International Journal of Artificial Intelligence in Education (2000), 11, 242-253

Overhearing Dialogues and Monologues in Virtual Tutoring Sessions: Effects on Questioning and Vicarious Learning

Scotty D. Craig, Barry Gholson, Matthew Ventura, Arthur C. Graesser and the Tutoring Research Group *Department of Psychology, The University of Memphis, Memphis, Tennessee, 38152, USA*
E-mail: scraig@memphis.edu, barry-gholson@mail.psyc.memphis.edu

Abstract. This research used vicarious-learning procedures in an attempt to increase students' knowledge of the domain and enhance the quality of their questions in an intelligent tutoring-system involving computer-controlled animated agents. Students who overheard a dialogue in which a virtual tutee asked a virtual tutor questions during acquisition wrote significantly more in free recall and asked significantly more questions in a transfer task than those who overheard a monologue. Students in the monologue condition asked significantly more shallow questions that require short answers than those in the dialogue condition. Students in the dialogue condition asked significantly more deep-level reasoning questions of the kind that were modeled by the virtual tutee during acquisition. It was concluded that modeling a brief dialogue involving question asking at the outset of a first learning session on an intelligent tutor system could enhance both the structure of the subsequent dialogue and knowledge acquisition.

INTRODUCTION

Many researchers have proposed that question generation is a fundamental component of the cognitive processes that operate at deep conceptual levels, such as text comprehension (e.g., Collins, Brown, & Larkin, 1980; Graesser, Person, & Huber, 1992; Graesser & Clark, 1985; Hilton, 1990; Kintsch, 1998), and the learning of complex materials (Collins, 1985; Graesser, Singer, & Trabasso, 1994; Palinscar & Brown, 1984; Schank, 1986). One of the primary challenges, however, is that student questions are infrequent and unsophisticated under most conditions (Buseri, 1988; Ohlsson, 1995; Good, Slavings, Harel, & Emerson, 1987; Nickel & Fenner, 1974). In a review of the relevant literature, Graesser and Person (1994) pointed out that the median number of students' questions per hour in the classroom is 3.0. Because an average classroom contains about 26.7 students (Dillon, 1988), the number of questions generated by any particular student is only about .11 (i.e., 3.0/26.7) question per hour of classroom instruction (Graesser & Person, 1994, p. 105).

The good news is that it is possible to explicitly train students to ask good questions, and that such training leads to improvements in the comprehension, learning, and memory of technical materials (Davey & McBride, 1986; Gavelek & Raphael, 1985; King, 1989, 1990; King, Staffieri, & Adelgais, 1998; Palincsar & Brown, 1984). The training procedures used to induce question asking were usually quite explicit, conducted as part of standard classroom activity, and spread across multiple sessions. Davey and McBride (1986), for example, used five training sessions distributed across two weeks, and King et al. (1998) used at least ten sessions per student spread across five weeks.

Graesser and Person (1994) have shown that, unlike their typical performance in the classroom, high school and college students do ask questions spontaneously when tutored individually. They also found, however, when they classified students' questions according to Bloom's (1956) taxonomy that 92% of the queries were at the shallow end of the educational objectives. Only 8% of students' questions spanned levels 2 through 6 of Bloom's taxonomy. Graesser and Person (1994) also classified student questions according to a scheme developed by Graesser, Person, and Huber (1992). Their question categories are presented in Table 1.

They reported that 44% of the students' questions fell into the first five categories (short answer) in the table, while another 10% were in the last two categories (assertion, request/directive). Half of the short-answer questions were verifications. Questions were pretty sparse in most of the 11 long answer categories that require more extended reasoning. The exceptions were about 9% interpretational (What's happening?) and 21% instrumental/procedural (How?). The remaining questions (16%) were about evenly distributed across the remaining nine long answer categories. Graesser and Person concluded (1994, p. 131) that, "Given that most students have not mastered effective question-asking skills, then there should be benefits in learning after they are taught how to ask good questions."

The shallow questions students ask during tutoring sessions are most conducive to fact learning, rather than the kind of knowledge construction labeled "generative learning" by Wittrock (1974, 1989). The latter kind of learning involves the construction of meaning through both analyzing ideas and integrating them. Learners presumably think about how the facts relate to each other and integrate those relations with what they already know (King et al., 1998; Webb, Troper, & Fall, 1995).

Our research was conducted in the context of an Intelligent Tutoring System (ITS), in which student learners interact with one or more computer-controlled animated agents (Graesser, Franklin, Wiemer-Hastings, & TRG, 1998; Graesser, Wiemer-Hastings, Wiemer-Hastings, Kreuz, & TRG, in press; Hu, Graesser, & TRG, 1998). Thus, we were interested in exploring procedures that would both induce students to ask deep-level questions (Bloom, 1956; Graesser et al., 1992; Ryan, 1971), and at the same time fit naturally into the context of the ITS under development (McCauley, Gholson, Hu, Graesser, & TRG, 1998; Wiemer-Hastings, Graesser, Harter, & TRG, 1998; Wiemer-Hastings, Wiemer-Hastings, Graesser, & TRG, 1999). In an attempt to accomplish these goals, we adopted a vicarious-learning perspective (Bandura, 1962; King, 1990; Lee, Dineen, & McKendree, 1998; Mayes, 1998; McKendree, Stenning, Mayes, Lee, & Cox, 1998; Rosenthal & Zimmerman, 1978).

The term vicarious learning has been used synonymously with observational learning, modeling, and social learning (Bandura, 1971; Rosenthal & Zimmerman, 1978). The perspective says that through simply observing some activity carried out by a model, the student learns how to perform that activity without overt practice of direct incentives (Rosenthal & Zimmerman, 1978, p. xi). In the present context the question addressed is (Lee et al.,1998, p. 17) "What benefits can students gain from dialogue as observers, not just as participants?"

Recent advances in educational technology, computer-based courses, distance learning (e.g., courses on the web), and ITSs (Anderson, Corbett, Koedinger, & Pelletier, 1995; Derry & Potts, 1998; Edelson, 1996; Holland, Kaplan, & Sams, 1995; Lesgold, Lajoie, Bunzo, & Eggan, 1992; Moore, 1994; Paulsen, 1995; Renwick, 1996; Scardamalia, Bereiter, Brett, Burtis, Calhoun, & Smith-Lea, 1992; Van Lehn, 1990; Weber, 1996), have created complex configurations of social agents. In addition to a teacher or tutor who has a direct conversation with the learner, there are tutors and secondary agents who overhear the language of the learner and conversations between the learner and other agents. Learners themselves are sometimes over-hearers of various conversations between agents. Therefore, learners are more and more likely to find themselves trying to understand talk in settings in which they are over-hearers (Schober & Clark, 1989; Brennan & Clark, 1996; Fox Tree, 1999), rather than active participants. These complex social configurations in the new learning technologies stimulate a host of questions for researchers in cognition and instruction. There is need for further empirical understanding of the conditions under which vicarious-learning procedures can be used effectively in language environments in which learners are relatively isolated (e.g., Lee et al., 1998; McKendree et al., 1998).

In the research reported below, we were mainly concerned with vicarious learning procedures (Bandura, 1962, 1977; Lee et al., 1998; Rosenthal & Zimmerman, 1978), or modeling. Can the procedures be used to improve the structure of subsequent discussion between a student and an animated agent in an ITS, and at the same time, possibly increase students' knowledge of the domain? The vicarious learning procedure under consideration is the asking of deep-level questions (Davey & McBride, 1986; Graesser & Person, 1994; King, 1989, 1990; Palincsar & Brown, 1984). During *acquisition*, student participants were exposed to a virtual tutoring session in which two computer-controlled embodied agents acted out the roles of tutor and tutee (cf. Aimeur & Fahmi, 1998). The information discussed by the embodied agents was from the domain of computer literacy. In a dialogue interaction condition, the informational content of the virtual tutoring session was presented in a lively dialogue format, in which the tutee asked a total of 66 questions distributed across eight subtopics. In a monologue condition, the content was embedded in a monologue format, in which the tutee asked only one question per subtopic. The content of the information delivered by the virtual tutor was identical in both formats. We assessed the success of the vicarious-learning formats through measures of free recall, and by the questions generated by the student learners in a *transfer* task.

Table 1. Question Categories in the Graesser, Person, and Humber (1992) Scheme

Question Category	Abstract Specification	Example
SHORT ANSWER		
Verification	Is a fact true? Did an event occur?	These are switches right?
Disjunctive	Is X or Y the case? Is X, Y, or Z the case	Is that volatile or nonvolatile?
Concept completion	Who? What? What is the referent of a noun argument slot?	What did he just say?
Feature Specification	What Qualitative attributes does entity X have?	What are the properties of the BUS?
Quantification	What is the value of a quantitative variable? How many?	How many electronic switches are in RAM?
LONG ANSWER		
Definition	What does X mean?	What is a CPU?
Example	What is an example label or instance of the category?	What is an example of a utility program?
Comparison	How is X similar to Y? How is X different from Y?	What is the difference between the RAM and the CPU?
Interpretation	What concept or claim can be inferred from a static or active pattern of data?	What is happening in this picture?
Causal antecedent	What state or event causally led to an event or state?	How did the computer crash?
Causal consequence	What are the consequences of an event or state?	What happens to RAM when the computer is turned off?
Goal orientation	What are the motives or goals behind an agent's action?	Why was the application loaded into RAM?
Instrumental/ Procedural	What instrument or plan allows an agent to perform an action?	How does the operating system open an application after you click on an icon?
Enablement	What object or resource allows an agent to perform an action?	What allows the operating system to translate things into binary number system?
Expectational	Why did some expected event not occur?	Why is this number a two and not a three?
Judgemental	What value does the answerer place on an idea or advice?	So, is a Macintosh or a PC better?
Assertion	The speaker makes a statement indicating he lacks knowledge or does not understand an idea?	I have no idea what it means.
Request/Directive	The speaker wants the listener to perform an action.	Can we move on?

METHOD

Subjects

A total of 48 students, whose participation met a course requirement, were drawn from introductory psychology classes at The University of Memphis. A checklist was used to screen potential participants. The participants included only students who reported that they had average or less knowledge of computers. The design of the experiment was a 2 (acquisition condition: monologue vs. dialogue) x 2 (tutorial content: software vs. hardware) factorial. The 48 participants were randomly assigned to the resulting four conditions, with 12 per group. Students were tested individually in one session averaging about 75 minutes.

Materials

The computerized materials were created using two different computer application packages. These two packages were Xtrain (Hu, 1998) and Microsoft Agent (Microsoft, 1998). Xtrain was created for the purpose of integrating Microsoft agents with text, pictures, HTML, Macromedia Flash, and Director files. Xtrain (Hu, 1998) was used to script and control the presentation of the virtual agents, narration, and the picture displays. Secondly, Microsoft Agent (Microsoft, 1998) was used for character creation, animation, and voice synthesis of the virtual tutor and tutee, which were displayed on a monitor.

Acquisition

Tutorial content for eight subtopics concerned with software and eight subtopics concerned with hardware was derived from curriculum scripts for an ITS created by the Tutoring Research Group at the University of Memphis (Graesser et al., 1998, in press; McCauley et al., 1998; Wiemer-Hastings et al., 1998). During acquisition half the participants were presented with a virtual tutoring session involving hardware content; for the remainder the content was software. In both the monologue and dialogue conditions each subtopic was introduced by a brief information delivery presented by the virtual tutor. In the monologue condition the virtual tutee then asked one question, and the virtual tutor followed this with a monologue presentation of the tutorial content for that subtopic. In the dialogue condition, after each information delivery the virtual tutee asked a series of questions, ranging from five to 14 across the eight subtopics (66 total). The exact words, phrases, and sentences used by the virtual tutor in response to the virtual tutee’s questions were identical in the dialogue and monologue conditions in each content domain (software and hardware).

During acquisition, the virtual tutor, the virtual tutee, and a picture relevant to the subtopic under discussion were displayed on the monitor at all times. An example, showing the two animated agents and a picture used for a subtopic on the operating system, as depicted on the monitor, are presented in Figure 1. The picture changed from subtopic to subtopic. At the end of a subtopic the virtual tutor indicated that they were going on to a new subtopic and said he would bring up a new picture. Immediately following acquisition each student was given a retention test that consisted of two open-ended questions concerning two of the subtopics covered by the virtual tutor and virtual tutee. The open-ended questions were selected to encourage deep reasoning. An example of the retention questions is presented in Appendix A. The questions were administered one at a time, and students were permitted to write until they said they were finished. For purposes of assigning retention-test questions to subjects, constrained randomizations were prepared so that among the 12 students in each of the four subgroups (e.g., dialogue training on hardware content) each subtopic was presented to three different students, and no two students were tested on exactly the same two subtopics.

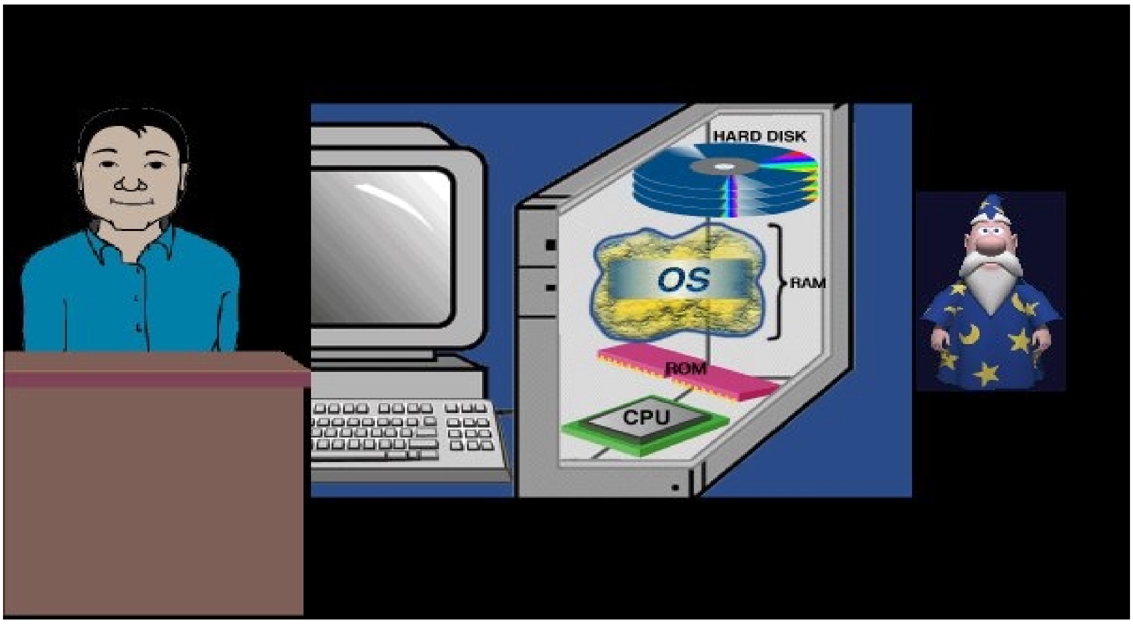

Figure 1. Screenshot of software subtopic on the Operating system.

The question taxonomy in Table 1 was used to classify the questions the virtual tutee asked the virtual tutor. In the dialogue condition, seven percent of the questions required short answers (1-5 in the table). Most of the remaining questions (88%) were drawn from six categories: comparison, interpretation, causal antecedent, causal consequence, instrumental/procedural, and enablement. The remaining five percent were in other long-answer categories. The questions asked by the virtual tutee in the monologue condition were drawn from five categories: comparison, causal antecedent, causal consequence, instrumental/procedural, and enablement.

Because the questions asked in the monologue condition were intended to justify the monologue that followed, questions in that condition were generally longer than those in the dialogue condition. For example, in a subtopic on the operating system, the first question asked by the virtual tutee following the information delivery in the dialogue condition was "Why does it load the operating system, what does that accomplish?" In the corresponding monologue condition, following the information delivery the question asked was "When you first turn on a computer they say it boots up. What I would like to know is, what booting up means? What is actually going on right after you first turn on a computer?" The information delivery spoken by the virtual tutor at the outset of the subtopic on the operating system, along with the complete dialogue that followed it, is given in Appendix B. In the monologue condition only the one question was asked, and it was followed only by the spoken text labeled "Virtual tutor" in Appendix B.

Transfer

Following acquisition each student was presented with a transfer task in which they were permitted to ask questions. They were told that the tutor would deliver some information on each subtopic, and then they (the student) could clarity any information that would help them understand the subtopic. In transfer only the virtual tutor and the appropriate picture were on the monitor for each of the eight subtopics. Students in the software content condition during acquisition were presented with hardware transfer, and those in hardware content condition in acquisition transferred to software. The contents of the information deliveries for each subtopic in transfer were identical to those presented for the corresponding subtopic in acquisition.

The procedure during transfer began with the virtual tutor giving the information delivery for the first subtopic. The virtual tutor ended the subtopic by inviting the students to clarity any information that would help them understand the material, but then the tutor stopped talking. The student then queried the experimenter until they said they were finished with that subtopic (i.e., said they had no more questions). The virtual tutor then brought up a new picture, gave the information delivery for the second subtopic, and so on. After the eight subtopics were presented in transfer, each subject received a second retention test that, parallel to acquisition, consisted of open ended questions about two of the subtopics covered in transfer.

RESULTS AND DISCUSSION

For purposes of analysis, data from the retention tests following acquisition and transfer were scored by computing the total number of propositions (or story units) that were written by each student. A 2 (monologue vs. dialogue) x 2 (hardware vs. software) analysis of variance revealed that students in the dialogue condition wrote significantly more content (46.2 propositions) than those in the monologue condition (36.3 propositions), $F(1,46) = 4.53$, $p < .05$. The difference did not reach significance in the retention test following acquisition (dialogue = 23.8, monologue = 19.6) but it was significant in the retention test that followed transfer (dialogue = 22.3, monologue = 16.7), $F(1,46) = 4.61$, $p < .05$. As will be seen below, this difference in retention following transfer may have resulted from the number and quality of the questions students asked in the dialogue condition during the transfer task.

These results are consistent with findings reported recently by Fox Tree (1999). She reported that over-hearers performed better in a referential communication task after they overheard dialogues than when they overheard monologues. In a preliminary task students were divided into directors and matchers. The goal was for the director to describe an ordered set of abstract shapes (tangrams) to the matcher so the matcher could place the shapes in the same order. Directors either gave instructions for the matcher to follow (monologue condition), or conversed freely with the matcher (dialogue condition). The sessions were taped, and only those sessions in which matchers correctly ordered the tangrams were used in the overhearing experiment. After ruling out a variety of possible explanations based on the experimental materials themselves, Fox Tree concluded that over-hearers performed better after listening to dialogues either (a) due to a greater number of discourse markers, or (b) because two interlocutors' perspectives are more informative than one perspective.

The structure of the transfer task in the current experiment involved turn-taking starting with the students. The students first queried the experimenter on their turn, the experimenter answered the question, and then it was the students' turn again, and so forth, until the students said they had no more questions. It was possible, however, for the student to ask more than one question during any given turn. Thus, the question asking data were analyzed both for the number of turns taken by the students, and the number of queries generated by the students. Regarding the total number of turns taken by the students, a preliminary analysis revealed that the students took a total of 1,169 turns, and that those in the dialogue condition averaged more turns than those in the monologue condition (dialogue mean = 29.7, monologue mean = 19.0), $\underline{F}(1,46) = 5.11$, $\underline{p} < .05$.

The data from each student's turn were further classified according to the scheme presented in Table 1. Following Graesser and Person (1994, p.119), when a question was a hybrid of two or more categories we gave low priority to verifications, assertions, and request/directive categories. In addition, short-answer questions had lower priority than long-answer questions. Two judges categorized the questions, and achieved an inter-rater reliability of .87. In their 1,169 turns the students actually generated 1,511 queries. Classification of these queries according to the scheme of Table 1 for the dialogue and monologue conditions is presented in Table 2. A test of independence contrasting the dialogue with the monologue condition revealed a significant difference, $X^2_{(df = 17)} = 165.55$, $\underline{p} < .001$. The types of questions asked by students in the monologue and dialogue conditions were, therefore, not drawn from the same distribution.

Table 2. The Number and Proportion of Questions Generated by Students in the Monologue and Dialogue Conditions

Question Category	Monologue		Dialogue	
SHORT ANSWER	Number	Proportion	Number	Proportion
Verification	63	.10	25	.03
Disjunctive	6	.06	3	.00
Concept completion	48	.08	19	.02
Feature Specification	28	.04	35	.04
Quantification	10	.02	16	.02
LONG ANSWER				
Definition	44	.07	63	.07
Example	13	.02	36	.04
Comparison	36	.06	81	.09
Interpretation	50	.08	117	.13
Causal antecedent	69	.11	94	.11
Causal consequence	81	.13	178	.20
Goal orientation	6	.01	18	.02
Instrumental/procedural	27	.04	39	.04
Enablement	39	.06	87	.10
Expectational	9	.01	29	.03
Judgemental	5	.01	7	.01
Assertion	75	.12	24	.03
Request/Directive	23	.02	8	.01
Total	632	1.00	879	1.00

Data from the 12 categories in Table 2 in which the 48 student participants averaged at least one observation per category were subjected to a 2 (training condition: monologue vs. dialogue) x 2 (tutoring content: hardware vs. software) x 12 (question category) multivariate analysis of variance. This analysis revealed a significant interaction between training condition and question category, $\underline{F}(1,44) = 5.70$, $\underline{p} < .001$. Means for each training condition by question category and the results of follow-up analyses are presented in Table 3. There were significant differences between the monologue and dialogue conditions in eight of the 12 question categories. Students in the monologue condition exhibited significantly more queries in the verification, concept completion, and assertion categories. These queries are relatively shallow, requiring short answers, or indicating the student lacks knowledge. Students in the dialogue condition generated significantly more queries in the example, comparison, interpretation, causal consequence, and enablement categories. Queries in these categories request long answers and at least some of them may reflect deep reasoning (Bloom, 1956). As we noted earlier (*Acquisition*), most of the questions (88%) asked by the virtual tutee were drawn from six categories: comparison, interpretation, causal antecedent, causal consequence, instrumental/procedural, and enablement. The proportion of queries in those six categories was .68 in the dialogue condition (Table 2), but only .48 in the monologue condition. This difference suggests that modeling deep-level questions during acquisition induced students, at least to some extent, to ask the same kinds of questions in transfer.

Graesser and Person (1994) have argued that six of the question categories in Table 1 involve deep reasoning: causal antecedent, causal consequence, goal-orientation, enablement, instrumental/procedural, and expectational. They showed that questions in these categories correlate reasonably well with levels 2-6 of Bloom's taxonomy, $\underline{r} = .64$, $\underline{p} < .05$ (Graesser & Person, 1994, p. 128). In the dialogue condition 50% of the questions were in those categories, compared to 36% in the monologue condition (Table 2). Students in the present study exhibited few queries in the goal-orientation or expectational categories (see Table 2), possibly because few were asked by the virtual tutee (one from each category) during acquisition. Those in dialogue condition did exhibit more queries in the remaining four categories than those in the monologue condition (Table 3), and the difference was significant in two of them.

Table 3. The Mean Number of Questions Generated by each Student in the Monologue and Dialogue Conditions, along with F Values and Significance Levels for the 12 Categories in which the 48 Students Averaged at least One Observation per Category.

Question Category	Means		F	Sig
SHORT ANSWER				
Verification	Monologue	2.54		
	Dialogue	1.04	7.62	.008
Concept completion	Monologue	2.00		
	Dialogue	.79	5.04	.030
Feature Specification	Monologue	1.17		
	Dialogue	1.46	.31	.58
LONG ANSWER				
Definition	Monologue	1.83		
	Dialogue	2.63	1.75	.20
Example	Monologue	.54		
	Dialogue	1.46	4.30	.04
Comparison	Monologue	1.50		
	Dialogue	3.21	6.83	.01
Interpretation	Monologue	1.50		
	Dialogue	4.63	4.11	.05
Causal antecedent	Monologue	2.96		
	Dialogue	3.92	1.43	.24
Causal consequence	Monologue	3.42		
	Dialogue	7.38	6.11	.02
Instrumental/procedural	Monologue	1.13		
	Dialogue	1.58	.84	.37
Enablement	Monologue	1.63		
	Dialogue	3.63	6.08	.02
Assertion	Monologue	3.08		
	Dialogue	.92	4.09	.05

SUMMARY AND CONCLUSIONS

The purpose of this research was to explore vicarious-learning procedures that might increase students' knowledge of the domain and enhance the quality and quantity of their questions in an ITS environment involving computer controlled embodied agents. Students in the dialogue condition, who overheard a discussion in which the virtual tutee asked the virtual tutor a series of 66 questions, learned more, took more turns in mixed-initiative dialogue, and asked more questions than those in the monologue-like condition. Although the research is preliminary, we believe the findings have implications both for ITSs (Anderson et al., 1995; Edelson, 1996; Holland et al., 1995; Lesgold et al., 1992; Scardamalia et al., 1992; Van Lehn, 1990) and for those concerned with improving students' question asking to enhance learning and comprehension (Collins et al., 1980; Davey & McBride, 1986; King, 1989, 1990; King et al., 1998; Hilton, 1990; Palinscar & Brown, 1984; Schank, 1986). On one hand, the generality of the findings beyond college students learning about computer hardware and software remains to be demonstrated. On the other hand, though, most of the queries (68%) in the dialogue condition were drawn from the categories that were most frequently modeled (88%) during acquisition. Thus it seems reasonable to conclude, minimally, that the procedures would generalize to young adults learning in a variety of other content domains.

The ITS under development (Graesser et al., 1998; McCauley et al., 1998; Wiemer-Hastings, Graesser, et al., 1998; Wiemer-Hastings, Wiemer-Hastings, et al., 1999) will be used in college-level computer literacy classes. Results of the current study indicate that getting students to ask the kinds of questions that improve comprehension and learning may not require explicit procedures and multiple training sessions (cf. Davey & McBride, 1986; King, 1989; King et al., 1998; Palinscar & Brown, 1984). Presenting a relatively brief dialogue (30-40 min), like the one used in this study, at the outset of a first tutoring session should enhance both knowledge acquisition and the quality of the subsequent dialogue between the student learner and the animated agent(s) serving in the role of tutor (cf. Aimeur & Fahmi, 1998).

References

Aimeur, E., & Fahmi, M. (1998). Pedagogical agents in DTL: The double learning strategy. *ITS '98 Workshop on Pedagogical Agents.* (pp. 14-19). San Antonio, Texas, August.

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

Bandura, A. (1962). Social learning through imitation. In M. R. Jones (Ed.), *Nebraska symposium of motivation* (pp. 211-269). Lincoln: University of Nebraska Press.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1. Cognitive domain*. New York: McKay.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, 1482-1493.

Buseri, J. C. (1988). Questions in Nigerian science classes. *Questioning Exchange*, 2, 275-280.

Collins, A. (1985). Teaching and reasoning skills. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Vol. 2* (pp. 579-586). Hillsdale, NJ: Erlbaum.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 385-407). Hillsdale, NJ: Erlbaum.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. *Journal of Educational Psychology*, *78*, 256-262.

Derry, S. J., & Potts, M. K. (1998). How tutors model students: A study of personal constructs in adaptive tutoring. *American Educational Research Journal*, *35*, 65-99.

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

Edelson, D. C. (1996). Learning from cases and questions: The socratic case-based teaching architecture. *The Journal of the Learning Sciences*, *5*, 357-410.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes*, *27*, 35-53.

Gavelek, J. R., & Raphael, T. E. (1985). Metacognition, instruction, and the role of questioning activities. In D. L. Forrest-Pressley, G. E. MacKinnin, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance* (Vol. 2, pp. 103-36). Orlando, FL: Academic Press.

Good, T. L., Slavings, R. L., Harel, K. H., & Emerson, H. (1987). Student passivity: A study of question asking in K-12 classrooms. *Sociology of Education*, *60*, 181-199).

Graesser, A. C., & Clark, L. F. (1985). *Structures and processes of implicit knowledge*. Norwood, NJ: Ablex.

Graesser, A. C., Franklin, S., Wiemer-Hastings, P., & The Tutoring Research Group. (1998). Simulating smooth tutorial dialogue with pedagogical value. *Proceedings of the American Association for Artificial Intelligence* (pp. 163-167). Menlo Park, CA: AAAI Press.

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

Graesser, A. C., Person, N. K., & Huber, J. D. (1992). Mechanisms that generate questions. In T. E. Lauer, E. Peacock, & A. C. Graesser (Eds.), *Questions and information systems* (pp. 67-187). Hillsdale, NJ: Erlbaum.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 37l-395.

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & the Tutoring Research Group. (in press). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research*.

Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin*, *107*, 65-81.

Holland, V. M., Kaplan, J. D., & Sams, M. R. (1995). (Eds.),_*Intelligent language tutors*. Hillsdale, NJ: Erlbaum.

Hu, X. (1998). Xtrain 1.0 [computer software]. The University of Memphis: Advanced Learning Technologies Laboratory.

Hu, X., Graesser, A. C., & The Tutoring Research Group. (1998). Using WordNet and latent semantic analysis to evaluate the conversational contributions of learners in tutorial dialogue. *Proceedings of the International Conference on Computers in Education* (Vol. 2, pp.337-341). Beijing, China: Springer.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology*, *14*, 366-381.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*, 664-687.

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology*, *90*, 134-152.

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

Lee, J., Dineen, F., & McKendree, J. (1998). Supporting student discussions: It isn't just talk. *Education and Information Technologies*, *3*, 17-29.

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.

Mayes, T. (1998). Teaching, technology and talk. *Psychologist*, *11*, 375-377.

McCaulley, L., Gholson, B., Hu, X., Graesser, A. C., & The Tutoring Research Group. Delivering smooth tutorial dialogue using a talking head. *Proceedings of the Workshop on Embodied Conversational Characters* (pp. 31-38). Tahoe City, CA: AAAI and ACM.

McKendree, J., Stenning, K., Mayes, T., Lee, J., & Cox, R. (1998). Why observing a dialogue may benefit learning. *Journal of Computer Assisted Learning*, *14*, 110-119.

Microsoft Agent. (June 11, 1998). Microsoft Corporation. <http://www.microsoft.com/intdev/agent>

Moore, J. D. (1994). *Participating in explanatory dialogues: Interpreting and responding to questions in context*. Cambridge, MA: The MIT Press.

Nickel, H., & Fenner, H. J. (1974). Direkte Und indirekte lenkung im unterricht in abhanggigkeit von fachspezifischen und methodisch-didaktisehen variablen sowie alter und geschlecht des lehrers (Direct and indirect classroom control and its dependence upon subject matter, instructional method, age, and sex of teachers. *Zeitschrift fur Entwicklungs-psychologie und Padagogische Psychologie*, *6*, 178-191.

Ohlsson, S. (1995). Learning to do and learning to understand: A lesson and a challenge for cognitive modeling. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science* (pp. 37-62). Oxford: Elsevier Science.

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

Paulsen, M. F. (1995). *Pedagogical Techniques for Computer-Mediated Communication*, <http://www.hs.nki.no/~morten/cmcped.htm>

Renwick, W. (1996). The future of face to face and distance technology in post-secondary education. *Informative Technology and the Future of Post Secondary Education*. Paris: OECD.

Rosenthal, R. L., & Zimmerman, B. J. (1978). *Social learning and cognition*. New York: Academic Press.

Ryan, F. L. (1971). *Exemplars for the new social studies*. Englewood Cliffs, NJ: Prentice-Hall.

Scardamalia, M., Bereiter, C., Brett, C., Burtis, P. J. Calhoun, C., & Smith-Lea, N. (1992). Educational applications of a network communal database. *Interactive Learning Environments*, *2*, 47-71.

Schank, R. C. (1986). *Explanation patterns: Understanding mechanically and creatively*. Hillsdale, NJ: Erlbaum.

Schober, M. F., & Clark, H. H. (1989). Understanding by adressees and overhearers. *Cognitive Psychology*, *21*, 211-232.

Van Lehn, K. (1990). *Mind bugs: The origins of procedural misconceptions*. Cambridge, MA: The MIT Press.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87*, 406-423.

Weber, G. (1996). Episodic learner modeling. *Cognitive Science*, *20*, 195-236.

Wiemer-Hastings, P., Graesser, A. C., Harter, D., & The Tutoring Research Group. (1998). The foundations and architecture of AutoTutor. *Proceedings of the 4th International Conference on Intelligent tutoring Systems* (pp. 334-343). Berlin, Germany: Springer-Verlag.

Wiemer-Hastings P., Wiemer-Hastings, K., Graesser, A. C., & The Tutoring Research Group. (1999). Approximate natural language understanding for an intelligent tutor. *Proceedings of the 12 International Florida Artificial Intelligence Research Society Conference* (pp. 192-196).

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, *1*, 87-95.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, *24*, 345-376.

Appendix A. Example of a Recall Question

Please write an explanation of how the various parts of the central processing unit interact with each other and with random access memory when an application program is running. Pretend you are writing to someone who does not know much about computers, or how they work. You may draw on anything you know about computers that will help you clarify your explanation.

Appendix B. Example Information Delivery and Dialogue

Tutor Information delivery: When you first turn on the computer it loads the operating system. The operating system plays a vital role in how efficiently you can use the computer. Most of what you see on the screen when you use an application program, and most of the common tasks you have the program perform are being performed by the operating system.

Virtual student: Why does it load the operating system, what does that accomplish?

Virtual tutor: The operating system is software that must be running in order for the computer to function.

Virtual student: Does turning on the computer have anything to do with what I've heard referred to as booting up?

Virtual tutor: Yes Sean, when you turn the computer on it immediately begins booting up.

Virtual student: What does it mean to boot up?

Virtual tutor: When you turn on the computer it automatically begins executing instructions that help the system boot. Booting up involves getting the operating system loaded into part of the system's memory.

Virtual student: What happens after the operating system is loaded into memory?

Virtual tutor: Once the operating system is loaded into memory, the computer is ready to respond to requests from the computer user.

Virtual student: Why does the computer need an operating system?

Virtual tutor: The operating system stands as an additional layer of insulation between the user and the computer hardware.

Virtual student: What would happen if you didn't have an operating system?

Virtual tutor: Well, you would have to provide input in terms of the zeros and ones of the binary language the computer hardware can read.

Virtual student: Wow, do you really mean all a computer's hardware can work with is zeros and ones?

Virtual tutor: Yes Sean, what the operating system does is make the computer more user friendly by avoiding the need to program in the binary language.

Virtual student: You said an additional layer of insulation. What other layers of insulation separate the user from the binary language of the hardware?

Virtual tutor: Well let's see, application programs, like word processors, spread sheets, or graphics packages are examples. They provide another layer between the user and the hardware, making computers easier to use, more user friendly.

Virtual student: Do all computers have operating systems?

Virtual tutor: No, the earliest computers were pretty primitive and didn't have operating systems. Instead those using the computer had to set switches and patch wires between switches.

Virtual student: That must have been a real nuisance. How did they get from there to operating systems?

Virtual tutor: Actually, that problem was overcome early in the development of computers.

Virtual student: How was it overcome?

Virtual tutor: John von Meumann laid the theoretical groundwork in 1945, and the first operating systems were being used by the fifties.

Virtual student: How is it that the computer is able to load the operating system as soon as it is turned on?

Virtual tutor: In most computers the operating system is stored on the hard disk. When the computer is turned on automatic commands are immediately executed that locate the operating system on the hard disk and load it into memory.

Virtual student: Where do those automatic commands it automatically executes come from?

Virtual tutor: Those commands are located on read only memory, or ROM. We can talk more about ROM later. For now lets talk more about how the operating system makes it possible for the hardware to read input and give you output. I'll bring up a new picture that includes translator programs.