CHAPTER 8

LEARNING WHILE HOLDING A CONVERSATION WITH A COMPUTER

Arthur C. Graesser, Natalie Person, Zhijun Lu, Moon Gee Jeon, and Bethany McDaniel
University of Memphis

ABSTRACT

Some of the recent electronic learning environments have moved beyond the conventional delivery of text, multimedia, and objective tests. There are systems with animated conversational agents, intelligent adaptive tutoring, interactive simulations, and other features designed to engage learners and promote deeper comprehension. One system is AutoTutor, a learning environment that tutors students by holding a conversation in natural language. AutoTutor's design was inspired by explanation-based constructivist theories of learning, intelligent tutoring systems that adaptively respond to student knowledge, and empirical research on dialogue patterns in tutorial discourse. AutoTutor presents challenging questions and then engages in mixed initiative dialogue that guides the student in building an answer. It provides feedback to the student on what the student types in (positive, neutral, negative feedback), pumps the student for more information, prompts the student to fill in missing words, gives hints, fills in missing information, identifies and

Technology-Based Education: Bringing Researchers and Practitioners Together, pages 143–167

corrects erroneous ideas, answers the student's questions, and summarizes answers. AutoTutor has produced significant learning gains in several experiments covering Newtonian physics and computer literacy. AutoTutor is currently being expanded to become more ubiquitous, more responsive to learner emotions, and compliant with electronic courseware standards.

There are many reasons for being optimistic about the revolution in online education and training. Courses are globally accessible in cities, rural areas, and underdeveloped countries. Learning experiences can be tailored to individuals. On-the-job access to training modules reduces the need to have corporate personnel travel for lengthy seminars. The optimistic projections are that there will be both savings in costs and training for any person, at any place, at any time (Wisher, Sabol, & Moses, 2002). These projections have motivated a dramatic reorganization of education and training facilities in schools, universities, the military, and virtually all government agencies (see Bruning, Horn, & PytlikZillig, 2003).

Nevertheless, the picture is not entirely optimistic in the arena of e-learning education and training. Much of the content has been unimaginative page-turning courseware (essentially books on the Web), devoid of deep pedagogical theoretical bases, interactive multimedia, or motivating environments (e.g., game-like environments). Very few of the intelligent tutoring systems and courseware with advanced interactive multimedia have solved some of the technical barriers that are needed before they can scale up to serve a large number of learners. The depth of interactions with real teachers has set the bar of learner expectations sufficiently high that many e-learning environments are perceived to be one-dimensional. Consequently, most e-learning courses have a high attrition rate, with learners giving up after one or a few sessions (Wisher et al., 2002).

This chapter addresses three limitations with existing courseware on the Internet. First, most humans prefer to communicate with others face to face, with appropriate pointing, gestures, speech intonation, and emotional responses. Although there are a small number of systems with animated conversational agents that emulate face-to-face interaction (Graesser, VanLehn, Rose, Jordan, & Harter, 2001; Gratch et al., 2002; Johnson, 2001), this form of communication is rarely implemented in Web and Internet courseware. Second, humans want immediate access to relevant, accurate information, particularly those in the "Google generation" who do not have the patience to sift through unreasonable numbers of irrelevant Web pages. There has been some progress in combining learning with information retrieval, where the learner can ask any question that comes to mind and receive informative relevant answers (Graesser, Hu, Person, Jackson, & Toth, 2004); however, these advances are rarely implemented in e-learning courseware. Third, instructors often desire more sophisticated learning environments than page-turning software with multimedia capabilities. Deeper learning of material is most likely to be achieved by intelligent tutoring systems that track the knowledge and misconceptions of students and that adaptively respond to these deficits at a fine-grained level. Intelligent tutoring systems have effectively achieved these pedagogical goals (Aleven & Koedinger, 2002; Anderson, Corbett, Koedinger, & Pelletier, 1995; Lesgold, Lajoie, Bunzo, & Eggan, 1992; Sleeman & Brown, 1982; VanLehn et al., 2002), but such systems are rarely on the Internet or Web.

The purpose of this chapter is to describe features of learning environments that hold promise in overcoming some of the disappointments with existing e-learning technologies. We describe some example systems that embody these features. One system that will be covered in some depth is AutoTutor, a system developed at the Institute for Intelligent Systems at the University of Memphis. AutoTutor is a learning environment that tutors students by holding a conversation in natural language. Whether these newer environments will be used and will promote learning gains remains an open question. Early evaluations have been very encouraging, but more empirical work is needed in field settings. Finally, we describe some future horizons of e-learning environments with conversational tutors.

FEATURES OF INTELLIGENT LEARNING ENVIRONMENTS THAT ENCOURAGE DEEPER LEARNING

As argued above, one approach to improving electronic learning environments is to develop courseware that promotes deeper learning. A contrast is frequently made between shallow and deep knowledge (Bransford, Brown, & Cocking, 2000; Chi, de Leeuw, Chiu, & LaVancher, 1994; Graesser & Person, 1994; Snow, 2002). Shallow knowledge includes lists of concepts, attributes of concepts, facts, rules, and procedures that are not tightly connected by an underlying conceptual system. Deep knowledge taps causal and functional explanations, logical justification of claims, hypothetical reasoning, and complex systems that are organized coherently. As an example, a student who memorizes the parts of the human body and their locations is acquiring shallow knowledge. Deep knowledge is acquired when a student learns how to explain the mechanisms of the circulatory system and how blood pressure is affected by the system.

Most of the learning environments that target deeper learning are guided by two core theoretical frameworks: Constructivism and/or inquiry learning. *Constructivists* view learners as actively constructing knowledge as opposed to passively registering information. Advocates of *inquiry learning* augment constructivism with inquiry. Learners are encouraged to ask questions, formulate hypotheses, plan tests of hypotheses, collect and analyze data, explain results, and communicate findings to peers (much like mini-scientists implementing the scientific method). Advanced tutoring environments attempt to implement principles underlying constructivism and inquiry learning, as discussed below.

Constructivism

Constructivism has had a historically strong foundation among learning theorists who take cognitive, social, and developmental approaches to education (Bransford et al., 2000; Chi et al., 1994, Piaget, 1952; Vygotsky, 1978). Constructivist theories are hardly monolithic in specific details about educational philosophies and practical recommendations, but all of them assume that learners actively construct knowledge representations while comprehending material or solving problems. Constructivism is the antithesis to theories that assume that learners learn by passively receiving material presented by information delivery systems. Furthermore, constructivist approaches have been so compelling that they have shaped major standards for curriculum and instruction in the United States during the last decade.

One challenge facing the designers of Web-based learning environments is to find ways to scale up constructive approaches to become part of widely used computer technologies. A major inspiration behind the design of AutoTutor was to make constructivist learning environments more pervasive. As we will elaborate later in this chapter, AutoTutor provides a form of student-centered learning that is guided by wise tutors, as opposed to mere information delivery systems or an unsupervised facility for Web surfing. AutoTutor also offers the learner formative assessment and feedback on mastery of the material, as opposed to unguided absence of testing or massive testing for summative evaluation. Finally, AutoTutor helps students build deep coherent explanations of the subject matter, as opposed to fragmentary shallow knowledge.

Inquiry Learning

As previously noted, constructivism is often augmented with inquiry learning methods. Educational researchers and teachers would unanimously agree that learning environments should encourage student questions. Those who have allegedly used Socratic teaching styles, for example, have attempted to guide their students through a process of interrogation in order to help them identify the limits of their own knowledge, to discover misconceptions, and to achieve genuine self-knowledge. Researchers in cognitive science and education have reported learning benefits for environments that encourage students to generate questions (Beck, McKeown, Hamilton, & Kucan, 1997; Dillon, 1988; King, 1994; Miyake & Norman, 1979; Pressley & Forest-Pressley, 1985). Question generation is one salient manifestation of active learning and reveals how deeply the learner has mastered the material (Graesser & Olde, 2003; Otero & Graesser, 2001; Scardamalia & Bereiter, 1985; Wisher & Graesser, in press). It is well documented that learning improves when learners are taught how to ask good questions, either through direct instruction on question asking (King, 1994; Rosenshine, Meister, & Chapman, 1996) or through the modeling of good question-asking skills by a person (Palincsar & Brown, 1984) or computer (Driscoll, Craig, Gholson, Ventura, Hu, & Graesser, 2003). Sophisticated learning environments should stimulate learner questions and facilitate the process of receiving answers to the questions that learners ask.

It is important to recognize that the asking of good questions is an acquired skill that does not come naturally to most students. The rate of student questions in classrooms has been estimated at 0.1 question per student per hour, whereas the rate in human-to-human tutoring is approximately 27 questions per hour (Graesser & Person, 1994). The upper bound rate of 125 questions per hour occurs in a computerized learning environment that forces students to learn exclusively by asking questions and reading answers (Graesser, Langston, & Baggett, 1993). Moreover, student questions are shallow and unsophisticated unless they are trained how to ask good questions (Wisher & Graesser, in press). The following are characteristics of learning environments that stimulate students' asking of good questions:

1. Learning environments stimulate questions when they place the student in cognitive disequilibrium, as in the case of challenges of entrenched beliefs, obstacles to goals, contradictions, anomalous events, deviations from norms, breakdown scenarios, salient contrasts, and decisions in the face of equally attractive options (Dillon, 1988; Graesser & McMahen, 1993; Graesser & Olde, 2003). Computational models have specified the particular questions that learners ask when confronted with different forms of cognitive disequilibrium, as in the case of SWALE (a question asking system named after a race horse, Kass, 1992; Schank, 1999) and PREG (a named derived from the Spanish morpheme for question, Otero & Graesser, 2001).
2. Learning environments stimulate questions when they didactically train or model the asking of questions (King, 1994). Modeling good question asking can be accomplished by expert human models or peers (Palincsar & Brown, 1984), animated conversational agents on

computers (Craig et al., 2002), or through a set of questions presented on a question menu (Graesser et al., 1993).

3. Inquiry-based learning environments stimulate student questions when they encourage hypothesis testing and experimentation for achieving long-term objectives on authentic science problems (Linn & Hsi, 2000; White & Frederiksen, 1998).

The hope is that student question-asking skills and the resulting learning will radically improve when students are immersed in learning environments that encourage question asking, that scaffold good question-asking skills, and that provide good answers that are tailored to the learner. As will be detailed in a later section, AutoTutor is a learning technology designed to meet these objectives.

Constructivism and inquiry learning have traditionally been two dominant theoretical frameworks for learning environments that encourage deeper learning. One of the chief difficulties, however, is how to get the learner to construct the knowledge, ask the questions, and seek the answers. There needs to be some form of scaffolding to guide the student in construction and inquiry. This scaffolding can be provided by a human tutor or by an animated conversational agent, as discussed in the remainder of this section.

Tutoring

One-to-one tutoring is a powerful method of promoting knowledge construction. There is substantial empirical evidence that human tutoring is extremely effective when compared to typical classroom environments (Bloom, 1984; Cohen, Kulik, & Kulik, 1982; Corbett, 2001). Cohen et al. (1982) performed a meta-analysis on a large sample of studies that compared human-to-human tutoring with classroom controls. The vast majority of the tutors in these studies were untrained in tutoring skills and had moderate domain knowledge; they were peer tutors, cross-age tutors, or paraprofessionals, but rarely accomplished professionals. These "unaccomplished" human tutors enhanced learning with an effect size of .4 sigma (i.e., .4 standard deviation units). Bloom (1984) reported that accomplished human tutors produce an even greater effect size of 2 sigma (2.0 standard deviations) for mathematics skill training. However, the reliability of this effect has been questioned due the relatively small number of studies on expert human tutors. Only two studies with this effect size were reported by Bloom and more recent studies in the tutoring literature are conspicuously absent. VanLehn, Graesser, Jackson, Jordon, Olney, and Rose (2004) reported only a 1-sigma learning gain for accomplished physics tutors who tutored college students through computer-mediated conversation. These tutors, who had doctoral degrees and several years of training experience, produced learning gains that were equivalent to the gains of computer tutors (including AutoTutor). Even if accomplished tutors do yield large learning gains (an issue still unsettled empirically), they are truly a rare resource (Cohen et al., 1982). This reduces the hope that accomplished human tutors offer a practical solution to the substantial gaps in education and training. Advanced computer systems may offer a more viable solution.

During the last 25 years, intelligent tutoring systems (ITSs) have implemented systematic strategies for promoting learning at deeper levels. ITSs have implemented a number of sophisticated pedagogical strategies, such as error identification and correction, building on prerequisites, frontier learning (i.e., expanding on what the learner already knows), student modeling (i.e., using inferences about student knowledge to guide tutoring), and building coherent explanations (Aleven & Koedinger, 2002; Anderson et al., 1995; Lesgold et al., 1992; Sleeman & Brown, 1982; VanLehn et al., 2002). The ITSs that have been successfully implemented and tested (such as VanLehn's Andes physics tutor and Anderson et al.'s Cognitive Tutor) have produced learning gains of approximately 1.0 sigma. It appears that the learning gains associated with sophisticated ITSs (1.0 sigma) are higher than those of unaccomplished human tutors (0.4 sigma), but perhaps not quite as good as the accomplished human tutors (1.0 to 2.0 sigma). As will be discussed later, the performance of AutoTutor is somewhere between an unaccomplished human tutor and an intelligent tutoring system.

Animated Pedagogical Agents

When tutoring, ITSs sometimes use pedagogical agents, or characters that occupy computer learning environments and facilitate learning by interacting with students or other agents. The medium of the agent's message delivery varies, especially given current technological advances. Early systems communicated primarily via printed text, whereas recent systems communicate in a variety of media. Pedagogical agents have been designed to exhibit a range of behaviors, functions, and cognitive capabilities. For example, they have been designed to generate multiple pedagogical strategies, assist instructors and students in virtual worlds, serve as knowledgeable navigational mentors or guides, reason about multiple agents in simulated environments, and act as a peer, co-learner, or competitor (Chan, 1996; Johnson, Rickel, & Lester, 2000).

Animated pedagogical agents, or what Johnson (2001) calls Guidebots, have recently appeared in learning environments and Websites. Animated pedagogical agents are lifelike personas that execute behaviors, emotionally respond with facial expressions, and communicate in natural language. They have the potential to bolster student learning outcomes by implementing reasonable pedagogical strategies and by exploiting both the auditory and visual channels of the learner. The advent of animated pedagogical agents is the result of recent advancements in multimedia interfaces, text-to-speech software, and agent-generation technologies.

Available evidence suggests that students prefer learning environments with animated agents over those that do not have agents. Experimental participants assigned randomly to learning conditions with animated agents (even ones that are not particularly expressive) perceive their learning experiences to be considerably more positive than those assigned to learning conditions that do not include animated agents. This recurring finding is known as the *persona effect* (Baylor, 2002; Reeves & Nass, 1996). The persona effect is somewhat enigmatic in that it frequently is not related to learning gains and other performance measures. That is, some researchers who have reported evidence for the persona effect also report little or no differences between agent and no-agent conditions for retention and learning measures (Graesser, Moreno et al., 2003; Moreno, this volume; Moreno, Mayer, Spires, & Lester, 2001).

To be fair, however, there are sporadic findings that animated pedagogical agents *do* sometimes promote learning on retention and/or transfer tasks. Atkinson (2002) reported that students who received explanations from an animated agent about how to solve proportion word problems outperformed other learning conditions on both near and far transfer problems. In a study conducted by R. Moreno et al. (2001), college students and seventh graders attempted to learn about how to design plants that could survive in a number of different environments. One group of students interacted with a pedagogical agent, Herman the Bug, while another group of students received identical graphics and textual explanations but no pedagogical agent. The results indicated that students in the pedagogical agent condition outperformed students in the no agent condition on transfer tests that tap deep knowledge, but not on retention tests. In our lab, Graesser, K. Moreno, et al. (2003) reported a significant but modest advantage of animated agents over printed messages. Given the results of these learning outcome studies and the fact that learners perceive their interactions with agents quite favorably, the future for pedagogical agents looks promising. However, researchers need to pin down the precise conditions and agent features that yield positive learning gains. When learning gains are modest or nonexistent, the best that can be hoped for is that the agents might improve students' motivation to learn. In that event, researchers would need to explore pedagogical techniques to improve motivation (e.g., perhaps teaming up with Disney or Hollywood).

Another successful example of pedagogical agents used to promote learning and transfer is the system developed by McNamara and her colleagues (McNamara, Levinstein, & Boonthum, 2004). They developed a system with an *ensembl*e of conversational agents that help students improve the depth of their reading skills. Tutor agents and peer agents exhibit reading skills at deeper levels and talk about the cognitive processes. Their system, iSTART (*Interactive Strategy Trainer for Active Reading and Thinking*), provides the learner with automated training with multiple agents who scaffold active reading strategies while explaining the content of difficult text. The iSTART system is based on a modified version of the learning strategy that trainers have implemented successfully in classroom settings (McNamara, 2004). Students learn to use techniques such as predicting, elaborating, paraphrasing, comprehension monitoring, forming bridging inferences, and logic. McNamara's research has shown that students with a better understanding of metacognitive reading strategies outperform less strategic learners on course exams, even several months after the training was delivered (McNamara, 2004).

How do the multiple agents provide training in iSTART? The ensemble of agents model and guide the learning process via their interactions with each other and with the learner during three phases of learning. In the first phase, the learner is provided with instruction on effective reading strategies. Whereas the human-delivered training uses a lecture format, iSTART uses interactions between an expert and two student agents that mimic a constructive, collaborative learning process. In the second phase, two agents demonstrate the use of different reading strategies and the learner identifies the strategies being used by the agent. The amount of support (e.g., hints) provided to the learner is adapted to learner's performance. In the third phase, the learner practices the strategies by reading texts and typing content. During this phase, the agent's interactions with the learner are moderated by the quality of what they type in. The iSTART system uses an ensemble of agents with different roles to implement modeling-scaffolding-fading, one of the popular techniques in apprenticeship learning (Collins, Brown, & Newman, 1989).

AUTOTUTOR: AN ANIMATED CONVERSATIONAL AGENT WITH TUTORIAL DIALOG

AutoTutor is an example of an advanced learning system that incorporates principles of constructivism and inquiry learning, and that uses animated agents to assist learners. This section discusses the features, functions, and performance of AutoTutor (Graesser, Lu et al., 2004; Graesser, Person, & Harter, 2001; Graesser, Wiemer-Hastings, Wiemer-Hastings, Kreuz, & Harter, 1999). AutoTutor is a fully automated computer tutor on the Internet that holds conversations with students in natural language and that simulates the discourse patterns of human tutors and some ideal tutoring strategies. AutoTutor speaks by utilizing a speech engine developed at Microsoft (www.microsoft.com/products/msagent) or SpeechWorks (www.speechworks.com). For some topics and versions of AutoTutor, there are graphical displays, animations of causal mechanisms, or interactive simulation environments, with AutoTutor talking about and pointing to various components. AutoTutor was designed to be a good conversational partner that comprehends, speaks, points, and displays emotions, all in a coordinated fashion. The initial versions of AutoTutor (versions 1.0, 1.1 and 2.0) were on the topic of computer literacy. A later version of AutoTutor, called "Why/AutoTutor," was designed to help college students learn Newtonian physics (Graesser, Jackson et al., 2003; Graesser, VanLehn et al., 2001) by asking them why-questions on difficult problems. One strength of AutoTutor is that the tutoring domain can be changed quickly with lesson authoring tools, without the need to rebuild any of the conversational components of the system.

Figure 8.1 shows a screen shot of AutoTutor on the topic of computer literacy. At the top window is the main question that requires deep reasoning to answer: "How is the packet switching model of message transmission like the postal system?" This question requires deep analytical reasoning and approximately seven sentences of information in an ideal answer. It

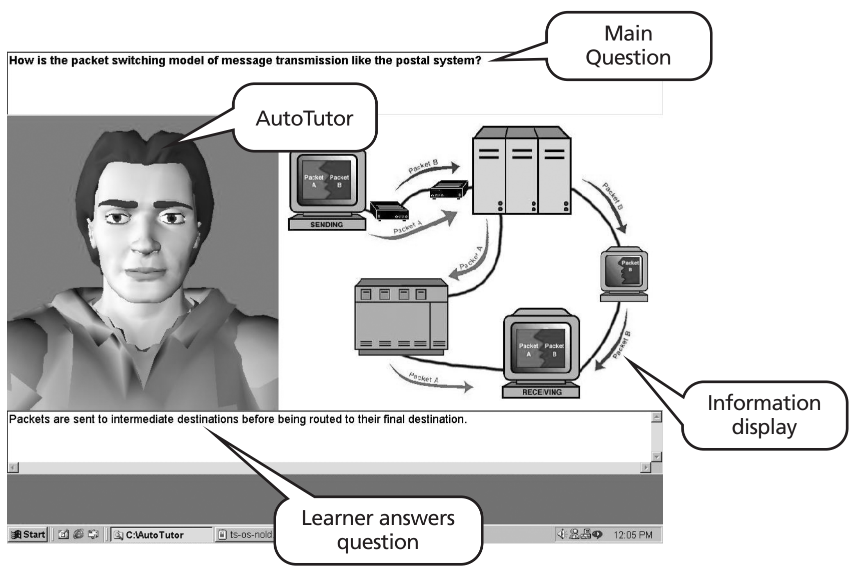

Figure 8.1. A computer screen of AutoTutor for the subject matter of introductory computer literacy.

typically takes 50–200 conversational turns in the dialogue to answer this main question. The student types in the answer content in the window at the bottom of the screen. The conversational agent at the left speaks the content of AutoTutor, with appropriate facial expressions and occasional gestures. This main question has an associated diagram for the learner to comprehend and that the two speech participants can reference during their conversation.

Dialog and Constructivism

AutoTutor produces several categories of *dialog moves* that facilitate information coverage anticipated by AutoTutor's *curriculum script* (i.e., the content of the subject matter). The dialog moves are fashioned to be sensitive to student input. AutoTutor provides *feedback* to the student (positive, neutral, negative feedback), *pumps* the student for more information ("What else?"), *prompts* the student to fill in missing words, gives *hints* to elicit lengthier idea units, fills in missing information with *assertions*, identifies and *corrects* bad answers, *answers* students' questions, and *summarizes* answers. AutoTutor also integrates student responses over time, allowing the student to refine or expand the answer. As the student expresses information over many turns, the information in the three to seven sentences is eventually covered and the question is answered. During the process of supplying the ideal answer, the student periodically articulates misconceptions and false assertions. If these misconceptions have been anticipated in advance and incorporated into the program, AutoTutor provides the student with information to correct the misconceptions. Therefore, as the student expresses information over the turns, this information is compared to anticipated correct information (called *expectations*) and incorrect information (called *misconceptions*). We refer to this tutoring mechanism as *expectation and misconception tailored dialog* (EMT dialog).

It is important to emphasize that the tutorial dialog patterns of AutoTutor were motivated by research in discourse processing and cognition. This design of AutoTutor was inspired by explanation-based constructivist theories of learning, by cognitive tutors that adaptively respond to student knowledge (e.g., Anderson et al., 1995), and by previous empirical research that has documented the collaborative constructive activities that routinely occur during human tutoring (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; Fox, 1993; Graesser & Person, 1994; Graesser, Person, & Magliano, 1995).

Somewhat surprising, most human tutors (as well as the EMT dialog moves displayed by AutoTutor) are not particularly sophisticated from the standpoint of ideal tutoring strategies that have been proposed in the fields of education and artificial intelligence (Graesser et al., 1995). Graesser and colleagues videotaped over 100 hours of naturalistic tutoring, transcribed the data, classified the speech act utterances into discourse categories, and analyzed the rate of particular discourse patterns. These analyses revealed that human tutors rarely implement intelligent pedagogical techniques such as bona fide Socratic tutoring strategies, modeling-scaffolding-fading, reciprocal teaching, frontier learning, building on prerequisites, or diagnosis/remediation of deep misconceptions (for more details on these methods, see Collins et al., 1989; Palincsar & Brown, 1984; Sleeman & Brown, 1982). Most tutors tend to coach students in constructing explanations, as captured in AutoTutor's EMT dialog patterns.

Interactive Simulations with Dialog

Interactive simulation is believed to be a powerful technology in the educational process, although the impact on learning remains an open empirical question. Before most simulations can be used, many students need training and scaffolding to use the simulation environments effectively. Otherwise, they remain confused about what to do next, or they meander unproductively through the space of alternative components that can be manipulated. However, the efficacy of simulation-based learning when coupled with dialog is uncharted territory.

We recently have developed a version AutoTutor (called "AutoTutor 3D") that grounds physics problems in a microworld and allows the student to manipulate aspects of the microworld to see what happens. That is, AutoTutor provides the student with the opportunity to see real-world physics through realistic animation and allows him/her to see what happens when altering parameters (e.g., an entity's mass, speed, acceleration, etc.) The goal is to develop a qualitative, workable, and verbalizable understanding of the physics involved. At the same time, the simulations are used to identify and correct any misconceptions that interfere with the analysis of real situations. The repertoire of dialog moves available to the tutor is expanded because the tutor can decide to show an animation, to invite the student to alter the parameters of the situation being modeled, or to direct the student to a particular choice of parameters in order to make a point or confront a misconception. The AutoTutor simulation environments are delivered on the Internet using a myriad of graphics utilities (3D Studio Max, Adobe Photoshop, Macromedia's Director and Shockwave).

Unlike most interactive simulation environments, however, we believe that learning is not optimized by merely having learners manipulate the physical parameters and observe what happens. We also believe it is important for them to articulate what they see and to use this knowledge to construct a coherent explanation that answers the main question. That is, interactive simulation is visualization and manipulation in service of articulation and explanation.

Answering Student Questions

AutoTutor handles a broad range of student questions and answers most questions by extracting the answers from electronic textbooks. In this sense, AutoTutor has successfully accommodated inquiry learning and mixed initiative dialog. The question answering system in AutoTutor produces answers to domain questions that are not hand-crafted by lesson planners who generate the curriculum script. That is, the answers are composed by first interpreting the question and then fetching a paragraph from the electronic textbook that includes an answer to the question. The question answering system classifies questions into 16 categories and can reasonably answer questions in most of these question categories, including definition, comparison, and deep comprehension questions (e.g., why, how, and what-if questions).

We have investigated the fidelity of the question answering system in providing relevant, good answers to students' questions as they learn. Learners pose questions during learning and then rate how relevant or informative the information is in the paragraph that gets returned. Available evaluations have revealed that over 90% of the answers are relevant and over 50% are rated as informative (Graesser, Hu et al., in press). This is substantially higher than randomly selected paragraphs in the subject matter, which yield percentages that vary from 0 to 8%. These systems are a very promising beginning in building conversational environments to support inquiry learning. In the future, we hope to improve the question answering facility by incorporating advances in computational and corpus linguistics (Jurafsky & Martin, 2000).

Data Bases and Linguistic Modules

The software residing on the AutoTutor server has a set of permanent databases and linguistic modules that do not get updated throughout the course of tutoring. These include (a) the curriculum script repository consisting of questions, answers, and associated dialog moves, (b) lexicons, syntactic parsers, speech act classifiers, and other computational linguistics modules, (c) a question answering facility, (d) a corpus of documents, including a textbook on conceptual physics, and (e) latent semantic analysis (LSA) vectors for words, curriculum content, and the document corpus. One advantage of AutoTutor over some of the existing ITSs with tutorial dialog is the use of LSA as a primary method of representing world knowledge (Graesser, Hu, & McNamara, 2005). LSA is a high-dimensional, statistical technique that, among other things, measures the conceptual similarity of any two pieces of text, such as a word, sentence, paragraph, or lengthier document (W. Kintsch, 1998; Landauer, Foltz, & Laham, 1998). LSA-based technology is currently being used in essay graders that grade essays as reliably as experts in English composition (Foltz, Gilliam, & Kendall, 2000) and in the Summary Street software that teaches learners how to summarize text (E. Kintsch, Steinhart, Stahl, & LSA Research Group, 2000).

We use LSA in AutoTutor to perform conceptual pattern matching operations when we compare student contributions to expectations and misconceptions. More specifically, each conversational contribution of the learner is compared with each expectation and misconception. LSA-based metrics vary from 0 to 1 when assessing how well the students' verbal input matches a particular expectation or misconception. If the LSA value is high enough, then the expectation or misconception is considered covered. In this fashion, the learner's profile on subject matter knowledge is tracked in a very detailed fashion.

Dynamic Processing Modules

AutoTutor also has a set of processing modules and dynamic storage units that maintain qualitative content and quantitative parameters. Unlike the previous modules, these processing modules and storage units are frequently updated as the tutoring process proceeds and are used to update values in the learner profile. For example, AutoTutor keeps track of student ability (through LSA evaluations of student assertions), student initiative (assessed as the incidence of student questions), student verbosity (number of words per turn), and the evolution of a question's answer in the dialog history. AutoTutor's dialog management module flexibly adapts to the student by virtue of these parameters, so no two conversations with AutoTutor are ever the same. Though it is beyond the scope of this chapter to detail AutoTutor's mechanisms, readers who are familiar with intelligent tutoring systems and natural language processing technologies may find it noteworthy that AutoTutor's dialog management module has an augmented finite state network, a set of fuzzy production rules, and a planning algorithm for selecting dialog moves to help fill in missing information to create an ideal answer. Other processing modules execute additional important functions, such as speech act classification, linguistic information extraction, evaluation of student assertions, and speech production by the animated agent.

Evaluations of AutoTutor

The performance of AutoTutor has been evaluated on a number of criteria. We have evaluated AutoTutor on learning gains in several experiments involving nearly 1000 students in computer literacy courses and physics courses. There were significant learning gains in all of these experiments, particularly at the level of deep explanations as opposed to shallow facts and ideas (Graesser, Jackson et al., 2003; Graesser, Lu et al., 2004; VanLehn et al., 2004). AutoTutor has produced learning gains of .4 to 1.5 sigma (a mean of .8), depending on the learning measure, comparison condition, subject matter, and version of AutoTutor. As previously noted, these tests place previous versions of AutoTutor somewhere between an unaccomplished human tutor and an intelligent tutoring system. It is important to point out that traditional ITSs normally take much longer to develop than does AutoTutor. The development time on AutoTutor for new subject matter is only a few months, whereas it would take years to develop a tutor on a new topic that uses the computational architecture of Andes or the Cognitive Tutor.

AutoTutor's LSA component successfully evaluates the quality of learner contributions in natural language. Its accuracy is on par with graduate-level research assistants (r =.5 to .7 when comparing LSA to expert judgments), but not quite as good as accomplished experts (Graesser, Hu, & McNamara, in press). In these evaluations, graduate students or experts rate the extent to which student essays express particular sentence-like expectations. Similarly, the LSA component evaluates the extent to which the expectations are covered. The correlations between computer and human are significant and impressive, ranging from .5 to .7, when computing the proportion of expectations covered in an essay. Thus, AutoTutor does a reasonable job tracking the knowledge of students at a fine-grained level.

AutoTutor is remarkably accurate (95% correct and d' of 3.7) in classifying student contributions into 19 different categories of speech acts, such as assertions, metacognitive comments ("I'm lost"), metacommunicative comments ("What did you say?"), and 16 question categories (Olney et al., 2003). Students can take initiative by asking questions, which are classified and directed to a question answering module.

Expert judges have evaluated AutoTutor with respect to conversational smoothness and the pedagogical quality of its dialog moves (Person, Graesser, Bautista, & Mathews, 2001). The experts' mean ratings are positive (i.e., smooth rather than awkward, good rather than bad pedagogical quality), but there is room to improve in the naturalness and pedagogical effectiveness of its dialog. Person and Graesser (2002) performed a *bystander Turing test* on the naturalness of AutoTutor's dialog moves. A total of 144 randomly selected tutor moves in the tutorial dialogs between students and AutoTutor were selected. Six human tutors (from the tutor pool on computer literacy at the University of Memphis) were asked to fill in what they would say at those 144 points. So, at each of the 144 tutor turns, the corpus contained what the human tutors generated and what AutoTutor generated. A group of computer literacy students was asked to discriminate between dialog moves generated by a human versus a computer; half in fact were by human and half were by computer. It was found that the bystander students were unable to discriminate whether particular dialog moves had been generated by a computer versus a human; the *d'* discrimination scores approached zero. This rather impressive outcome supports the claim that AutoTutor is a good simulation of human tutors.

Authoring Tools for Content Development

One of the chief challenges in the building of intelligent learning environments is the development of structured content repositories for use during learning sessions. To facilitate the development of new lessons and topics, there needs to be authoring tools that bridge the gap between subject matter experts and technology-driven systems. Authoring tools are crucial for the future scalability of any widely used ITS or computer-based training system. A perfect tutoring system is very limited if it can only provide instruction for a few topics and/or there are only a handful of people capable of programming it to work with new domains and problems. Unfortunately, there are a limited number of authoring tools that enable nearly effortless content generation (Murray, Blessing, & Ainsworth, 2003; Susarla et al., 2003). One such system, called REDEEM, has a graphical user interface that allows teachers with little technical knowledge to mold course material and teaching strategies to an individual student's needs. Teachers can be trained to use REDEEM in approximately 90 minutes and create ITSs from computer-based teaching materials at a rate of about three hours per hour of instruction.

The e-learning industry needs to establish standards with respect to the structure of content packages and the learning management systems (LMS). The industry has essentially been producing target delivery software with authoring tools, content, and LMSs that are idiosyncratic to particular software providers. Consequently, a major problem for computer-based education is the lack of software reusability–a tremendous waste of resources. The most successful solution to this problem has been the Sharable Content Object Reference Model (SCORM) that was introduced by the Advanced Distributed Learning (ADL) initiative of the Department of Defense (http://www.adlnet.org). SCORM requires metadata for learning objects, called Sharable Content Objects (SCOs). With such metadata, each SCO can be shared by any delivery software and LMS. The curriculum scripts of AutoTutor were therefore developed to be compliant with ADL/ SCORM standards.

The next generation of ITSs will not only allow individuals to manage and create content for the system, but also to manipulate and change the mechanisms which the ITS uses in interacting with a learner. Many ITSs solely adhere to static, unchangeable pedagogical strategies, which has the potential liability of alienating teachers who wish to remain actively involved in the learning process. Building upon previous research and development pertaining to dialog planning systems (Rich & Sidner, 1998; Larsson & Traum, 2000), we have created a dialog management authoring tool that allows users to alter the conversational and pedagogical properties of AutoTutor. The authoring tool allows an expert to change the rules and mechanisms of AutoTutor for different domains and learning purposes.

FUTURE HORIZONS

The previous section described the versions of AutoTutor that have been developed and tested at this point in our long-term project. At this point we shift to some new features that will be incorporated into versions of AutoTutor currently under development. Towards the end of this section, we will identify some limitations of AutoTutor as a learning technology.

Responsiveness to Learner Emotions

We are developing a version of AutoTutor that perceives and responds to learner emotions in addition to the learner's knowledge states. AutoTutor will have sensing devices and signal processing algorithms that classify affective states of learners. Emotions will be classified on the basis of dialog patterns during tutoring, the content covered, facial expressions, body posture, mouse haptic pressure, and keyboard pressure. The first phase of the project will evaluate how accurately AutoTutor classifies learner affect on the basis of these channels. The affect states relevant to learning that we have in mind are confusion, frustration, boredom, interest, excitement, and insight (eureka). The second phase will be to modify AutoTutor's planning of dialogue moves to be sensitive to learner emotions. We have a number of hypotheses about relationships between emotions and learning that optimize motivation and/or learning. If the learner is extremely frustrated, it might be important to generate a prompt or hint, so that the learner gets back on a productive track of learning. If the learner is bored, the presentation of an engaging problem may motivate the learner. If the learner is engaged or experiencing eureka, it may be best for AutoTutor to stay out of the learner's way or to celebrate with the student. We will evaluate whether learning is enhanced by an AutoTutor that is adaptive to the learner's emotions.

There is already some empirical evidence that emotions might be intimately interwoven with complex learning (Craig, Graesser, Sullins, & Gholson, in press). Craig et al. recently conducted an experiment in which six different affect states (frustration, boredom, flow, confusion, eureka, and neutral) were observed during the process of learning introductory computer literacy with AutoTutor. Learner emotions were observed by expert judges at random points during a session with AutoTutor. These observational analyses revealed significant relationships between learning gains (posttest-pretest scores on multiple choice tests) and the affective states of boredom ($r = -.39$), flow ($r = .29$), and confusion ($r = .33$). Correlations with eureka ($r = .03$), and frustration ($r = -.06$) were near zero. The positive correlation between confusion and learning is perhaps counterintuitive, but is actually consistent with a model that assumes that *cognitive disequilibrium* is one precursor to inquiry and deep learning (Graesser & Olde, 2003; Otero & Graesser, 2001). Cognitive disequilibrium occurs when the learner experiences contradictions, discrepancies, novel input, obstacles to goals, decision deadlocks, and major knowledge gaps. The findings that learning correlates negatively with boredom and positively with flow are consistent with predictions from Csikszentmihalyi's (1990) analysis of *flow* experiences. Conscious flow occurs when the student is so absorbed in the material that time disappears, fatigue disappears, and extraneous interruptions get unnoticed.

We have assembled and installed most of the emotion sensing technologies with AutoTutor. We are in the process of having humans or the computer analyze: (1) the AutoTutor log file with speech acts of student and tutor turns, as well as knowledge states achieved from the tutorial dialog, (2) the body posture pressure measurement system purchased from Tekscan, (3) the upper facial sensor device developed by Roz Picard's Affective Computing Lab at MIT (Picard, 2000), (4) a haptic pressure sensor for the mouse (supplied by MIT), and (5) a keyboard pressure sensor purchased from Tekscan. Affect states will be interpreted and/or classified on the basis of these five input channels together with computational models. We anticipate that most of these sensing technologies will be integrated with the learners' workstations of the future.

ePAL: An Electronic Personal Advisor for Learning

We are convinced that intelligent agent technology has advanced sufficiently to use as the core architecture for designing a 24/7 e-learning system. In essence, AutoTutor would always be available as a conversation partner to assist the learner accessing learning objects, using them, and monitoring learning objectives. We plan on developing ePAL, the electronic Personal Advisor for Learning, to accomplish the following goals: (a) provide a unified framework for incorporating a large landscape of Web-enabled learning objects (courseware, simulation environments, ITSs, etc.), (b) use the framework and available learning objects to provide informative, high-quality content that is delivered in pedagogically effective ways, (c) conduct extensive learner profiling for the purpose of recording and tailoring the interaction of learners over long periods of time (months and years), (d) be fully interactive with the learner using natural simulated face-to-face communication via animated agents, (e) provide effective mentoring, and (f) be scaleable to a large number of learners. In order to accomplish these ambitious objectives, we would need a number of major components, as described below.

A *learner profile* is a record of the knowledge, skills, and history of the learner at varying levels of detail that is declared by the system designers. In order for a particular learning module to be incorporated into ePAL, it must maintain its section of the learner profile. The profile includes an evolving assessment of the student's particular knowledge and skills, as well as other historical data that are useful for making pedagogical decisions (e.g., time on task, percentage of entries preceded by help requests). Many of the existing learning modules that are available in the e-learning enterprise maintain learner profiles through a variety of techniques. A key extension in ePAL is that the global learner profile must accommodate these learning modules, as well as a set of cognitive abilities, emotions, personality traits, and historical attributes that are continuously consulted by ePAL as it plans what to do next. It is already feasible to accomplish these goals computationally. However, prior to implementing such a feature, it is essential to resolve the complications that arise from the standpoint of learner privacy and the representation of the learning profiles at varying grains of detail.

A full-blown ePAL will feature an *intelligent mentor.* Each learner will in essence have an agent working for them as a mentor that guides the learner on what to accomplish next. The mentor agent will maintain a profile of the learner's interests, needs, abilities, personality, and preferred learning styles. The mentor will select learning objects and dynamically sequence the content in a fashion that is tailored to the learner profile. The mentor will recommend that the learner stop from becoming overly involved with one topic at the expense of the others, will suggest learning strategies, will supervise the learner's interaction with the tutor, and will offer suggestions on controlling the learner's emotions.

Based upon the long and successful history of the building of AutoTutor, ePAL will have a *conversational interface*. The interface includes an animated mentor agent that is controlled by dialog and interface components, that scaffolds the learning process through tutoring, and that guides the learner's navigation. The basic principle behind this system component is that a human interacting with a computer agent should be just like a human interacting with a person. The agent will converse in natural language, attempt to say the right thing at the right time, and be capable of pointing, gesturing, and expressing emotion. Whether our vision is realistic or pure science fiction is very much an open question.

The tone of this chapter has been uniformly encouraging with regard to building animated conversational agents that hold conversations with the learner in natural language. However, it is appropriate to end this chapter by pointing out some limitations with this technology that might mute some of the enthusiasm. One problem is that a conversational agent appears mighty awkward when it misunderstands the learner and produces inappropriate dialogue moves. Such breakdowns in comprehension run the risk of eroding the learner's confidence in the intelligence of the agent. A second problem is that these conversational agents are not well equipped for precise subject matters, such as mathematics, statistics, and symbolic logic. Instead, the scope of these conversational agents is confined to verbal content and reasoning. A third problem is that the coordination of different discourse production technologies has not advanced sufficiently to adequately mimic humans. The timing of speech synthesis, speech intonation, pauses, facial expressions, eye blinks, pointing, and other gestures is far from settled. A fourth problem applies to the ePAL, but not AutoTutor. There are a bewildering number of legal questions that arise when exploring the prospects of tracking learner characteristics and accessing third-party learning and software modules. Matters of privacy and intellectual property are yet to be worked out in these large-scale cyber infrastructures. These challenges are duly noted, but do not prevent us from embarking on the adventures of these new technologies that simulate human conversation.

AUTHOR NOTES

The Tutoring Research Group (TRG) is an interdisciplinary research team comprised of approximately 35 researchers from psychology, computer science, physics, and education (visit http://www.autotutor.org). The research on AutoTutor was supported by the National Science Foundation (SBR 9720314, REC 0106965, REC 0126265, ITR 0325428) and the DoD Multidisciplinary University Research Initiative (MURI) administered by ONR under grant N00014-00-1-0600. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DoD, ONR, or NSF.

Correspondence concerning this chapter should be addressed to Art Graesser, Psychology Department, 202 Psychology Building, University of Memphis, Memphis, TN, 38152-3230, or via email at a-graesser@memphis.edu.

REFERENCES

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*, 147–179.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4,* 167–207.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*, 416–427.

Baylor, A. L. (2002). Agent-based learning environments for investigating teaching and learning. *Journal of Educational Computing Research, 26*, 249–270.

Beck, I.L., McKeown, M.G., Hamilton, R.L., & Kucan, L. (1997). *Questioning the author: An approach for enhancing student engagement with text.* Delaware: International Reading Association.

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.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.). (2000). *How people learning: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Bruning, R., Horn, C., & PytlikZillig, L.M. (Eds.). (2003). *Web-based learning: What do we know? Where do we go?* Greenwich, CT: Information Age Publishing.

Chan, T. W. (1996). Learning companion systems, social learning systems, and the global social learning club. *Journal of Artificial Intelligence in Educational, 7,* 125–159.

Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439–477.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science, 25*, 471–533.

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., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Erlbaum.

Corbett, A.T. (2001). Cognitive computer tutors: Solving the two-sigma problem. In M. Bauer, P.J. Gmytrasiewicz, & J. Vassileva (Eds.), *User Modeling: Proceedings of the Eighth International Conference* (pp. 137–147). Berlin: Springer-Verlag.

Craig, S.D., Graesser, A.C., Sullins, J., & Gholson, B. (in press). Affect and learning: An exploratory look into the role of affect in learning. *Journal of Educational Media*.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper-Row.

Dillon, T.J. (1988). *Questioning and teaching: A manual of practice*. New York: Teachers College Press.

Driscoll, D.M.,Craig, S.D., Gholson, B., Ventura, M.,Hu, X., & Graesser, A.C. (2003). Vicarious learning: Effects of overhearing dialog and monolog-like discourse in a virtual tutoring session. *Journal of Educational Computing Research, 29*, 431–450.

Foltz, P. W., Gilliam, S., & Kendall, S. (2000). Supporting content-based feedback in on-line writing evaluation with LSA. *Interactive Learning Environments, 8*, 111–128.

Fox, B. (1993). *The human tutorial dialogue project*. Hillsdale, NJ: Erlbaum.

Graesser, A. C., Hu, X., & McNamara, D. S. (in press). Computerized learning environments that incorporate research in discourse psychology, cognitive science, and computational linguistics. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications: Fetschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer*. Washington, D.C.: American Psychological Association.

Graesser, A.C., Hu, X., Person, P.,Jackson, T., & Toth, J. (in press). Modules and information retrieval facilities of the Human Use Regulatory Affairs Advisor (HURAA). *International Journal on eLearning*.

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*, 180–192.

Graesser, A. C., Jackson, G. T., Mathews, E. C., Mitchell, H. H., Olney, A., Ventura, M., et al. (2003). Why/AutoTutor: A test of learning gains from a physics tutor with natural language dialogue. In R. Alterman & D. Hirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 1–6). Mahwah, NJ: Erlbaum.

Graesser, A. C., Langston, M. C., & Baggett, W. B. (1993). Exploring information about concepts by asking questions. In G. V. Nakamura, R. M. Taraban & D. Medin (Eds.), *The psychology of learning and motivation: Vol. 29. Categorization by humans and machines* (pp. 411–436). Orlando, FL: Academic Press.

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology, 85*, 136–151.

Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., Person, N., & et al. (2003). AutoTutor improves deep learning of computer literacy: Is it the dialogue or the talking head? In U. Hoppe, F. Verdejo & J. Kay (Eds.), *Proceedings of Artificial Intelligence in Education* (pp. 47- 54). Amsterdam: IOS Press.

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., Harter, D., & Tutoring Research Group (2001). Teaching tactics and dialogue in AutoTutor. *International Journal of Artificial Intelligence in Education*, *12*, 257–279.

Graesser, A.C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology*, *9*, 359–387.

Graesser, A C., VanLehn, K., Rose, C., Jordan, P., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine*, *22*, 39–51.

Graesser, A.C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & Tutoring Research Group (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research, 1*, 35–51.

Gratch, J., Rickel, J., Andre, E., Cassell, J., Petajan, E., & Badler, N. (2002). Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems*, *17*, 54–63.

Johnson, W. L. (2001). Pedagogical agent research at CARTE. *AI Magazine, 22*, 85–94.

Johnson, W. L., & Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, *11*, 47–78.

Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice Hall.

Kass, A. (1992). Question asking, artificial intelligence, and human creativity. In T. W. Lauer & E. Peacock (Eds.), *Questions and information systems* (pp. 303–360). Hillsdale, NJ: Erlbaum.

King A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31,* 338–368.

Kintsch, E., Steinhart, D., Stahl, G., & LSA Research Group. (2000). Developing summarization skills through the use of LSA-based feedback. *Interactive learning environments*, *8*, 87–109.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Landauer, T., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*, 259–284.

Larsson, S., & Traum, D. (2000)., Information state and dialogue management in the TRINDI Dialogue Move Engine Toolkit. *Natural Language Engineering, 6*(3–4), 323–340.

Lesgold, A., Lajoie, S., Bunzo, M., & Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. H. Larkin & R. W. Chabay (Eds.), *Computer-assisted instruction and intelligent tutoring systems* (pp. 201–238). Hillsdale, NJ: Erlbaum.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Erlbaum.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38,* 1–30.

McNamara, D.S., Levinstein, I.B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers, 36*, 222–233.

Miyake, N. & Norman, D. A. (1979). To ask a question one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior, 18*, 357–364.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition & Instruction, 19*, 177–213.

Murray, T., Blessing, S., & Ainsworth, S. (2003). *Authoring Tools for Advanced Technology Learning Environments*. Amsterdam: Kluwer.

Olney, A., Louwerse, M. M., Mathews, E. C., Marineau, J., Mitchell, H. H., & Graesser, A. C. (2003). Utterance classification in AutoTutor. In J. Burstein & C. Leacock (Eds.), *Building Educational Applications using Natural Language Processing: Proceedings of the Human Language Technolog* (pp. 1–8). Philadelphia: Association for Computational Linguistics.

Otero, J., & Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition & Instruction, 19*, 143–175.

Palincsar, A. S., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction, 1*, 117–175.

Person, N. K., Graesser, A. C., Bautista, L., Mathews, E. C., & Tutoring Research Group (2001). Evaluating student learning gains in two versions of AutoTutor. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Artificial intelligence in education: AI-ED in the wired and wireless future* (pp. 286–293). Amsterdam: IOS Press.

Person, N. K., Graesser, A. C., & Tutoring Research Group (2002). Human or computer? AutoTutor in a bystander Turing test. In S. A. Cerri, G. Gouarderes & F. Paraguacu (Eds.), *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems* (pp. 821–830). Berlin: Springer-Verlag.

Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.

Pressley, M., & Forest-Pressley, D. (1985). Questions and children's cognitive processing. In A. C. Graesser, & J. B. Black (Eds.), *The psychology of questions* (pp. 277–296). Hillsdale, NJ: Erlbaum.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, televisions, and new media like real people and places*. Cambridge, UK: Cambridge University Press.

Rich, C., & Sidner, C. L. (1998). COLLAGEN: A collaborative manager for software interface agents. *User Modeling and User-adapted Interaction, 8,* 315–350.

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.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S.F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills* (Vol. 2, pp. 563–577). Hillsdale, NJ: Erlbaum.

Schank, R. C. (1999). *Dynamic memory revisited*. Cambridge, UK: Cambridge University Press.

Sleeman, D., & Brown, J. (Eds). (1982). *Intelligent tutoring systems*. New York: Academic Press.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Corporation.

Susarla, S., Adcock, A., Van Eck, R., Moreno, K., Graesser, A.C., & Tutoring Research Group (2003). Development and evaluation of a lesson authoring tool for AutoTutor. In V. Aleven, U. Hoppe, J. Kay, R. Mizoguchi, H. Pain, F. Verdejo, et al. (Eds.), *AIED2003 Supplemental Proceedings* (pp. 378–387). Sydney: University of Sydney School of Information Technologies.

VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., & Rosé, C.P. (2004). Natural language tutoring: A comparison of human tutors, computer tutors and text. Manuscript submitted for publication.

VanLehn, K., Jordan, P., Rosé, C. P., Bhembe, D., Bottner, M., Gaydos, A., et al. (2002). The architecture of Why2-Atlas: A coach for qualitative physics essay writing. In S.A. Cerri, G. Gouarderes & F. Paraguacu (Eds.), *Proceedings of the Sixth International Conference on Intelligent Tutoring* (pp.158–167). Berlin: Springer-Verlag.

Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*, 3–117.

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

Wisher, R. A., Sabol, M. A., & Moses, F. L. (2002). *Distance learning: The soldier's perspective* (Special Rep. No. 49). Alexandria, VA: U.S. Army Research Institute.