Interactive Metacognition:

Monitoring and Regulating a Teachable Agent

Daniel L. Schwartz, Catherine Chase, Doris Chin, Marily Oppezzo, Henry Kwong
Stanford University

Sandra Okita
Teachers College, Columbia University

Rod Roscoe, Hoyeong Jeong, John Wagster, & Gautam Biswas
Vanderbilt University

To appear in D.J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.), *Handbook of Metacognition in Education*.

Metacognition is typically characterized by monitoring and regulating thought processes to make sure those processes are working as effectively as possible (ref list). Good teachers are extremely metacognitive. They monitor student understanding and they regulate the processes students use to learn and solve problems (Shulman, PCK*). Whether or not good teachers apply metacognition to themselves, they do apply it interactively to their students' thinking. The proposal of this chapter is that asking children to apply metacognition interactively towards another can help science learning as well as the development of metacognitive skills. This proposal is examined in the context of an interactive technology called a Teachable Agent, where students teach a computer agent that can answer questions and show its reasoning.

Research on learning-by-teaching has found that teaching another person can lead to superior learning. For instance, when people prepare to teach pupils to take a test, they learn more compared to when they prepare to take the test themselves (Bargh & Schul, *; Biswas et al., *). Moreover, during the act of teaching, people learn by clarifying the confusions of their tutees (Chi et al., 2001; Palinscar & Brown, 1984; Uretsi, 2000). In these cases, students are anticipating or experiencing their tutees' cognition. Interestingly, when people move into didactic mode, and stop attending to their tutee's thoughts, they learn less (Fuchs, Fuchs, Bentz, Phillips, & Hamlett, 1994; Graesser, Person, & Magliano, 1995; Chi, Roy and Hausmann, *).

The interactive quality of other-directed metacognition can help resolve two psychological challenges. One challenge is the dual-task demand of metacognition. During metacognition, people need to think their thoughts, and they simultaneously need to think about their thinking about those thoughts. When problem solving becomes difficult, there can be less free capacity for metacognition. For example, when trying to recall a person's name under pressure, people may not consider other ways to search for the name, and simply keep circling their hand, presumably, to prime the cognitive pump. Teaching can help alleviate the dual-task demand of metacognition. The tutee has the responsibility of problem solving, which frees up resources for the teacher's metacognition. Gelman and Meck (1983*), for example, found that young children could monitor errors in adult counting better than their own counting, when the counting task reached the edge of the children's abilities.

The second challenge of metacognition is motivational. Because metacognition takes extra work, people will tend to "get by" if they can, rather than take the extra cognitive effort needed to go beyond "good enough" (Martin & Schwartz, accepted). Students often skim readings, because they think it is not worth checking their understanding. Teachers, however, are responsible for their students' performance, not to mention their public competence. This increase in responsibility can motivate students to engage in metacognition.

Ideally, the affordances of interactive metacognition encourage students to practice metacognitive abilities on the external plane so they can develop and eventually be turned inward to their own cognition (Vygostsky, *). For example, in a series of studies by Okita (2008*), elementary school children learned tricks for mentally solving complex arithmetic problems. In half of the cases, students practiced the problems on their own. In the other half of the cases, students took turns. On one turn, they would try a problem, and on the next turn, they would monitor another person solving a problem. They had to stop the person if they thought there was a mistake. Students who monitored the other person demonstrated a U-shaped curve in their own problem solving. After monitoring the other person, they would go slower and were less accurate when solving their own problems. Over time, however, they sped up and became more accurate than the students who never monitored the other person. Presumably, by monitoring the other person, the students were learning to monitor themselves, which caused a temporary drop in efficiency, but a better payoff in the long run.

The research below demonstrates that a computer agent can help students engage in metacognition. This engagement helps them learn science content better, and it eventually helps them learn to use metacognition more effectively for themselves. The first two sections focus on the monitoring side of metacognition. They show that students treat the agent as having cognition, and the students are interactively engaged in metacognition towards their agent. The studies also show that the agent's knowledge is similar to the students, so that the distance between monitoring the agent and their own thoughts is relatively small. The consequence is that the students start to think like the agent. The remaining sections focus on the regulation side of metacognition. Here, interactive metacognition is expanded. Students not only interactively monitor their agents, but they also make interactive choices about the best external resources to use to help improve the agent's learning. This further helps the children learn science and to eventually make better decisions about how to monitor and regulate their own cognition.

An Interactive Technology for Applying Metacognition on Another

This section explains how Teachable Agents (TA) naturally integrate learning and metacognition. Betty's Brain, the TA shown in Figure 1 and the focus of the chapter, was designed for knowledge domains where qualitative causal chains are a useful structural abstraction (e.g., the life sciences). Students teach Betty by creating a concept map of nodes connected by qualitative causal links; for example, 'burning fossil fuels' increases 'carbon dioxide'. Betty can answer questions based on how she was taught. For instance, Betty includes a simple query feature. Using generic artificial intelligence techniques (see Biswas et al., *), Betty animates her reasoning process as she answers questions. In Figure 1, Betty uses the map she was taught to answer the query, "What happens to 'heat radiation' if 'garbage' increases?" Students can trace their agent's reasoning, and then remediate their agents' knowledge (and their own), if necessary. A version of the Betty's Brain environment and classroom management tools can be found at <aaalab.stanford.edu/svBetty.html>. Betty is not meant to be the only means of instruction, but rather, she provides a way to help students organize and reason about the content they have learned through other lessons.

[Figure 1 about here – Betty's Brain]

In reality, when students work with Betty, they are programming in a high-level, graphical language. However, Betty's ability to draw inferences gives the appearance of sentient behavior. Betty also comes with narratives and graphical elements to help support the mindset of teaching; for example, students can customize their agent's appearance and give it a name. ("Betty's Brain" is the name of the software, not a student's specific agent.) Moreover, as described below, Betty can take quizzes or play games. The goal of a TA is not to fool children or adults, so much as to enlist their social imagination so they will engage in the processes of monitoring and regulating their agent's knowledge.

A key element of TAs is that they externalize thought processes. This differs from most simulations, which portray situations and not thoughts. Betty visually animates causal thinking about a situation and literally makes thinking visible. Thus, students are applying metacognition to the agent's thinking, and the thinking is in an easily accessible format.

Monitoring One's Own Thoughts in Another

For students to practice metacognition on their agent, they need to view Betty as exhibiting cognitive processes. This section shows that students do treat their agent as sentient, which leads them to take responsibility for monitoring and regulating their agents' knowledge. Moreover, it shows that Betty's knowledge is a fair representation of the students' own knowledge, which shortens the distance between monitoring the agent and monitoring themselves.

Students Treat a TA as Sentient

When programming and debugging their agents, students are also monitoring and regulating their agents' knowledge and reasoning. A study with 5th-graders demonstrated that students treat the agent as having and using knowledge. Students also monitor their agents' failures and share responsibility, which leads them to revise their own understanding so they can teach better. By this age, children know the computer is not really alive, but they suspend disbelief enough to treat the computer as possessing knowledge and feelings (e.g., Reeves and Nass, *; Turkle, *).

The study used the Triple-A Gameshow, which is an environment where multiple TAs, each taught by a different student, can interact and compete with one another (Figure 2). Students can log on from home to teach their agents, chat with other students, and eventually have their agents play in a wagering game. The Gameshow was developed to make homework more interactive, social, and fun. In the study, however, the focus was on student attitudes towards Betty during game play, and students worked alone. During game play, (1) the game host poses questions to the agents; (2) the students choose a wager that their agent will answer correctly; (3) the agents answer based on what they have been taught; (4) the host reveals the correct answer; and finally, (5) wager points are awarded. In addition to boosting engagement, the wagering feature was intended to lead students to think through how their agent would answer the question, thereby monitoring their agent's understanding.

[Figure 2 about here – Gameshow podium]

The study included two conditions. In both, students received a text passage on the mechanisms that sustain a fever, and they taught their TA. The treatment difference occurred when playing the Gameshow. In the TA condition, the agents answered six questions, and the character on the podium represented the agent. In the Student condition, the students answered the questions, and the character represented the student. To capture students' thoughts and feelings towards the agent, students in both groups thought aloud.

In the TA condition, students treated the agent as having mental states. Students' attributions of mental states were coded as being directed to themselves, their agents, or both. Examples of self-attributions include, "It's kind of confusing to me," "I have a really good memory," and "No, actually, I don't know." Examples of agent-attributions include, "He doesn't know it," and "He knows if shivering increases…." Sometimes, a single statement could include both self and agent attributions; for example, "I'm pretty sure he knows this one," and, "I guess I'm smarter than him."

During game play, students in both treatments made about two mental state attributions per question. For the TA condition, over two-thirds of these attributions were towards the agent or a combination of agent and student. Thus, students treated the agent as a cognitive entity, and in fact, they sometimes confused who was doing the thinking, as in the case of one boy, who stated, "'cause I don't… 'cause he doesn't know it."

The TA students also took an "intentional stance" (Dennett, *) towards their agents, by apportioning responsibility to the agent for success and failure. They could have behaved as though all successes and failures were theirs, because the agent is simply a program they had written, but they did not. Table 1 indicates the number of attribution-of-credit statements made in response to successful and unsuccessful answers. Examples of success attributions include, "I'm glad I got it right" (self), "He got it right!" (agent), or "We got it!" (both). Examples of failure attributions include, "I didn't teach him right" (self), "He said *large* increase when it was only increase" (agent), or "Guess we were wrong" (both).

[Table 1 about here – attributions of sentience]

Students in the TA condition liberally attributed responsibility to the agent. Importantly, the TA condition exhibited more attention to failure, which is a key component of monitoring (Attention to Failure reference?*). They made nearly three times as many attributions in a failure situation relative to the Student condition. The attributions were spread across themselves and their agents. They often made remarks about flaws in their teaching such as, "Whoa. I really need to teach him more." Thus, at least by the verbal record, the TA condition led the students to monitor and acknowledge errors more closely than the Student condition.

The study also bears on one of the challenges of metacognition; namely, caring enough to engage in the extra work that metacognition often entails. After completing the game play, students were told there would be a more difficult round. They were given the opportunity to revise their maps and re-read the passage in preparation. While all the children in the TA condition chose to go back and prepare for the next round, only two-thirds of the Student condition prepared. Of those who did prepare, the TA students spent significantly more time at it. The protocol data from the game play help indicate one possible reason. The Student condition exhibited nearly zero negative responses to failure (e.g., "Ouch!). Given an unsuccessful answer, the Student condition averaged 0.02 negative affective responses. In contrast, the TA condition averaged 0.62 expressions of negative affect. Much of this negative affect was regarding their agent's feelings. For example, one student said "Poor Diokiki… I'm sorry Diokiki" when his agent, Diokiki, answered a question incorrectly. The TA students felt the responsibility for their agents' failures, which likely caused them to spend more time preparing for the next round of game play.

Overall, these data indicate that the children treated their agents as if they were sentient, which had subsequent effects on student learning behaviors. The children were "playing pretend." They knew their agent was not a sentient being. Regardless, their play involved the important features of metacognition – thinking about mental states and processes, noticing and taking responsibility for mistakes, and experiencing sufficient affect that they found it worth the effort to do something about the mistakes when given a chance to revise. Monitoring another, in this case an agent one has taught, can lead to more metacognitive behaviors than completing a task oneself.

<u>The Agents Knowledge Reflects the Students' Knowledge</u>

Schoenfeld (1987, *), discussing the importance of monitoring, states that "… the key to effective self-regulation is being able to accurately self-assess what is known and not known" (p. *). In Betty, students are assessing what their agent does and does not know. The agent's knowledge is a reflection of their own knowledge, so that working with the agent indirectly entails working on an externalized version of their own knowledge. This was demonstrated by correlating the test scores of the students and their agents.

A TA can be automatically tested on the complete population of questions in a concept map. By using a hidden expert map that generates the correct answers, the program can successively test the TA on all possible questions of the form, "If node <X> increases, what happens to node <Y>?" The results produce an *APQ Index* (all possible questions) that summarizes the overall test performance of the TA. A study with 30 sixth-grade students compared the agents' APQ index with how well students did on their own tests. Students completed three cumulative units by teaching their agent about global warming and climate change. At the end of each unit, the agents were tested to derive an APQ Index, and students took a short answer, paper-and-pencil test. The student tests each had four questions that included content from the expert map (TA-like questions), and four questions that depended on other content from the lessons (Non-TA Questions). The Non-TA Questions helped to determine whether Betty correlated with student knowledge more broadly, and not just questions that Betty could answer.

[Table 2 about here – APQ x Student test scores]

Table 2 indicates that the TA scores were positively correlated with students' test scores. These correlations compare favorably with the correlations between students' scores on the TA-like questions and the Non-TA questions for each unit test, which are .47 (Test 1), .46 (Test 2), and .14 (Test 3). Thus, the APQ Index correlated better with student performance on the TA and Non-TA questions than these two types of paper-and-pencil items correlated with each other. (The low correlations involving Test 3 are due to one of the TA-like paper-pencil questions, which exhibited different properties from other questions across the student tests.) Conceivably, with the further development and evaluation, it will be possible to test agents instead of students, thereby saving valuable instructional time.

The results indicate that when students monitor their agent's knowledge, for example, by asking it a question, they are likely to be monitoring a fair externalization of their own knowledge. This helps to dissolve the gap between self and other, so that the task of working with the agent is a proxy for the task of reflecting upon their own knowledge.

Adopting the Cognition of Another

Given that students treat the TA as exhibiting mental states and the TA reflects the student's knowledge, the next question is whether these have any effect on student learning. Ideally, by taking up another's cognition, one learns from it. Siegler (1996*), for example, found that young children learned number conservation more effectively when prompted to explain the experimenter's reasoning rather than their own. Betty reasons by making inferences along causal chains. When students teach Betty, they learn to simulate her causal reasoning.

Learning to simulate Betty's cognition about a situation is different from learning to simulate the situation itself. Many times, people create a mental model of a situation that helps them imagine the behavior of the system and make predictions (Gentner & Gentner, *; more mental model refs). For example, when reasoning about how gears work, people can simulate an internal image of the gears to solve problems (Schwartz, *). To run their mental model, people imagine the forces and movements of the gears, and they observe the resulting behaviors in their minds eye. With Betty, students create a mental model of the agent's reasoning. So, rather than simulating spatial movement, the students simulate chains of declarative reasoning. This way, the agent's cognition becomes internalized as a way of reasoning for the student.

Relevant data come from a pair of 6th-grade classes that learned about global warming. Over two weeks, students learned the mechanisms of the greenhouse effect, the causes of greenhouse gasses, and finally, the effects of global warming. The students completed hands-on activities, saw film clips, received lectures, and completed relevant readings. Betty was used to help one class organize the many concepts and relations. Figure 3 shows a finished "expert" version of Betty.

[Figure 3 about here – Global Warming Map]

The study had two conditions. All the students in the study completed the same basic lessons. At regular points, students were asked to create concept maps to organize their learning, and they all learned how to make causal relations in a concept map. One class was assigned to the Betty condition; the students used the Betty software. The other condition was assigned to the Self condition; these students used *Inspiration,* a popular, commercial concept-mapping program.

Students in both conditions received multiple opportunities for feedback with an important difference. In the Betty condition, agents answered the questions, and the feedback was directed towards the agents. In the Self condition, the students answered the questions, and the feedback was directed towards them. This difference occurred across several feedback technologies. For example, the agents took quizzes or the students took quizzes. For homework, the agents answered questions in the Gameshow or the students answered the questions in the Gameshow. Thus, the main difference between conditions was that in the Betty condition, the learning interactions revolved around the task of teaching and monitoring the agent, whereas in the Self condition, the learning interactions revolved around the task of creating a concept map and answering questions and monitoring oneself.

[Figure 4 about here – accuracy by inference chain length]

The students in the Betty condition adopted Betty's reasoning style. After each unit – mechanisms, causes, effects – all the students completed short-answer, paper-pencil tests. The tests included questions that required short, medium, or long chains of inference. Figure 4 shows that overtime the Betty students separated themselves from the Self students in their abilities to complete longer chains of inference. After the first unit, the two groups overlapped, with the Betty students showing a very modest advantage for the longer inferences. After the second unit, the TA students showed a strong advantage for the medium-length inferences. By the final unit, the TA students showed an advantage for short, medium, and long inferences.

This study used intact classes, so the results are promissory rather than conclusive. Nevertheless, the steady improvement in length of inference is exactly what one would expect the Betty software to yield, because this is what the agent's reasoning models and enforces. By monitoring the feedback generated by Betty's chains of reasoning, students adopted her style of thinking.

Regulating Cognition for Another

In addition to monitoring cognition, metacognition involves taking steps to guide cognition, or as it is often termed "regulating" cognition (refs – Azevedo, Pintrich, Zimmerman Brown**). Regulating another can help students learn to regulate for themselves.

Thus far, Betty's features have supported monitoring, but there were few features to help students chose what to do if they detected a problem. For example, one student's agent was performing poorly in the Gameshow and the student did not know what to do.

Fortunately, another student used chat to provide support, "Dude, the answer is right there in the reading assignment!"

To help students learn to self-regulate their thinking, Betty also comes in a self-regulated learning (SRL) version. For example, when students add new concepts or links, Betty can spontaneously reason and remark that the answer she is deriving does not seem to make sense. This prompts students to reflect on what they have just taught Betty and to appreciate the value of checking understanding. SRL Betty also includes Mr. Davis, a mentor agent shown in Figure 5. Mr. Davis complements the teaching narrative, because he grades Betty's quiz performance. Mr. Davis also provides motivational support and strategies to help the student improve the TA's knowledge. In SRL Betty, the computer characters are more interactive because they can take initiative instead of just reacting to the student.

[Figure 5 – Mr. Davis]

SRL Betty implements regulation goals that are borrowed from Zimmerman's (1989*) list of metacognitive elements. The system monitors for specific patterns of interaction, and when found, Betty or Mr. Davis provide relevant suggestions (also see Tan, Biswas, & Schwartz, 2006*, Jeong, et al., 2008*). Table 3 provides a sample of triggering patterns and responses used by the SRL system; there are many more than those shown in Table 3.

[Table 3 about here – SRL Patterns and Responses]

<u>Self –Regulation Support Improves Student Learning</u>

The self-regulation support in SRL Betty helps students learn science content better. Fifty-six 5th-grade students learned about interdependence in a river ecosystem with a special focus on the oxygen cycle. The students worked over seven class periods starting with the food chain, then photosynthesis and respiration, and finally the waste cycle. To help the students learn, there were quizzes and reading resources built into the system.[1]

There were three conditions. The Regulated-Teaching (RT) condition used SRL Betty, per Table 3. Students could submit Betty to take a quiz, and Mr. Davis provided metacognitive tips about resources and steps the students could use to teach Betty better. Mr. Davis did not dictate specific changes to Betty's knowledge, for example, to add a particular concept or change a link. Instead, he told students strategies for improving Betty's knowledge (e.g., set goals based on the answers that Betty had got wrong).

In the Intelligent Coach (IC) condition, students used the software to make concept maps of their own knowledge. There was no teaching cover story. Instead, of asking Betty to answer a question, students could ask Mr. Davis to answer a question using the concept map or to explain how the map gave a certain answer. When they wanted, students could also submit their concept map to Mr. Davis for a quiz. Mr. Davis scored the maps by indicating which questions it would get right and wrong. Unlike the RT condition, Mr. Davis provided direct instructions for how to fix the concept map (see Biswas et al., 2005, *). For example, Mr. Davis could tell the students to "consider how macro-invertebrates might affect algae and add an appropriate link." Thus, the IC

[1] In the studies using the Gameshow, the students received the nodes, and their task was to determine the links. In the SRL system, the students had to decide which nodes to include in their maps based on the reading, so they could develop strategies for identifying key concepts.

condition removed the teaching narrative, and replaced SRL guidance with directives for how to fix the map.

The final condition was Learning-by-Teaching (LT). The system provided the same directive support as the IC condition, except that students thought they were teaching Betty. So, for example, instead of asking Mr. Davis to show how their map would answer a specific question, students asked Betty to show how she would answer a specific question.

After the seven days, the students' concept maps were scored for the inclusion of correct nodes and links based on the reading materials. The left-side of Table 4 shows that the RT condition did the best, indicating that the metacognitive prompting was valuable for content learning. The LT condition also did better than the IC condition, even though the only treatment difference was whether students thought they were teaching and monitoring Betty, instead of being monitored by Mr. Davis. This latter finding reaffirms the previous findings in a tighter experimental design. If students believe they are teaching an agent, it leads to superior learning even when they are using the same concept mapping tool and receiving equivalent feedback.

[Table 4 about here – Concept Map Scores]

The study included a second transfer phase. Six weeks later, students left their original conditions to spend five class periods learning about the land-based nitrogen cycle. The logic of this transfer phase was that if students had developed good metacognition, they would be more prepared to learn the new content on their own (Bransford & Schwartz, 1999). All the students worked with a basic Betty version; there were on-line reading resources; Betty could answer questions; and, students could see how well Betty did on quizzes. There was no extra support, such as how to improve Betty's map or their teaching. The right-hand column of Table 4 shows the results. Students who had been in the RT treatment learned the most from the transfer task, followed closely by the LT students, with the IC students doing significantly worse.

The SRL Betty version used by the RT condition led to better learning in the first phase of the study, but once students left the SRL system, they performed about the same as the LT students who had never received metacognitive support for regulation. By these data, it is an open question whether SRL Betty provided any extra benefit for developing a lasting metacognition compared to just teaching Betty. The RT students may not have internalized the metacognitive prompts sufficiently to help them regulate their cognition in the transfer phase. Moreover, the evidence on the role of metacognition in the transfer performance is circumstantial. It is possible that students who taught Betty had developed a stronger base of content knowledge on the oxygen cycle, which made it easier for them to learn the analogous nitrogen cycle. To get more direct evidence of lasting changes in metacognition, it is necessary to look at the students' interactive metacognitive behaviors and not just their knowledge products.

Adopting Interactive Metacognition

As mentioned at the outset, in an interactive context, metacognition can help people make choices about what actions are most likely to support their learning. Looking at students' choices provides a way to evaluate, and conceivably, support metacognitive development. In a technology environment that permits students to choose learning activities, machine learning techniques can be a powerful analysis ally. Using a new analysis technique, the log files from the preceding study show that the SRL version of Betty helped students in the RT condition better regulate their choices when they moved into the transfer phase.

[Table 5 about here – Possible student choices]

In the preceding study, students could make a number of choices about which activities to pursue. Table 5 summarizes the possibilities. For example, one student read the resources, and then made a number of edits to the map. Afterwards, the student submitted the map to a quiz, made some more edits, and then asked a pair of questions of the map. In raw form, the sequence was: RA → EM → EM → EM → AQ → EM → EM → EM → EM → EM → EM → EM → EM → EM → RQ → EM → EM → AQ → AQ. To make sense of these complex choice sequences, a new methodology (Li and Biswas, 2002*; Jeong & Biswas, 2008) automated the derivation of Hidden Markov Models. Hidden Markov Models (HMM) represent probabilities of transitioning between different "aggregated" activity states (Rabiner, 1989*). An aggregated activity state represents common interaction patterns that comprise sequences of recurring smaller activity transitions. As described below, this is useful for identifying high-level choice patterns, much in the way that factor analysis is useful for identifying clusters of survey items that reflect a common underlying psychological property.

The HMM analysis generated three interpretable patterns relevant to metacognition. The *Basic Map Building* pattern involves editing the map, submitting the map for a quiz, and occasionally referring to the reading resources. It reflects a basic and important metacognitive strategy. Students work on their maps, check the map with a quiz to see if there are errors, and occasionally refer to the readings. In the *Map Probing* pattern, students edit their maps, and then they ask a question of their map to check for specific relations between two concepts (e.g., if fish increase, what happens to algae?).

This pattern exhibits a more proactive, conceptually driven strategy. Students are targeting specific relations rather than relying on the quiz to identify errors, and students need to formulate their own questions to check their maps. The *Map Tracing* pattern reflects students asking Betty or Mr. Davis (depending on the system) to explain the steps that led to an answer. When Betty or Mr. Davis initially answers a question, they only state the answer and show the paths they followed. To see the increases and decreases in the intermediate links of a long inference, students have to request an explanation. These decomposing explanations are particularly useful when maps become complex, and there are multiple paths between two concepts. Map Tracing is a sophisticated metacognitive strategy, because it involves decomposing a chain of reasoning step-by-step, even after the answer has been generated.

[Figure 6 about here. HMM transition probabilities]

Figure 6 shows how students transitioned from one state to another. Tracing the probabilities of transitions yields the proportion of time students are in a specific activity state. In the first phase of the study, students in all three conditions spent a significant proportion of time in Basic Map Building. However, the RT (Regulated Teaching) students often transitioned into Map Probing and Map Tracing. The SRL version included two features to make this happen. The increase in Map Probing occurred because Betty would not take a quiz if students had not checked her reasoning by asking her a question. The increase in Map Tracing occurred because Betty and Mr. Davis suggested that the students ask Betty to explain her reasoning, so the students could trace her reasoning and look for errors. As a result, the RT group used the Map Probing and Tracing strategies 2 to 1 over the other two groups. The SRL version caused a more versatile repertoire of interactive strategies for learning, which cashed out in better content learning, as indicated their map quality.

The metacognitive strategies practiced in the initial learning phase transferred somewhat when students had to learn the nitrogen cycle on their own. At transfer, when all students had to learn about the nitrogen cycle without any special feedback or tips, the interactive differences between conditions were much smaller. However, there was a meaningful difference that involved transitions into Map Tracing. Students who had used the SRL system were most likely to transition into Map Tracing, and the LT students who had taught Betty without SRL were the next most likely. These differences are modest, but they may explain why the LT and RT treatments learned more at posttest. The students were more likely to check how their agent was reaching its conclusion, which conceivably, could have caused the superior learning.

Deriving patterns through HMM is a new approach to understanding students' metatcognitive choices in an interactive environment. The main promise of analyzing these patterns is that it can help improve the design of interactive technologies for learning. By identifying better and worse interactive patterns for learning, it will be possible to design the computer system to identify those patterns in real-time and provide prompts (a) to move students away from ineffective metacognitive patterns, or (b) to keep them in good patterns. Thus, an important new step will be to correlate interactive patterns with specific learning outcomes, so it is possible to determine which choice patterns do indeed lead to better learning.

CONCLUSION

The leading hypothesis in this chapter has been that monitoring another person, or in this case an agent, can engage productive metacognitive behaviors. This interactive metacognition can lead to better learning, and ideally, if given sufficient practice, students will eventually turn the metacognition inwards. The first empirical section demonstrated that students do take their agent's behavior as cognitive in nature, and that the agent's reasoning is correlated with the students' own knowledge. Thus, when students work with their agent, they are engaging in metacognition. It is interactive metacognition directed towards another. The second empirical section demonstrated that monitoring an agent can lead to better learning, because students internalize the agent's style of reasoning. In the final empirical sections, the Teachable Agent was enhanced to include support for regulating the choices that students make to improve learning. Again, the results indicated that working with an agent led to superior content learning, especially with the extra metacognitive support. Moreover, students who had worked with an agent made a near transfer to learn a new topic two months later. An analysis of the interactive patterns indicated that the students who had taught agents exhibited a more varied repertoire of interactive choices for improving their learning. They also exhibited some modest evidence of transferring these metacognitive skills by checking intermediate steps within a longer chain of inference.

In a separate study not reported here, an Intelligent Coaching condition included self-regulated learning support, similar to the Regulated Teaching condition. (Mr. Davis gave prompts for how to improve the concept map by consulting resources, etc.). In that study, the IC + metacognitive support condition did no better than an IC condition, whereas the RT condition did. So, despite similar levels of metacognitive prompting, the prompting was more effective when directed towards monitoring and regulating one's Teachable Agent. This result also supports the claim that taking on the task of monitoring and regulating another's thoughts can be more beneficial than monitoring and regulating one's own thought, even when there is metacognitive help for monitoring oneself.

It is informative to contrast Betty with other technologies designed as objects-to-think-with. Papert (*), for example, proposed that the programming language Logo would improve children's abilities to plan. Logo involved programming the movement of a graphical "turtle" on the computer screen. Evidence did not support the claim that Logo supported planning (Pea, *). One reason might be that students had to plan the behavior of the turtle, but the logical flow of the program did not resemble human planning itself. For example, the standard programming construct of a "do-loop" involves iterating through a cycle and incrementing a value until a criterion is reached. The execution of the logic of this plan does not resemble many human versions of establishing and managing a plan. Therefore, programming in Logo is an interactive task, but it is not a task where one interacts with mental states or processes. In contrast, the way Betty reasons through causal chains is similar enough to human reasoning that the interactive programming of Betty can be treated as working with her mental states, and students can internalize her cognitive structure and their own metacognitive behaviors towards those structures.

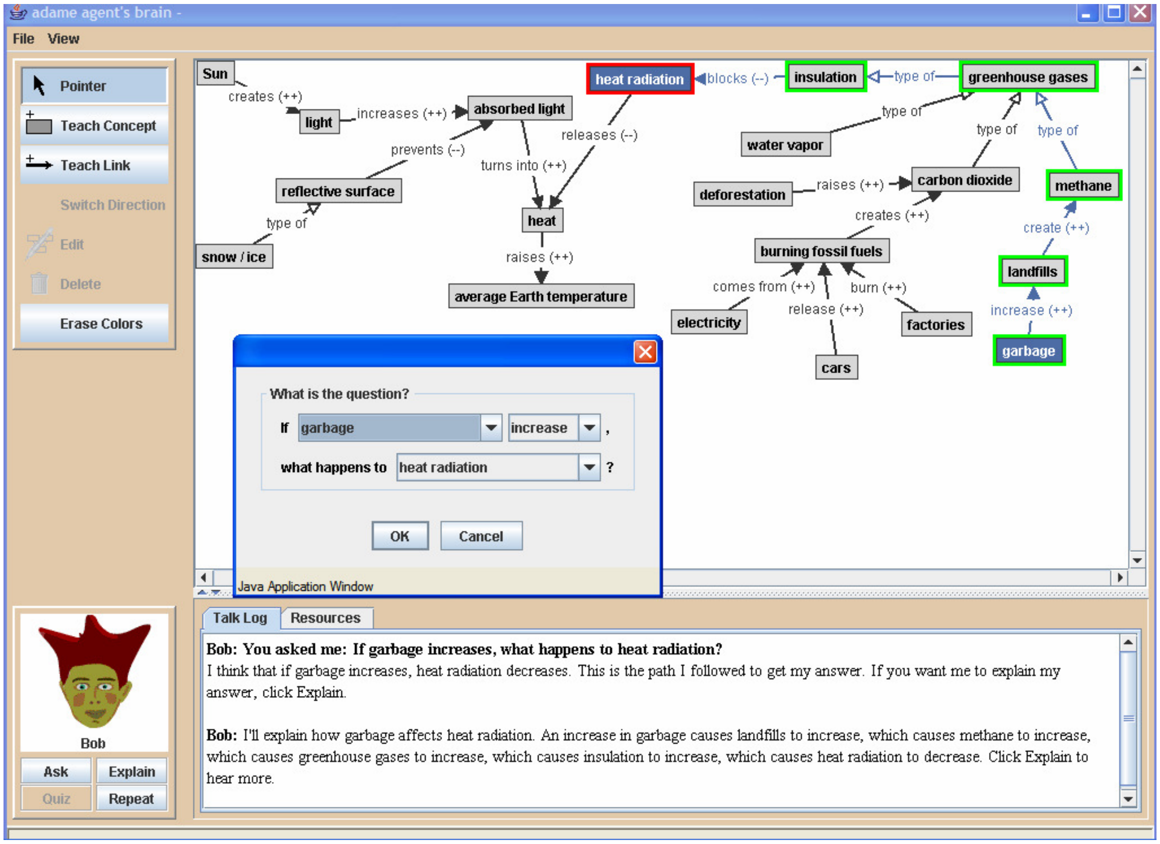
adame agent's brain -
File View
Pointer
Teach Concept
Teach Link
Switch Direction
Edit
Delete
Erase Colors
Sun
creates (++)
light
increases (++)
absorbed light
heat radiation
blocks (--)
insulation
type of
greenhouse gases
type of
water vapor
type of
type of
releases (--)
prevents (--)
turns into (++)
reflective surface
type of
snow / ice
heat
raises (++)
average Earth temperature
deforestation
raises (++)
carbon dioxide
methane
creates (++)
create (++)
burning fossil fuels
landfills
comes from (++)
burn (++)
release (++)
increase (++)
electricity
factories
garbage
cars
What is the question?
If garbage increase ,
what happens to heat radiation ?
OK
Cancel
Java Application Window
Bob
Ask
Explain
Quiz
Repeat
Talk Log
Resources
Bob: You asked me: If garbage increases, what happens to heat radiation?
I think that if garbage increases, heat radiation decreases. This is the path I followed to get my answer. If you want me to explain my answer, click Explain.
Bob: I'll explain how garbage affects heat radiation. An increase in garbage causes landfills to increase, which causes methane to increase, which causes greenhouse gases to increase, which causes insulation to increase, which causes heat radiation to decrease. Click Explain to hear more.

Figure 1


Way to go, Bob! You now have 500 points.
-350
0
175
0
wager: 350
wager: 0
wager: 175
wager: 500
If cars decrease, what happens to burning fossil fuels?
send

Figure 2

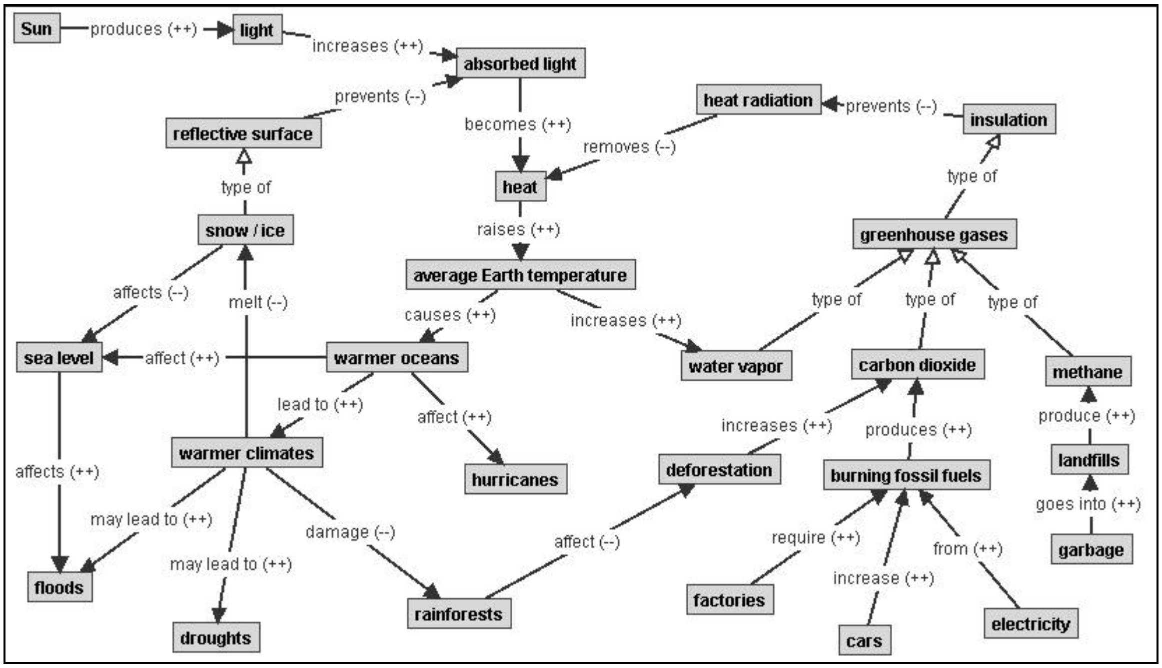
Sun
produces (++)
light
increases (++)
absorbed light
prevents (--)
reflective surface
becomes (++)
heat radiation
prevents (--)
insulation
removes (--)
heat
type of
type of
snow / ice
raises (++)
greenhouse gases
average Earth temperature
affects (--)
melt (--)
causes (++)
increases (++)
type of
type of
type of
sea level
affect (++)
warmer oceans
water vapor
carbon dioxide
methane
lead to (++)
affect (++)
increases (++)
produces (++)
produce (++)
warmer climates
hurricanes
deforestation
burning fossil fuels
landfills
affects (++)
may lead to (++)
damage (--)
affect (--)
require (++)
from (++)
goes into (++)
may lead to (++)
increase (++)
garbage
floods
rainforests
factories
electricity
droughts
cars

Figure 3.

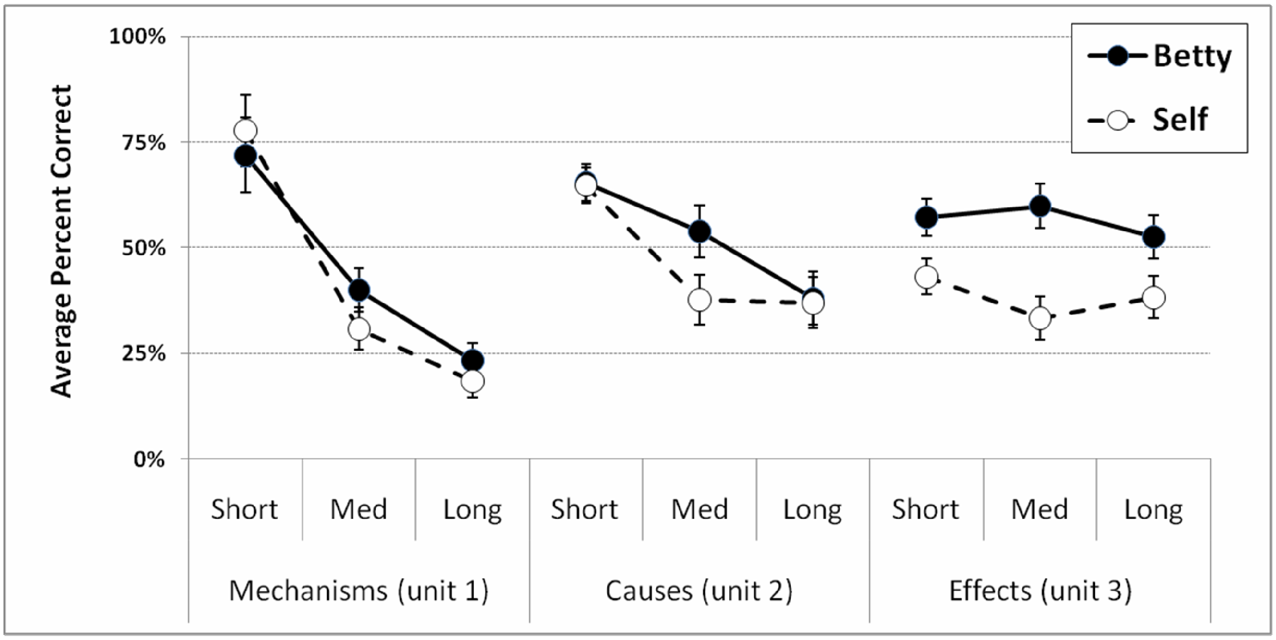
Betty
Self
100%
75%
50%
25%
0%
Average Percent Correct
Short
Med
Long
Short
Med
Long
Short
Med
Long
Mechanisms (unit 1)
Causes (unit 2)
Effects (unit 3)

Figure 4.

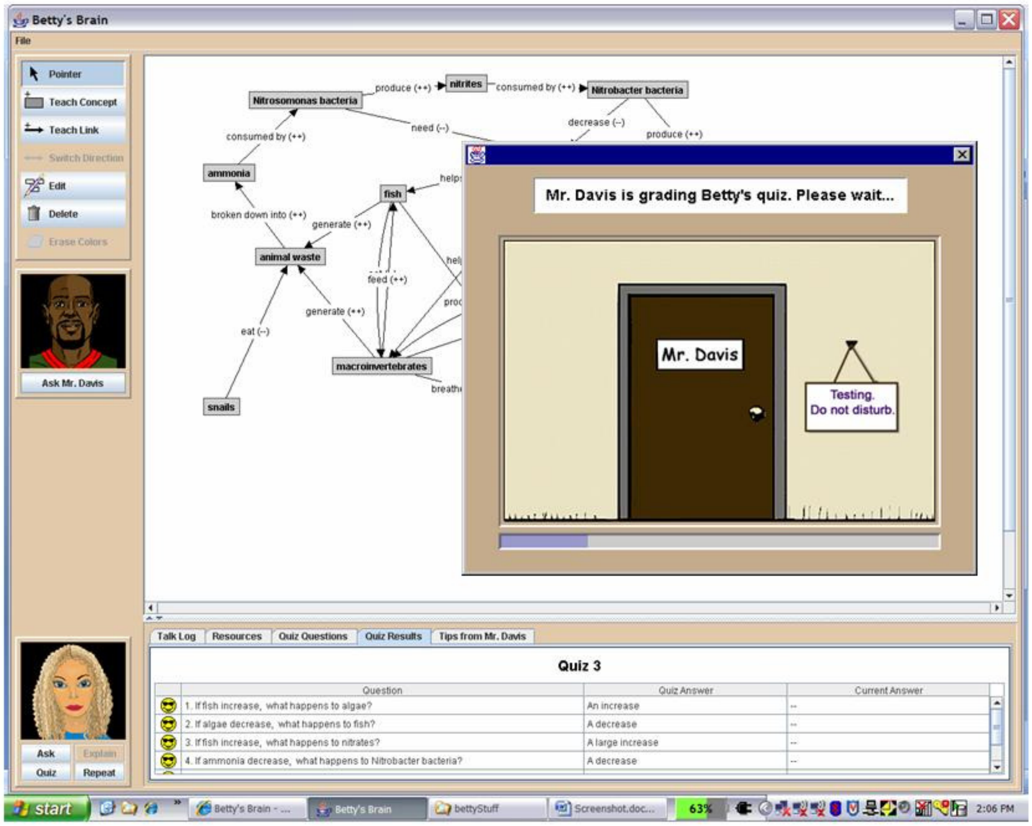
Betty's Brain
File
Pointer
Teach Concept
Teach Link
Switch Direction
Edit
Delete
Erase Colors
Ask Mr. Davis
Ask
Explain
Quiz
Repeat
Nitrosomonas bacteria
produce (++)
nitrites
consumed by (++)
Nitrobacter bacteria
need (--)
decrease (--)
produce (++)
consumed by (++)
ammonia
broken down into (++)
fish
generate (++)
animal waste
feed (++)
generate (++)
eat (--)
macroinvertebrates
snails
Mr. Davis is grading Betty's quiz. Please wait...
Mr. Davis
Testing.
Do not disturb.
Talk Log
Resources
Quiz Questions
Quiz Results
Tips from Mr. Davis
Quiz 3
Question
Quiz Answer
Current Answer
1. If fish increase, what happens to algae?
An increase
2. If algae decrease, what happens to fish?
A decrease
3. If fish increase, what happens to nitrates?
A large increase
4. If ammonia decrease, what happens to Nitrobacter bacteria?
A decrease
start
Betty's Brain - ...
Betty's Brain
bettyStuff
Screenshot.doc...
63%
2:06 PM

Figure 5.

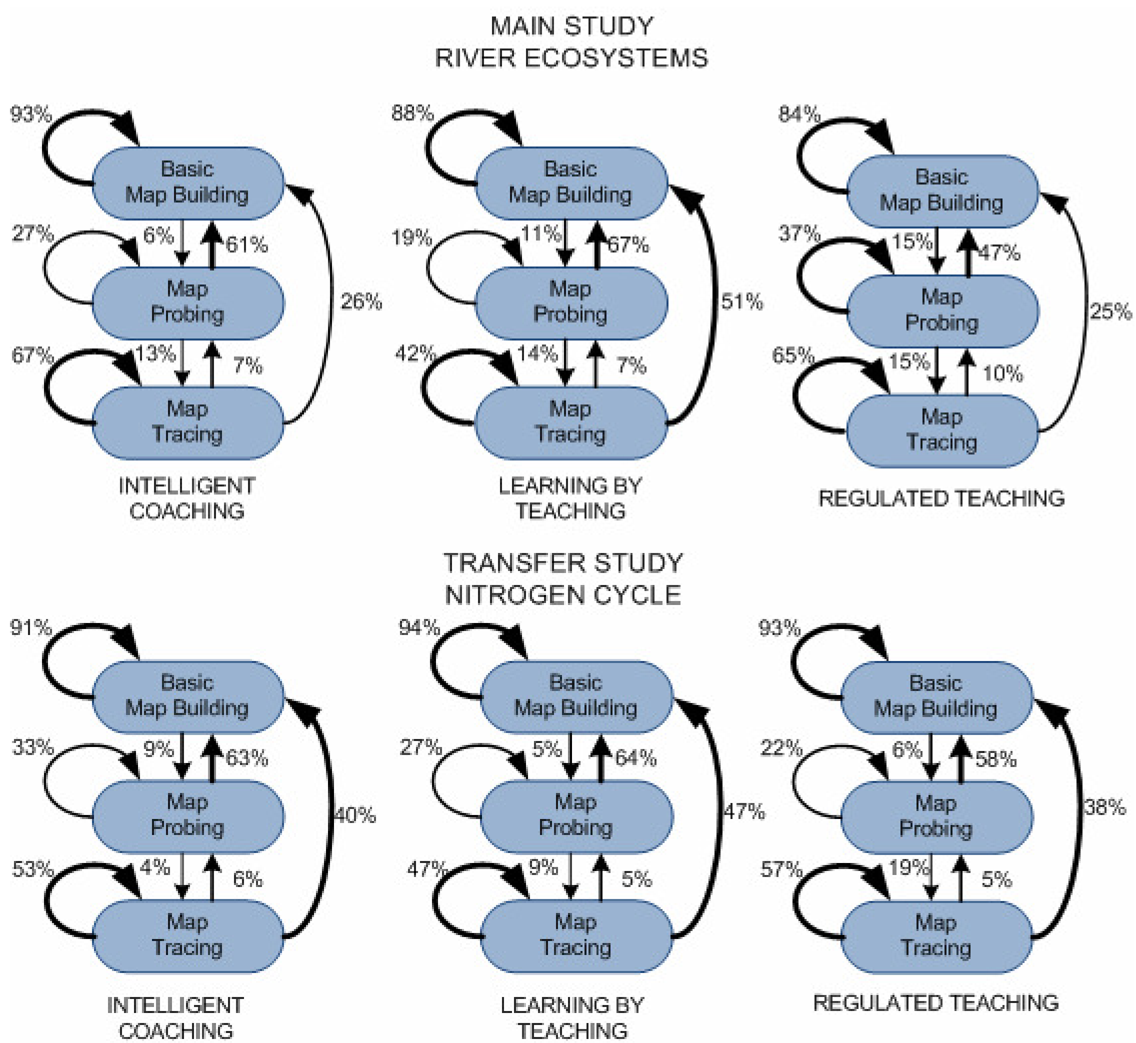
MAIN STUDY
RIVER ECOSYSTEMS
Basic
Map Building
Map
Probing
Map
Tracing
93%
27%
6%
61%
67%
13%
7%
26%
INTELLIGENT
COACHING
88%
19%
11%
67%
42%
14%
7%
51%
LEARNING BY
TEACHING
84%
37%
15%
47%
65%
15%
10%
25%
REGULATED TEACHING
TRANSFER STUDY
NITROGEN CYCLE
91%
33%
9%
63%
53%
4%
6%
40%
INTELLIGENT
COACHING
94%
27%
5%
64%
47%
9%
5%
47%
LEARNING BY
TEACHING
93%
22%
6%
58%
57%
19%
5%
38%
REGULATED TEACHING

Figure 6.

Figure CAPTIONS

Figure 1. A Teachable Agent. The student has (a) named his agent "Bob" instead of Betty, (b) customized Bob's look, (c) taught Bob about global warming, and (d) asked Bob what happens to heat radiation if garbage increases.

Figure 2. Triple-A-Challenge Gameshow. Students log on to for homework. After teaching their agents, the agents play against (and with) one another. A host asks questions of each agent. Students wager on whether they think their agent will give the right answer. There is a chat window so students can communicate with one another during the game.

Figure 3. Target Knowledge Organization for Global Warming Curriculum.

Figure 4. Effects of Betty versus Self. Each test included questions that depended on short, medium, or long chains of causal inference to answer correctly. With more experience across the lesson units, Betty students showed an increasing advantage for longer causal inferences. The Self condition used the concept mapping software *Inspiration* instead of Betty.

Figure 5. Adding Self-Regulated Learning to Betty's Brain. The student has submitted Betty to take a Quiz given by Mr. Davis, and the results are shown in the bottom panel. Mr. Davis and Betty provide tips and encouragement for engaging in metacognitive behaviors.

Figure 6. Transitional Probabilities between Interactive States. The numbers beside the arrows indicate the probability that students would transition from one state to another. For example, in the Intelligent Coaching condition in the main study, students had a very low probability (6%) of leaving Basic Map Building, and therefore, the

entailed frequency of Map Probing and Map Tracing is extremely low. Interactive states and transitional probabilities were derived using Hidden Markov Model statistical learning.

Table 1. Average Number of Attributions to Success and Failure per Successful and Unsuccessful Questions Answers (and standard errors of the mean).

	Attributions to Success				Attributions to Failure			
Condition	Self	Agent	Both	Total	Self	Agent	Both	Total
TA Answers	.17(.12)	.27(.12)	0.0 (.0)	.44 (.16)	.54(.13)	.47(.21)	.66(.19)	1.67(.28)*
Student Answers	.53(.10)	n/a	n/a	.53(.10)	.65(.22)	n/a	n/a	.65 (.22)

Note: * Comparison of condition means, $p < .05$.

Table 2. Correlations between Betty and Student Test Scores.

	Student Test Scores								
	All Questions			**TA Questions**			**Non-TA Questions**		
Betty APQ Index	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Test 1	.60**			.51**			.56**		
Test 2		.66**			.47*			.66**	
Test 3			.34			.12			.48*

Note: ** $p < .01$; * $p < .05$

Table 3. Sample of Interactive Patterns and Responses by the SRL System

Regulation Goal	Pattern Description	Betty Response	Mr. Davis Response
MONITORING THROUGH EXPLANATION	**Multiple requests for Betty to give an answer but no request for explanation**	Let's see, you have asked me a lot of questions, but you have not asked for my explanations lately. Please make me explain my answers so you will know if I really understand.	Without asking Betty to explain her answers, you may not know whether she really understands the chain of events that you have been trying to teach her. Click on the Explain button to see if she explains her answer correctly.
SELF-ASSESSMENT	**Repeated quiz request but no updates have been made to the map.**	Are you sure I understand what you taught me? Please ask me some questions to make sure I got it right. I won't take the quiz otherwise. Thanks for teaching me about rivers!	You have not taught Betty anything new. Please, spend some time teaching her new links and concepts and make sure she understands by asking her questions. Then she can take the quiz again. If you need help learning new things, check the resources.
TRACKING PROGRESS	**The most recent quiz score is significantly worse than the previous**	I would really like to do better. Please check the resources, teach me, and make sure I understand by asking me questions that are on the quiz. My explanation will help you find out why I am making mistakes in my answers. Also, be sure to check out the new tips from Mr. Davis.	Betty did well on the last quiz. What happened this time? Maybe you should try rereading some of the resources and asking Betty more questions so that you can make sure she understands the material**.**
SETTING LEARNING GOALS	**Betty is asked a question that she cannot answer for the second time**	I just don't know the relationships yet, maybe we should ask Mr. Davis what we need to learn.	I've seen this kind of difficultly with teaching some of my own students in the past. You should try looking for missing link connections or links that are in the wrong direction.

Table 4. Average concept map scores at the end of the main treatment (oxygen cycle) and the transfer treatment (nitrogen cycle).

	Study Phase			
	Main Treatment (1st)		Transfer for Learning (2nd)	
	Map Score		Map Score	
Condition	M	(SD)	M	(SD)
Intelligent Coach (IC)	22.8	(5.3)	22.7	(13.7)
Learning-by-Teaching (LT)	25.7*	(6.3)	31.8#	(12.0)
Regulated Teaching (RT)	31.6*+	(6.6)	32.6*	(9.9)

Note: * Greater than IC, $p < .05$; # Greater than IC, $p < .1$; + Greater than LT, $p < .05$.

Table 5. Possible choices of activities.

Activity Name	Student Actions
Edit Map (EM)	adding, modifying, or deleting concepts and links
Resource Access (RA)	accessing the resources
Request Quiz (RQ)	submitting map to take a quiz
Ask Query (AQ)	querying Betty or Mentor to use map to answer a question
Request Explanation (RE)	asking Betty or Mentor to explain an answer to a query
Continue Explanation (CE)	asking for a more detailed explanation