

Running head: PEDAGOGICAL AGENTS

Self-regulated Learning in Learning Environments with Pedagogical Agents
that Interact in Natural Language

Arthur Graesser

Department of Psychology and Institute for Intelligent Systems, University of Memphis
and Danielle McNamara

Department of Psychology and Institute for Intelligent Systems, University of Memphis

Send correspondence to:

Art Graesser
Department of Psychology & Institute for Intelligent Systems
202 Psychology Building
University of Memphis
Memphis, TN 38152-3230
901-678-5847
901-678-2579 (fax)
a-graesser@memphis.edu

Abstract

This article discusses the occurrence and measurement of self-regulated learning (SRL) both in human tutoring and in computer tutors with agents that hold conversations with students in natural language and help them learn at deeper levels. One challenge in building these computer tutors is to accommodate, encourage, and scaffold SRL because these skills are not adequately developed for most students. Automated measures of SRL are needed to track progress in meeting this challenge. A direct approach is to train students on fundamentals of metacognition and SRL, which is the approach taken by iSTART, MetaTutor, and other agent environments. An indirect approach to promoting SRL, taken by AutoTutor, is to track the student's knowledge and SRL based on the student's language and to respond intelligently with discourse moves to promote SRL. This fine-grained adaptivity considers the student's cognitive states, the discourse interaction, and the student's emotional states in a recent AutoTutor version.

Self-regulated Learning in Learning Environments with Pedagogical Agents
that Interact in Natural Language

This article discusses the occurrence and measurement of self-regulated learning (SRL) both in human tutoring and in computer tutors with pedagogical agents that attempt to promote deeper learning by holding conversations with students in natural language. Some of the computer systems with pedagogical agents adopt a direct approach to training students on the fundamentals of metacognition and SRL. However, others take an indirect approach to promoting SRL by tracking the student's knowledge and SRL on the basis of the student's language and by generating discourse moves that encourage SRL. One-to-one tutoring is presumably well suited to promote SRL because the tutor can track the particular knowledge and skills of individual students and can tailor the tutoring session to their idiosyncrasies. In order to achieve this fine-grained adaptation to the student, it is necessary to measure the occurrence of SRL throughout the tutoring process and to infer the SRL skills of particular students. The SRL measures of concern in this article are those that can be gleaned from the conversational interaction between student and tutor.

There is an ideal vision of self-regulated learners. They formulate learning goals, track progress on these goals, identify their own knowledge deficits, detect contradictions, ask good questions, search relevant information sources for answers, make inferences when answers are not directly available, and initiate steps to build knowledge at deep levels of mastery. Their "meta" knowledge of cognition, emotions, communication, and social interaction is well honed, so self-regulated learning (SRL) is easy. On a parallel track, there is the vision of the ideal tutors who encourage SRL. They do not lecture incessantly and steamroll the sessions. Like a Socratic tutor, they ask the students clever questions that lead the students to explore important content and to self-

discover their own misconceptions. They provide the student a long leash as the student actively explores the learning environment, but the leash is not so long that the student flounders extensively.

These visions of the ideal learner and tutor are indeed inspiring. However, they are also exceptionally rare in most educational environments, as we document in this article. Sophisticated SRL is rarely encouraged by most human tutors who allegedly tailor the tutorial session to individual students (Chi, Roy, & Hausmann, 2008; Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; Graesser & Person, 1994; Graesser, D'Mello, & Person, 2009). Instead, it is the tutor who drives the agenda and conversation. Sophisticated SRL strategies are also rare in hypermedia environments in which students have substantial freedom in exploring the hyperspace of content at their own pace (Azevedo, 2005, Azevedo & Cromley, 2004; Greene, Moos, Azevedo, & Winters, 2008). The vast majority of students need scaffolding to build effective SRL strategies.

This article focuses on a class of computer tutors that use *pedagogical agents* to help students learn about science and technology at deeper levels of understanding. Subject matters in science and technology are particularly difficult for students to master when the mechanisms are complex and deep comprehension of the material is necessary. Deep comprehension is particularly important for students in college, the developmental level that is the primary target of the systems described in this article. We have also tested our systems on students in middle or high school, but it is the college student or high school senior that we have in mind in this article. Our central assumption is that one way of promoting deeper comprehension is for an expert to hold conversations in natural language with the student. Dialogue-based apprenticeship learning was indeed the primary form of learning for millennia until classrooms emerged in the industrial revolution (Collins & Halverson, 2009; Resnick, in press; Rogoff & Gardner, 1984). Our belief is that deep comprehension can also

be achieved by pedagogical agents that hold conversations with the student in natural language, that systematically apply pedagogical strategies, and that encourage SRL when appropriate.

Differences between deep and shallow comprehension have often been identified by researchers in cognitive science, education, and discourse processing (Bloom, 1956; Graesser & McNamara, in press; Graesser & Person, 1994; Mosenthal, 1996). In our view, deep comprehension of topics in science and technology occur when students build mental models of causal structures, dynamic processes, tradeoffs between variables, and other challenging aspects of complex systems. A student who has achieved deep comprehension can generate inferences, accurately scrutinize the validity of claims, and apply their knowledge to new situations. In contrast, shallow comprehension is limited to lists of components, lists of properties of entities, definitions of terms, and imprecise associations between ideas.

The nature of the dialogue between the student and expert tutor (whether human or virtual) should undoubtedly influence comprehension and learning. All conversations consist of a turn-by-turn interaction between the two parties, with one or more *dialogue move* within each turn. A dialogue move is a verbal expression or action that serves a discourse function (Graesser, D’Mello, & Person, 2009), such asking a question, giving a command, making an assertion, giving short feedback (“very good,” “uh huh,” “no quite”), or even performing an action. How might the conversations differ between successful and unsuccessful tutorial interactions? The contrast has been made between tutor-centered, student-centered, and interaction-centered tutoring environments (Chi, Roy, & Hausmann, 2008; Chi et al., 2001). In the prototypical tutor-centered session, the tutor lectures to the student as the student nods to acknowledge understanding, even when the student does not understand what the tutor is talking about. This extreme case of the lecturing tutor is as limited in effectiveness as normal classrooms with teacher monologues, which are known to be

inferior to human tutoring (Cohen, Kulik, & Kulik, 1982; Graesser, D'Mello, & Cade, in press; Graesser et al., 2009). At the other end of the continuum, the student takes control in student-centered tutoring by selecting topics and problems, actively working on solutions, and asking questions – all vestiges of SRL. This approach is limited by the fact that students are not very good at calibrating their comprehension of material (Dunlosky & Lipko, 2007; Maki, 1998), rarely take the initiative to ask questions (Graesser & Person, 1994), and do not smartly set their agendas in unguided inquiry (Klahr, 2002).

There are intermediate positions between the two extremes of radical tutor-centered and student-centered tutoring. For example, just-in-time instruction (Schwarz & Bransford, 1998) can be quite effective when the student is ready for instruction. A student is particularly open for instruction after the student fails at solving a problem or encounters some other impasse. In open learning environments (Bull & Kay, 2007), the student and instructor have open access to the knowledge and skills that the student has mastered, as well as the landscape to be mastered. This landscape can be depicted in a spreadsheet with particular content and skills, along with the student's associated level of mastery. The motivated student pursues mastery of the content and skills that need work through self-guided selections. Yet another approach to brokering the wishes of the tutor and the student is that of interaction-centered tutoring (Chi et al., 2008; Graesser et al., 2009; VanLehn et al., 2007). In this case, the dialogue interactions between tutor and student converge on a sweet spot of optimal scaffolding; the student is actively engaged as the tutor subtly steers the interaction to productive learning activities. Tutoring researchers are currently exploring how to manage the interactive dialogue in a fashion that optimizes learning, motivation, and adaptation to individual learners.

We are convinced that pedagogical agents hold considerable promise in optimizing interaction-centered tutoring and training. The dialogue strategies of the agent can be consistent, precise, complex, adaptive, and durable. This is in sharp contrast to human tutors who rarely possess these desirable features (Graesser, Person, & Magliano, 1995; Graesser, D’Mello, & Cade, in press). It is extremely difficult to train a human tutor to systematically apply a strategy that goes against the grain of his or her natural conversational inclinations. It is nearly impossible to train a human to perform complex quantitative computations that precisely track student characteristics and that formulate dialogue moves that optimally adapt to the learner. And of course, human tutors get fatigued, whereas computers are tireless. Although there allegedly are advantages of human over computer tutoring in promoting learning, the gap has appreciably narrowed, if not been eliminated, between human tutors and some intelligent tutoring systems with pedagogical agents (Graesser, Jeon, & Dufty, 2008; Graesser et al., 2007; VanLehn et al., 2007).

The theoretical perspective we have adopted with respect to SRL is both eclectic and dialectic. It is eclectic in the sense that it incorporates SRL components that have been proposed by goal-based theories (Pintrich, 2000), information-processing theories (Winne, 2001), social theories that emphasize achievement, attributions, and self-efficacy (Schunk & Zimmerman, 2008), and hybrid theoretical frameworks (Azevedo & Cromley, 2004). Most SRL theories assume that the learners construct goals and plans, monitor metacognitive activities, implement learning strategies, and reflect on their progress, but they emphasize different characteristics of these processes. Instead of adopting only one of these prominent SRL frameworks, we incorporate components that are relevant to the dialectical mechanisms in tutorial dialogues. These dialectical mechanisms include discussion and reasoning as a method of intellectual inquiry, collaborative exchanges during problem solving, question asking and answering,

argumentation to resolve conflicting ideas, and Socratic methods of eliciting truth and exposing false beliefs. These dialectical mechanisms are regarded as a fundamental form of scaffolding to narrow the wide gap between ideal SRL strategies and the unimpressive state of SRL in most students. The hope is that the dialogue with agents will eventually be internalized in the students' habits of mind.

Where is the SRL in Human Tutoring?

Graesser, D'Mello, and Person (2009) documented many of the limitations of human tutoring. Some of these limitations can be rectified to some extent by computer tutors. One counterintuitive conclusion has been that both expert and untrained human tutors have unspectacular "meta" knowledge of pedagogical strategies, cognition, communication, and emotions. Graesser, Person, and Magliano (1995) videotaped over 100 hours of naturalistic tutoring in a corpus with typical unskilled tutors, transcribed the data, classified the spoken utterances into discourse categories, and analyzed the rate of particular discourse patterns. Cade, Copeland, Person, and D'Mello (2008) conducted a similar analysis on outstanding tutors in middle school and high school. These analyses revealed that the tutors rarely implement intelligent pedagogical techniques that have been discussed in the education literature for at least two decades, such as *bona fide* Socratic tutoring strategies, modeling-scaffolding-fading, reciprocal teaching, building on prerequisites, or diagnosis/remediation of deep misconceptions. When a student expresses a bug, misconception, or some form of error that the tutor can detect, the tutor typically corrects it immediately. Contrary to the spirit of SRL and Socratic tutoring, tutors rarely back away and wait for the students to correct their own errors.

Most tutors do foster SRL by trying to get the student to do the talking and acting rather than merely lecturing to the student (Cade et al., 2008; Graesser et al., 2009, in press). More

specifically, the structure of the dialogue in human tutoring (Chi et al., 2001, Graesser et al., 2009) typically follows an *Expectation and Misconception Tailored* dialogue. That is, human tutors have a list of expectations (anticipated good answers) and a list of anticipated *misconceptions* associated with each problem or main question. As the learner expresses information over many turns, the list of expectations is eventually covered and the main question is deemed as answered. Consider the example physics problem below and a representative expectation E and misconception M.

PHYSICS QUESTION: If a lightweight car and a massive truck have a head-on collision, upon which vehicle is the impact force greater? Which vehicle undergoes the greater change in its motion, and why?

E. The magnitudes of the forces exerted by A and B on each other are equal.

M: A smaller object exerts no force on a larger object.

Good tutors try to get the student to articulate the expectations, which is aligned with the metacognitive and pedagogical principal of active learning. They do so by three categories of dialogue moves: *pumps* (“what else?”, “keep going”), *hints*, and *prompts* for the student to fill in important words in the expectation that have not been expressed. Hints and prompts are normally expressed as questions. Particular tutor questions are composed to elicit content in the answers that fill in missing words, phrases, and propositions. For example, a hint to get the student to articulate expectation E might be “What about the forces exerted by the vehicles on each other?” This hint would ideally elicit the answer “The magnitudes of the forces are equal.” A prompt to get the student to say “equal” would be “What are the magnitudes of the forces of the two vehicles on each other?” If the tutor fails to get the student to articulate an expectation through multiple hints and prompts, then the tutor simply *asserts* the correct answer.

These discourse patterns make it possible to measure the extent to which students exhibit SRL. Students exhibit more SRL to the extent that they contribute more correct information and the tutor's dialogue moves drift toward the following precedence ordering: pump > hint > prompt > assertion. For the self-regulated learner, all the tutor needs to do is express a pump and perhaps a hint, whereas the student articulates most of the expectations. Learning gains are indeed correlated with SRL when defined in this way (Graesser et al., 2007; Jackson & Graesser, 2006). From the standpoint of misconceptions, the self-regulated students correct their own errors after the tutor signals scepticism through short negative feedback or asking a question that challenges the misconception. These discourse sequences or patterns are diagnostic of SRL and have been tracked and quantitatively measured in both human tutoring (Graesser et al., 2009) and the AutoTutor system described later (Graesser et al., 2007; Jackson & Graesser, 2006).

Some analyses of tutorial dialogue present a pessimistic picture of SRL and metacognition. For example, students rarely take charge of the tutoring session by selecting problems to work on, asking questions, and initiating changes in topics (Chi et al., 2001, 2008; Graesser et al., 2009; Graesser et al., 1995; Graesser, McNamara, & VanLehn, 2005); this is confirmed when we measure the relative frequency (i.e., instances per hour, per 1000 words, or per session) of students selecting problems (Graesser et al., 1995), students asking questions (Graesser et al., 1995; Graesser, McNamara, & VanLehn, 2005), and students initiating topic changes (Graesser et al., 1995). Such findings once again underscore the need to train students how to acquire SRL by selecting problems, asking questions, and initiating new topics.

Another example of the poverty of SRL and metacognition is the students' responses to the tutors' comprehension-gauging questions, such as *Do you understand?*, *Are you following?*, and *Does that make sense?* If the student's comprehension calibration skills were accurate, then the

student should answer YES when the student understands and NO when there is little or no understanding. Tutoring data do not reliably exhibit this pattern, however (Graesser et al., 1995). There often is a positive correlation between a students' knowledge of the material (based on pretest or posttest scores) and their likelihood of saying NO rather than YES to the tutors' comprehension-gauging questions (Graesser et al., 1995). So it is the knowledgeable students who tend to say *No, I don't understand*. This result supports the claim that deeper learners have higher standards of comprehension (Baker, 1985) and that many students have poor comprehension calibration skills. The fact that many students have subpar comprehension calibration is well documented in the metacognition literature, where meta-analyses have shown only a 0.27 correlation between comprehension scores on expository texts and the students' judgments on how well they understand the texts (Maki, 1998). Metacognition is a prerequisite for self-regulated learning, so there is a need to train both students and tutors on the fundamentals of metacognition.

The tutor ideally builds an accurate and detailed model of the cognitive states of the student, or what is called the *student model* by researchers who develop intelligent tutoring systems. The realistic picture, however, is that human tutors have only an approximate appraisal of the cognitive states of students and that they formulate responses that do not require fine tuning of the student model (Chi, Siler, & Jeong, 2004; Graesser et al., 1995; Graesser et al., 2009, in press). There are three sources of evidence for this claim. First, the short feedback that tutors give to students on the quality of the students' contributions is often faulty. After the student expresses something in a conversational turn, the tutor normally gives short feedback that is either positive (*great, that's right*), neutral (*okay, uh huh*), or negative (*no, not quite*). This short feedback is often incorrect; the feedback has a higher likelihood of being positive than negative after student contributions that are vague or error-ridden (Graesser et al., 1995). Second, tutors

do not have a high likelihood of detecting misconceptions and error-ridden contributions of students (Chi et al., 2004; VanLehn et al., 2007). Third, tutors do not select new cases or problems to work on that are sensitive to the abilities and knowledge deficits of students (Chi et al., 2008). The selection of problems should be tailored to the student's profile according to the zone of proximal development, i.e., not too easy or not too hard, but just right. But available evidence does not support the claim that human tutors are strategic problem selectors (Graesser et al., 2009, in press).

We have so far addressed the cognitive components of tutoring because cognition is presumably the core of learning. However, it is important to also consider the emotions of students when tutors generate pedagogical strategies and foster SRL (D'Mello, Picard, & Graesser, 2007; Graesser et al., 2009; Lepper & Woolverton, 2002). For example, Lepper, Drake, and O'Donnell-Johnson (1997) proposed an INSPIRE model to coordinate pedagogy, cognition, and emotions. This model encourages the tutor to be empathetic and attentive to the student's needs, to assign tasks that are not too easy or difficult, to give indirect feedback on erroneous student contributions rather than harsh feedback, to encourage the student to work hard and face challenges, to empower the student with useful skills, and to pursue topics the student is curious about. One of Lepper's tutoring strategies is to have the tutor assign an easy problem to a struggling student, but the tutor claims that the problem is difficult and encourages the student to give it a try. When the student readily solves the problem, the student builds self-confidence and self-efficacy in conquering difficult material (Schunk & Zimmerman, 2008). Self-efficacy is of course a facilitator of SRL. Unfortunately, our analysis of human tutoring revealed that tutors rarely exhibited such sophisticated affect-motivated strategies (Graesser et al., 2009). However,

it is possible to design computer tutors to implement such strategies because they can quantitatively track the students' knowledge, skills, and emotions.

The sketch of tutoring we have painted so far suggests that both the students and human tutors have limited “meta” knowledge of pedagogy, cognition, communication, emotion, and SRL. It clearly will take some training before these capabilities emerge for the vast majority of students and tutors. In spite of these limitations, it is well documented that tutoring is one of the best learning environments available (Graesser et al., in press). Meta-analyses show learning gains from typical, non-expert human tutors of approximately 0.4 sigma (effect size in standard deviation units) compared to classroom controls and other suitable controls (Cohen, Kulik, & Kulik, 1982). Non-expert tutors include paraprofessionals, cross-aged tutors (i.e., students who are older than the tutee), or same-age peers who have had little or no tutor training and have modest subject-matter expertise. There have not been many systematic studies on learning gains from expert tutors because they are expensive and difficult to recruit in research projects. However, available studies show effect sizes of 0.8 to 2.0 (Bloom, 1984; Chi, et al., 2008; Graesser et al., in press; VanLehn et al., 2007).

Where is SRL in Computer Tutors with Pedagogical Agents?

Some of the above limitations with human tutoring might be corrected in computerized learning environments with pedagogical agents. These animated conversational agents are being incorporated in an increasing number of learning environments (Baylor & Kim, 2005; Biswas, Leelawong, Schwartz, Vye, & TAGV, 2005; Graesser et al., 2008; McNamara, O'Reilly, Rowe, Boonthum, & Levinstein, 2007; McQuiggan & Lester, 2007; Moreno & Mayer, 2004). The agents interact with students and help them learn either by modelling good pedagogy and learning processes or by holding an interactive conversation. The agents may take on different roles, such as mentors, tutors, peers, players in multiparty games, or avatars in the virtual worlds. The students

communicate with the agents through speech, keyboard, gesture, touch panel screen, or conventional input channels. In turn, the agents express themselves with speech, facial expression, gesture, posture, and other embodied actions. Single agents model individuals with different knowledge, personalities, physical features, and styles. Ensembles of agents model social and pedagogical interactions.

There are two ways that these agent-based learning environments might promote SRL. One way is the direct approach, where the agents directly train the students' SRL and different forms of meta-knowledge. Progress on acquiring SRL strategies can be measured and tracked throughout the interaction. The second approach, the one adopted by AutoTutor, is more indirect. The discourse patterns of AutoTutor may model or encourage student SRL, but does not directly train the student on SRL strategies. The students are expected to induce these strategies from the dialogue patterns in the tutoring experience. Once again, progress on acquiring SRL strategies can be measured and tracked throughout the conversations.

Agents with Direct Training of SRL and Meta-cognitive Strategies

We have been involved in the design and testing of several learning environments that have pedagogical agents that directly train students what these strategies are and how to use the strategies. A brief description of some of these projects conveys the potential of these technologies. There is evidence to support the effectiveness of most of these trainers to improve the students' awareness of these trained strategies after short retention intervals and such awareness sometimes improves content learning. However, available research has not settled on how much training is needed and on how well the training transfers to new applications, to novel tasks, and to various tasks after long retention intervals.

(1) *MetaTutor*. MetaTutor trains students on 13 strategies that are theoretically important for SRL (Azevedo et al., 2009). The process of SRL theoretically involves the learners' constructing a plan, monitoring metacognitive activities, implementing learning strategies, and reflecting on their progress and achievements (Azevedo, 2005, Azevedo & Cromley, 2004; Pintrich, 2000; Winne, 2001). There is a main agent (Gavin) that coordinates the overall learning environment and three satellite agents that handle three phases of SRL: Planning, monitoring, and applying learning strategies. Each of these phases can be decomposed further, under the guidance of the assigned conversational agent. For example, metacognitive monitoring can be decomposed into judgments of learning, feeling of knowing, content evaluation, monitoring the adequacy of a strategy, and monitoring progress towards goals. Examples of learning strategies include searching for relevant information in a goal-directed fashion, taking notes, drawing tables or diagrams, re-reading, elaborating the material, making inferences, and coordinating information sources (text and diagrams). Each of these metacognitive and SRL skills have associated measures that are based on the student's actions, decisions, ratings, and verbal input. The frequency and accuracy of each measured skill is collected throughout the tutoring session and hopefully increases as a function of direct training.

(2) *iSTART (Interactive Strategy Trainer for Active Reading and Thinking)*. This strategy trainer helps students become better readers by constructing self-explanations of the text (McNamara, O'Reilly, Rowe, Boonthum, & Levinstein, 2007). The construction of self-explanations during reading is known to facilitate deep comprehension when there is some context-sensitive feedback on the explanations that get produced. The iSTART interventions focus on five reading strategies that are designed to enhance self-explanations: *monitoring comprehension* (i.e., recognizing comprehension failures and the need for remedial strategies),

paraphrasing explicit text, making *bridging inferences* between the current sentence and prior text, making *predictions* about the subsequent text, and *elaborating* the text with links to what the reader already knows. The accuracy of applying these metacognitive skills is measured and tracked throughout the tutorial session.

Groups of agents scaffold these strategies in three phases of training. In an *Introduction Module*, a trio of agents (an instructor and two students) collaboratively describe self-explanation strategies with each other. In a *Demonstration Module*, two agents demonstrate the use of self-explanation in the context of a science passage and then the student identifies the strategies being used. A measure of metacognitive skill is the accuracy of the students' identifying the correct strategy exhibited by the student agent. In a final *Practice* phase, an agent coaches and provides feedback to the student one-to-one while the student practices self-explanation reading strategies. That is, for particular sentences in a text, the agent reads the sentence and asks the student to self-explain it by typing a self-explanation. The iSTART system then attempts to interpret the student's contributions, gives feedback, and asks the trainee to modify unsatisfactory self-explanations (McNamara, Boonthum, Levinstein, & Millis, 2007). Measures of metacognitive skill consist of the quality of the students' self-explanations, of the relative frequency of positive feedback given by iSTART agents on the students' self-explanations, and of the relative frequency of requests for the students to modify the self-explanations.

(3) *SEEK (Source, Evidence, Explanation, and Knowledge) Web Tutor*. Critical thinking about science requires learners to actively evaluate the truth and relevance of information, the quality of information sources, and the implications of evidence and claims (Halpern, 2002). A critical stance toward scientific information is especially important in the internet age, an era when there are millions of web pages but no control over the quality of the scientific information. The

SEEK web tutor was designed to improve college students' critical stance while they search web pages on the topic of plate tectonics (Graesser, Wiley et al., 2007; Wiley et al., 2010). The SEEK Tutor fosters a critical stance with three main facilities. The first is a *Hint* button on the Google search engine page which contains suggestions on how to effectively guide the student's search. The second facility is *Pop-up Ratings and Justifications*, which asks students to evaluate the expected reliability of the information in a site. The reliability of these judgments can be scaled because some sites have reliable information but others convey pseudoscience. The third facility consists of a *Pop-up Journal* with five questions about the reliability of the site that the learner had just visited. These questions were designed to address some of the core aspects of critical stance: *Who authored this site? How trustworthy is it? What explanation do they offer for the cause of volcanic eruptions? What support do they offer for this explanation? Is this information useful to you, and if so, how will you use it?* The system forces the learner to think about each of the five core aspects of critical stance and also to articulate verbally the reasons for their ratings. The learning of the metacognitive skill of taking a critical stance can be measured from the relative frequency and the accuracy of their verbal answers. That is, what is the relative frequency of verbal expressions that match experimenter-defined codes on critical stance for each page? Some objective measures consist of the likelihood that students select web sites to read and the amount of study time per web site. More time should be devoted to high quality web sites than to websites that are poor information sources.

(4) *iDRIVE (Instruction with Deep-level Reasoning questions In Vicarious Environments)*.

iDRIVE has dyads of animated agents train students to learn science content by modeling deep reasoning questions in question-answer dialogues (Craig, Sullins, Witherspoon, & Gholson, 2006; Gholson & Craig, 2006; Gholson et al., 2009). A student agent asks a series of deep

questions about the science content and the teacher agent immediately answers each question. There is evidence that learning gains are higher and students ask better questions for those students who have the mindset of asking deep questions (*why, how, what-if, what-if-not*) that tap causal structures, complex systems, and logical justifications (Graesser & Olde, 2003). When students are trained how to ask good questions, the frequency of good questions increases and their text comprehension improves (Gholson & Craig, 2006; Rosenshine, Meister, & Chapman, 1996; Wisher & Graesser, 2007).

The above four computer environments directly train the strategies associated with SRL and “meta” knowledge. Empirical tests of these systems have investigated the extent to which the students become more aware or competent in the SRL and metacognitive strategies after direct training, at least when assessed through verbal protocols, subjective ratings, decisions on options presented by the computer, actions performed, and time spent studying information sources. Most of the assessments have involved relatively short training sessions (1-3 hours) and retention intervals (shortly after training). In more rigorous tests of these trainers, the researchers assess the students on new tasks several days or weeks later to see how well these strategies transfer (Jackson, Guess, & McNamara, 2010). If there is a successful response to intervention, then the transfer SRL behavior will differ from the baserate SRL behavior. It is too early to make strong conclusions about the effectiveness of these trainers in rigorous assessments of transfer. Research is needed to determine how much SRL training is needed before there is a facilitation of deep learning of subject matters in science and technology.

Agents with Indirect SRL Training: AutoTutor and Operation ARIES!

As discussed earlier, most human tutors do not directly train students on meta-knowledge and SRL strategies (Graesser et al., 2009, in press; Graesser et al., 1995). The students need to

induce whatever strategies are reflected in the tutorial dialogue and there is no guarantee that induction of strategies will be successful. The emphasis is on content rather than strategies. A number of agent-based systems have been developed that simulate human tutors by holding a conversation in natural language. These include *AutoTutor* (Graesser et al., 2005; Graesser, Jeon, & Dufty, 2008) and *Operation ARIES!* (Millis et al., 2009). AutoTutor and ARIES are briefly described below, with a particular concentration on aspects of the conversational dialogue that promote SRL. There is substantial evidence that AutoTutor improves learning (Graesser et al., 2004; VanLehn et al., 2007) compared to reading a textbook for an equivalent amount of time; the mean effect size is 0.8 over 18 studies in physics and computer literacy. However, the focus of this article is on SRL and meta-knowledge strategies rather than learning gains on subject matter content.

(5) *AutoTutor*. AutoTutor's dialogues are organized around difficult questions and problems that require reasoning and explanations in the answers. These questions require the student to articulate in natural language approximately 3-7 sentences (expectations). Coverage of these expectations typically requires 30-100 conversational turns between the tutor and student. AutoTutor simulates human tutors by implementing the Expectation and Misconception Tailored dialogue that was described earlier. That is, AutoTutor attempts to get the student to articulate the expectations by generating pumps, hints, prompts, and feedback, but resorts to generating assertions with correct information when the student has trouble. The tutor corrects misconceptions that students express during the dialogue.

Most students do not have a high degree of articulate knowledge about physics or computer literacy so it takes many conversational turns before the student can construct a good answer. When students are asked these challenging questions, their initial answers are typically

short, only 1 word to 2 sentences in length. However, this is insufficient information to adequately answer the question so tutorial dialogue is needed to flesh out a complete answer. AutoTutor engages the student in a limited form of mixed-initiative dialogue that draws out more of what the student knows and that assists the student in the construction of an improved answer. The dialogue moves of the tutor include those that give short feedback to the students on their contributions (positive, negative neutral, as discussed earlier) and those that advance the coverage of the expectations (pumps, hints, prompts, and assertions). The *short feedback* monitors the quality of the information generated by the student whereas the *coverage advancer* encourages a higher quantity of student contributions. .

Table 1 presents an example physics problem on gravitational forces between the sun and the earth, followed by a dialogue between the tutor and the student. Conversational turns by AutoTutor include the main question (turn 1), pumps (turn 3), hints (turns 3 and 9), prompts (turn 5), feedback (turns 7 and 11), corrections (turn 11), assertions and answers to student questions (turns 5 and 9), and comprehension gauging questions (turn 11). Discourse markers connect multiple dialogue moves within a turn and sometimes cue the student on the function of the subsequent dialogue move (*Okay, I bet you can get this*, before the hint in turn 9). As discussed earlier in the context of human tutoring, the tutor would generate a high proportion of pumps and hints to the extent that a student takes the initiative and does the talking whereas many pumps and assertions would be needed to extract knowledge from students with lower SRL. The accuracy of student knowledge is manifested in the relative frequency of positive feedback rather than negative feedback or corrections. The discourse moves of AutoTutor are used to scale the students' SRL and knowledge (Graesser, Penumatsa et al., 2007; Jackson & Graesser, 2006).

The discourse of the student also provides measures of SRL. A high proportion of student questions (turns 4 and 8) reflects student initiative. Metacognition is reflected in the feeling-of-knowing (turn 2) and meta-comprehension judgments (turn 12). A student who articulates more information, particularly accurate information, is an indicator of SRL as well as knowledge.

Assessments of the quality of the students' contributions during computer tutoring have required special quantitative models. It is easy to reliably track and respond to the verbosity of students, which is typically computed as the number of words or number of content words (nouns, main verbs, adjectives). However, the quality of the students' contributions presents some challenges because the majority of the students' contributions are ungrammatical and not semantically well-formed. Our agent-based learning environments have therefore adopted statistical representations of meaning, world knowledge, and the subject matter content. In particular, most of our systems use *latent semantic analysis* (Graesser, Penumatsa et al., 2007; Landauer, McNamara, Dennis, & Kintsch, 2007; McNamara, Boonthum et al., 2007) and other statistical measures in computational linguistics to evaluate the matches between student verbal input and expectations or misconceptions. Measures of the quality of student input include the expectation-input match scores for individual expectations, sets of expectations for a problem, and sets of problems. These match scores have been validated in studies that compare computer-generated match scores with scores of expert judges (Graesser, Penumatsa et al., 2007).

AutoTutor has other dialogue facilities that promote mixed-initiative dialogue and thereby transfer more control into the hands of the student. One of these is its attempts to answer the student's questions. The answers to the questions are retrieved from glossaries or from paragraphs in textbooks via intelligent information retrieval. AutoTutor asks a counter-

clarification question (e.g., “I don’t understand your questions, so could you ask it in another way?”) when it does not understand the student’s question. Most versions of AutoTutor explicitly invite questions from students, such as “Do you have any questions?” or “This is a good point to answer questions you might have.” A reasonable index of SRL is the incidence of student questions, which increase with AutoTutor in quantity and depth (Graesser et al, 2005), and also acts of students seeking help (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003). However, student question/help metrics need to be treated with some caution or need to be adjusted when students “game the system” by mechanically gathering answers through questions or help, without attempting to understand the material (Baker, Corbett, Roll, & Koedinger, 2009). Fortunately, there are metrics for detecting students who game the system, such as quick and repeated requests for help without an associated progress in learning.

AutoTutor currently does not allow the student to select questions/problems and to change topics. These acts of student initiative would no doubt provide excellent measures of SRL. It would not be difficult to allow students to select topics, questions, and problems in AutoTutor, but that would require a larger curriculum and repertoire of material than what is available. Students rarely take this initiative in human tutoring, so an open selection of learning objects would be an important advance (Bull & Kay, 2007).

(6) *Operation ARIES! (Acquiring Research Investigative and Evaluative Skills)*. This system helps students acquire scientific critical thinking by interacting in natural language with two animated pedagogical agents (Millis, Cai, Graesser, Halpern, & Wallace, 2009). One agent in ARIES, called the guide-agent, is an expert on scientific inquiry and serves as a knowledgeable tutor. The other agent is a fellow student, but could potentially take on other roles (e.g., a neighbor, another scientist, an evaluator of research). A 3-way conversation

transpires between the human student, the expert agent, and the student agent. The human students interact with both agents by holding mixed-initiated “trialogs” in natural language. The agents provide a large number of options and learning activities in this system, often under student control. For example, the agents give the students options on what texts to read, ask the students questions about the text, invite student questions, answer student questions, prompts students to critique experiments on methodology, and scaffolds learners on how to ask good questions in an interrogation module (Millis et al., 2009). ARIES is also implemented in a game environment to optimize engagement.

The trialogs are dynamically selected in a fashion that is sensitive to the student’s performance. When the student is performing poorly, the students will observe the two agents interacting with each other -- a form of vicarious learning (Gholson & Craig, 2006). When the student is performing extremely well, the trialog manager gets the human student to teach the student agent, with the tutor agent periodically stepping in – a form of teachable agent (Biswas et al., 2006). When the student’s performance is intermediate, there is a trialog based on AutoTutor. These alternative trialog architectures allow the learner to take on different roles and thereby scaffold the student to acquire strategies of SRL.

The complexity and diversity of ARIES provide more sensitive measures of SRL than does AutoTutor. There is an eBook with 22 chapters. Students can select what chapters to study and can decide whether to take a pre-test to test out of reading a chapter. Students can see which chapters they have completed and receive feedback on their performance on specific content areas. A metric of SRL can be computed that assesses the extent to which students linearly go through the eBook and read each chapter, as opposed to skipping over chapters they know or

taking a quick pretest to verify their mastery. The dialogue-based metrics of SRL can be computed in the same fashion as they are computed for AutoTutor.

Implications and Future Directions

This article has made the case that pedagogical agents have a promising future in the design of learning environments to promote SRL. Human tutors have limitations in promoting SRL because of unspectacular knowledge of pedagogy, SRL, and “meta” knowledge about cognition, communication, and emotions. Computerized agent-based systems can rectify some of these limitations in addition to simulating normal conversational patterns of humans (such as Expectation plus Misconception Tailored dialogue). It is conceivable that these agent-based systems have the best of both worlds. They can have human-like conversations, but can go a step further by being consistent, precise, complex, adaptive, and durable.

This article has identified a number of measures of SRL that the computer can record. For example, these include (a) the relative rates of student questions, help seeking, problem selection, and topic changes, (b) the proportion of tutor pumps and hints rather than prompts and assertions, (c) the proportion of expectations covered by students, and (d) patterns of tutor and student dialogue moves. It is easy to document the normalized frequency distributions of these signals of SRL. A promising next stage of research will conduct detailed data mining on sequences and combinations of various categories of states (cognition, discourse, emotions).

One direction for the future is to explore the role of emotions in these learning environments with pedagogical agents (D’Mello, Picard, & Graesser, 2007). One version of AutoTutor detects learner emotions automatically on the basis dialogue patterns, facial expressions, body posture, speech patterns, and a combination of these channels (D’Mello & Graesser, in press). For example, AutoTutor’s negative feedback and hints are diagnostic of

student frustration and confusion, whereas the student conversational turns that have high discourse cohesion are diagnostic of flow (i.e., extreme engagement). Facial actions associated with the eyes and mouth are very diagnostic of delight, surprise, and confusion, whereas the face is not a major channel for boredom, frustration, and flow. Body posture is a robust channel for the expression of boredom; AutoTutor's body pressure measurement system has revealed that bored students either fidget or have a large distance between their face and the screen. The computer automatically senses the students' emotions, but it is beyond the scope of this article to discuss these sensing and classification algorithms (see D'Mello & Graesser, in press).

Connections between complex learning and emotions have received increasing attention in psychology and education (Dweck, 2002; Pekrun, 2006; Meyer & Turner, 2006). It is important to understand affect-learning connections in order to design engaging educational artifacts that range from responsive intelligent tutoring systems like AutoTutor to entertaining media and games (Conati, 2002; McQuiggan & Lester, 2007). The technologies would accommodate the frequent affective states during tutoring of technical material, which have been identified in a number of studies as confusion, frustration, boredom, anxiety, and flow/engagement, with delight and surprise occurring less frequently (Baker, D'Mello, Rodrigo, & Graesser, 2010; D'Mello & Graesser, in press). Students would learn how to regulate their emotions in productive ways when they experience negative emotions such as boredom, anxiety, frustration, and confusion. AutoTutor could be helpful in scaffolding emotion regulation by generating encouraging remarks or high-quality hints when the students are faced with frustration, or by pointing out how important it is to put in more effort when the student is bored or confused.

Agent researchers have only begun to scratch the surface on their potential. Individual agents can have an endless number of dialogue styles, strategies, personalities, and physical features. For example, we recently developed an AutoTutor version that is emotionally supportive and another version that tries to shake up the emotions of the student by being rude and telling the student what emotion the student is having (which is very engaging to interact with for many students). Instead of giving earnest short feedback, the rude AutoTutor gives positive feedback that is sarcastic (e.g., “Aren’t you the little genius”) and negative feedback that is derogatory (e.g., “I thought you were bright, but I sure pegged you wrong”). This simple substitution of short feedback ended up dramatically changing AutoTutor’s personality. No doubt, some students would rather interact with the polite tutor than a rude tutor. The agents can be matched to the cognitive, personality, emotional, and social profiles of individual learners in a number of ways and possibly help students develop “meta” knowledge of emotions and social interaction.

It is important to acknowledge that agent-based systems have some noteworthy limitations in meeting the goals of promoting SRL. Most of these systems cover a narrow spectrum of the curriculum so they cannot accommodate any direction that students might pursue. The agents cannot interpret any change of topic that the student may want to pursue and cannot answer any question the student may ask. Students typically give up asking questions and trying to impose their own agenda very quickly when it is clear that student options are limited and their contributions are misunderstood. The hope is these challenges will be overcome when learning environments grow to the point of covering a large amount of the curriculum. This limitation may be minimized by giving the student free rein to the curriculum in an open learning forum, by giving the student choices on what to learn next, and by resisting the temptation to push an agenda.

Acknowledgements

The research on was supported by the National Science Foundation (SBR 9720314, REC 0106965, REC 0126265, ITR 0325428, REESE 0633918, ALT-0834847, DRK-12-0918409), the Institute of Education Sciences (R305G020018, R305H050169, R305B070349, R305A080589, R305A080594), and the Department of Defense Counter Intelligence Field Activity (H9C104-07-0014). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, IES, or DoD. Requests for reprints should be sent to Art Graesser, Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152-3230, a-graesser@memphis.edu.

References

- Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73*, 277-320.
- Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist, 40*, 193-198.
- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia. *Journal of Educational Psychology, 96*, 523-535.
- Azevedo, R., Witherspoon, A., Graesser, A.C., McNamara, D., Chauncey, A., Siler, E., Cai, Z., Rus, V., & Lintean, M. (2009). MetaTutor: Analyzing self-regulated learning in a tutoring system for biology. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Artificial Intelligence in Education* (pp. 635-637). Amsterdam: IOS Press.
- Baker, L. (1985). Differences in standards used by college students to evaluate their comprehension of expository prose. *Reading Research Quarterly, 20*, 298-313.
- Baker, R., Corbett, A., Roll, I., Koedinger, K. (2009). Developing a generalizable detector of when students game the system. *User Modeling and User-Adapted Interaction*.
- Baker, R.S., D'Mello, S.K., Rodrigo, M.T., & Graesser, A.C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*, 223-241.
- Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education, 15*, 95-115.

- Biswas, G., Leelawong, K., Schwartz, D., Vye, N. & The Teachable Agents Group at Vanderbilt (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, 19, 363-392.
- Bloom, B.S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive Domain*. New York: McKay.
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13, 4-16.
- Bull, S. & Kay, J. (2007). Student models that invite the learner in: The SMILI open learner modeling framework. *International Journal of Artificial Intelligence in Education* 17, 89-120.
- Cade, W., Copeland, J. Person, N., and D'Mello, S. K. (2008). Dialogue modes in expert tutoring. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the Ninth International Conference on Intelligent Tutoring Systems* (pp. 470-479). Berlin, Heidelberg: Springer-Verlag
- Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008) Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, 32(2), 301-341.
- Chi, M.T.H., Siler, S.A. & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition and Instruction*, 22, 363-387.
- Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471-533.
- Clifford, M. M. (1991). Risk taking: Theoretical, empirical, and educational considerations. *Educational Psychologist*, 26, 263-298.

- Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, 19, 237-248.
- Collins, A., & Halverson, R. (2009). *Rethinking education in the age of technology: The digital revolution and schooling in America*. New York: Teacher College Press.
- Conati C. (2002). Probabilistic assessment of user's emotions in educational games. *Journal of Applied Artificial Intelligence*, 16, 555-575.
- Craig, S. D., Sullins, J., Witherspoon, A. & Gholson, B. (2006). Deep-level reasoning questions effect: The role of dialog and deep-level reasoning questions during vicarious learning. *Cognition and Instruction*, 24(4), 563-589.
- D'Mello, S., & Graesser, A.C. (in press). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-adapted Interaction*.
- D'Mello, S.K., Picard, R., & Graesser, A.C. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, 22, 53-61.
- Dunlosky, J., & Lipko, A. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, 16, 228-232.
- Dweck, C. S. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 61-87). Orlando, FL: Academic Press.
- Gholson, B., & Craig, S.D. (2006). Promoting constructive activities that support learning during computer-based instruction. *Educational Psychology Review*, 18, 119-139.

Gholson, B., Witherspoon, A., Morgan, B., Brittingham, J. K., Coles, R., Graesser, A. C.,

Sullins, J., & Craig, S. D. (2009). Exploring the deep-level reasoning questions effect during vicarious learning among eighth to eleventh graders in the domains of computer literacy and Newtonian physics. *Instructional Science*, 37, 487-493.

Graesser, A. C., D'Mello, S.K., Cade, W. (in press). Instruction based on tutoring. In R.E. Mayer and P.A. Alexander (Eds.), *Handbook of Research on Learning and Instruction*. New York: Routledge Press.

Graesser, A. C., D'Mello, S., & Person, N. K. (2009). Metaknowledge in tutoring. In D. Hacker, J. Donlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education*. Mahwah, NJ: Taylor & Francis.

Graesser, A. C., Jeon, M., & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes*, 45, 298–322.

Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M.M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers*, 36, 180-193.

Graesser, A.C., & McNamara, D.S. (in press). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*.

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist*, 40, 225-234.

Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95, 524–536.

- Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, 31, 104-137.
- Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9, 1-28.
- Graesser, A. C., Penumatsa, P., Ventura, M., Cai, Z., & Hu, X. (2007). Using LSA in AutoTutor: Learning through mixed initiative dialogue in natural language. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 243–262). Mahwah, NJ: Erlbaum.
- Graesser, A. C., Wiley, J., Goldman, S. R., O'Reilly, T., Jeon, M., & McDaniel, B. (2007). SEEK Web tutor: Fostering a critical stance while exploring the causes of volcanic eruption. *Metacognition and Learning*, 2, 89–105.
- Greene, J. A., Moos, D. C., Azevedo, R., & Winters, F. I. (2008). Exploring differences between gifted and grade-level students' use of self-regulatory learning processes with hypermedia. *Computers & Education*, 50, 1069–1083.
- Halpern, D.F. (2002). *An introduction to critical thinking* (4th edition). Mahwah, NJ: Erlbaum.
- Jackson, G. T., & Graesser, A. C. (2006). Applications of human tutorial dialog in AutoTutor: An intelligent tutoring system. *Revista Signos*, 39, 31–48.
- Jackson, G.T., Guess, R.H., & McNamara, D.S. (2010). Assessing cognitively complex strategy use in an untrained domain. *Topics in Cognitive Science*, 2, 127-137.
- Klahr, D. (2002). *Exploring science: The cognition and development of discovery processes*. Cambridge, MA: MIT Press.
- Landauer, T., McNamara, D. S., Dennis, S., & Kintsch, W. (2007)(Eds.). *Handbook of Latent Semantic Analysis*. Mahwah, NJ: Erlbaum.

- Lepper, M. R., Drake, M., & O'Donnell-Johnson, T. M. (1997). Scaffolding techniques of expert human tutors. In K. Hogan & M. Pressley (Eds), *Scaffolding student learning: Instructional approaches and issues* (pp. 108-144). New York: Brookline Books.
- Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 135-158). Orlando, FL: Academic Press.
- Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117-144). Mahwah, NJ: Erlbaum.
- McNamara, D. S., Boonthum, C., Levinstein, I. B., & Millis, K. (2007). Evaluating self-explanations in iSTART: comparing word-based and LSA algorithms. In Landauer, T., D.S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis*. Mahwah, NJ: Erlbaum.
- McNamara, D. S., O'Reilly, T., Rowe, M., Boonthum, C., & Levinstein, I. B. (2007). iSTART: A web-based tutor that teaches self-explanation and metacognitive reading strategies. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies*. Mahwah, NJ: Erlbaum.
- McQuiggan, S., & Lester, J. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies*, 65, 348-360.
- Meyer, D. K., & Turner, J. C. (2006). Re-conceptualizing Emotion And Motivation To Learn In Classroom Contexts. *Educational Psychology Review*, 18 (4), 377-390.
- Millis, K., Cai, Z., Graesser, A., Halpern, D. & Wallace, P. (2009). Learning scientific inquiry by asking questions in an educational game. In T. Bastiaens et al. (Eds.), *Proceedings of World*

Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 2951-2956). Chesapeake, VA: AACE.

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology, 96*, 165-173.

Mosenthal, P. (1996). Understanding the strategies of document literacy and their conditions of use. *Journal of Educational Psychology, 88*, 314-332.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315-341.

Person, N. K., Kreuz, R. J., Zwaan, R., & Graesser, A. C. (1995). Pragmatics and pedagogy: Conversational rules and politeness strategies may inhibit effective tutoring. *Cognition and Instruction, 13*, 161-188.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). New York: Academic Press.

Resnick, L. B. (in press). Nested learning systems for the thinking curriculum. *Educational Researcher*.

Rogoff, B. & Gardner, W., (1984). Adult guidance of cognitive development. In: Rogoff, B. and Lave, J., Editors, 1984. *Everyday cognition: Its development in social context*, Harvard University Press, Cambridge, MA, pp. 95–116.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*, 181-221.

Schwartz, D. L. & Bransford, J.D. (1998). A time for telling. *Cognition & Instruction*, 16, 475-522.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31, 3-62.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (in press). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 153-189). Mahwah, NJ: Erlbaum.

Wisher, R. A., & Graesser, A. C. (2007). Question asking in advanced distributed learning environments. In S. M. Fiore & E. Salas (Eds.), *Toward a science of distributed learning and training* (pp. 209–234). Washington, DC: American Psychological Association.

Table 1. *Example AutoTutor-student dialogue with classified dialogue moves of AutoTutor, classified speech acts of the student, and comments.*

AutoTutor-1: *The sun exerts a gravitational force on the earth as the earth moves in its orbit around the sun. Does the earth pull equally on the sun? Explain why.* (MAIN QUESTION)

Student-2: *I don't know.* (A Metacognitive speech act that expresses what the student knows.)

AutoTutor-3: *Could you at least try to give me an answer? That way, we'll have something as a starting place to work from.* (PUMP. This dialogue move encourages the student to say more, but does not direct the student to express particular information.)

How does Newton's third law of motion apply to this situation? (HINT. This dialogue move leads the student to articulate a particular sentence, clause, or proposition.)

Student-4: *Does Newton's law apply to opposite forces?* (This is a YES/NO Verification Question; the student wants confirmation that the idea is correct.)

AutoTutor-5: *Yes, right!* (ANSWER. AutoTutor answers the student's question.)

Hum. (DISCOURSE MARKER. This connects different dialogue moves within a turn.)

There's still something missing from your explanation. (DISCOURSE MARKER.

AutoTutor lets the student know more is expected.)

Newton's Third Law refers to the forces exerted by one body on another _____?

(PROMPT. This dialogue move encourages the student to fill in a particular word or phrase.

Prompts are needed for students who have lower knowledge and verbal ability).

Student-6: *body* (Short Answer)

AutoTutor-7: *Bravo!* (POSITIVE FEEDBACK)

Let's move on. (DISCOURSE MARKER)

What can you say about the magnitude of the force that the earth exerts on the sun? (HINT)

Student-8: *What is the difference between a force and a vector?* (Student question.)

AutoTutor-9: *Force is any influence that tends to accelerate an object; a push or pull; measured in Newtons. Force is a vector quantity and vector is an arrow whose length represents the magnitude of a quantity and whose direction represents the direction of the quantity.*

(ANSWER. This answer is fetched from the textbook)

Okay, I bet you can get this! (DISCOURSE MARKER)

What is going on in this situation between the earth and sun? (HINT)

Student-10: *The sun is pulling on the earth because it is bigger.* (Student misconception).

AutoTutor-11: *No.* (NEGATIVE FEEDBACK) *the forces are equal in magnitude. Um, well, for every force there is a reaction force, such that the two forces constitute an action reaction force pair. The two forces have opposite direction and equal magnitude. No matter how different the masses of the two objects are, they will exert reaction forces on each other in the same magnitude.* (CORRECTION AutoTutor corrects the student's misconception.)

Do you understand? (COMPREHENSION GAUGING QUESTION)

Student-12: *Yeah. I see.* (Metacomprehension judgement)