To Tutor the Tutor: Adaptive Domain Support for Peer Tutoring

Erin Walker[1], Nikol Rummel[2], Kenneth R. Koedinger[1]

[1] Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[2] Department of Psychology, Albert-Ludwigs-Universitat Freiburg, Germany
erinwalk@andrew.cmu.edu, rummel@psychologie.uni-freiburg.de, koedinger@cmu.edu

Abstract. The effectiveness of intelligent tutoring systems at increasing learning might be improved if the systems were combined with collaborative activities that encouraged conceptual elaboration. We extended the Cognitive Tutor Algebra, an intelligent tutoring system for high-school mathematics, with a peer tutoring activity that was designed to encourage interaction, reflection, and accountability. Two types of domain support were provided: adaptive support, which used the intelligent tutor domain models to provide feedback to the peer tutor, and fixed support, which simply consisted of answers to the problems. We compared the two peer tutoring conditions (adaptive or fixed support) to individual use of the cognitive tutor (without peer-tutoring activities). Even though students in the individual condition solved more problems during instruction, we did not find significant differences between the individual and collaborative conditions on learning. However, we found a correlation between tutee impasses and tutor learning.

Keywords: Peer tutoring, cognitive tutor, algebra, in-vivo experimentation, adaptive collaborative learning system

1 Introduction

The Cognitive Tutor Algebra (CTA) has been shown to increase student learning by roughly one standard deviation over traditional classroom instruction [1], and is used by about 475,000 students a year [2]. However, the impact of the intervention still falls short of the effectiveness of good human tutors, who can improve student learning by two standard deviations over classroom practice [3]. As students may acquire shallow conceptual knowledge while using tutoring systems, researchers augment cognitive tutors with activities that encourage conceptual elaboration such as self-explanation [4] and scripted collaboration [5]. However, it appears that in order for significant improvement over CTA instruction to occur, students must be able and motivated to apply the metacognitive skills targeted by an intervention [6].

In our work, we augment individual use of the CTA with a collaborative peer tutoring activity. Instead of the computer tutoring the student, students take turns tutoring each other (see Figure 1). Tutees can ask their tutors questions and self-explain, and tutors can then provide their tutees with elaborated help. Ideally, students

in the tutee role should benefit from the peer instruction at least as much as students using the CTA individually, and students in the tutor role should benefit even more from the additional conceptual demands of tutoring. However, because peer tutors are also in the process of learning the domain material, they may not be able to provide the tutee with feedback that is timely or correct. The tutee may then be unable to successfully complete the curriculum problems, and will not benefit from the instruction. Therefore, we implemented a meta-tutor that provides adaptive domain support to the peer tutor. In this paper, we discuss how students learn from tutoring and how an adaptive system might support this process, describe the design of the meta-tutor, and compare the adaptive system to a fixed support system and to typical use of the CTA.
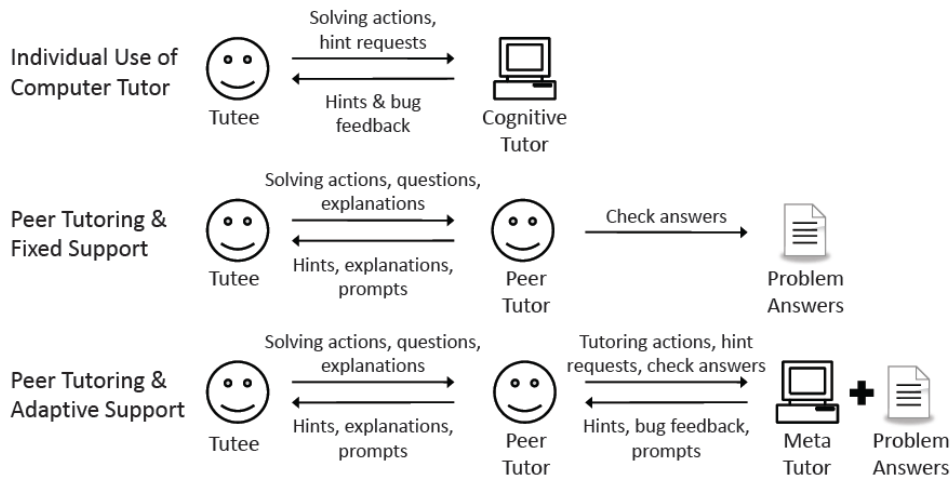

Figure 1. Three tutoring scenarios used in the study.

1.1 Peer Tutoring: Learning by Teaching

Incorporating peer tutoring into the CTA might be a way to encourage deep learning. Roscoe and Chi conclude that peer tutors benefit due to *knowledge-building,* where they reflect on their current knowledge and use it as a basis for constructing new knowledge [7]. Because these positive effects are independent of tutor domain ability, researchers implement reciprocal peer tutoring programs, where students of similar abilities take turns tutoring each other. This type of peer tutoring has been shown to increase academic achievement and positive attitudes in long-term classroom interventions [8]. Biswas et al. [9] described three properties of peer tutoring related to tutor learning: tutors are accountable for their tutee's knowledge, they reflect on tutee actions, and they engage in asking questions and giving explanations. Tutee learning is maximized at times when the tutee reaches an impasse, is prompted to find and explain the correct step, and is given an explanation if they fail to do so [10].

Peer tutors rarely exhibit knowledge-building behaviors spontaneously [7], and thus successful interventions provide them with assistance in order to achieve better learning outcomes for them and their tutees. This assistance can target *tutoring behaviors* through training, providing positive examples, or structuring the tutoring activities. For example, training students to give conceptual explanations had a significantly positive effect on learning [11]. It is just as critical for assistance to target *domain expertise* of the peer tutors, in order to ensure that they have sufficient knowledge about a problem to help their partner solve it. Otherwise, there may be cognitive consequences (tutees cannot correctly solve problems) and affective consequences (students feel that they are poor tutors and become discouraged [12]). Domain assistance can take the form of preparation on the problems and scaffolding during tutoring [e.g., 8]. Although assistance for peer tutoring has generally been fixed, providing adaptive support may be a promising approach.

1.2 Adaptive Collaborative Learning Systems

In order to benefit from collaboration students must interact in productive ways, and collaborative activities can be structured (*scripted*) to encourage these behaviors [e.g., 13]. However, fixed scripts implemented in a one-size-fits-all fashion may be too restrictive for some students and place a high cognitive demand on others [13, 14]. An adaptive system would be able to monitor student behaviors and provide support only when needed. Preliminary results suggest that adaptive support is indeed beneficial: Adaptive prompting realized in a Wizard of Oz fashion has been shown to have a positive effect on interaction and learning compared to an unscripted condition [15]. An effective way to deliver this support would be to use an adaptive collaborative learning system, where feedback on collaboration is delivered by an intelligent agent.

Work on adaptive collaborative learning systems is still at an early stage. One approach is to use machine learning to detect problematic elements of student interaction in real-time and trigger helpful prompts. Although implementations have lead to significant learning gains, the adaptive feedback appears to be disruptive to dyadic interaction [16]. Another promising approach has explored using an intelligent agent as one of the collaborators; students teach the agent about ecosystems with the help of a mentoring agent [9]. However, the agents do not interact with the students in natural language, one of the primary benefits of collaboration.

With respect to peer tutoring, intelligent tutoring technology could be applied either to supporting tutor behaviors or domain knowledge of peer tutors. As it is very difficult to build an intelligent tutor for collaborative processes, we decided to develop a general script for the peer tutoring interaction and then focus on providing *adaptive domain assistance* to peer tutors by leveraging the existing domain models of the CTA. A condition where students tutor each other with adaptive domain support provided to the peer tutor is likely to be better than a condition where the peer tutor merely has access to an answer key, because the support would be tailored to each individual tutor's needs. It is also likely to be better than a condition where students use the CTA individually, because the students in the collaborative condition would be able to interact deeply about the domain material.
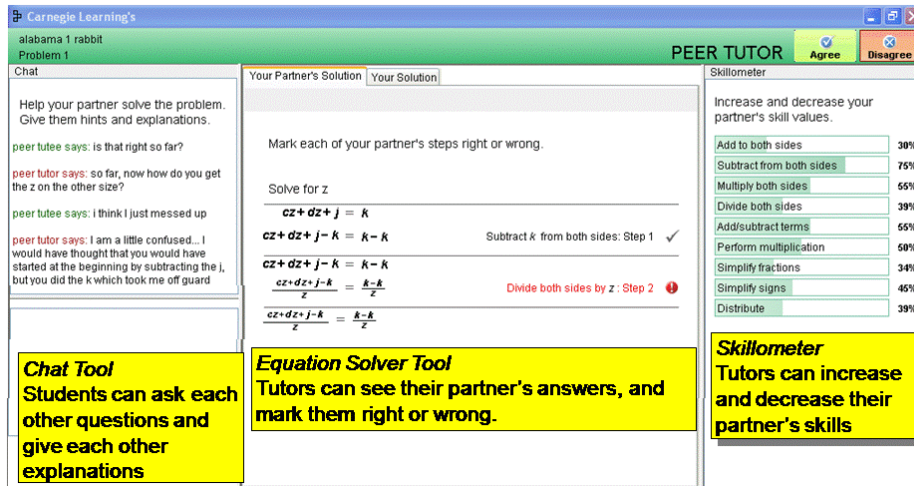
alabama 1 rabbit
Problem 1

PEER TUTOR   Agree   Disagree

Chat

Help your partner solve the problem.
Give them hints and explanations.

peer tutee says: is that right so far?

peer tutor says: so far, now how do you get
the z on the other size?

peer tutee says: i think I just messed up

peer tutor says: I am a little confused... I
would have thought that you would have
started at the beginning by subtracting the j,
but you did the k which took me off guard

Chat Tool
Students can ask each other questions and give each other explanations

Your Partner's Solution | Your Solution

Mark each of your partner's steps right or wrong.

Solve for z

$cz + dz + j = k$

$cz + dz + j - k = k - k$ Subtract k from both sides: Step 1 ✓

$cz + dz + j - k = k - k$

$\frac{cz + dz + j - k}{z} = \frac{k - k}{z}$ Divide both sides by z : Step 2 ⊘

$\frac{cz + dz + j - k}{z} = \frac{k - k}{z}$

Equation Solver Tool
Tutors can see their partner's answers, and mark them right or wrong.

Skillometer

Increase and decrease your
partner's skill values.

Add to both sides	30%
Subtract from both sides	75%
Multiply both sides	55%
Divide both sides	39%
Add/subtract terms	55%
Perform multiplication	50%
Simplify fractions	34%
Simplify signs	45%
Distribute	39%

Skillometer
Tutors can increase and decrease their partner's skills

Figure 2. Peer tutor's interface.

2 Method

2.1 System Design

Peer Tutoring Script. We extended the CTA for peer tutoring using a literal equation solving unit, where students are given a prompt like "Solve for x," and then given an equation like "ax + by = c." Students went through a preparation and collaboration phase. In the *preparation phase*, students individually solved the problems they would later tutor. They used an equation solver to perform operations on the equation, were given immediate feedback from the CTA when making a mistake, and could ask for a hint from the CTA at any time. They were also given feedback on their progress by a Skillometer. After each problem in the preparation phase, we gave students reflection questions to prepare them for tutoring (e.g., "A good question asks why something is done, or what would happen if the problem was solved a certain way. What is a good question to ask about the problem?").

During the *collaboration phase*, students in the same class were grouped into pairs of similar abilities and collaborated with each other at different computers, taking turns being peer tutors and tutees on alternating problems. Although they were located in the same room, they were seated far apart and discouraged from talking to each other out loud. Peer tutees solved the same problems as their tutor had solved in the preparation phase, using the same interface. The peer tutor was able to see the peer tutee's actions, but could not solve the problem themselves (see Figure 2). Instead, the peer tutor took the role of the cognitive tutor, marking the peer tutee's actions right or wrong and adjusting the values of the tutee's skill bars. There was also a chat tool, where tutees could ask questions and tutors could provide hints and feedback.
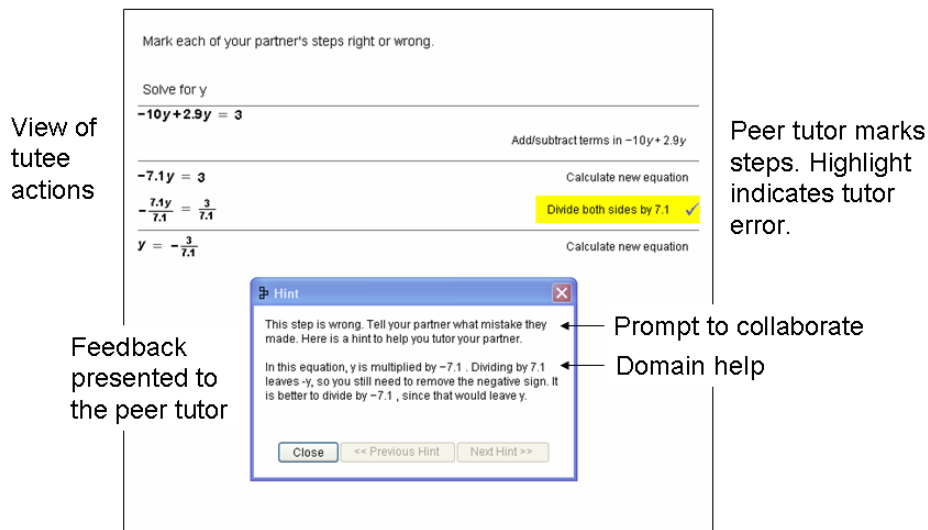
Figure 3. Feedback presented to the peer tutor.

Domain Support. We implemented two different support conditions for peer tutors: fixed domain support and adaptive domain support. In the *fixed support* condition, answers to the problem were located in a separate tab in the interface. Peer tutors could access the tab at any time, but viewing the tab they could no longer see what the tutee was doing. If both the tutee and tutor agreed that the problem was finished the students could move to the next problem, even if they were not actually done.

In the *adaptive support* implementation, peer tutors were given feedback by the intelligent tutoring system in two cases. If the peer tutee asked for a hint, the peer tutor could request it from the cognitive tutor and relay it to the tutee. If the peer tutor marked something incorrectly in the interface (e.g., they marked a wrong step by the tutee correct), the intelligent tutor would highlight the answer in the interface, and present the peer tutor with an error message. Hints and error messages were composed of a prompt to collaborate and the domain help the tutees would have received had they been solving the problem individually (see Figure 3). If both students agreed the problem was done, and were incorrect, the peer tutor would be notified and told to ask for a hint about how to complete the problem. In general, messages provided by the intelligent tutoring system were presented only to the peer tutor, and it was the peer tutor's responsibility to explain them to the tutee. Feedback was based on the peer tutor's actions, and not solely on the peer tutee's actions. As with the fixed support, peer tutors had access to the problem answers in the interface.

2.2 Experimental Design

We compared three conditions: (1) students tutored each other with adaptive domain support in addition to the peer tutoring script (*adaptive collaboration condition*), (2) students tutored each other with fixed domain support in addition to the peer tutoring

script (*fixed collaboration condition*), and (3) students used the CTA individually (*individual condition*). As argued above, we expected the adaptive collaborative condition to learn more than the fixed collaboration and individual conditions because of the combination of tutoring interaction and adaptive domain support.

Participants. Participants were 62 high school students from five algebra classes at a vocational high school in the United States, taught by the same teacher. The high school used the individual version of the CTA as part of regular classroom practice. Students from each class were randomly assigned to one of the three conditions. 11 students were excluded from analysis because either they or their partner were absent during a collaborative part of the intervention, and they were not re-paired with another student. Another 12 participants did not take the delayed posttest, but were included in all other analyses. The total number of students included in the analysis was 51 (20 in the individual condition, 14 in the fixed collaboration condition, and 17 in the adaptive collaboration condition). 39 students took the delayed posttest (18 in the individual condition, 10 in the fixed collaboration condition, and 11 in the adaptive collaboration condition).

Procedure. The study took place over five weeks. Students were given a 15 minute pretest during the first week. The intervention took place during two 70 minute class periods, each one week apart. On both intervention days, students in the collaborative conditions spent half the period in the preparation phase and the remaining time taking turns tutoring each other in the collaboration phase. Students in the individual condition used the CTA alone during both phases. The week after the intervention, students were given a 15 minute posttest. Two weeks later, students were given a 15 minute delayed test to assess their long-term retention. The pre-, post-, and delayed tests were counterbalanced, contained 8 questions, and were administered on paper.

3 Results

3.1 Learning Gains

We scored answers on the pre-, post-, and delayed tests by marking whether the solutions were correct or incorrect. If students got a completely correct solution or reached a nearly correct solution with just a copying error, they received a 1. If students performed at least one important conceptual step incorrectly they received a 0. Points on all the questions were summed, with a maximum score of 8 points. We conducted a two-way (condition x test-time) repeated-measure ANOVA, with test-time (pretest, posttest, or delayed test) as the repeated measure. There was a significant effect for test-time ($F(2,72) = 41.303$, $p < .001$), but there were no significant differences between conditions, and no interaction. A priori contrasts revealed that the effect was due to the difference between the pretest and the other two tests ($t(36) = 69.541$), $p < .001$) and not due to the difference between the posttest and the delayed posttest ($t(36) = 2.544$, $p = .119$). Table 1 contains the scores of the

students who took all three tests. For the correlational analysis in the remainder of this section, we computed normalized gain scores for the posttest and the delayed test.

Our hypothesis that the adaptive collaboration condition would lead to more learning than the other two conditions was not supported by the data. We next investigated how process related to learning outcomes, by examining student progress through the unit, the effect of tutee impasses, and the feedback that students received.

Table 1. Pre, post and delayed test scores.

	Pretest		Posttest		Delayed Posttest	
Condition	M	SD	M	SD	M	SD
Individual	1.28	1.60	3.00	1.75	3.67	1.78
Fixed	.90	.876	3.50	2.17	3.60	2.17
Adaptive	.82	1.08	2.36	1.57	2.82	1.78

3.2 Student Progress

We expected that the collaboration conditions might complete fewer problems than the individual condition because students spend more time interacting. However, tutees in all conditions should make similar numbers of incorrect problem-solving actions per problem if they receive comparable support from their tutors. We conducted a one-way (condition: individual, fixed, adaptive) ANOVA on the number of problems completed per hour in the collaboration phase of the study (see Table 2). For this analysis, we grouped the students in the collaborative conditions by dyad, as the number of problems that one pair member completes (and the time that they take) is dependent on the number of problems the other pair member completes. Condition was indeed significantly related to problems solved ($F(2,34) = 8.764$, $p = .001$). We then conducted a one-way (condition: individual, fixed, adaptive) ANCOVA on the average number of incorrect actions per problem (see Table 2). We included pretest as a covariate because it was significantly correlated with incorrect attempts. Because we wanted a comparable measure of individual progress across conditions, we looked at incorrect attempts per problem *for each tutee*, rather than by dyad. Pretest was significantly predictive of incorrect attempts ($F(1,47) = 5.449$, $p = .024$). Condition marginally affected incorrect attempts per problem ($F(2,47) = 2.480$, $p = .095$).

Table 2. Problems completed, incorrect attempts, help requested, and help given.

		Problems Completed per hour			Incorrect Attempts per problem		Help Requested per problem		Help Given per problem	
Condition	N	M	SD	N	M	SD	M	SD	M	SD
Individual	20	47.0	30.2	20	1.46	1.26	.648	.806	1.41	1.41
Fixed	8	13.3	7.71	14	1.81	1.04	.929	.647	.943	.839
Adaptive	9	17.7	5.69	17	2.46	1.87	1.32	1.60	1.96	1.63

Table 3. Types of help requested and given.

Role	Utterance Type	Utterance Content	Measure
Tutor	Explanation	you need to factor f to get it by itself	Help Given
Tutor	Hint	you need to get the 4 away from the t	Help Given
Tutor	Prompt	but where is the -1.3333 come from?	Help Given
Tutor	Instruction	divide y+r to both sides	Help Given
Tutor	Confusion	I'm not sure I going to check the hint	None
Tutee	Specific Request	now do I divide both sides by (3-x)	Help Requested
Tutee	General Request	what do I do next	Help Requested
Tutee	Specific Description	I still have the m-n on the y side	None
Tutee	General Description	I think I did it right	None
Tutee	Confusion	I have nooooooo clue	Help Requested

3.3 Tutee Impasses

The collaborative conditions differed on how easy it was for students to move to the next problem. In the adaptive condition, students could not continue unless they had successfully completed the problem, making it possible for students to get "stuck", where they repeatedly tried incorrectly to move to the next problem. The number of these incorrect done tries was negatively correlated with tutee gain scores on the delayed test ($r = -.591$, $p = .056$), but positively correlated with tutor gain scores on the delayed test ($r = .463$, $p = .115$). In the fixed condition, students were not notified when their attempts to continue were incorrect, and thus could "skip" to the next problem even if the previous problem was not done. Problems skipped were negatively correlated with tutee learning ($r = -.614$, $p = .059$) and tutor learning ($r = -.369$, $p = .329$). If problems were skipped tutors did not benefit from tutee impasses.

To further investigate, we looked at how incorrect attempts might be related to tutee and tutor learning. In the adaptive collaboration condition, total incorrect problem-solving attempts were negatively correlated with tutee gain scores on the delayed test ($r = -0.614$, $p = .044$), but *positively* correlated with the delayed gain score of the tutor ($r = .428$, $p = .190$). Posttest correlations with tutee learning ($r = -.206$, $p = .427$) and tutor learning ($r = .320$, $p = .210$) were not as strong. In the fixed collaboration condition, the pattern was still present, but slightly weaker for the tutees, perhaps because students could sidestep impasses by skipping to the next problem; incorrect attempts were negatively correlated with tutee gains on the delayed test ($r = -.378$, $p = .281$) and positively correlated with tutor gains on the delayed test ($r = .472$, $p = .199$). In this condition, posttest scores were not correlated with incorrect attempts for the tutee ($r = -.046$, $p = .876$) or the tutor ($r = .034$, $p = .917$).

3.4 Feedback Received

Even though the number of incorrect attempts made across conditions was not significantly different, there may have been differences in the way tutors gave feedback. We computed a measure for help requested by counting the number of

times the students in the individual condition clicked on the hint button and the number of times peer tutees in the collaborative conditions expressed confusion or asked a question (see Table 3). We then computed a measure of help given by the peer tutor by counting the number of times the cognitive tutor gave a feedback message and the number of times the peer tutor gave advice (see Table 2). Help given was not significantly different across conditions ($F(2,48) = 2.16$, $p = .127$), nor was help requested ($F(2,48) = 1.794$, $p = .191$).

4 Discussion

Both individual use of the CTA and peer tutoring activities lead to significant learning gains. The fact that the collaborative conditions performed just as well as the individual condition is encouraging. They achieved similar gains even though they solved fewer problems, suggesting that interacting in depth about a small number of problems might be as efficient for learning as solving a large number. Further, peer tutoring can have social and attitudinal benefits [8], so students may gain more from peer tutoring than from working individually.

It is surprising that quantitative measures of student progress and feedback exchange within problems were so similar across all three conditions. Although one would expect hint requests to be the same across tutees and individual problem-solvers, one might expect that peer tutors would have difficulty giving hints to a tutee compared to an intelligent system, either delivering more or less help than necessary. However, the results indicate that on a broad level the mechanisms of intelligent tutoring and novice tutoring are similar. It appears that to improve the effects of the tutoring, the best approach may be to focus on the details of the interaction by coding utterances for the type of help given or requested (e.g., using Webb's coding scheme [17]). We can look more closely at how different feedback provided to peer tutors affects quality and timing of their help, and how those elements might relate to tutor and tutee learning.

The tutors' apparent benefit from tutee impasses, which were negatively correlated with tutee learning gains, is problematic since it suggests that in order for the tutor to improve, the tutee must struggle. This result occurred mainly on the delayed test, which is a measure of long-term retention, and therefore a measure of deeper learning than the posttest. It is important to be cautious in interpreting this correlation, but it is consistent with the result that viewing erroneous worked examples may improve student learning [18]. If this is the case, it is important to give students the opportunity to experience these impasses. However, we need to examine in greater detail why tutors could not help tutees benefit from impasses, and provide assistance for these circumstances to better support the peer tutor in explaining the relevant concepts to the tutee. Focusing on this aspect, we may be able to use collaborative learning to improve on individual use of the CTA.

References

1. Koedinger, K., Anderson, J., Hadley, W., & Mark, M. Intelligent tutoring goes to school in the big city. International Journal of Artificial Intelligence in Education, 8, 30-43 (1997)
2. Carnegie Learning, www.carnegielearning.com
3. Bloom, B. S. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13, 3-16 (1984)
4. Aleven, V., & Koedinger, K.R. An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. Cognitive Science 26, 147-179 (2002)
5. Diziol, D., Rummel, N., Spada, H., & McLaren, B. Promoting learning in mathematics: Script support for collaborative problem solving with the Cognitive Tutor Algebra. In: C. A. Chinn, G. Erkens & S. Puntambekar (eds.), CSCL 2007, pp. 39-41. (2007)
6. Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. Can help seeking be tutored? Searching for the secret sauce of metacognitive tutoring. In: R. Luckin, K. Koedinger, & J. Greer (eds.), AIED 2007, pp. 203-10. (2007)
7. Roscoe, R. D. & Chi, M. Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. Review of Educational Research 77(4), 534-574 (2007)
8. Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. Journal of Educational Psychology 81(2), 173-177 (1989)
9. Biswas, G., Schwartz, D. L., Leelawong, K., Vye, N., & TAG-V. Learning by teaching: A new agent paradigm for educational software. Applied Artificial Intelligence 19, 363–392 (2005)
10. VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. Why do only some events cause learning during human tutoring? Cognition and Instruction 21(3), 209-249 (2003)
11. Fuchs, L., Fuchs, D., Hamlett, C., Phillips, N., Karns, K., & Dutka, S. Enhancing students' helping behaviour during peer-mediated instruction with conceptual mathematical explanations. The Elementary School Journal 97(3), 223-249 (1997)
12. Medway, F. & Baron, R. Locus of control and tutors' instructional style. Contemporary Educational Psychology, 2, 298-310 (1997).
13. Rummel, N. & Spada, H. Can people learn computer-mediated collaboration by following a script? In F. Fischer, I. Kollar, H. Mandl &, J. Haake, Scripting computer-supported communication of knowledge. Cognitive, computational, and educational perspectives (pp. 47-63). New York: Springer. (2007)
14. Dillenbourg, P. Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), Three worlds of CSCL. Can we support CSCL (pp. 61-91). Heerlen: Open Universiteit Nederland. (2002)
15. Gweon, G., Rosé, C., Carey, R. & Zaiss, Z. Providing Support for Adaptive Scripting in an On-Line Collaborative Learning Environment. Proc. of CHI 2006, pp. 251-260. (2006)
16. Kumar, R., Rosé, C. P., Wang, Y. C., Joshi, M., Robinson, A. Tutorial dialogue as adaptive collaborative learning support. Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED 2007), Amsterdam: IOSPress. (2007)
17. Webb, N., Troper, J., & Fall, R. Constructive activity and learning in collaborative small groups. Journal of Educational Psychology, 87(3), 406-423 (1995)
18. Große, C. S. & Renkl, A. Finding and fixing errors in worked examples: Can this foster learning outcomes? Learning & Instruction, 17, 612-634 (2007)