Investigating the Effect of Meta-cognitive Scaffolding for Learning by Teaching

Noboru Matsuda[1], Cassondra L. Griger[1], Nikolaos Barbalios[1], Gabriel J. Stylianides[3], William W. Cohen[2], and Kenneth R. Koedinger[1]

[1] Human-Computer Interaction Institute, Carnegie Mellon University, USA
{mazda,grigercl,nbarba,krk}@cs.cmu.edu
[2] Machine Learning Department, Carnegie Mellon University, USA
wcohen@cs.cmu.edu
[3] Department of Education, University of Oxford, Oxford, UK
gabriel.stylianides@education.ox.ac.uk

Abstract. This paper investigates the effect of meta-cognitive help in the context of learning by teaching. Students learned to solve algebraic equations by tutoring a teachable agent, called SimStudent, using an online learning environment, called APLUS. A version of APLUS was developed to provide meta-cognitive help on what problems students should teach, as well as when to quiz SimStudent. A classroom study comparing APLUS with and without the meta-cognitive help was conducted with 173 seventh to ninth grade students. The data showed that students with the meta-cognitive help showed better problem selection and scored higher on the post-test than those who tutored SimStudent without the meta-cognitive help. These results suggest that, when carefully designed, learning by teaching can support students to not only learn cognitive skills but also employ meta-cognitive skills for effective tutoring.

Keywords: Learning by teaching, teachable agent, SimStudent, Algebra equation solving, meta-cognitive help.

1 Introduction

The effect of learning by teaching has been well known [1, 2] in many disciplines for diverse student populations and skill levels. Many empirical studies observe that when students tutor each other, not only tutees but also tutors learn—often called the *tutor-learning effect*. Yet it is only recently that researchers have started to investigate why and how students learn by teaching. This scholarly development is largely due to the growing maturity of advanced learning technologies that allow students to interactively tutor a synthetic peer, commonly called a *teachable agent* [3]. The teachable agent technology allows researchers to collect detailed interaction data to understand the relationship between tutoring activities and the tutor-learning outcome [4, 5].

Learning by teaching is a complicated phenomenon that includes many factors to be considered, which are often hard to control [2, 6]. Therefore, researchers conduct exploratory studies that focus on particular aspects of tutor learning and the functionalities of the learning by teaching environment. The current paper focuses on the effect of the *meta-cognitive help* for learning by teaching.

S. Trausan-Matu et al. (Eds.): ITS 2014, LNCS 8474, pp. 104–113, 2014.

Biswas et al. examined the effect of the meta-cognitive assistance for tutor learning [7]. Students taught Betty's Brain, the teachable agent, about river ecosystems. There was a mentor agent who provided both cognitive help (e.g., corrective feedback on the errors that Betty's Brain made on the quiz) and meta-cognitive help (e.g., how to gauge what Betty's Brain knows about the river ecosystems). In the classroom study, they found no effect of the mentor agent on tutor learning. In the current study, however, since students need to learn both procedural skills and conceptual knowledge, we might see different effect of the meta-cognitive help.

Walker et al. [8] compared "adaptive" and "fixed" meta-cognitive help for tutor learning in Algebra equations where pairs of students teach each other. The "adaptive" help was contextualized, whereas the "fixed" help was provided randomly. The results from a classroom study showed that the "adaptive" meta-cognitive help is more effective for tutor learning than the "fixed" meta-cognitive help. The current study will build on these findings to further investigate the effect of the meta-cognitive help for tutor learning.

Our previous studies showed that students often failed to select appropriate problems to tutor [4]. Therefore, we hypothesized that providing students with scaffolding on how to select problems to tutor would facilitate tutor learning. On the other hand, to select appropriate problems to tutor, students need to gauge their tutees' proficiency. Therefore, we further hypothesized that providing students with scaffolding on how to gauge tutee's proficiency would amplify the effect of the meta-cognitive help on problem selection, which would result in better tutor learning. To test these hypotheses, we used the online learning environment (called APLUS) where students learn to solve algebra equations by teaching a teachable agent called SimStudent.

2 SimStudent and APLUS

2.1 SimStudent

SimStudent is a computational model of learning, realized as a machine-learning agent, which can be interactively tutored. It is implemented with various AI techniques including programming by demonstration in the form of inductive logic programming, version space, and iterative-deepening search [4].

SimStudent learns cognitive skills in the form of production rules by generalizing *positive examples* (showing when to apply a particular skill, e.g., adding a constant to both sides) and *negative examples* (showing when not to apply a particular skill).

When SimStudent is used as a teachable agent, the affirmative feedback from the student for steps performed by SimStudent and the steps demonstrated by the student as a hint become positive examples, whereas the negative feedback becomes negative examples. A hint from the student on how to perform the next step also becomes a positive example. The next section provides details about the interaction between the student and SimStudent. See [4] for more technical details.

2.2 APLUS

Figure 1 shows an example screenshot of APLUS. To teach SimStudent, shown as an avatar (g), a student enters an equation in the first row, e.g., 2x+4 = 2 (c). When a problem is entered, SimStudent attempts to solve it step-by-step (d) by applying already learned productions. If SimStudent was able to perform a step, it asks the student about the correctness of the step performed. The student then provides *yes/no feedback*. When the student's feedback is negative (i.e., "no," which means the student thinks that the step performed by SimStudent is incorrect), then SimStudent makes another attempt, if able.

If there is no production applicable to perform a step, SimStudent asks the student what to do for the next step. The student provides help by actually performing the next step in the tutoring interface (i.e., entering the text in the next empty cell) (e).

Resources are available at the top left corner of the APLUS interface for students to review in order to prepare for tutoring (b). The [Introduction Video] tab shows a 10-minute video clip explaining how to use APLUS. The [Unit Overview] tab summarizes how to solve equations with worked-out examples. It also shows suggestions of problems to be used for tutoring. The [Examples] tab shows worked-out examples in the tutoring interface with detailed explanations about how to perform each step.

The quiz has four sections ordered by difficulty: 1 one-step equation, 3 two-step Equations, 4 equations with variables on both sides, and Final Challenge that has 8 equations with variables on both sides. Quiz sections are "locked" until the previous section is passed (i.e., all equations in the section are solved correctly). When the

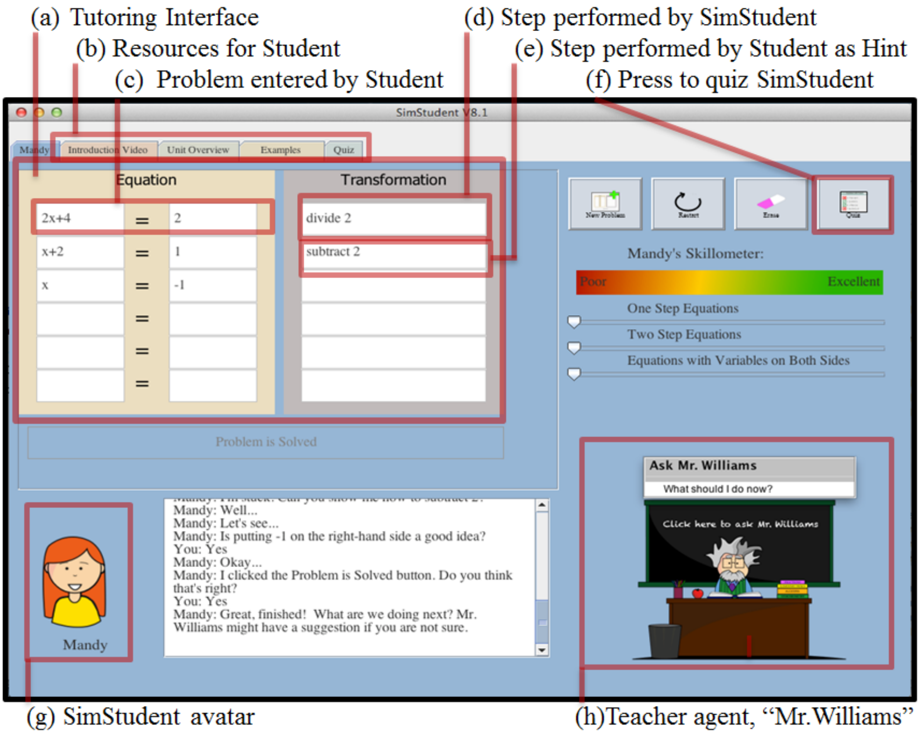

Fig. 1. Annotated sample screenshot of APLUS

student quizzes SimStudent (f), SimStudent attempts to solve quiz problems by applying learned productions. Mr. Williams, the teacher agent shown in the lower right corner (h), then summarizes SimStudent's performance on the quiz. The student can review the exact solutions made by SimStudent one by one in the tutoring interface.

In the meta-tutor version of APLUS, students can click on Mr. Williams to ask him for help. The next section explains details about the meta-cognitive help.

2.3 Meta-tutor with Meta-cognitive Help

In a version of APLUS, Mr. Williams performs as a *meta-tutor* who provides meta-cognitive help when asked. For the current version, two types of meta-cognitive help are available: (1) the *quiz help* suggests to students when to quiz their SimStudent and explains why (e.g., "It's a good strategy to quiz Mandy, because it would help you to understand what Mandy already knows. Click the Quiz button."), and (2) the *problem help* suggests to students what problem should be tutored next and explains why (e.g., "Since Mandy was wrong on the quiz, you may want to give 4y-8=10 to Mandy.").

Meta-tutor's help is thus available only when a problem is completed or a quiz is done. When the student asks for help by clicking on Mr. Williams, Mr. Williams shows only one menu item saying, "What am I supposed to do now?" (Figure 1-h); otherwise, Mr. Williams says, "You should complete the problem."

We use a model-tracing technique [9] to control the meta-tutor. That is, we have a (meta)cognitive model of how to tutor SimStudent, written as a set of production rules. Each production has associated hint messages. A student's tutoring activities are model-traced using the (meta)cognitive model so that when the student asks for help, the meta-tutor can provide just-in-time suggestions. Currently, there are six production rules: three for quiz and three for problem help.

3 Evaluation Study

3.1 Research Questions and Hypotheses

The goal of the evaluation study was to understand the effect of the meta-cognitive help provided by the meta-tutor. In particular we address the following two research questions: (1) Does the meta-tutor providing the quiz and problem help facilitate tutor learning? (2) If so, how does each type of help affect tutor learning?

We hypothesized that selecting problems based on the quiz results is an effective strategy, because it allows students to address specific weaknesses of their SimStudent's learning. Therefore, providing a meta-cognitive hint on problem selection based on quiz results should facilitate tutor learning—the *problem hint* hypothesis. To make the quiz-based problem selection work, students need to quiz SimStudent with appropriate timing. Thus, we also hypothesized that a meta-cognitive hint on when to quiz, in combination with the problem hint, should further facilitate tutor learning—the *quiz hint* hypothesis.

3.2 Methods

A classroom (in-vivo) study in the normal Algebra I classes at an urban public middle school in Pittsburgh, Pennsylvania was conducted with assistance of Pittsburgh

Science of Learning Center. The study was a randomized controlled trial with two conditions. The Meta-Tutor condition used the version of APLUS with the meta-tutor described in section 2.3. The baseline condition used the basic version of APLUS, which also had Mr. Williams but it did not provide the meta-cognitive help.

The study was five 42 minutes classroom periods over five consecutive days. On the first day, all students took an online pre-test (section 3.4) and then watched the introduction video available in APLUS. Students were then randomly assigned to a study condition. On the second through the fourth day, students used the assigned version of APLUS. On the fifth day, students took an online post-test.

Students were told that their goal was to have SimStudent pass the quiz, and SimStudent must learn how to solve equations with variables on both sides to pass the quiz (which is also mentioned in the Unit Overview). We will therefore call equations with variables on both sides as the *target equation* hereafter.

3.3 Participants

One hundred seventy-three (173) 7th through 9th grade students in nine Algebra-I classes participated in the study. A classroom-level randomization was applied to eight classes, and a within-class randomization for the remaining class. Out of those 173 students, 151 were present in the class on the first day and took the pre-test, 127 participated all three days for tutoring SimStudent, and 121 took the post-test.

As the result, 112 out of 173 students took both pre- and post-tests and participated in all three days of tutoring sessions. Those 112 students (53 in the Meta-Tutor condition and 59 in the Baseline condition) are included in the following data analysis. No other criteria for inclusion were used.

3.4 Measure

The online test consisted of two parts—Procedural Skill Test and Conceptual Knowledge Test. The *Procedural Skill Test* consisted of three sections: (1) The equation section had 10 equation problems with four one-step equations, two two-step equations, and four target equations. (2) The effective next step section had two problems each showing an equation with four options for a next step: add or subtract a term from both sides, or multiply or divide both sides by a constant. Students were asked to indicate whether each option was correct or not. (3) The error detection section had three problems each showing an incorrect solution for a given equation with multiple intermediate steps that contained one (and only one) incorrect step. Students were asked to identify the incorrect step. 53% (8 out of 15) of Procedural Skill Test items were about the target type of equation (with variables on both sides).

The *Conceptual Knowledge Test* consisted of 24 true/false items with seven items asking about variable terms, six items asking about constant terms, six items asking about like terms, and five items asking about equivalent terms.

After the study, the reliability of the test items was evaluated using Cronbach's alpha. For the Procedural Skill Test, the equation section showed $\alpha = .87$, the effective next step section had $\alpha = .76$, and the error detection section had $\alpha = .57$. Due to the low reliability index, we decided to exclude the error detection section from the analysis (and refer the average of other two sections as the score for the Procedural Skill Test). For the Conceptual Knowledge Test, $\alpha = .89$.

Table 1. Means (and standard deviations) for the Conceptual Knowledge Test (CKT) and the Procedural Skill Test (PST) by condition

	CKT		PST	
	Pre-test	Post-test	Pre-test	Post-test
Baseline	.43(.25)	.54(.21)	.69(.24)	.71(.25)
Meta-tutor	.43(.30)	.49(.22)	.71(.23)	.78(.19)
Total	.43(.27)	.52(.21)	.70(.24)	.74(.23)

In the analysis below, we also used the process data in addition to the learning outcome data (i.e., test scores). APLUS automatically logged detailed interaction between the student and the system included the problems used for tutoring, frequency of quiz, status of the resource and meta-tutor usage, and suggestions from the meta-tutor, etc. The correctness of steps suggested by SimStudent, and the accuracy of feedback and hints that students provided to SimStudent were also logged. Cognitive Tutor Algebra-1 [10] was embedded into the system to compute accuracy of feedback and hints for the purposes of logging.

4 Results

4.1 Test Scores

Table 1 shows the test scores. For the Procedural Skill Test, there was a reliable condition difference on the post-test scores—a one-way ANCOVA with the pre-test score as a covariate revealed a statistically significant difference on the post-test; $F(1,110) = 3.99$, $p < 0.05$. The effect size, Cohen's d, was 0.30. A post-hoc analysis revealed that only the Meta-Tutor condition showed a significant increase from pre- to post-test; $paired\text{-}t(52) = -2.96$, $p < 0.01$. No pre- and post-test difference was observed for the Baseline condition; $paired\text{-}t(58) = -0.68$, $p = 0.45$.

For the Conceptual Knowledge Test, there was no reliable condition difference observed, but the difference between pre- and post-test scores (when aggregated across all students in the two conditions) was statistically significant; $M_{pre} = .43$ ($SD = 0.27$) vs. $M_{post} = .52$ ($SD = 0.22$). A two-way repeated measures ANOVA with test-time (pre vs. post) as a within-subject variable and condition as a between-subject variable revealed a main effect of test-time; $F(1,110) = 18.32$, $p < 0.001$.

4.2 Meta-tutor Help

On average, students in the Meta-Tutor condition (N=53) asked Mr. Williams for help 5.5 times ($SD = 7.1$). The distribution was very skewed—11 (21%) students did not ask Mr. Williams at all, while 24 (45%) of students asked up to three times. Data also showed that different students apparently had different biases on the timing of hint requests—45% of students did not receive meta-tutor's message for the problem help at all, whereas 49% did not receive the message for the quiz help at all.

Despite the surprisingly low frequency of meta-tutor use, there was a reliable difference on the Procedural Skill post-test between conditions. Since the meta-tutor only provided quiz help and problem help, we predicted a difference in the way students quizzed SimStudent and selected problems to tutor that affected tutor learning.

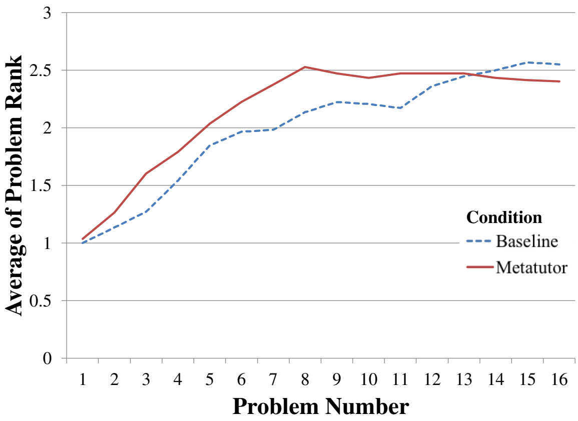

Fig. 2. The transition of problem types. The y-axis shows the problem "rank." The x-axis shows the number of problems tutored. On the 8th problem, the majority of the Meta-Tutor students had tutored the rank 3 problems, i.e. the target equations.

A regression analysis showed the number of help asked was not a reliable predictor of the Procedural Skill post-test score; $F(1,50) = 0.05$, $p = 0.82$. The probability of following meta-tutor's advice had no correlation with the post-test score either.

On average, students in each condition tutored 32.9±9.7 (Baseline) and 29.8±7.3 (Meta-tutor) problems. The number of problems tutored per se was not a reliable predictor of the Procedural Skill post-test. However, the type of problems (i.e., one- step, two-step, and target equations, which are equations with variables on both sides) tutored was a reliable predictor for the Procedural Skill post-test. The percent ratios of each problem type to all problems tutored were used as independent variables to predict the Procedural Skill post-test score. All three independent variables turned out to be statistically reliable predictors: $PST_{post} = -0.48 \times P_{ONE} + 0.99 \times P_{TWO} + P_{TGT} \times 0.63$ ($r^2 = 0.93$) where PST_{post} means the Procedural Skill post-test score; P_{ONE}, P_{TWO}, and P_{TGT} show the percent of one-step, two-step, and target equations tutored, respectively; for P_{ONE}, $F(1,109) = 902.97$, $p < 0.001$; for P_{TWO}, $F(1, 109) = 580.20$, $p < 0.001$; and for P_{TGT}, $F(1,109) = 117.98$, $p < 0.001$.

We then hypothesized that the meta-tutor's advice affected the way students selected problems, and in particular, students in the Meta-Tutor condition made quicker transitions from one-step equations to more advanced types of equations than the Baseline students. The data in Figure 2 support our hypothesis. In the figure, we "ranked" the types of problems that students used for tutoring: the rank is "1" for one-step equations, "2" for two-step equations, and "3" for target equations. The x-axis shows the chronological number of problems tutored. The y-axis shows the average "rank" of the problem tutored aggregated across all students in each condition. As we hypothesized, the Meta-Tutor condition showed a steeper slope that reached to 2.5 on the 8th problem, meaning that the majority of the students started to tutor target problems on and then after the 8th problem. On the other hand, it was around the 14th problem before the Baseline students started tutoring the target problems.

A regression analysis confirmed that our hypothesis was supported. Since two conditions did not reach the 2.5 rank-level in the same way, we computed the regression slope for the first 8 problems for the Meta-Tutor (MT) condition and the first 15 problems for the Baseline (BL) condition. The regression analysis revealed a significant difference between the slopes for the two conditions: $\beta_{MT} = 0.22$ vs. $\beta_{BL} = 0.11$,

$F(1, 1298) = 33.60$, $p < 0.001$. $r^2 = 0.25$. The Meta-Tutor students made a quicker transition from the entry-level problems to the target problems than the Baseline students.

We also examined the effect of quiz help. Since the meta-tutor (if asked) suggested quizzing SimStudent before tutoring, we hypothesized that the Meta-Tutor (MT) students showed a higher probability of starting the tutoring session with quiz than the Baseline (BL) students. This hypothesis was not supported. There was no difference in the probability of starting with quiz; $M_{MT} = .08$ ($SD = .07$) vs. $M_{BL} = .07$ ($SD = .06$), $t(107) = 1.98$, p = 0.88. We also computed the probability of "appropriate" tutoring actions, which, by definition, is the ratio of selecting problems based on the quiz results and quizzing after tutoring to all tutoring activities (which the meta-tutor also suggested upon a request). Again, there was no condition difference in their averages: $M_{MT} = 0.33$ (SD = .16) vs. $M_{BL} = 0.33$ (SD = .13), $t(104) = 0.22$, p = 0.83.

4.3 Accuracy of Tutoring

On average, 70% (SD = 22%) of Hints and 73% (SD = 10%) of Feedback that students provided to SimStudent were correct. To measure the overall accuracy of tutoring, we computed the Response Accuracy as 2×HA×FA/(HA+FA), where HA means the accuracy of Hints and FA means the accuracy of Feedback. The overall mean Response Accuracy was .70 (SD = .17).

It turned out that the Response Accuracy (RA) was a reliable predictor of the Procedural Skill post-test score (PST); $F(1,109) = 10.7$, $p = 0.001$; even when the PST pre-test score was controlled; $F(1,109) = 56.9$, $p < 0.001$; the model equation $PST_{post} = 0.55 \times PST_{pre} + 0.34 \times RA + 0.12$ ($r^2 = 0.56$). The Response Accuracy was also a reliable predictor of the Conceptual Knowledge post-test score (CKT); $F(1,109) = 30.42$, $p < 0.001$; even when the CKT pre-test score was controlled; $F(1, 109) = 5.55$, $p < 0.05$; the model equation $CKT_{post} = 0.40 \times CKT_{pre} + 0.26 \times RA + 0.16$ ($r^2 = 0.41$).

4.4 Resource Usage

There was no notable condition difference in the way students used the resources—in general, students did not use resources as often as we expected. **Table 2** shows average frequency and duration. Example problems were reviewed 29 times on average per student, but the average total duration on examples was only about 10 seconds per student. Regression analyses revealed that both frequency and duration of resource usage were not reliable predictors of the post-test score for the Procedural Skill and Conceptual Knowledge Tests.

Table 2. Average frequency (top) and duration (bottom) of resource usage

	Video	Unit Overview	Examples
Frequency	2.2 (3.4)	4.1 (8.3)	29.3 (33.6)
Duration	7.3s (38.2s)	6.6 (11.5s)	10.6s (13.1s)

5 Discussion

The data showed that the ability to tutor the *target* problems *correctly* (operationalized as the ratio of target problems tutored and the response accuracy as shown in sections 4.2 and 4.3) had a strong predictive power for the Procedural Skill post-test score, regardless of the availability of the meta-tutor. This finding is a replication of our previous study [4] that used the same version of APLUS that was used in the control condition of the current study.

The data also showed that the meta-cognitive help provided by the meta-tutor positively affected tutor learning. In particular, suggestions provided by the meta-tutor allowed students to make appropriate transition in tutoring from entry-level equations to the target equations. This finding supports the previous observation that learning by teaching is not an automated process, but rather requires careful scaffolding [4].

Despite the meta-tutor's assistance, many Meta-Tutor students failed to tutor a sufficient number of target equations. Ironically, it might be the case that the lack of teaching a sufficient number of target equations was due to the advice of the meta-tutor—since students were not able to manage tutoring the entry-level problems correctly, their SimStudents did not pass the entry-level quiz sections (i.e., one- and two-step equations), hence why the meta-tutor kept suggesting to students to continue teaching those entry level equations.

The challenge for the meta-tutor, therefore, is how to encourage students to teach a sufficient number of target equations with appropriate accuracy. For those students who have trouble teaching entry-level equations, the meta-tutor should provide assistance on skills to solve those equations (which are a prerequisite for learning the target equations). We have recently started to extend the meta-tutor (for our future studies) to provide *cognitive help* on feedback and hints that students provide to their SimStudents. With this extension, when students are not sure about the correctness of the steps performed by SimStudent, they will be able to ask Mr. Williams if their judgments are correct (before providing feedback to SimStudent). Additionally, when students do not know how to perform a next step for which SimStudent asks for help, students will be able to ask the meta-tutor what they should do next.

The meta-tutor should also encourage students to use resources more often as needed. For example, when students continue to ask for help on what to do next, then the meta-tutor might suggest that student should review the unit overview. If students repeatedly fail to have their SimStudents pass the quiz, then the meta-tutor might suggest that students should review example problems.

6 Conclusion

We found that the availability of the meta-tutor facilitated tutor learning on procedural skills for solving algebra equations. The data suggested that the meta-cognitive help given by the meta-tutor positively allowed students to select appropriate problems that affected both SimStudents' and hence students' learning.

Our data suggest that learning by teaching with meta-cognitive tutoring supports students in employing meta-cognitive skills on how to better tutor their peers that may not be available in traditional classroom instructions. At the same time, the data also suggest that to make learning by teaching more effective, the learning environment must be carefully designed so that students can tutor their tutees appropriately, which involves scaffolding both on how to teach (meta-cognitive help) and what to teach (cognitive help).

Acknowledgements. The research reported here was supported by National Science Foundation Awards No. DRL-0910176 and DRL-1252440; and the Institute of Education Sciences, U.S. Department of Education, through Grant R305A090519 to Carnegie Mellon University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. This work is also supported in part by the Pittsburgh Science of Learning Center, which is funded by the National Science Foundation Award No. SBE-0836012.

References

1. Gartner, A., Kohler, M., Riessman, F.: Children teach children: Learning by teaching. Harper & Row, New York (1971)
2. Roscoe, R.D., Chi, M.T.H.: Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. Review of Educational Research 77(4), 534–574 (2007)
3. Brophy, S., et al.: Teachable agents: Combining insights from learning theory and computer science. In: Lajoie, S.P., Vivet, M. (eds.) Proceedings of the International Conference on Artificial Intelligence in Education, pp. 21–28. IOS Press, Amsterdam (1999)
4. Matsuda, N., et al.: Cognitive anatomy of tutor learning: Lessons learned with SimStudent. Journal of Educational Psychology 105(4), 1152–1163 (2013)
5. Leelawong, K., Biswas, G.: Designing Learning by Teaching Agents: The Betty's Brain System. International Journal of Artificial Intelligence in Education, 18–3 (2008)
6. Foot, H., et al.: Theoretical issues in peer tutoring. In: Foot, H., Morgan, M., Shute, R. (eds.) Children Helping Children, pp. 65–92. Wiley, New York (1990)
7. Biswas, G., et al.: Measuring Self-Regulated Learning Skills through Social Interactions in a teachable Agent Environment. Research and Practice in Technology Enhanced Learning, 123–152 (2010)
8. Walker, E., Rummel, N., Koedinger, K.R.: Adaptive support for CSCL: Is it feedback relevance or increased student accountability that matters? In: Proceedings of the International Conference on CSCL (2011)
9. Corbett, A.T., Koedinger, K.R., Hadley, W.S.: Cognitive tutors: From the research classroom to all classrooms. In: Goodman, P.S. (ed.) Technology Enhanced Learning: Opportunities for Change, pp. 235–263. Erlbaum, Mahwah (2001)
10. Ritter, S., et al.: Cognitive tutor: Applied research in mathematics education. Psychonomic Bulletin & Review 14(2), 249–255 (2007)