Using Teachable Agent Feedback to Support Effective Learning-by-Teaching

Rod D. Roscoe (rod.d.roscoe@vanderbilt.edu) **John Wagster (john.wagster@vanderbilt.edu)**

Gautam Biswas (gautam.biswas@vanderbilt.edu)

Department of EECS/ISIS, Box 351824, Sta B, Vanderbilt University
Nashville, TN 37235 USA

Abstract

Teachable agents build upon research showing that students can learn by teaching their peers. In these systems, students learn by teaching simulated pupils who can be designed to support productive types of feedback and teacher-pupil interactions. We found that students learned better when they taught an agent designed to mimic a self-regulated learner (i.e. pushed the teacher to ask questions, and probe the agent for explanations to check her understanding). Analyses of log file data showed that interacting with the agent helped students engage in beneficial learning and teaching behaviors.

Keywords: teachable agents; log file analysis; learning-by-teaching, metacognition; self-regulated learning; feedback

Introduction

Teachable agents are computer-based learning environments designed to support learning-by-teaching in middle school classrooms (Biswas, Leelawong, Schwartz, Vye, & the TAG-V, 2005). The students do not instruct an actual peer, but instead teach a simulated pupil that accepts input from the teacher and uses that information to answer questions. These agents do not simulate the full range of actions available to a human pupil. They may also not engender the same commitment to "good teaching" that a human pupil would; the stakes are not as high. Despite such differences, the design of teachable agents is necessarily informed by research on learning-by-teaching in human teaching settings (e.g. peer tutoring, Roscoe & Chi, 2007). The focus of our research has been on how students learn by teaching and to look for interesting points of control for manipulating interactions with the teachable agent to enhance the learning process.

Previous research on learning-by-teaching has established that teaching a peer offers meaningful learning opportunities for the teachers, which emerge directly from the teaching process (Roscoe & Chi, 2007). For example, striving to explain key principles may help peer teachers evaluate and perhaps reorganize their own knowledge (Coleman, Brown, & Rivkin, 1997). Interactions between the teacher and pupil are a vital aspect of this process (Roscoe & Chi, in press). Teachers are not explaining the material for their own benefit; the explanations need to be understandable and educational for their audience. This may require teachers to attend to the errors of their pupil, and perhaps revise their explanations or generate new ones.

Another useful interaction stems from the pupils' questions (Graesser & Person, 1994; Roscoe & Chi, 2007). Pupils' questions can bring up ideas that the teachers omitted or introduce new ideas for discussion. Pupils' questions might also offer unintended metacognitive cues that the teachers' knowledge is incomplete or flawed. Pupils might ask questions because the teachers' explanations do not make sense and need to be reconsidered and improved.

The impact of pupils' actions and questions on the teacher can be characterized as a form of feedback (Schwartz, Blair, Biswas, Leelawong, & Davis, in press). A teachers' general goal is to get the pupil to understand the concepts and related problems, and so the pupils' actual or apparent understanding provides indirect feedback on how well the teacher is doing. In regards to learning-by-teaching, this form of feedback can also signal to teachers that their own knowledge may be lacking. The pupils' performance is partly a reflection or product of the teachers' domain knowledge and strategies.

From this perspective, two major stumbling blocks for learning-by-teaching might arise. One problem is that many students are not very good at gauging their own comprehension (Azevedo & Cromley, 2004). They may falsely judge that they understand when they do not, or fail to detect contradictions between their beliefs and other sources of information. Because learning-by-teaching requires a significant degree of self-monitoring and self-regulation, student teachers may not always appreciate or utilize pupil feedback effectively. This may be especially true for participants in our studies, who tend to be middle school students with low prior domain knowledge and little experience with teaching.

Another obstacle could be inconsistency or poor quality of the feedback from the pupil. Human pupils may also monitor their own learning ineffectively. Thus, when peer teachers ask "Do you get it?" a response like "Yeah, I think so" could be very misleading. Another nontrivial problem is pupil passivity. In teaching and tutoring situations, some pupils do not say much and ask few questions (Graesser & Person, 1994), thus providing very little feedback to the teacher. Pupils also possess their own prior knowledge and misconceptions. Some pupil errors will not actually be indicative of any problem on the part of the peer teacher.

These issues suggest that teachable agents need to be designed in ways that simultaneously support metacognition and self-regulation for the student teachers, and limit problems of inaccurate, lacking, or irrelevant agent feedback. We wish to avoid teachable agents that employ their own learning strategies or possess prior knowledge of the material, which can confuse or frustrate the student teacher (Leelawong & Biswas, in press). The agents should also be designed to respond to the student teacher in an active, consistent manner that accurately reflects what the agent "knows," and is thus more reflective of the teachers' performance. We hypothesize that feedback of this nature will support more effective learning-by-teaching with teachable agents, even for middle school students.

We have developed a teachable agent system called Betty's Brain in which students teach an agent about river ecosystems. In this paper, we describe a study that manipulated the kinds of metacognitive feedback that the system provided to the students. Student learning was assessed based the quality of student-generated concept maps and written pretest/posttest measures We also analyzed log file data (i.e. records of students' interactions with the system) to test whether the variations feedback influenced the students' learning and teaching behaviors.

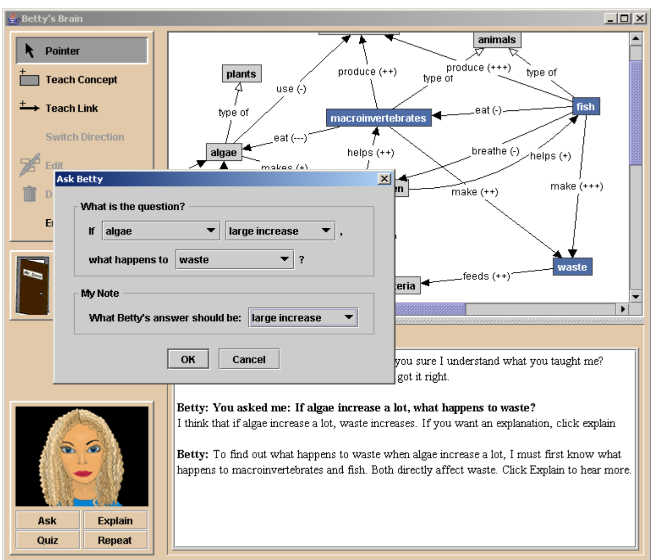

Figure 1. Betty's Brain interface with query window.

Betty's Brain

Teaching Betty

The teaching process in Betty's Brain is organized around three activities: teach, query, and quiz (Biswas et al., 2005). Throughout this process, students can also access text resources that provide information on river ecosystems.

To *teach* Betty, students generate a representation called a concept map (Novak, 1998) consisting of concepts and causal relationships. Students can teach Betty concepts such as "fish" and "carbon dioxide," and relationships such as "fish exhale carbon dioxide." Students also indicate whether the first entity causes an increase or decrease in the second; when fish exhale carbon dioxide, carbon dioxide increases.

Independent of learning-by-teaching, this process of editing their concept map could also help students to better understand the material (Nesbit & Adesope, 2006; Novak, 1998). Similar to explaining, concept mapping requires students to externalize their knowledge. This may facilitate self-evaluation (e.g. checking whether links make sense) and knowledge revision (e.g. modifying the concepts). Furthermore, the visual representation of the concept map structure (Figures 1 and 2) makes chains of reasoning more explicit. For example, a map may indicate that "fish eat macroinvertebrates, and macroinvertebrates eat algae." This representation may help students realize the implicit idea that fish indirectly affect the amount of algae.

The *query* function allows students to ask Betty causal questions about entities in the river. These questions are generated using a pull-down menu. Students specify two entities and a change in the first entity. An example might be "If carbon dioxide increases, what happens to algae?"

Betty uses qualitative reasoning methods to follow all of the chains of concepts and links connecting the two concepts. Betty initially answers questions by giving the output of her reasoning process, such as "I think if carbon dioxide increases, algae increases." These answers can be correct or incorrect depending on the correctness or completeness of what she has been taught. In addition, students can request Betty to explain her answer. Betty will begin to articulate the paths she followed to produce her answer. In order to get Betty to explain her entire reasoning process, the student teacher may have to ask her to continue her explanation several times.

The *quiz* function allows student teachers to have Betty take a quiz consisting of pre-defined questions. After taking the quiz, Betty's answers are graded by a mentor agent. Grading is done by comparing the answer produced by the students' map to the answer produced by an underlying expert map that has all of the required concepts and links. Thus, students know which questions Betty answered correctly or incorrectly and what her answer was for each.

Feedback from Betty

The basic version of the system enables some of the features of desirable pupil feedback that we argued should support learning-by-teaching. When asked a question, Betty will always try to answer, and will always justify her answer (based on the concept map she has been taught) if probed. Thus, passivity is less of a problem. In addition, Betty only knows what she is taught about the domain; she has no prior knowledge or preconceptions. This means that her knowledge and mistakes are almost always relevant to the students' own knowledge and errors.

For this project, we extended another version of the system in which Betty provided more active and metacognitive forms of feedback (Leelawong & Biswas, in press). This version employs a self-regulated learning (SRL) framework. Self-regulated learning describes a set of comprehensive skills such as setting learning goals, selecting appropriate strategies, and monitoring one's learning progress and strategies (Azevedo & Cromley, 2004). Betty's SRL persona incorporates aspects of this metacognitive knowledge (Wagster, Tan, Wu, Biswas, &

Schwartz, 2007). For example, as students add concepts and links, Betty sometimes spontaneously re-explains what she has been taught. This is similar to how self-regulated pupils might restate and draw further inferences from what a teacher has said in order to make sure they understand it.

The SRL version of the system does more than just produce spontaneous statements from Betty. We have identified patterns of teacher actions where metacognitive feedback might be useful. For example, sometimes students ask Betty to take the quiz but have not taught her anything new since the last quiz, or probed her understanding with a question. Betty "knows" that she is probably not ready and will actually refuse to take the quiz. The response given to the student would be "Are you sure I understand what you taught me? Please ask me some questions to make sure I got it right." The students cannot progress with the quiz unless they teach her more or ask a question.

However, asking lots of queries is not sufficient either. If students do not occasionally ask Betty to explain her reasoning, she will ask the teacher to do so; "You have not asked me for my explanations lately. Please make me explain my answers to you know if I really understand." Thus, Betty strongly urges her student teacher to probe the complex chains of reasoning in the map more deeply.

In essence, SRL Betty captures teacher-pupil interactions where the pupil is active and aware of when she may not be learning. Her feedback further encourages the teachers to evaluate Betty's knowledge (i.e. their own knowledge) and make revisions when necessary. Our prediction is that this enhanced, SRL-inspired form of teachable agent feedback will improve learning-by-teaching outcomes.

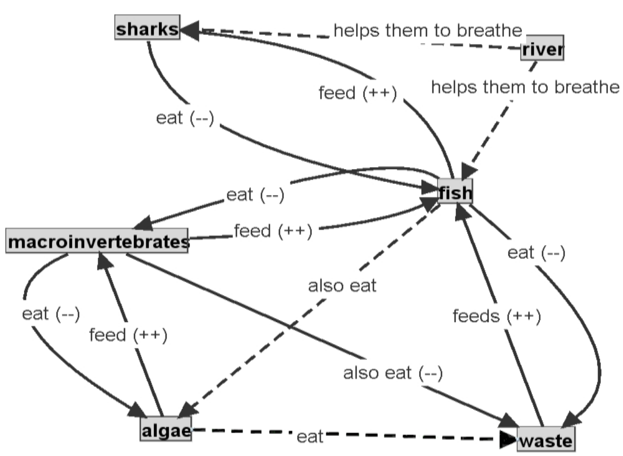

Figure 2. Example student concept map.

Method

Participants and Conditions

Our participants were 56 students in two 5th grade science classrooms, taught by the same teacher. The study took place in a relatively high-achieving public middle school during the 2005-2006 school year.

Students were assigned to one of three conditions using stratified random assignment based on standardized test scores. Students in the learning-by-teaching (LBT, n = 17) condition taught Betty using the basic version of the system. Students in the self-regulated learning (SRL, n = 19) condition taught Betty using the SRL version of the system. In the intelligent coaching system (ICS, n = 18) control condition, students created river ecosystem concept map for themselves and did not teach the Betty agent at all. These students could still use all of the system features for adding concepts and links, asking queries, and requesting a quiz. System responses were delivered by Mr. Davis, the mentor agent. They did not receive any SRL feedback.

These experimental manipulations occurred during the main phase of the study (seven 45-minute sessions). After an eight-week delay, students participated in a transfer phase (five 45-minute sessions) in which they learned about the nitrogen cycle and taught Betty. All students used an identical version of the LBT system in this phase. For this paper, we focus on data from the main phase.

Learning Assessments

Student learning was assessed through two measures. One measure was the final map generated by students at the end of the main phase. These maps were coded to identify correct inclusions of concepts and links based on the text resources provided. The number of correct concepts and links was tabulated to produce an overall map score. This measure is an indicator of how well students could model the network of causal relationships that explain the functioning of a normal river ecosystem. The Betty's Brain system is designed to support the learning of such models.

A second measure was a written assessment that tested students' understand of underlying principles and processes. This test included eight free response and ten multiple-choice questions, which covered concepts such as interdependence, balance, photosynthesis, and food chains. Each free response question was worth six points to allow students to earn partial credit for their responses. Each multiple choice question was worth one point. We included this measure because the text resources discussed these principles related to the river ecosystem. Thus, by studying the resources in order to teach Betty, students might also acquire deeper conceptual knowledge. However, although understanding these ideas would definitely help students teach Betty better, the current system does not require students to teach Betty this kind of information explicitly.

Log File Analyses

As students interact with the system, log files are generated that record each action. Such actions can be captured in five basic categories: editing the map (EM), accessing the text resources (RA), asking Betty a question (AQ), requesting Betty to take a quiz (RQ), and asking Betty to explain her reasoning (EX). We analyzed these behaviors as proportions of their overall number of actions.

These data help us to determine if system feedback influenced the students learning and teaching behaviors. If feedback is of a useful nature, the students should not only learn more effectively, but this learning should be associated with differences in their actions. For example, students might probe Betty's knowledge more carefully, thus probing their *own* understanding more carefully.

Results

Learning Outcomes

We first examined the concept maps produced by students. As an objective measure of the quality of the maps, we compared students' maps to the expert map used by the grading system. This map had 10 concepts and 19 links, and thus had a score of 29. The average map scores (Table 1) for the ICS students; $t(17) = -4.92$, and LBT students; $t(16) = -2.12$, $p = .05$, were significantly lower than the expert map score. The average SRL score was slightly higher than the expert map score, meaning that these students actually added more correct information than was required by the expert map. This difference was not significant, however.

Table 1: Mean (SD) map scores and assessment test gains.

	Condition		
Measure	ICS	LBT	SRL
Final map	22.8 (5.3)	25.6 (6.5)	31.6 (6.6)
Free response	2.5 (2.5)	5.1 (4.6)	5.2 (3.4)
Mult. choice	0.4 (2.4)	1.1 (2.3)	0.4 (1.5)

We next compared map scores across conditions (Table 1). Not surprisingly, given the results above, there was a significant effect for condition; $F(2,51) = 9.74$, $p < .001$. The SRL students generated significantly better maps than both the ICS students; $p < .001$, and LBT students; $p = .02$. The LBT and ICS map scores did not significantly differ from each other, although LBT scores were slightly higher.

Our analysis of students' maps suggests that receiving feedback from SRL Betty helped students to develop a more correct and complete model of the river ecosystem. Teaching Betty without such feedback was only slightly better than not teaching at all. This is consistent with the hypothesis that effective learning with teachable agents is supported by receiving active, metacognitive feedback that is reflective of the teachers' own knowledge.

We next considered students' gain in scores from pre-test to post-test on the written assessments (Table 1). We first tested whether students in each condition gained significantly on the free response and multiple choice questions. Gains for the multiple choice questions were not significant for any of the groups. However gains for the free response questions were significant for the ICS students; $t(16) = 3.95$, $p = .001$, LBT students; $t(15) = 6.22$, $p < .001$, and SRL students; $t(16) = 4.55$, $p < .001$. Thus, all groups improved in their ability to define key ecosystem principles.

Although all groups improved on the free response questions, it appeared that the LBT and SRL students may have improved somewhat more than the ICS students. A comparison of the groups revealed a marginally significant effect; $F(2,51) = 2.92$, $p = .06$, but pair-wise comparisons between groups were not significant.

Overall, our assessments of student learning and performance show a relatively straightforward benefit for the SRL-inspired pupil feedback from Betty.

Student Actions

The feedback provided by the SRL version of Betty is designed to give the student teacher better feedback about their own understanding, as reflected in Betty's knowledge and performance. Receiving this feedback should also affect how students interact with Betty, hopefully leading to more productive learning and teaching behaviors.

To test this hypothesis, we considered the occurrence of the five primary actions students could take while teaching Betty or creating their own map, such as editing the map and asking a query. These behaviors are reported as proportions of the overall number of behaviors (Table 2).[1] Table 2 also provides a correlation between the given action and students' map scores.

Group comparisons revealed a number of interesting differences between the conditions. We found significant main effects of condition for editing the map; $F(2,51) = 26.85$, $p < .001$, asking queries; $F(2,51) = 28.98$, $p < .001$, requesting quizzes; $F(2,51) = 14.12$, and asking for an explanation; $F(2,51) = 8.74$, $p < .001$.

Table 2. Mean (SD) proportions of student actions.

	Condition			
Action	ICS	LBT	SRL	$r(52)$
EM*	.65 (.12)	.48 (.10)	.41 (.08)	-.48*
RA	.11 (.08)	.13 (.08)	.08 (.04)	.09
AQ*	.08 (.07)	.12 (.06)	.22 (.05)	.44*
RQ*	.09 (.04)	.20 (.09)	.12 (.05)	.03
EX*	.05 (.06)	.04 (.04)	.13 (.10)	.31*

The pattern that we observe is that the ICS students spent most of their time editing their map (i.e. adding and deleting concepts and links). The LBT and SRL students spent over 40% of their time editing their maps, but also engaged in a more diverse set of actions. The LBT students utilized the quizzes more often than the ICS students; $p < .001$, and SRL students; $p < .001$. In contrast, the SRL students spent significantly more time asking questions and requesting explanations than did the ICS and LBT students (all comparisons significant at the .01 or .001 levels). These actions were specifically encouraged by the system.

[1] The reported proportions do not always add up to 1.00 because there was a small percentage of "other" behaviors that were occurred infrequently or as the result of minor system bugs.

We also found that certain student actions correlated positively or negatively with map scores. Editing the map was negatively and significantly correlated with map scores, $p < .001$. Positive correlations were found for asking queries; $p < .001$, and requesting explanations; $p < .05$.

In sum, the picture that emerges is that SRL students, who received metacognitive feedback and prompts, were more likely to probe Betty's understanding by asking her questions and having her explain her reasoning. These actions were positively associated with map quality. Thus, Betty's SRL feedback was successful in helping the SRL condition students to attend to and evaluate her knowledge and understanding, which in turn helped them build a good model of the ecosystem.

The LBT students tended to rely more on the quizzes, which also provided feedback on the agent's performance. However, the quizzes only showed Betty's answers and not how she got those answers (i.e. provided shallower pupil feedback). This may explain why requesting a quiz was not linked to map scores, and the LBT group produced lower quality maps than the SRL students. Finally, ICS group seemed to mostly ignore the various feedback mechanisms available while they created their maps. This strategy was actually negatively related to map quality scores.

Overall, feedback from the system seemed to be a vital component in this learning task, and self-regulated learning inspired feedback was the most effective.

Student Responses to Feedback

To further probe the impact of the feedback students received, we also considered students' subsequent actions after asking a query and requesting a quiz. These functions allowed the students to test whether Betty (i.e., their map) can answer questions about the river ecosystem. Varying the types of feedback available may also influence what students do after getting this information. For these analyses, we tabulated the total number of AQ or RQ actions and the proportions of subsequent actions (Table 3).

These analyses suggest that students did indeed differ in their responses to system feedback. Significant group effects were observed for AQ-EM sequences; $F(2,51) = 3.91$, $p = .03$, and AQ-AQ sequences; $F(2,51) = 15.90$, $p < .001$. After asking a question, ICS students were more likely to immediately edit their map than were SRL students, $p = .04$. In contrast, SRL students were more likely to ask *another* question to further probe Betty's understanding. SRL students followed a query with a second query about twice as often as LBT or ICS students, $p < .001$.

Further differences were observed with the RQ sequences. Significant effects were found for RQ-EM sequences, $F(2,51) = 16.30$, $p < .001$, RQ-RA sequences; $F(2,51) = 4.92$, $p = .01$, RQ-AQ sequences; $F(2,51) = 46.45$, $p < .001$, and RQ-RQ sequences; $F(2,51) = 8.63$, $p = .001$.

Given the earlier results, it is not surprising that the ICS students were more likely to edit their map immediately after taking a quiz. ICS students were more likely to do this than LBT; $p = .007$, and SRL students; $p < .001$.

Table 3. Mean (SD) proportions of AQ and RQ sequences.

	Condition		
Sequence	ICS	LBT	SRL
AQ Sequences			
AQ-EM*	.42 (.30)	.40 (.12)	.25 (.11)
AQ-RA	.03 (.08)	.08 (.07)	.06 (.04)
AQ-AQ*	.15 (.14)	.17 (.08)	.34 (.10)
AQ-RQ	.07 (.23)	.14 (.09)	.15 (.09)
AQ-EX	.14 (.18)	.16 (.16)	.18 (.11)
RQ Sequences			
RQ-EM*	.70 (.20)	.47 (.18)	.31 (.24)
RQ-RA*	.15 (.10)	.20 (.15)	.08 (.09)
RQ-AQ*	.06 (.11)	.07 (.07)	.43 (.18)
RQ-RQ*	.04 (.04)	.20 (.12)	.13 (.15)

Note: Sequences of RQ-EX were not possible unless a bug occurred, and so are not included in this table.

LBT students were more likely than SRL students to access the resources after a quiz, $p = .009$. This response makes sense because if Betty's answers are incorrect, then reading the text could help student learn correct ideas to teach her. One problem that might arise is if the student teacher has not understood the source of Betty's error. A search through the text could be fruitless in that case. The SRL students were more likely than ICS and LBT students, $p < .001$, to ask Betty a question after getting her quiz results. This suggests that these students were attempting to more specifically diagnose Betty's quiz performance.

In sum, our analyses of students' actions subsequent to using the query and quiz functions support the conclusions drawn from our previous analyses. The SRL students continued to probe Betty's understanding more effectively than did the LBT and ICS students. Not only did these students benefit from receiving the SRL pupil feedback from Betty, they might have begun to learn the value of such feedback. This could explain why they seemed to seek out additional pupil feedback when it was available.

Discussion and Conclusions

Learning-by-teaching provides many learning opportunities for teachers, but taking advantage of these opportunities places a heavy metacognitive demand on our student teachers. Further complicating the matter is that critical cues for self-monitoring, the feedback from the pupil, may not always be sufficient or reliable. We have applied these lessons to the design of a teachable agent system, Betty's Brain, which provides students with self-regulated learning feedback. This feedback is designed to make the metacognitive cues from the agent (e.g., her knowledge and answers to questions) clear and more salient, and also support the student teachers' metacognition. We hypothesized that such agent feedback would support more effective learning in the teachable agent environment.

Our results indicated that the SRL-inspired feedback helped middle school students develop a more correct and complete causal model of a river ecosystem than students who taught Betty without feedback or did not teach. These results were paralleled by analyses of students' interactions with Betty. SRL students were more likely to probe Betty's understanding by asking questions and having her explain her reasoning. Thus, the feedback may have supported student learning by improving their use of metacognitive strategies. These results build upon our prior research demonstrating the effectiveness of teachable agents systems for learning (e.g. Biswas et al., 2005; Leelawong & Biswas, in press; Schwartz et al., in press; Wagster et al., 2007).

Although teachable agents differ from teaching a human peer, these results may offer a tentative implication regarding learning in those settings. Our results show that teaching an agent designed to be active and metacognitive improved learning. Peer tutors might also benefit from tutoring tutees who are more self-regulated in their learning. One way to achieve this might be to offer self-regulated learning strategy training to the tutees before they enter the tutoring setting. This is somewhat different from the usual approach, which focuses on training the tutors to use knowledge-building strategies (Roscoe & Chi, 2007). Of course, self-regulated strategy training would also likely have benefits for the tutees (Azevedo & Cromley, 2004).

These results also have implications for computer-based tutoring systems. Traditional computer tutors focus on tracking the students' knowledge (student modeling) and tailoring feedback to fill knowledge gaps. Our TA environments use students' interaction patterns to provide self-regulation feedback that is directed toward monitoring and probing the map to ensure understanding. One might say that our students are at a disadvantage because they miss out on domain knowledge feedback, but our results imply that the SRL-based feedback makes them better learners than those who just receive domain content feedback.

One finding was somewhat disappointing. Although all students improved in their understanding of underlying river ecosystem principles (e.g. interdependence), students who received SRL feedback did not gain significantly more in this area than other students. We had hoped that supporting students' self-regulated learning would improve their conceptual knowledge in addition to their ability to model the river ecosystem. However, one limitation of the current system is that it does not require students to teach this kind of information explicitly. For example, students teach about causal relationships involved in photosynthesis. They teach that sunlight provides energy for plants and algae, which use this energy and carbon dioxide to make food, and release dissolved oxygen as a byproduct. These relationships are clearly described in the text as part of photosynthesis, but students do not explicitly indicate this to be the "photosynthesis process" in their maps.

In future versions of the system, it may be worthwhile to expand the kinds of information that can be taught to Betty (i.e. included in the maps). For example, in addition to adding concepts and links, students may have to flag the processes these concepts and links participate in. A similar possibility would be to allow students to connect entities on their map to headings or passages in the text resources. To support the student teachers' beneficial use of these new functions, the quiz and query functions could be expanded to directly test Betty's knowledge of this deeper conceptual information. More importantly, the Betty agent's feedback could be expanded to prompt students to teach her these ideas. Based on our prior and current results, increasing the scope of information that can be taught to Betty, in conjunction with her metacognitive feedback support, should lead to even stronger learning gains for the students.

References

Azevedo, R. & Cromley, J. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523-535.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N. & the TAG-V (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, *19*, 363-392.

Coleman, E., Brown, A. & Rivkin, I. (1997) The effect of instructional explanations on learning from scientific texts. *Journal of the Learning Sciences, 6*(4), 347-365.

Graesser, A. & Person, N. (1994). Question asking during tutoring. *American Educational Research Journal, 31*(1), 104-137.

Leelawong, K. & Biswas, G. (in press). Designing learning by teaching systems: The Betty's Brain System. *International Journal of AI in Education*.

Nesbit, J. & Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*(3), 413-448.

Novak, J. (1998). Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations. Lawrence Erlbaum and Associates, Mahwah, NJ.

Roscoe, R. & Chi, M. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77*(4), 534-574.

Roscoe, R. & Chi, M. (in press). Tutor learning: The role of explaining and responding to questions. *Instructional Science*.

Schwartz, D., Blair, K., Biswas, G., Leelawong, K., Davis, J. (in press). Animations of thought: Interactivity in the teachable agent paradigm. To appear in R. Lowe & W. Schnotz, (Eds.). *Learning with animation: Research and implications for design.* Cambridge Univ. Press: UK.

Wagster, J., Tan, J. Wu, Y., Biswas, G., & Schwartz, D. (2007). Do learning by teaching environments with metacognitive support help students develop better learning behaviors? In the *Proceeding of the 29th Annual Meeting of the Cognitive Science Society,* 695-700. Nashville, TN.