) and the number of *combination problems that were solved correctly* (i.e. students had solved correctly both the conceptual and the procedural problem-solving step).

Participants

Thirty students participated in the study: five students per individual condition and five dyads per collaborative condition. The participants had been recruited from a local high school (Realschule). They were in grade 8 and already had basic experience with solving equations and story-problems. We randomly assigned students to conditions. In the collaborative conditions, students were allowed to choose one of their class mates as their collaboration partner. Most students chose a partner with a similar prior knowledge, thus worked in homogenous dyads. Seventeen students were male, thirteen students were female. Their mean age was 13.77 (.52) years. Conditions did not differ with regard to their prior knowledge assessed as their grade in mathematics and the scores received in the pre-test. In the procedural pre-test, students solved 4.60 out of 8 tasks correctly; in the conceptual pre-test, students solved 3.37 out of 6 tasks correctly.

Results

We compared student behavior during the learning phase, and their individual learning outcomes (for means and standard deviations, see Table 2). Due to the small study sample, we chose an α-level of .10. The differences in the instructional material presented to students in the procedural and the conceptual conditions during the *learning phase* do not allow for direct comparison of student performance across all four conditions. Therefore, for this phase, we compared individual and collaborative performance (error rate) in the procedural conditions, and individual and collaborative performance in the conceptual conditions separately, using ANCOVA analysis (covariate: prior knowledge as assessed in the pre-test). For the *post-tests*, we compared all four conditions with a two-factorial covariance analysis with instruction (procedural vs. conceptual) as factor 1, learning situation (individual vs. collaborative) as factor 2, and the pre-test score as covariate. Factor 1 allows us to evaluate the effect of the instruction on students' learning outcome (e.g. whether procedural instruction improved the outcome in the procedural post-test more compared to conceptual instruction). Factor 2 allows us to evaluate the effect of collaboration on the two knowledge types. Furthermore, we analyzed the differential effect of collaboration with two a priori contrasts: individual vs. collaborative learning within the procedural condition, and individual vs. collaborative learning within the conceptual condition (for the contrasts, only significant results are reported).

Different covariates were included for the analyses of the various parts of our tests in the learning and the post-test phase: For the procedural variables (error rate of the procedural conditions, procedural near and far transfer), we included the pre-test scores from the procedural problem set as covariate; for the conceptual variables (error rate of conceptual conditions, conceptual near and far transfer), we included the pre-test scores from the conceptual problem set as covariate. For the combination problem set, the covariate combined procedural and conceptual pre-test scores: We z-transformed both test scores and merged them to a new variable. Student performance during the learning phase was compared using individual student data from the individual conditions and dyadic student data from the collaborative conditions; the covariate in the collaborative conditions was the dyads' average pre-test scores. For the test phase, both partners of each dyad were included in the analysis as an interclass correlation analysis had confirmed the independency of partners' results (for all correlations $p > .10$).

Table 2: Means and standard deviations of students' performance during learning phase and post-test

	Procedural conditions		Conceptual conditions	
	Individual	Collaborative	Individual	Collaborative
Learning Phase				
Error rate	.19 (.10)	.18 (.05)	.52 (.15)	.47 (.04)
Time before action	17.88 (6.47)	16.85 (1.74)	34.37 (13.27)	46.45 (5.45)
Time after error	23.86 (7.44)	19.27 (3.16)	24.49 (14.67)	26.10 (11.66)
Post-test: procedural problem-set				
Near transfer	6.60 (3.65)	5.80 (3.36)	5.60 (3.44)	5.30 (2.41)
Far Transfer: errors detected	5.00 (1.31)	3.40 (1.06)	3.40 (1.14)	3.20 (1.87)
Far Transfer: errors corrected	3.60 (2.07)	2.60 (1.17)	1.80 (1.30)	1.80 (1.14)
Post-test: conceptual problem-set				
Near transfer	5.20 (2.68)	4.20 (1.48)	5.00 (2.74)	4.90 (1.23)
Far Transfer: errors detected	2.00 (.71)	2.10 (.74)	2.00 (1.00)	2.30 (.67)
Far Transfer: errors corrected	.40 (.89)	.30 (.48)	1.20 (.84)	.60 (.84)
Post-test: combination problem-set				
Derive equation	.80 (.84)	.60 (.52)	.60 (.89)	1.30 (.82)
Correct solution	.80 (.84)	.20 (.42)	.40 (.55)	.90 (.57)

Note: Time before action and time after error are measured in seconds.

Student Behavior in the Learning Phase

Comparison of Performance

First, we compared the performance of students in the individual and the collaborative *procedural condition*. The analysis showed a significant influence of procedural prior knowledge as assessed in the pre-test (covariate) on students' *error rate*, $F(1,7) = 3.55$, $p = .10$, $\eta^2 = .17$. However, no difference between conditions was found, $F < 1.00$. Second, we compared the performance of students in the individual and the collaborative *conceptual condition*. Again, prior knowledge assessed as the conceptual pre-test score showed a significant influence on students' *error rate,* $F(1,7) = 14.72$, $p < .01$, $\eta^2 = .68$, while no difference between conditions was found ($F <$

1.00). As the analysis revealed, students only solved about half of the problems correctly (average error rate across conditions .49). Since the problems in the conceptual instruction were multiple choice, i.e., students had to judge for each story problem if it was consistent with the equation or not, student performance was only on par with the random statistical expectation.

Comparison of Process Variables

The two-factorial analysis of the process variables *time before action* and *time after error* revealed interesting differences between conditions. First, conditions differed regarding the variance of the variable *time before action,* $F(3,16) = 11.95$, $p < .01$: Particularly the individual conditions showed a high variance, while in the collaborative conditions, the dyads spent similar times before entering the next problem-solving action. As the analysis revealed, conceptual conditions spent significantly more time prior to actions than the procedural conditions, $F(1,16) = 42.33$, $p < .01$, $\eta^2 = .72$; however, this result is less surprising due to the differences in the learning material during instruction. We could not establish a significant difference of the factor learning situation, $F(1,16) = 2.43$, $p = .14$, $\eta^2 = .13$. The analysis revealed a significant interaction between instruction and learning situation, $F(1,16) = 3.42$, $p = .08$, $\eta^2 = .17$: While the individual and the collaborative procedural condition showed similar average times, the contrast comparing the two conceptual conditions revealed that dyads spent significantly more time before judging the concordance between the algebraic equations and the story problems than students learning individually, $F(1,16) = 5.81$, $p = .03$. A subsequent correlation analysis revealed that the time spent prior to a student action had a different meaning depending on the type of instruction material. In the procedural conditions, longer times correlated positively with higher error rates ($r = .60$, $p = .07$), in other words, particularly weak students needed time to decide on their next action; in the conceptual conditions, longer times correlated negatively with higher error rates ($r = -.73$, $p = .02$), indicating that elaboration on the translation between algebraic equations and story problems improved problem-solving.

For the variable *time after error*, the Levene test again revealed different variances, $F(3,16) = 5.13$, $p = .01$. Both for the procedural and the conceptual instruction, the process variable showed a higher variance in the individual condition than in the collaborative condition. The variance analysis did not reveal differences between conditions (for all factors $F < 1.00$). As for time before actions, we found a differential impact of conditions on the time spent after errors: In the procedural conditions, there was a trend for a positive correlation between time after errors and error rate ($r = .40$, $p = .25$); in the conceptual conditions, we again found a negative correlation between time after errors and error rate ($r = -.75$, $p = .01$), in other words, students and dyads that showed a better performance during the learning phase spent more time elaborating on errors. This is particularly interesting as the learning material of the conceptual instruction was multiple choice with two options, thus if the first answer was wrong, it was clear which choice to select. This yields the conclusion that good learners did not elaborate on how to correct the answer, but actually tried to understand why their initial choice was wrong. The increased elaboration also had a positive impact on learning: Students in the conceptual conditions who spent more time elaborating before actions and after errors, showed better results in the conceptual near transfer problems (for time before action, $r = .62$; $p = .06$; for time after error, $r = .64$; $p = .05$). In the procedural conditions, we again found a trend in the opposite direction with longer times being related to worse results in the procedural near transfer problems (for time before action, $r = -.52$, $p = .12$; for time after error, $r = -.45$, $p = .20$).

Student performance in the post-test

Comparison of Performance in the Near Transfer Problem Sets

In the *procedural near transfer* problem set, we found a significant influence of the covariate prior knowledge, assessed as the procedural pre-test scores, on the number of problems solved correctly, $F(1,25) = 11.06$, $p < .01$, $\eta^2 = .31$. Students in the procedural conditions solved significantly more tasks correctly than students in the conceptual conditions, $F(1,25) = 2.98$, $p = .10$, $\eta^2 = .11$. Neither the factor learning situation, $F(1,25) = 1.97$, $p = .17$, nor the interaction effect ($F < 1.00$) was significant. In the *conceptual near transfer* problem set, we could not establish significant differences between conditions (for the covariate $F(1,25) = 1.47$, $p = .24$; for all other factors, $F < 1.00$). As during the learning phase, students only solved about half of the problems correctly (4.72 out of 9 problems), thus their performance did not exceed statistical chance.

Comparison of Performance in the Far Transfer Problem Sets

In the *procedural far transfer* problem set, we found a significant influence of the covariate prior knowledge on students' *error detection,* $F(1,25) = 8.46$, $p = .01$, $\eta^2 = .25$. Students in the procedural conditions found significantly more erroneous problem-solving steps than students in the conceptual conditions, $F(1,25) = 6.41$, $p = .02$, $\eta^2 = .20$, and students in the individual conditions outperformed students in the collaborative conditions, $F(1,25) = 6.57$, $p = .02$, $\eta^2 = .21$. As the analysis of the contrasts revealed, the difference between individual and collaborative conditions was due to a better performance of students that had learned with procedural instruction individually when compared to students that had learned with procedural instruction in the collaborative condition, $F(1,25) = 7.81$, $p = .04$. Also with regard to *error correction*, we found a significant influence of the covariate prior knowledge on student performance, $F(1,25) = 6.89$, $p = .02$, $\eta^2 = .22$. Again, the procedural conditions outperformed the conceptual conditions, $F(1,25) = 11.66$, $p < .01$, $\eta^2 = .32$. Neither the factor learning situation,($F(1,25) = 2.26$, $p = .15$ nor the interaction effect were significant, $F(1,25) = 1.18$, $p = .29$.

In the *conceptual far transfer* problem set, the covariate conceptual prior knowledge did not show a significant influence on the number of *errors detected* by students, $F(1,25) = 1.63$, $p = .21$. We neither found a significant influence of the factors learning situation and instruction nor a significant interaction effect (all $F < 1.00$). Also with regard to *error correction,* prior knowledge did not show a significant influence, $F(1,25) = 1.69$, $p = .21$; however, we found a positive impact of the instruction: Students that had received conceptual instruction during the learning phase corrected significantly more erroneous problem-solving steps than students that had solved procedural tasks, $F(1,25) = 3.47$, $p = .07$, $\eta^2 = .12$. Neither the factor learning situation, $F(1,25) = 1.75$, $p = .20$ nor the interaction effect ($F < 1.00$) showed significant results.

Comparison of performance in the combination problem sets

The analysis of the *number of equations correctly derived* (i.e. the conceptual problem-solving step) revealed a significant influence of the covariate that combined the procedural and conceptual prior knowledge as assessed in the pre-test, $F(1,25) = 2.99$, $p = .10$, $\eta^2 = .11$. Neither the factor instruction nor the factor learning situation showed significant on the conceptual problem-solving step (for both factors $F < 1.00$). There was a trend for an interaction effect, $F(1,25) = 2.64$, $p = .12$, $\eta^2 = .10$, that derived from a significant difference between the conceptual conditions: Students that had learned collaboratively with conceptual instruction were significantly better in setting up the equation than students that had learned individually with conceptual instruction, $F(1,25) = 3.18$, $p = .09$. This result indicates a positive impact of collaboration on conceptual knowledge acquisition.

For the number of *problems solved correctly* (i.e. both conceptual and procedural knowledge were required), the covariate did not show a significant influence ($F < 1.00$). However, the analysis revealed a significant interaction of instruction and learning situation, $F(1,25) = 6.11$, $p = .02$, $\eta^2 = .20$: In the procedural conditions, students that had learned individually outperformed students of the collaborative condition, $F(1,25) = 5.15$, $p = .03$, while in the conceptual conditions, the contrary effect was found, $F(1,25) = 3.12$, $p = .09$. Interestingly, all of the students in the procedural individual condition that had found the correct algebraic equation also solved the equation correctly (for both variables, the mean is .80). However, the overall performance of students was quite low: Only a few students were successful in setting up the equation, and on average, they merely solved 0.57 out of 3 tasks correctly.

Discussion

The analysis of students' performance and problem-solving behavior during the learning phase and the analysis of their learning outcome assessed in the post-test provide some support for our hypotheses. In the following, we will shortly summarize the study results. First, we found significant differences between procedural and conceptual instruction on student learning outcome. As hypothesized, the procedural instruction had a positive impact on student performance in the procedural problem sets of the post-test. First, students of the procedural conditions solved more procedural near transfer problems correctly than students of the conceptual conditions; second, they also outperformed students of the conceptual conditions in error detection and correction in the far transfer problem-set. In contrast, the conceptual instruction only had a minor impact on the post-test performance: Students of the conceptual conditions neither outperformed students of the procedural conditions in the near transfer problem set nor in the error detection of the far transfer problem set. However, we found a positive impact of conceptual instruction on students' error correction in the conceptual far transfer problem set. As indicated by the low performance of students in the conceptual problem sets, the conceptual tasks might have been too difficult for students, and the conceptual instruction might not have been enough to compensate for the missing prior knowledge.

The analysis of the process variables of the learning phase also revealed differences between procedural and conceptual conditions. Students in the conceptual conditions spent more time prior to actions than students in the procedural conditions. While this can be explained by the high amount of text to be read (story problems vs. algebraic equations), the meaning of the time spent prior to actions and after errors revealed different between conditions: In the procedural conditions, weaker students did not know how to approach the problems and thus needed more time for the problem-solving steps; good learners needed less time to solve a problem step, indicating that the problem-solving procedures were already automatized (cf. Anderson 1983), that is, students had already reached higher skill fluency. In contrast, in the conceptual conditions, good learners spent more time elaborating on the translation between the story problem and the equation, while weaker learners showed less time prior to solution attempts. Particularly interesting, good learners also spent more time after errors even though in these multiple choice problems, it was clear which choice to select if the first answer was wrong. This indicates that in the conceptual conditions, good learners engaged in deeper cognitive processes than weak learners in order to understand the corrections. Interestingly, this elaborative learning behavior was also positively correlated with a better learning outcome in the conceptual near transfer problems. The differences between the procedural and the conceptual instruction support our assumption that different processes are relevant for the acquisition of these knowledge types: While sufficient practice is most important for procedural knowledge acquisition, the conceptual problems require students to engage in deeper cognitive processes to translate between the different representation formats of the concepts – and these elaborative processes take time.

Since different processes are relevant for the acquisition of these knowledge types, it is likely that collaboration might indeed have a differential effect. Results from the post-tests further confirm this hypothesis. As the comparison of the individual and the collaborative procedural condition revealed, collaborative learning impeded students' procedural knowledge acquisition. Although neither in the performance during the learning phase nor in the procedural near transfer problem set did we find differences between the individual and the collaborative procedural condition, the procedural individual condition outperformed the procedural collaborative condition in error detection in the far transfer test. Furthermore, all students of the procedural individual condition that derived the correct equation in the combination problems (conceptual problem-solving step) were also able to find the right answer (procedural problem-solving step), while this was only the case for a third of the students in the procedural collaborative condition (.60 equations derived correctly, .20 problems solved correctly). In contrast, the study confirmed the effectiveness of collaborative learning for conceptual knowledge acquisition: Although we neither found differences between the individual and the collaborative conceptual condition in the performance during the learning phase and in the near and far transfer problem sets, students who had worked with a partner on conceptual material during the learning phase were significantly better in the conceptual problem-solving step of the combination problem – deriving the equation. The analysis of the learning processes showed that the advantage of the collaborative condition with regard to conceptual knowledge might be explained by students' learning processes: The collaborative conceptual condition spent more time elaborating and discussing the concordance between story problems and equations than the individual conceptual condition; as the correlation analysis revealed, the longer elaboration times were related to better performance and learning outcome.

Overall, the study revealed major difficulties of students in all conditions in solving the conceptual tasks. First, student performance both during learning and post-test phase was only on par with the random statistical expectation. Second, in the combination problem set, only a few students were successful in setting up the equation, and on average, they merely solved 0.57 out of 3 tasks correctly. This result is consistent with the difficulties of students regarding story problems which are often reported in the literature (e.g. Brenner et al., 1997) and indicates an increased need in supporting students' acquisition of conceptual knowledge.

So far, the analysis of the learning phase concentrated on quantitative variables. In the future, we also plan to analyze students' interaction during the learning phase with a qualitative approach. Particularly, this can help us to gain a better understanding of the differences between the learning processes during procedural versus conceptual instruction. For instance, one could hypothesize that the collaboration on the different task types evoke different interaction processes (cf. Dillenbourg et al., 1996): Procedural instruction might rather yield interactions on a low level of elaboration (e.g. discussing what to do next) that are less beneficial for learning, while with conceptual instruction, students might engage in deeper cognitive processes (e.g. discussing why a solution is correct or incorrect). Furthermore, the analysis of students' interaction can reveal differences in the learning processes of students that show good versus bad performance in the conceptual tasks. This might help to develop conceptual instruction that facilitate students' conceptual knowledge acquisition more effectively.

This study was an initial small scale study to establish basic effects for knowledge acquisition in the domain of algebra. Thus, the results are only suggestive, and we will have to see if they can be replicated with a larger study sample and a more stringent alpha. Nevertheless, the results are quite promising and have important implications for the school context. As the study revealed, collaboration is not always equally effective to support student learning. While we have seen that collaborative practice in applying procedures has a negative impact on procedural knowledge acquisition, we found indications that the collaborative elaboration of underlying concepts benefits students' conceptual knowledge acquisition. The analysis of the processes during the learning phase revealed that these differences can be explained by the different learning processes necessary for procedural and conceptual knowledge acquisition. In order to ensure benefits of collaborative learning in the school classroom, it thus is important to introduce collaborative learning more selectively for tasks that require students to elaborate on the mathematical concepts. Particularly for story problems that demand a solid understanding of mathematical concepts and which have shown particularly challenging for students, collaborative learning can be beneficial.

Acknowledgements

This research is supported by the Virtual PhD Program (VGK) of the German Science Foundation (DFG) and by the Landesstiftung Baden-Württemberg.

References

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (in press). Example-tracing tutors: A new paradigm for intelligent tutoring systems. *International Journal of Artificial Intelligence and Education.*

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge: Harvard University Press.

Berg, K. F. (1994). *Scripted cooperation in high school mathematics: Peer interaction and achievement.* Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, Louisana.

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durán, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: The role of Multiple representations in learning algebra. *American Educational Research Journal, 34*(4), 663-689.

Booth, J.L., Koedinger, K.R., & Siegler, R.S. (2007). *The effect of prior conceptual knowledge on procedural performance and learning in algebra.* Poster presented at the 29th Annual Cognitive Science Society conference in Nashville, TN.

Dillenbourg, P., Baker, M., Blaye, A., & O'Molley, C. (1995). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science* (pp. 189-211). Oxford: Elsevier/Pergamon.

Diziol, D., & Rummel, N. (2008). *Evaluating the impact of scripted and unscripted collaboration on students' interactions and learning with the Cognitive Tutor Algebra: A contrasting case analysis.* Manuscript submitted for publication.

Ellis, S., Klahr, D., & Siegler, R. (1993). *Effects of feedback and collaboration on changes in children's use of mathematical rules.* Paper presented at the meeting of the Society for Research in Child Development, New Orleans.

Hausmann, R. G. M., Chi, M. T. H., & Roy, M. (2004). Learning from collaborative problem solving: An analysis of three hypothesized mechanisms. In K. D. Forbus, D. Gentner & T. Regier (Eds.), *26nd Annual Conference of the Cognitive Science Society* (pp. 547-552). Mahwah, NJ: Lawrence Erlbaum.

Hiebert, J & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and instruction, 14*(3), 251-283.

Koedinger, K. R. (1998, June 5-6, 1998). *Intelligent cognitive tutors as modeling tool and instructional model.* Paper presented at the NCTM Standards 2000 Technology Conference.

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.

Mevarech, Z. R. & Stern, E. (1997). Interaction between knowledge and contexts on understanding abstract mathematical concepts. *Journal of Experimental Child Psychology, 65*, 68-95.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction, 9*(4), 329-389.

Nathan, M. J., Mertz, K., & Ryan, R. (1994). *Learning through self-explanation of mathematics examples: Effects of cognitive load.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

O'Donnell, A. M. (1999). Structuring dyadic interaction through scripted cooperation. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning.* (pp. 179-196): Lawrence Erlbaum Associates, Publishers.

OECD (2003.). *The PISA 2003 assessment framework – Mathematics, reading, science and problem solving knowledge and skills.* Organisation for Economic Co-Operation and Development.

Rittle-Johnson, B., & Alibali, M.W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*(1), 175-189.

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

Souvignier, E., & Kronenberger, J. (2007). Cooperative learning in third graders' jigsaw groups for mathematics and science with and without questioning training. *British Journal of Educational Psychology, 77,* 755–771.

Staub, F. C., & Reussser, K. (1995). The role of presentational structures in understanding and solving mathematical word problems. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse Comprehension. Essays in Honor of Walter Kintsch* (pp. 285-305). Hillsdale, NJ: Lawrence Erlbaum.

Walker, E., Rummel, N., & Koedinger, K. R. (2008). To tutor the tutor: Adaptive domain support for peer tutoring. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. P. Lajoie (Eds.), *Lecture Notes in Computer Science, Vol. 5091* (pp. 626-635). Berlin / Heidelberg: Springer Verlag.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Education Research, 13*, 21-39.