Student Question-Asking Patterns in an Intelligent Algebra Tutor

Lisa Anthony[1], Albert T. Corbett[1], Angela Z. Wagner[1], Scott M. Stevens[1], Kenneth R. Koedinger[1]

[1] Human Computer Interaction Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15217, USA
{lanthony, corbett, awagner, sms, koedinger}@cs.cmu.edu
http://www.cs.cmu.edu/~alps/

Abstract. Cognitive Tutors are proven effective learning environments, but are still not as effective as one-on-one human tutoring. We describe an environment (ALPS) designed to engage students in question-asking during problem solving. ALPS integrates Cognitive Tutors with Synthetic Interview (SI) technology, allowing students to type free-form questions and receive pre-recorded video clip answers. We performed a Wizard-of-Oz study to evaluate the feasibility of ALPS and to design the question-and-answer database for the SI. In the study, a human tutor played the SI's role, reading the students' typed questions and answering over an audio/video channel. We examine the rate at which students ask questions, the content of the questions, and the events that stimulate questions. We found that students ask questions in this paradigm at a promising rate, but there is a need for further work in encouraging them to ask deeper questions that may improve knowledge encoding and learning.

1 Introduction

Intelligent tutoring environments for problem solving have proven highly effective learning environments [2,26]. These environments present complex, multi-step problems and provide the individualized support students need to complete them: step-by-step accuracy feedback and context-specific problem-solving advice. Such environments have been shown to improve learning one standard deviation over conventional classrooms, roughly a letter grade improvement. They are two or three times as effective as typical human tutors, but only half as effective as the best human tutors [7].

While intelligent problem-solving tutors are effective active *problem-solving* environments, they can still become more effective active *learning* environments by engaging students in active knowledge construction. In problem solving, students can set shallow *performance* goals, focusing on getting the right answer, rather than *learning* goals, focusing on developing knowledge that transfers to other problems (*c.f.,* [10]). Some successful efforts to foster deeper student learning have explored plan scaffolding [18], and self-explanations of problem-solving steps [1]. We are developing an environment intended to cultivate active learning by allowing students to ask open-ended questions. Encouraging students to ask deep questions during problem solving may alter their goals from performance-orientation toward learning-orientation, perhaps ultimately yielding learning gains. Aleven & Koedinger [1] showed that getting students to explain what they *know* helps learning; by extension, getting students to explain what they *don't know* may also help.

In this project, we integrate Cognitive Tutors, a successful problem-solving environment, with Synthetic Interviews, a successful active inquiry environment, to create ALPS, an "Active Learning in Problem Solving" environment. Synthetic Interviews simulate face-to-face question-and-answer interactions. They allow students to type questions and receive video clip answers. While others [4,12,13,21] are pursuing various tutorial dialogue approaches that utilize natural language processing technology, one advantage of Synthetic Interviews over these methods is that their creation may be simpler. A long-term summative goal in this line of research is whether or not this strategy is as pedagogically-effective as it is cost-effective. Before addressing this goal, however, we first must address two important formative system-design goals, which have not been explored in detail in the context of computer tutoring environments: to what extent will students, when given the opportunity, ask questions of a computer tutor to aid themselves in problem solving, and what is the content of these questions? This paper briefly describes the ALPS environment and then focuses on a Wizard-of-Oz study designed to explore these formative issues.

1.1 Cognitive Tutors

Cognitive Tutors are intelligent tutoring systems designed based on cognitive psychology theory and methods, that pose complex, authentic problems to students [2]. In the course of problem solving, students represent the situation algebraically in the worksheet, graph the functions, and solve equations with a symbol manipulation tool. Each Cognitive Tutor is constructed around a cognitive model of the knowledge students are acquiring, and can provide step-by-step accuracy feedback and help. Cognitive Tutors for mathematics, in use in over 1400 US schools, have been shown to raise student achievement one standard deviation over traditional classroom instruction [8].

Cognitive Tutors provide a help button, which effectively answers just one question during problem solving: "What do I do next?" The tutor provides multiple levels of advice, typically culminating in the actual answer. This help mechanism is sufficient for students to solve problems successfully, but may limit student opportunities to engage in active learning. In fact, students can abuse this help system. For instance, Aleven & Koedinger [1] found that 85% of students' help-seeking events in one geometry tutor unit consisted of quickly "drilling down" to the most specific hint level without reading intermediate levels. Answer-seeking behavior like requesting these "bottom-out" hints may be characteristic of an orientation toward near-term performance rather than long-term learning [3].

Cognitive Tutors might be even more effective if they provided the same "learning by talking" interactions as effective human tutors, by supporting active-learning activities like making inferences, elaborating, justifying, integrating, and predicting [6]. The ALPS environment employs *active inquiry* Synthetic Interview technology to open a channel for students to ask questions as the basis of such active-learning activities.

1.2 Synthetic Interviews

The Synthetic Interview (SI) [25] is a technology that provides an *illusion* of a face-to-face interaction with an individual: users ask questions as if they were having a conversation with the subject of the interview. For example, SIs have been created for asking Albert Einstein about relativity and for asking medical professionals about heart murmurs. This simulated dialogue effect is achieved by indexing videotaped answers based on the types of questions one can expect from the users of that particular SI. Users type a question, and the Synthetic Interview replies with a video clip of the individual answering this question. The SI performs this mapping from query to answer via an information retrieval algorithm based on "TFIDF" (term-frequency, inverse document frequency, *e.g.*, [23]). Question-matching occurs statistically based on relative word frequency in the database of known questions and in the user query, rather than through knowledge-based natural-language processing (NLP). Systems using knowledge-based NLP often suffer an implementation bottleneck due to the knowledge engineering effort required to create them [20]. Unlike the reliance of such NLP systems on *explicit* domain knowledge authoring, SIs possess *implicit* domain knowledge via what questions are answered and how. Any given answer has many question formulations associated with it. Several rounds of data collection may be required to obtain a sufficient query-base for the SI algorithm; past SIs have had up to 5000 surface-form-variant questions associated with 200 answers. This need for multiple rounds of data collection is similar to that needed to create other textual classification systems, and on the whole, purely statistical approaches (like Synthetic Interviews) still require less development effort than NLP systems [20].

1.3 ALPS: Active Learning in Problem Solving

The ALPS environment is an adaptation of the Cognitive Tutor to include a Synthetic Interview. The current version is a high school Algebra I lesson covering linear function generation and graphing. In addition to the normal Cognitive Tutor windows, the student sees a web browser pointing to the Synthetic Interview server. This browser shows the video tutor's face at all times, with a text box in which the student may type in a question for the tutor. We hypothesize that formulating questions rather than just pressing a hint button can help engage students in learning and self-monitoring**.**

This paper describes a design study employing a Wizard-of-Oz simulation of the ALPS environment in which a human tutor plays the Synthetic Interview. The study examines how students take advantage of the opportunity to ask open-ended questions in a computer-based problem solving environment, by looking at the following issues: the rate at which students ask questions; the contexts in which students ask questions; the extent to which tutor prompting elicits questions; and the content of student questions with respect to learning- or performance-orientation. These results will help guide design of question-scaffolding in the ALPS environment. The study also serves to collect student questions to populate the ALPS question and answer databases.

2 Student Questions in Other Learning Environments

Past research on question-asking rates in non-computer environments provides reasonable benchmarks for gauging ALPS' usability and effectiveness. Graesser and Person [14] report that, in conventional classroom instruction, the rate of questions per student per hour is 0.11. This extremely low number is due to the fact that students share access to the teacher with 25 to 30 other students, and is also due to the lecture format of typical classroom instruction. At the other extreme, in one-on-one human tutoring, students ask questions at the average rate of 26.5 questions per hour [14]. Of these, 8.5 questions per hour are classified as deep-reasoning questions.

The nature of student questions in *intelligent tutoring systems* is largely unexplored. ITSs that allow natural language student inputs generally embody Socratic tutorial dialogues (*c.f.*, AutoTutor [13], CIRCSIM-Tutor [12], Atlas [11]). By nature, Socratic dialogues are overwhelmingly driven by questions from the tutor. Although there are problem-solving elements in many of these systems, the tutor-student dialogue is both the primary activity and the primary mode of learning. Because Socratic dialogues are tutor-controlled, students in these systems tend to ask relatively few questions. Therefore, these ITSs vary in how fully they attempt to process student questions and question rate and content are largely unreported. A few studies have examined student questions in *computer-mediated* Socratic tutoring, however, in which the student and human tutor communicate through a textual computer interface. In a study by Jordan and Siler [16], only about 3% of (typed) student utterances were questions, and in Core et al [9], only 10% of student moves were questions. Shah et al [24] found that only about 6% of student utterances were questions; students asked 3.0 questions per hour, well below that of human face-to-face tutoring.

In contrast to such tutor-controlled dialogues, the study reported in this paper examines student question-asking in the Cognitive Tutor, a mathematics problem-solving environment with greater *learner* control. The student, not the tutor, is in control of his progress; students work through the problem-solving steps at their own pace. The program provides accuracy feedback for each problem-solving step, but the students must request advice when they encounter impasses. Therefore, we expect that student question-asking rates will be higher in ALPS than in the systems reported above.

Graesser and Person [14], in a study on human tutoring, found a positive correlation between final exam scores and the proportion of student questions during tutoring sessions that were classified as "knowledge-deficit" or "deep-reasoning" utterances. Therefore, we believe that getting students to ask questions, *to the extent that they are asking deep-reasoning questions*, may alter student goals, and yield learning gains.

3 Wizard-of-Oz Design Study

In the Wizard-of-Oz (WOZ) study, a human played the role of the Synthetic Interview while students worked in the Cognitive Tutor. The students were able to type questions in a chat window and receive audio/video responses from the human tutor (Wizard). Our research questions concerned several characteristics of the questions students might ask: (1) **Frequency**—at what rate do students ask questions to deepen their knowledge; (2) **Prompting & Timing**—what elicits student questions most; and (3) **Depth**—what learning goals are revealed by the content of student questions.

3.1 Methods

Participants. Our participants were 10 middle school students (nine seventh graders, one eighth grader; eight males, two females) from area schools. Two students had used the standard Cognitive Tutor algebra curriculum in their classrooms that year, three students had been exposed to Cognitive Tutors in a previous class session, and five had never used Cognitive Tutors before.

Procedure. The study took place in a laboratory setting. The students completed algebra and geometry problems in one session lasting one and a half hours. During a session, the student sat at a computer running the Cognitive Tutor with a chat session connected to the Wizard, who was sitting at a computer in another room. The students were instructed to direct all questions to the Wizard in the other room via the chat window. In a window on his own computer screen, the Wizard could see the student's screen and the questions the student typed. The Wizard responded to student questions via a microphone and video camera; the student heard his answer through the computer speakers and saw the Wizard in a video window onscreen. Throughout problem solving, if the student appeared to be having difficulty (e.g., either he made a mistake on the same problem-solving action two or more times, or he did not perform any problem-solving actions for a prolonged period), the Wizard prompted the student to ask a question by saying "Do you want to ask a question?"

Measures. The data from the student sessions were recorded via screen capture software. All student mouse and keyboard interactions were captured, as well as student questions in the chat window and audio/video responses from the Wizard. The sessions were later transcribed from the captured videos. All student actions were marked and coded as "correct," "error," "typo," or "interrupted" (when a student began typing in a cell but interrupted himself to ask a question). Student utterances were then separately coded by two of the authors along three dimensions based on the research questions mentioned above: initiating participant (student or tutor); question timing in the context of the problem-solving process (*i.e.*, before or after errors or actions); and question depth. After coding all 10 sessions along the three criteria, the two coders met to resolve any disagreements. Out of 431 total utterances, disagreement occurred in 12.5% of items; the judges discussed these to reach consensus.

3.2 Qualitative Results and Discussion

We classified each problem-solving question at one of the following three depths: answer-oriented, process-oriented, or principle-oriented. Answer-oriented questions can be thought of as "what" questions. The student is asking about the problem-solving process for a particular problem, usually in very specific terms and requesting a very specific answer (*e.g.*, "what is the area of this triangle [so I can put it in the cell]?"). Process-oriented questions can be thought of as "how" questions. The student is asking how to perform a procedure in order to solve a particular problem, but the question represents a more general formulation of the request than simply asking for the answer (*e.g.*, "how do I figure out the area of this triangle?"). Principle-oriented questions can be thought of as "why" questions and are of the most general type. The student is asking a question about a mathematical concept or idea which he is trying to understand (*e.g.*, "Why is the area of a triangle 1/2*b*h?") These three categories form a continuum of question depth, with answer-oriented lying at the shallow end of knowledge-seeking, principle-oriented lying at the deep end, and process-oriented lying somewhere in the middle. We include here an illustrative example from the WOZ of interaction sequences from each category. In each example, input from the student is denoted with **S** and the Wizard, with **W**.

Answer-oriented: These questions ask about the answer to a problem step or about a concrete calculation by which a student may try to get the answer. The following interaction occurred in a problem asking about the relationship among pay rate, hours worked and total pay. An hourly wage of "$5 per hour" was given in the global problem statement, and the student was answering the following question in the worksheet: "You normally work 40 hours a week, but one particular week you take off 9 hours to have a long weekend. How much money would you make that week?" The student correctly typed "31" for the number of hours worked, but then typed "49" (40 + 9) for the amount of money made. When the software turned this answer red, indicating an error, the student asked, "Would I multiply 40 and 9?" The Wizard asked the student to think about why he picked those numbers. The student answered, "Because they are the only two numbers in the problem."

Asking "Would I multiply 40 and 9?" essentially asks "Is the answer 360?" The student wants the Wizard to tell him if he has the right answer, betraying his performance-orientation. The student is employing a superficial strategy: trying various operators to arithmetically combine the two numbers ("40" and "9") that appear in the question. After the first step in this strategy (addition) fails, he asks the Wizard if multiplication will yield the correct answer (he likely cannot calculate this in his head). Rather than ask how to reason about the problem, he asks for the answer to be given to him.

Process-oriented: These student questions on how to find an answer frequently take the form of "how do I find…" or "how do I figure out…" The following occurred when a student was working on a geometry problem involving the area of a 5-sided figure composed of a rectangle plus a triangle. He had already identified the given information in the problem and was working on computing each subfigure's area. He typed "110" for the area of the rectangle and asked, "How do you find the area of a triangle?" The Wizard told him the general formula. In this case, the student correctly understood what he was supposed to compute, but did not know the formula. He is not asking to be told the answer, but instead *how to find* it. The Wizard's general answer can then help the student on future problems.

Principle-oriented: General principle-oriented questions show when the student is moving beyond the current problem context and reasoning about the general mathematical principles involved. We saw only one example of this type of question. It took place after the student had finished computing the area and perimeter of a square of side length 8 (area = 64, perimeter = 32). The student did not need help from the Wizard while solving this problem. He typed "2s+2s" for the formula of a square's perimeter, and typed "s*s" for the formula of a square's area. He then asked, "Is area always double perimeter?" The student's question signified a reflection on his problem-solving activities that prompted him to make a potential hypothesis about mathematics. A future challenge is to encourage students to ask more of these kinds of questions, actively engaging them in inquiry about domain principles.

3.3 Quantitative Results & Discussion

Figures 1, 2, and 3 show the results from the analysis along three dimensions: initiating participant, question timing, and question depth. Error bars in all cases represent the 95% confidence interval. Figure 1 shows the mean number of utterances per student per hour that are prompted, unprompted, or part of a dialogue. "Unprompted" (M=14.44, SD=7.07) means the student asked a question without an explicit prompt by the tutor. "Prompted" (M=3.49, SD=1.81) means the student asked after the Wizard prompted him, as in by saying "Do you want to ask a question?" "Dialogue response" (M=11.80, SD=12.68) means the student made an utterance in direct response to a question or statement by the Wizard, and "Other" (M=8.23, SD=5.04) includes statements of technical difficulty or post-problem-solving discussions initiated by the Wizard. The latter two categories are not included in further analyses.

Figure 1 shows that students asked questions at a rate of 14.44 unprompted questions per hour. Students ask approximately four times more unprompted than prompted questions ($t(18)=4.74$, $p<.01$). The number of prompted questions is bounded by the number of prompts from the Wizard, but note that the number of Wizard prompts per session (M=9.49, SD=2.65) significantly outnumbers the number of prompted questions ($t(18)=5.92$, $p<.01$). Even when the Wizard *explicitly* prompts students to ask questions, they often do not comply. This suggests that a question-encouraging strategy in ALPS simply consisting of prompting will not be sufficient.

Figure 2 shows question timing with respect to the student's problem-solving actions. "Before Action" (M=8.62, SD=6.26) means the student asked the question about an action he was about to perform. "After Error" (M=8.46, SD=2.55) means the student asked about an error he had just made or was in the process of resolving. "After Correct Action" (M=0.85, SD=1.26) means the student asked about a step he had just answered correctly. The graph shows that students on average ask significantly fewer questions after having gotten a step right than in the other two cases ($t(28)=5.09$, $p<.01$), revealing a bias toward treating the problem-solving experience as a performance-oriented task. Once they obtain the right answer, students do not generally reflect on what they have done. This suggests that students might need encouragement after having finished a problem to think about what they have learned and how the problem relates to other mathematical concepts they have encountered.

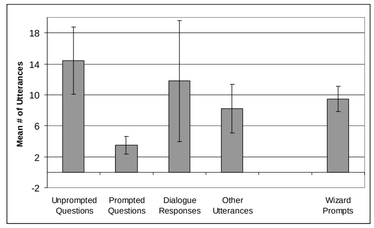

Fig. 1. Mean number of utterances per hour

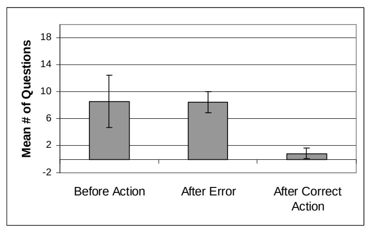

Fig. 2. Mean number of unprompted and prompted questions per hour by question timing

Figure 3 shows the mean number of questions grouped by question topic. "Interface" (M=10.21, SD=5.60) means the question concerned how to accomplish something in the software interface or how to interpret something that happened in the software. "Definition" (M=0.97, SD=1.09) questions asked what a particular term meant. "Answer" (M=4.98, SD=3.58), "Process" (M=1.68, SD=1.60), and "Principle" (M=0.07, SD=0.23) questions are defined above. Figure 3 shows an emphasis on interface questions; although one might attribute the high proportion of student interface questions to the fact that half the participants were students who had not used the Cognitive Tutor software before, the data show no reliable difference between the two groups in question rate or content. Yet even among non-interface questions, one can see that students still focus on "getting the answer right," as shown by the large proportion of answer-oriented questions. The difference between the number of "shallow" questions (answer-oriented) and the number of "deep" questions (process-oriented plus principle-oriented) is significant ($t(28)=4.55$, $p<.01$).

While Figure 2 shows that students on average ask questions before actions and after errors at about the same rate, the type of question asked varies across the two contexts. The distinction between the distributions of these two question contexts may be revealing: asking a question before performing an action may imply forethought and active problem solving, whereas asking only after an error could imply that the student was not thinking critically about what he understood. Figure 4 displays a breakdown of the interaction between question timing and the depth or topic. Based on the data, when students ask questions before performing an action, they are most likely to be asking about how to accomplish some action in the interface which they are intending to perform. When they ask questions after an error, they are most often asking about how to get the answer they could not get right on their own. The one principle-oriented question was asked after a correct action and is not represented in Figure 4.

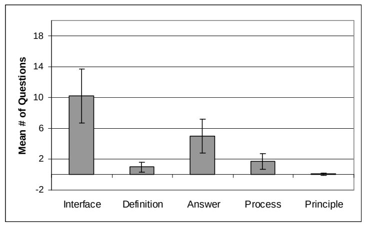

Fig. 3. Mean number of unprompted or prompted questions per hour by perceived depth

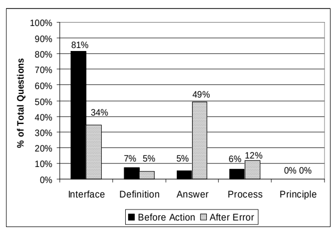

Fig. 4. Comparison of distributions of “Before Action” and “After Error” questions based on question depth. “After Correct Action” is not included due to low frequency of occurrence

Additional analysis shows that, of the questions that are “After Error” (102 total), 100% are directly about the error that the student has just made or is in the process of resolving (*i.e.*, through several steps guided by the Wizard). Of those that are “After Correct Action” (9 total), 4 (44%) are requests for feedback about progress (*e.g.*, “am I doing ok so far?”), 4 (44%) are clarifications about how the interface works (*e.g.*, “can I change my answers after I put them in?”) and only one (11%) is a process- or principle-oriented query about general mathematics (*e.g.*, “is area always double perimeter?”). Thus it seems that, although students do take the opportunity to ask questions, they do not generally try to elaborate their knowledge by asking deep questions.

4 Current and Future Work

Database Seeding: A Preliminary ALPS Pilot. The Wizard-of-Oz study was also designed to populate the ALPS question and answer databases. The ten students generated 208 total questions variations, for which we recorded 47 distinct video clip answers. Recently we conducted a preliminary pilot of the ALPS environment in which five middle school students used ALPS at home. The Synthetic Interview technology processed student questions and presented video clip answers. The five students asked 23 total questions in about 100 minutes total use; all are effectively "unprompted," as the pilot system was not capable of prompts like those in the Wizard-of-Oz study. Students in the pilot asked 12.94 questions per student per hour, slightly lower than the unprompted question rate observed in the WOZ.

A concern has been the clear tendency of the students in the WOZ toward engaging the human Wizard in dialogues, especially when trying to repair errors. However, as Nass and Reeves showed, people treat computers like they treat people [19], implying that the kinds of interactions we will see with the SI-enabled system will be similar to those in the WOZ. A point in favor of this view is that the unprompted question-asking rates reported in our pilot with the computer SI, are similar to those in the WOZ with the human Wizard. Therefore, we do not believe that applying the results from the WOZ to the SI is problematic.

Question-Asking Rate and Content. Students in the Wizard-of-Oz study asked 14.44 unprompted questions per hour. The Wizard's prompts to ask questions yielded an additional 3.49 questions per hour, bringing the question-asking rate to about 2/3 of that observed with human tutors. However, the 1.75 deep questions (process- and principle-oriented questions) that students asked is only about 1/5 the rate observed with human tutors. Hausmann and Chi [15] report a similar result for a computer-mediated self-explanation environment in which students read instructional text and typed self-explanations of the text as they read. In this environment students typed superficial paraphrases of the text sentences at a far higher rate than deeper self-explanations of the sentences, and self-explanations were generated at a far lower rate than in earlier studies of spoken self-explanations [5].

Increasing the rate of deep questions in the ALPS environment is an important challenge. Hausmann and Chi suggest that the additional cognitive load of typing versus spoken input may inhibit students' self-explanation rate. They did succeed in raising students' self-explanation rate somewhat in the computer-mediated environment with content-free prompts designed to elicit explanations, for instance, "Could you explain how that works?" By analogy the first step in raising the rate of deep questions in the ALPS environment may be to replace the generic Wizard prompt ("Do you want to ask a question") with an analogous prompt designed to elicit deeper questions, such as "Do you want to ask how to find this answer?" In the long run, the integration of a speech recognizer that allows students to ask questions orally may be necessary to achieve the highest rate of deep questions, but we plan first to explore several types of question scaffolding strategies.

First, prior instruction on how to structure deep questions can be designed. It has been shown that training students to self-explain text when working on their own by asking themselves questions improves learning [22]. By analogy, training students on how to ask questions of a tutor may be effective in ALPS. Second, it may be possible to progressively scaffold question-asking by initially providing a fixed set of appropriate questions in menu format, and later providing direct feedback and advice on the questions students ask. It may also be possible to capitalize on shallow questions students ask as raw material for these scaffolds; the system could suggest several ways in which a student question is shallow and could be generalized. Finally, it may be useful to emphasize post-problem review questions as well as problem-solving questions. Katz and Allbritton [17] report that human tutors often employ post-problem discussion to deepen understanding and facilitate transfer. Since students do not have active performance goals at the conclusion of problem solving, it may be an opportune time not just to invite, but to actively encourage and scaffold deeper questions.

5 Conclusions

The Wizard-of-Oz study allowed us to evaluate ALPS' viability and identify design challenges in supporting active learning via student-initiated questions. The study successfully demonstrated that students ask questions in the ALPS environment at a rate approaching that of one-on-one human tutoring. However, based on student question content, we can conclude that students are still operating with performance goals rather than learning goals. It may be that the students did not know how to ask deep questions, or that the question-asking experience was too unstructured to encourage deep questions. There may be ways in which we can promote learning goals, including using prompts specifically designed to elicit deeper questions, implementing various deep-question scaffolds, encouraging reflective post-problem discussions, and adding a speech recognizer to reduce cognitive load.

Acknowledgments

Supported in part by National Science Foundation (NSF) Grant EIA0205301 "ITR: Collaborative Research: Putting a Face on Cognitive Tutors: Bringing Active Inquiry into Active Problem Solving." Thanks for support and effort: Brad Myers, Micki Chi, Sharon Lesgold, Harry Ulrich, Chih-Yu Chao; Timm Mason, Pauline Masley, Heather Frantz, Jane Kamneva, Dara Weber; Alex Hauptmann; Carolyn Penstein Rosé, Nathaniel Daw and Ryan Baker. Thanks very much to our ALPS video tutor Bill Hadley.

References

1. Aleven, V.A.W.M.M., Koedinger, K.R.: An Effective Metacognitive Strategy: Learning by Doing and and Explaining with a Computer-Based Cognitive Tutor. Cognitive Science 26 (2002) 147–149
2. Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R.: Cognitive Tutors: Lessons Learned. Journal of the Learning Sciences 4 (1995) 167 –207
3. Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z.: Off-task Behavior in the Cognitive Tutor Classroom: When Students "Game the System." Proc. CHI (2004) to appear
4. Carbonell, J.R.: AI in CAI: Artificial Intelligence Approach to Computer Assisted Instruction. IEEE Trans. on Man-Machine Systems 11 (1970) 190–202
5. Chi, M.T.H., DeLeeuw, N., Chiu, M.-H., LaVancher, C.: Eliciting Self-explanations Improves Understanding. Cognitive Science 18 (1994) 439–477

6. Chi, M.T.H., Siler, S.A., Jeong, H., Yamauchi, T., Hausmann, R.G.: Learning from Human Tutoring. Cognitive Science 25 (2001) 471–533
7. Corbett, A.T.: Cognitive Computer Tutors: Solving the Two-Sigma Problem. Proc. User Modeling (2001) 137–147
8. Corbett, A.T., Koedinger, K.R., Hadley, W.H.: Cognitive Tutors: From the Research Classroom to All Classrooms. In: P. Goodman (ed.): Technology Enhanced Learning: Opportunities for Change. L. Erlbaum, Mahwah New Jersey (2001) 235–263
9. Core, M.G., Moore, J.D., Zinn, C.: Initiative in Tutorial Dialogue. ITS Wkshp on Empirical Methods for Tutorial Dialogue Systems (2002) 46–55
10. Elliott, E.S., Dweck, C.S.: Goals: An Approach to Motivation and Achievement. Journal of Personality and Social Psychology 54 (1988) 5–12
11. Freedman, R.: Atlas: A Plan Manager for Mixed-Initiative, Multimodal Dialogue. AAAI Wkshp on Mixed-Initiative Intelligence (1999)
12. Freedman, R.: Degrees of Mixed-Initiative Interaction in an Intelligent Tutoring System. AAAI Symposium on Computational Models for Mixed-Initiative Interaction (1997)
13. Graesser, A., Moreno, K.N., Marineau, J.C., Adcock, A.B., Olney, A.M., Person, N.K.: AutoTutor Improves Deep Learning of Computer Literacy: Is it the Dialog or the Talking Head? Proc. AIEd (2003) 47–54
14. Graesser, A.C., Person, N.K.: Question Asking During Tutoring. American Educational Research Journal 31 (1994) 104–137
15. Hausmann, R.G.M., Chi, M.T.H.: Can a Computer Interface Support Self-explaining? Cognitive Technology 7 (2002) 4–14
16. Jordan, P., Siler, S.: Student Initiative and Questioning Strategies in Computer-Mediated Human Tutoring Dialogues. ITS Wkshp on Empirical Methods for Tutorial Dialogue Systems (2002)
17. Katz, S., Allbritton, D.: Going Beyond the Problem Given: How Human Tutors Use Post-Practice Discussions to Support Transfer. Proc. ITS (2002) 641–650
18. Lovett, M.C.: A Collaborative Convergence on Studying Reasoning Processes: A Case Study in Statistics. In: Carver, S., Klahr, D. (eds.): Cognition and Instruction: Twenty-five Years of Progress. L. Erlbaum, Mahwah New Jersey (2001) 347–384
19. Reeves, B., Nass, C.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and & Places. Cambridge University Press, Cambridge UK (1996)
20. Rosé, C.P., Gaydos, A., Hall, B.S., Roque, A., VanLehn, K.: Overcoming the Knowledge Engineering Bottleneck for Understanding Student Language Input. Proc. AIEd (2003) 315–322
21. Rosé, C.P., Jordan, P., Ringenberg, M., Siler, S., VanLehn, K., Weinstein, A.: Interactive Conceptual Tutoring in Atlas-Andes. Proc. AIEd (2001) 256–266
22. Rosenshine, B., Meister, C., Chapman, S.: Teaching Students to Generate Questions: A Review of the Intervention Studies. Review of Educational Research 66 (1996) 181–221
23. Salton, G., Buckley, C.: Term Weighting Approaches in Automatic Text Retrieval. Technical Report #87-881, Computer Science Dept, Cornell University, Ithaca, NY (1987)
24. Shah, F., Evens, M., Michael, J., Rovick, A.: Classifying Student Initiatives and Tutor Responses in Human Keyboard-to-Keyboard Tutoring Sessions. Discourse Processes 33 (2002) 23–52
25. Stevens, S.M., Marinelli, D.: Synthetic Interviews: The Art of Creating a ‘Dyad’ Between Humans and Machine-Based Characters. IEEE Wkshp on Interactive Voice Technology for Telecommunications Applications (1998)
26. VanLehn, K., Lynch, C., Taylor, L., Weinstein, A., Shelby, R., Schulze, K., Treacy, D., Wintersgill, M.: Minimally Invasive Tutoring of Complex Physics Problem Solving. Proc. ITS (2002) 367–376