ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
2014, Vol. 23(5) 374–380

DOI: 10.1177/0963721414540680
cdps.sagepub.com
SAGE

Learning by Communicating in Natural Language With Conversational Agents

Arthur C. Graesser, Haiying Li, and Carol Forsyth
University of Memphis

Abstract

Learning is facilitated by conversational interactions both with human tutors and with computer agents that simulate human tutoring and ideal pedagogical strategies. In this article, we describe some intelligent tutoring systems (e.g., AutoTutor) in which agents interact with students in natural language while being sensitive to their cognitive and emotional states. These systems include one-on-one tutorial dialogues, conversational trialogues in which two agents (a tutor and a "peer") interact with a human student, and other conversational ensembles in which agents take on different roles. Tutorial conversations with agents have also been incorporated into educational games. These learning environments have been developed for different populations (elementary through high school students, college students, adults with reading difficulties) and different subjects spanning science, technology, engineering, mathematics, reading, writing, and reasoning. This article identifies some of the conversation patterns that are implemented in the dialogues and trialogues.

Keywords

conversational agents, intelligent tutoring systems, learning technologies

The positive impact of human tutoring on student learning is a well-established empirical finding that has motivated policies in many educational systems. Meta-analyses that have compared tutoring to classroom teaching and other suitable comparison conditions have reported effect sizes between $\sigma = 0.20$ and $\sigma = 1.00$ (Cohen, Kulik, & Kulik, 1982; Graesser, D'Mello, & Cade, 2011). Effect sizes are computed by comparing mean test scores of participants in treatment and comparison conditions and dividing those means by the pooled standard deviations. Expertise varies substantially among tutors, who include student peers, students slightly older than their tutees, community volunteers, paraprofessionals, college students with some pedagogical training, and experienced tutors with substantial subject-matter knowledge and pedagogical training. Sometimes one tutor simultaneously handles two or more students who are experiencing similar problems.

The question of why tutoring is effective in helping learning is far from settled, but detailed analyses have been conducted on the discourse, language, facial expressions, gestures, and actions used in tutorial conversations (Graesser et al., 2011; Graesser, Person, & Magliano, 1995). Researchers are currently investigating which conversational components are likely to explain learning gains. Tutor effectiveness does not simply consist in lecturing the student, which can be done in a classroom, but rather in attempts to get the student to construct answers and solutions to problems (Chi, Siler, Yamauchi, Jeong, & Hausmann, 2001). Surprisingly, tutor effectiveness cannot be attributed to a fine-grained diagnosis of what the student knows, to high shared knowledge (i.e., common ground; Clark, 1996) between the tutor and student, or to accurate feedback given to the student. Tutors have limited abilities to diagnose student knowledge because their shared knowledge is minimal. Instead, most human tutors follow a systematic conversational structure that has been termed *expectation- and misconception-tailored (EMT) dialogue*.

Corresponding Author:
Arthur C. Graesser, Department of Psychology and Institute for Intelligent Systems, Psychology Building Room 202, University of Memphis, Memphis, TN 38152-3230
E-mail: graesser@memphis.edu

What is EMT dialogue? Human tutors anticipate particular correct answers (called "expectations") and particular misconceptions when they ask the students challenging questions (or present them with challenging problems) and trace the students' reasoning. As the students express their answers, which are distributed over multiple conversational turns, tutors compare these contributions with these expectations and misconceptions. The tutors give feedback to the students that is sensitive to how well the students' contributions match the expectations or misconceptions. The tutors also produce *dialogue moves* to encourage the students to generate content and to improve their answers to challenging questions or problems. The following dialogue moves are prevalent in the EMT dialogue of most human tutors, including expert tutors:

Short Feedback. The feedback is either *positive* ("Yes"; "Very good"; a head nod; a smile), *negative* ("No"; "Not quite"; a head shake; a pregnant pause; a frown), or *neutral* ("Uh-huh"; "I see").

Pumps. The tutor gives nondirective pumps ("What else?"; "Tell me more") to get the student to do the talking.

Hints. The tutor gives hints to get the student to do the talking or perform the actions, but directs the student along some conceptual path. The hints vary from being generic statements or questions ("What about . . . ?"; "Why not?") to speech acts that more directly lead the student to a particular answer. Hints are the ideal scaffolding move to promote active student learning while directing the student to focus on relevant material.

Prompts. The tutor asks a leading question in order to get the student to articulate a particular word or phrase. Sometimes students say very little, so these prompts are needed to get them to say something specific.

Assertions. The tutor expresses a fact or state of affairs.

Pump-hint-prompt-assertion cycles are frequently used by tutors to extract or cover particular expectations. Eventually, all of the expectations are covered and the exchange is finished for the main question or problem. During this process, students occasionally ask questions, which are immediately answered by the tutor, and occasionally express misconceptions, which are immediately corrected by the tutor. Consequently, there are other categories of tutor dialogue moves: *answers* to student questions, *corrections* of student misconceptions, *summaries, mini-lectures*, and off-topic *comments*.

Simulating Human Tutors With AutoTutor

AutoTutor is an intelligent tutoring system that is designed to simulate the discourse moves of human tutors and also to implement some ideal tutoring strategies (Graesser et al., 2012; Graesser et al., 2004). In the program, there is an animated conversational agent that generates speech, facial expressions, and some gestures. Conversational agents have been increasingly popular in contemporary advanced learning environments (Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010; Gholson et al., 2009; McNamara, O'Reilly, Best, & Ozuru, 2006; Olney et al., 2012; Ward et al., 2013). All of the components of EMT dialogue can be implemented in AutoTutor because of advances in computational linguistics (Jurafsky & Martin, 2008) and statistical representations of world knowledge (Landauer, McNamara, Dennis, & Kintsch, 2007). AutoTutor is capable of classifying the students' contributions into different categories of *speech acts*, such as questions, statements, metacognitive expressions ("I don't know"; "I see"), short responses ("Okay"), and expressive evaluations ("This is terrible"). AutoTutor responds adaptively, in a fashion that is sensitive to the students' speech-act categories and the quality of their statements.

We emphasize that AutoTutor cannot interpret all speech acts that students produce, but it can simulate EMT dialogue, and that does help students learn. As in the case of human tutors, AutoTutor has only an approximate "understanding" of what a student knows and expresses, much of which is vague, ungrammatical, and semantically ill-formed. Nevertheless, AutoTutor can assess how well the student's contributions in the various conversational turns match the expectations and misconceptions by using *semantic-pattern-matching algorithms* that are sufficiently accurate. AutoTutor gives short feedback that depends on the quality of student contributions in the previous turn. AutoTutor generates pumps, hints, prompts, and assertions to fill in missing words of expectations, following *semantic-pattern-completion algorithms*. AutoTutor can answer a subset of student questions, but students do not frequently ask questions in tutoring (Chi et al., 2001; Graesser et al., 1995), so this is not a major limitation. An example AutoTutor conversation is presented in Table 1.

Empirical evidence supports the claim that AutoTutor and similar computer tutors that use natural-language dialogue yield learning gains comparable to those of trained human tutors, with effect sizes averaging 0.8 and ranging from 0.6 to 2.0 (Graesser et al., 2012; Graesser et al., 2004; Hu & Graesser, 2004; McNamara et al., 2006; Olney et al., 2012; VanLehn, 2011; VanLehn et al., 2007).

Indeed, direct comparisons between these computer tutors and human tutors have shown no significant differences. These assessments have covered subject matters and skills in the areas of science and technology (e.g., physics, biology, computer literacy, and scientific reasoning) as well as reading comprehension. The quality of the dialogue in AutoTutor is also reasonably coherent, although not perfect. In fact, the dialogue is sufficiently tuned so that a bystander who reads the tutorial dialogue in print cannot tell whether a particular tutor turn was generated by AutoTutor or by an expert human tutor.

Several versions of AutoTutor have been developed since 1997, when the system was created. The conversational agents with talking heads have been compared with conversational agents using speech alone, chat messages in print, and multiple communication channels. Students communicating in text versus spoken utterances have also been compared. For the most part, it is the semantic, conceptual content in the conversation that predicts learning, not the medium of communication (D'Mello, Dowell, & Graesser, 2011; Graesser, Jeon, & Dufty, 2008; VanLehn et al., 2007). One version of AutoTutor, called AutoTutor-3D, guides learners on using interactive simulations of physics microworlds (Graesser, Chipman, Haynes, & Olney, 2005). An interactive simulation world with people, objects, and a spatial setting is created for each physics problem. Students manipulate parameters of the situation (e.g., mass of objects, speed of objects, distance between objects) and then ask the system to simulate what will happen. Students are also prompted to describe what they see. Their actions and descriptions are evaluated with respect to matching expectations or misconceptions. AutoTutor manages the dialogue with hints and suggestions that scaffold the interactive simulation experience. AutoTutor-3D yields a significant improvement in learning (σ = 0.2) over the strictly conversational AutoTutor for those students who run the simulations multiple times rather than only once (Jackson, Olney, Graesser, & Kim, 2006).

Another version of AutoTutor is sensitive to students' emotions and responds with appropriate emotional expressions (D'Mello & Graesser, 2012). Students' emotions are detected by the computer through sensing channels on the basis of dialogue patterns during tutoring, the content covered, and the students' facial expressions, body posture, and speech intonation. The primary emotions that occur during learning with AutoTutor are frustration, confusion, boredom, and flow (engagement); surprise and delight also occur occasionally (Graesser & D'Mello, 2012). An AutoTutor that is supportive and emotionally empathetic toward low-knowledge students helps learning more than an AutoTutor that is not emotionally sensitive.

Agents in Trialogues

Versions of AutoTutor have recently been developed for group conversations. The incremental value of multiple agents is that the human learner can learn by observing how the agents interact. A student can learn vicariously by observing one agent communicating with another agent, showing how actions are performed, and reasoning collaboratively with the other agent. Interactions of two agents with one human student are called *trialogues*. The two agents can disagree, contradict each other, and hold an argument, periodically turning to the student to solicit his or her perspective (Lehman et al., 2013). Such experiences put students in cognitive disequilibrium, which encourages problem solving, reasoning, and, ultimately, deep learning. This section describes some systems using trialogues that help students learn in a social world.

It is possible to extend the horizons of intelligent tutoring systems beyond trialogues. Researchers have developed systems in which a human student interacts with more than two agents, as well as systems in which one tutor agent serves several students in multiparty collaborative learning (Kumar & Rosé, 2011). Researchers are also developing multiple-agent configurations for use in high-stakes assessment environments, including those of the Educational Testing Service. However, it is beyond the scope of this article to cover these larger group configurations and high-stakes assessments.

Trialogues with agents for adult literacy

As an illustration, AutoTutor trialogues are currently being developed for the Center for the Study of Adult Literacy (CSAL; http://csal.gsu.edu/content/homepage). The goal is to help adults with low literacy read better so they can improve their lives. Figure 1 shows a snapshot of the interface for AutoTutor-CSAL. In this setup, there is a teacher agent (Cristina, top left), a student agent (Jordan, top right), and the human who interacts with the agents (Haiying). The program is a competition between the human learner and the student agent, with guidance from the teacher agent. The learner's task is to apply a *clarifying strategy* to identify the meaning of words with multiple meanings in context (in this case, the two meanings of *fan*). A window at the top shows a sentence in which the focal word ("fans") is underlined. Below this sentence are two images and associated sentences that represent two meanings of *fan*. Two scoreboards at the bottom of the screen display the names of the human learner and the student agent and their current scores in the competition.

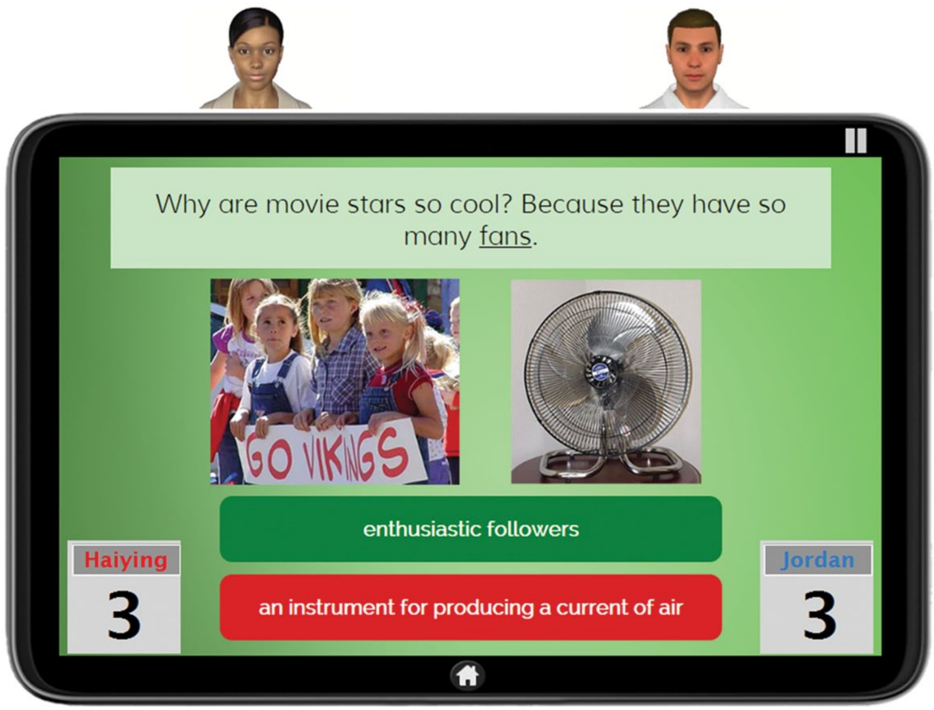

Fig. 1. Screen shot of conversational agents in an AutoTutor trialogue designed to help adult learners read.

How are the trialogues played out? Cristina first asks a question and asks the human learner, Haiying, to choose the correct answer by clicking on one of the two answers. When clicked, the answer changes its color. Cristina then asks Jordan to choose the answer. When Jordan gives an answer, either agreeing or disagreeing with Haiying, Cristina gives the two of them feedback, shows the correct answer in green, and updates the scores. In summary, the agents navigate the human learner through the experience, present him or her with competition, give feedback and explanations for why answers are correct or incorrect, and keep score on points in the competition. Across all of the lessons in CSAL, both agents and human learners play many roles: teacher, helper, collaborator, ideal model, judge, and critic.

Trialogues with agents in scientific reasoning

A detailed example of a trialogue is presented in this section. This trialogue is based on an instructional game initially called Operation ARIES! (Millis et al., 2011), which was subsequently commercialized by Pearson Education as Operation ARA (Halpern et al., 2012). ARIES is an acronym for Acquiring Research Investigative and Evaluative Skills, whereas ARA is an acronym for Acquiring Research Acumen. The game teaches high school and college students how to critically evaluate aspects of scientific investigations, such as the need for control groups, the distinction between correlation and causation, and operational definitions. In one game phase, the two agents and student collaboratively answer difficult questions based on a text the student has read. In another phase, the trialogue involves critiquing case studies of scientific research with respect to scientific methodology.

Table 1 shows a conversation that demonstrates an EMT trialogue in Operation ARA. Tom is the name of the human student, whereas Dr. Quinn is the teacher agent and Glass is a student agent. Dialogue-move categories are shown in brackets and italics. Sometimes Glass gives incorrect answers, whereas Dr. Quinn has the ground truth. The trialogue encourages Tom to fill in information, answer questions, and resolve disagreements. Therefore, the human learner is an active participant rather than merely a passive, vicarious observer.

The trialogues in ARIES and ARA have been shown to help students learn (Halpern et al., 2012; Millis et al., 2011), but we are still exploring what characteristics of the conversational interaction account for this learning (Forsyth et al., 2012). We do know that AutoTutor dialogues and trialogues increase learning substantially compared with reading texts on the content for an equivalent amount of time (Graesser et al., 2004; Kopp, Britt, Millis, & Graesser, 2012; VanLehn et al., 2007). However, we need to better understand why and how.

In closing, this article has described how AutoTutor dialogues and trialogues help students learn by engaging

Table 1. Example Trialogue Conversation With Operation ARA

Dr. Quinn: I really think we should talk about this topic a little further.
Dr. Quinn: Why do you think it is that you can't make a causal inference when you do a correlational study? *[Main question]*
Tom: Because the data aren't collected scientifically.
Glass: I agree Tom.
Glass: It is hard to make a causal conclusion with correlational research because the data aren't collected scientifically.
Dr. Quinn: That is not right, Glass. *[Negative feedback]*
Dr. Quinn: The data are collected scientifically in correlational research. This is not the problem. *[Correction]* Think about this some more. *[Pump]*
Glass: So, what would be the reason that causal inferences are not appropriate when doing correlational research? This is making me think about how it is that participants are assigned to groups in true experiments. *[Peer hint]* Tom, can you help me with this question?
Tom: It has to do with the assignment of the participants.
Dr. Quinn: O.k. *[Neutral feedback]*
Dr. Quinn: Causal inferences can really only be made in true experiments when participants are assigned to groups in what way? *[Prompt]*
Tom: Randomly.
Dr. Quinn: Great! *[Positive feedback]*
Dr. Quinn: When participants are assigned to groups randomly. *[Correct prompt completion]*
Dr. Quinn: Think about this.
Dr. Quinn: Remember, you're trying to explain why it is that causal inferences are not appropriate when doing correlational research. It might help you to think about the various possible explanations for a significant correlation, and then answer the question.
Tom: I think it has to do with the sample size.
Glass: No, I don't think that is the answer.
Glass: It seems to me that with correlational research, relationships are identified, but it is hard to know which variable caused the change in the other variable . . .
Dr. Quinn: In summary, correlations tell us about relationships . . . *[Summary]*

Note: In this example trialogue, Tom is the human learner, Dr. Quinn is the tutor agent, and Glass is the student agent. Dialogue-move categories are annotated with italics in brackets.

them in conversations in natural language. These computer agents exhibit conversation patterns of human tutors in addition to ideal pedagogical strategies. Learning is facilitated by these conversational-agent environments, compared with students reading textbooks and engaging in other non-interactive learning environments. Perhaps this is not surprising, given that human tutors are also more effective than classrooms and textbook reading. Indeed, learning has occurred for millennia in apprenticeship contexts in which the learner communicates with the tutor, master, or mentor in natural language. This is the moment in history when researchers in computational linguistics, artificial intelligence, intelligent tutoring systems, discourse processes, and the learning sciences are simulating many of these conversation patterns. The agents are not perfect simulations of humans, but they are good enough to help students learn.

Recommended Reading

Graesser, A. C. (2011). Learning, thinking, and emoting with discourse technologies. *American Psychologist*, *66*, 743–757. An overview of the agent-based learning environments and the computer systems that automatically analyze discourse and emotions that were developed at the Institute of Intelligent Systems.

Graesser, A. C., Conley, M., & Olney, A. (2012). Intelligent tutoring systems. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook: Vol. 3. Applications to learning and teaching* (pp. 451–473). Washington, DC: American Psychological Association. A handbook chapter that reviews the research, development, and testing of contemporary intelligent tutoring systems.

Rus, V., D'Mello, S., Hu, X., & Graesser, A. C. (2013). Recent advances in intelligent systems with conversational dialogue. *AI Magazine*, *34*, 42–54. An article that describes recent intelligent tutoring systems using natural-language conversation that are sensitive to the knowledge and emotions of learners.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Financial support was provided by grants and contracts from the National Science Foundation (Grants SBR 9720314, REC 0106965, REC 0126265, ITR 0325428, REESE 0633918, ALT-0834847, DRK-12-0918409, 1108845), the Institute of Education Sciences (Grants R305H050169, R305B070349, R305A080589, R305A080594, R305A090528, R305A100875, R305C120001), the Army Research Laboratory (W911NF-12-2-0030-P00002), and the Office of Naval Research (N00014-12-C-0643). Pearson Education created a Web version of Operation ARA, but the authors have received no royalties or payments from Pearson Education for ARA. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the Institute of Education Sciences, or the Department of Defense.

References

Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning*, *5*, 123–152.

Chi, M. T. H., Siler, S., Yamauchi, T., Jeong, H., & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science*, *25*, 471–534.

Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

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.

D'Mello, S. K., Dowell, N., & Graesser, A. C. (2011). Does it really matter whether students' contributions are spoken versus typed in an intelligent tutoring system with natural language? *Journal of Experimental Psychology: Applied*, *17*, 1–17.

D'Mello, S. K., & Graesser, A. C. (2012). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*, *2*, 1–39.

Forsyth, C. M., Pavlik, P., Graesser, A. C., Cai, Z., Germany, M., Millis, K., . . . Dolan, R. (2012). Learning gains for core concepts in a serious game on scientific reasoning. In K. Yacef, O. Zaïane, H. Hershkovitz, M. Yudelson, & J. Stamper (Eds.), *Proceedings of the 5th International Conference on Educational Data Mining* (pp. 172–175). Chania, Greece: International Educational Data Mining.

Gholson, B., Witherspoon, A., Morgan, B., Brittingham, J. K., Coles, R., Graesser, A. C., . . . 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., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions in Education*, *48*, 612–618.

Graesser, A. C., & D'Mello, S. (2012). Emotions during the learning of difficult material. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 57, pp. 183–226). San Diego, CA: Elsevier.

Graesser, A. C., D'Mello, S. K., & Cade, W. (2011). Instruction based on tutoring. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 408–426). New York, NY: Routledge Press.

Graesser, A. C., D'Mello, S. K., Hu, X., Cai, Z., Olney, A., & Morgan, B. (2012). AutoTutor. In P. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing: Identification, investigation, and resolution* (pp. 169–187). Hershey, PA: IGI Global.

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. *Behavior Research Methods, Instruments, & Computers*, *36*, 180–193.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*, 495–522.

Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity*, *7*, 93–100.

Hu, X., & Graesser, A. C. (2004). Human Use Regulatory Affairs Advisor (HURAA): Learning about research ethics with intelligent learning modules. *Behavior Research Methods, Instruments, & Computers*, *36*, 241–249.

Jackson, G. T., Olney, A., Graesser, A. C., & Kim, H. J. (2006). AutoTutor 3-D simulations: Analyzing user's actions and learning trends. In R. Son (Ed.), *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (pp. 1557–1562). Mahwah, NJ: Erlbaum.

Jurafsky, D., & Martin, J. H. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.

Kopp, K., Britt, A., Millis, K., & Graesser, A. (2012). Improving the efficiency of dialogue in tutoring. *Learning and Instruction*, *22*, 320–330.

Kumar, R., & Rosé, C. P. (2011). Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies*, *4*, 21–34.

Landauer, T., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.

Lehman, B., D'Mello, S. K., Strain, A., Mills, C., Gross, M., Dobbins, A., . . . Graesser, A. C. (2013). Inducing and tracking confusion with contradictions during complex learning. *International Journal of Artificial Intelligence in Education*, *22*, 85–105.

McNamara, D. S., O'Reilly, T., Best, R., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, *34*, 147–171.

Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A., & Halpern, D. (2011). Operation ARIES! A serious game for teaching scientific inquiry. In M. Ma, A. Oikonomou, & J. Lakhmi (Eds.), *Serious games and edutainment applications* (pp. 169–196). London, England: Springer-Verlag.

Olney, A., D'Mello, S. K., Person, N., Cade, W., Hays, P., Williams, C., . . . Graesser, A. C. (2012). Guru: A computer tutor that models expert human tutors. In S. Cerri, W. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of Intelligent Tutoring Systems (ITS) 2012* (pp. 256–261). Berlin, Germany: Springer.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*, 197–221.

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.

Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., & Weston, T. (2013). My science tutor: A conversational multimedia virtual tutor. *Journal of Educational Psychology*, *105*, 1115–1125.