AutoTutor Improves Deep Learning of Computer Literacy: Is it the Dialog or the Talking Head?

Arthur C Graesser, Kristen Moreno, Johanna Marineau
Department of Psychology and the Institute for Intelligent Systems
University of Memphis
202 Psychology Building, Memphis, TN 38152-3230, USA
{a-graesser, kmoreno,jmarinea}@memphis.edu

Amy Adcock
College of Education and the Institute for Intelligent Systems
University of Memphis
Ball Hall 307A, Memphis, TN 38152 USA
aadcock@memphis.edu

Andrew Olney
Department of Computer Science and the Institute for Intelligent Systems
University of Memphis
202 Psychology Building, Memphis, TN 38152-3230, USA
aolney@memphis.edu

Natalie Person
Department of Psychology
Rhodes College, Memphis, TN 38112, USA
person@rhodes.edu

Abstract. AutoTutor is a tutoring system that helps students construct answers to deep-reasoning questions by holding a conversation in natural language. AutoTutor delivers its dialog moves with an animated conversational agent whereas students type in their answers via keyboard. We conducted an experiment on 81 college students who learned topics on computer literacy (hardware, operating systems, internet) with AutoTutor or control conditions, and were assessed on learning gains. There was an experimental design that allowed us to assess the impact of learning condition (AutoTutor, read-text control, versus nothing) and the medium of presenting AutoTutor's dialog moves (print only, speech only, talking head, versus talking head + print). All versions of AutoTutor improved performance in assessments of deep learning, but not shallow learning. Effects of the medium were more subtle, which suggests that it is the message (the dialog moves of AutoTutor) that is more important than the medium.

1. Introduction

AutoTutor is an intelligent tutoring system that holds dialogues with students in natural language [13, 15, 17]. AutoTutor is not merely an information delivery system, but serves as a discourse scaffold that responds appropriately to the students' conversational turns and coaches the students in actively constructing knowledge. AutoTutor presents challenging questions from a curriculum repository. A good response would have about a paragraph of information (3-7 sentences) because these questions are selected to elicit explanations, justifications, procedures, elaborations, comparisons, and contrasts. When college students answer one of these questions, however, they typically type in only 1 or 2 sentences even though they have more knowledge that is relevant to an answer. This is where AutoTutor is particularly helpful. AutoTutor engages in a mixed initiative dialog that assists in the evolution of an improved answer and that draws out more of what the students know. For example, AutoTutor gives feedback to the student on what the student types in (positive, neutral, negative feedback), pumps for information ("What else?"), prompts for the student to fill in missing words, gives hints, fills in missing information with assertions, corrects misconceptions, answers students questions, and summarizes answers. A complete answer is eventually constructed during this collaborative dialog, which typically extends to about 50 conversational turns.

The design of AutoTutor was inspired by explanation-based constructivist theories of learning [1, 7, 28] and by previous empirical research that has documented the collaborative constructive activities that routinely occur during human tutoring [8, 14, 26]. Learners actively construct explanations and elaborations of the learning material by interacting with the world and other people. These active constructive activities allegedly produce better learning than merely presenting information to students. Human tutors are excellent collaborators in these constructive efforts because they guide the students in productive constructive processes and simultaneously respond to what the student's information needs are. Rather interestingly, the dialog moves of most human tutors are not particularly sophisticated from the standpoint of today's pedogogical theories and intelligent tutoring systems. They normally coach the student in filling in missing pieces of information in an expected answer and they fix bugs and misconceptions that are manifested by the student during tutoring. The argument has been made [14] that it is conversational properties of tutorial dialog, not sophisticated tutoring tactics, that explain why human tutors facilitate learning. Therefore, AutoTutor was designed to simulate the dialog moves of human tutors.

Empirical studies have confirmed that one-on-one tutoring is a powerful method of promoting knowledge construction. The vast majority of the tutors in these studies have had moderate domain knowledge and little or no training in pedogogy or tutoring; they were peer tutors, cross-age tutors, or paraprofessionals, but rarely accomplished tutors. Unaccomplished human tutors enhance learning with an effect size of .4 standard deviation units (called sigma's), which translates to approximately a half a letter grade [11]. Accomplished human tutors produce effect sizes of 2 sigma according to Bloom [5], although there are reasons for questioning the magnitude of this affect because precious few studies have assessed learning gains with accomplished tutors. By way of comparisons, intelligent tutoring systems produce effect sizes of approximately 1 sigma [12, 29]. So how does AutoTutor compare? Four previous experiments on AutoTutor have produced gains of .4 to 1.5 sigma (a mean of .7), depending on the learning measure, the comparson condition, the subject matter, and version of AutoTutor [13, 23, 27]. This would place AutoTutor somewhere between an unaccomplished human tutor and an intelligent tutoring system with sophisticated pedagogical tactics but no natural language dialog.

We are currently exploring what it is about AutoTutor that explains its facilitation of learning [27]. One possible explanation is AutoTutor's dialog mechanisms. AutoTutor knows what to say next, as it composes its turns in response to learner contributions and in service of its pedagogical goals. These dialog moves somehow help learners construct better explanations and elaborations of the material. Our tests of the quality of AutoTutor's dialog moves have indeed been very encouraging. When experts have rated AutoTutors dialog moves on conversatioal smoothness and pedagogical quality, the mean ratings lean more to the good side than the bad side [24]. We have also conducted experiments that use a bystander Turing test for assessing AutoTutor's conversational quality [22]. The bystander Turing test proceeds by periodically stopping AutoTutor and recording the printed transcripts of the conversation. Expert human tutors subsequently view these transcripts and generate the next turn (N+1) of the dialog after reading turns 1 to N. We can thereby compare what AutoTutor verus a human tutor generates for turn N+1. The subsequent step in the methodology was to have college students (bystanders) read the transcripts (turns 1 to N+1) and rate whether N+1 was generated by a human or a computer; half of these had been generated by a human and half by AutoTutor. We performed signal detection analyses on the bystander's decisions by computingh hit rates, false alarm rates and *d*' discrimination scores. We found that the mean *d'* score was zero, which means that bystanders are entirely unable to discriminate whether a dialog move was generated by a computer or a human. These results support the claims that AutoTutor is a good simulation of a human tutor and that it is feasible to to build a conversation-based tutoring system that students will use. And more to the point of the present study: perhaps it is the dialog mechanisms that account for AutoTutor's facilitation of learning. This will be called the **dialog facilitation** hypothesis.

An alternative hypothesis, called the **media facilitation** hypothesis, is that learning is explained by the media through which the dialog moves are expressed. Simply put, it is the medium, not the message that is important in these dialogs. AutoTutor has always delivered its dialog moves with an animated conversational agent. Animated conversational agents have facial features synchronized with speech and in some cases appropriate gestures [6,10, 18]. The computer controls the eyes, eyebrows, mouth, lips, teeth, tongue, cheekbones, and other parts of the face in a fashion that is ideally meshed appropriately with the speech and gestures. The text-to-speech engines can articulate any text that gets passed to the agent, with timing and intonation that is surprisingly similar to human speech. The agents provide an anthropomorphic human-computer interface that simulates having a conversation with a human.

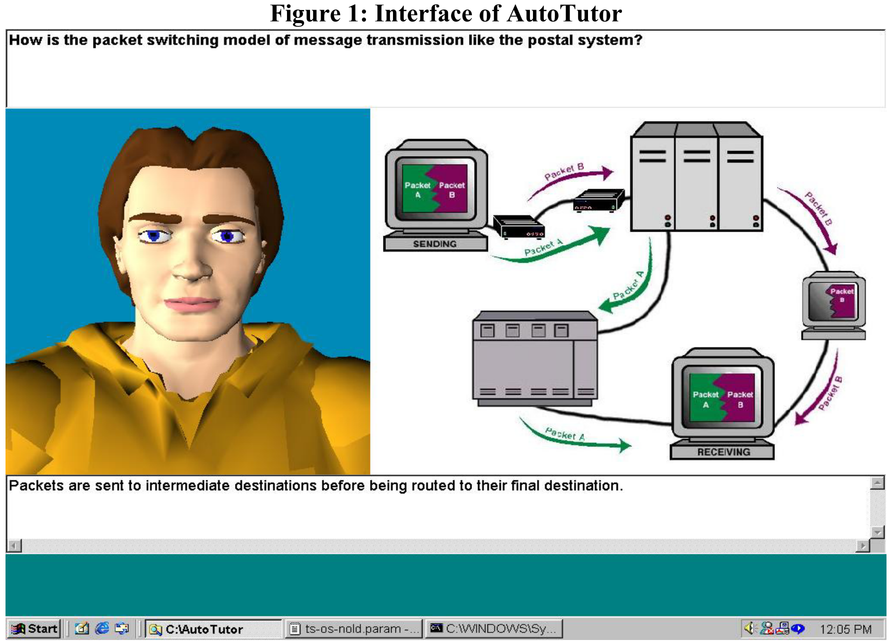
In recent years, educational researchers have explored how these conversational agents can be effectively integrated with learning environments. There is some evidence that these agents have a positive impact on learning or on the learner's perceptions of the learning experience when subject matter is presented with the agents, compared with speech alone or text controls [2, 3,9,20, 30]. In contrast, null effects were found when the agent served as a navigational guide for using a web site [16]. However, additional research is needed to determine the precise conditions, agent features, and levels of representation that are associated with learning gains. One rather provocative result is that there is a near zero correlation between learning gains and how much the students like the conversational agents [19]. Therefore, it is important to distinguish liking from learning in this area of research.

The present study conducted an experiment that tested the dialog facilitation hypothesis and the media faciulitation hypothesis. College students learned about topics in compute literacy either by **AutoTutor** or by reading text on the same content (**Read-text**); these conditions were compared to a **Control** condition in which they did not receive any materal on computer literacy. There were four different media conditions that delivered AutoTutor's dialog moves: **Print** only, **Speech** only, **Talking Head** (with speech, face, and some gesture), and **Talking Head+Print**. Learning gains were assess be administering a pretest and a post test that tapped shallow knowledge and deep knowledge. Our design allowed us to trace the independent contributions of AutoTutor's dialog facilities and the alternative media on learning.

2. A Brief Sketch of AutoTutor

Figure 1 shows the interface of AutoTutor. The major question is selected and presented at the top window, in this case “How is the message switching model of message transmission like a postal system?” This question remains at the top of the web page until it is finished being answered by a multiturn dialog between the learner and AutoTutor. The students type in their contributions for each turn via keyboard, with the content being reflected at the bottom window. Diagrams accompany some questions and are presented at the right. The talking head in the left window speaks the content of AutoTutor’s turns during the process of answering the main question. The speech engine is Microsoft Agent. The present study manipulates the media that presents the content of AutoTutor’s turn (i.e., print, sound alone, talking head, vs. Talking head with print).

Figure 1: Interface of AutoTutor



AutoTutor’s mechanisms have been described in previous publications [3 of them] so a succinct description of its key components and its functionality should be adequate. AutoTutor has a set of permanent databases that do not get updated throughout the course of tutoring. These include (a) the curriculum repository of questions, answers, diagrams, and associated dialog moves, (b) lexicons, syntactic parsers, and other computational linguistics modules, (c) a gloassary of technical terms and their definitions, (d) a corpus of documents, including a text book on computer literacy and several articles, and (e) latent semantic analysis vectors for words, curriculum content, and the document corpus. AutoTutor uses LSA as the backbone for representing world knowledge about computer literacy, or any other subject matter that is tutored, such as conceptual physics [21, 27].

The quality of student answers is successfully computed from the Assertions expressed during the student’s turns. Assertions of high quality have (a) a high cosine match with expectations (sentences that are part of the ideal answer) in the curriculum content for a question and (b) a low match with anticipated bad answers. LSA is successful only to the extent it compares student verbal input with the expected information in the curriculum repository. In this sense, it is a statistical pattern matcher that attempts to capture world knowledge. LSA is totally useless for interpreting verbal input from scratch.

AutoTutor has a set of processing modules and dynamic storage units that maintain qualitative content and quantitative parameters. These storage registers are frequently updated as the tutoring process proceeds. For example, AutoTutor keeps track of student ability (as evaluated by LSA from student Assertions), student initiative (such as the incidence of student questions), student verbosity (number of words per turn), and the progress in having a question answered by virtue of the the dialog history. The dialog management module of AutoTutor flexibly adapts to the student by virtue of these parameters, so no two conversations with AutoTutor are ever the same. The dialog management module has an aumented finite state network, a set of fuzzing production rules, and a special algorithm for selecting dialog moves to help fill in missing information in an ideal answer. Other processing modules execute other imortant functions: speech act classification, linguistic information extraction, evaluation of student assertions, speech production with the animated conversational agent, and others which need be described in this report.

AutoTutor was written in Java, resides on a Pentium computer server, and is delivered on the web.

3. An Experiment that Evaluates whether Students Learn from AutoTutor

3.1 Methods

Students enrolled in a computer literacy course were tutored with AutoTutor in an experiment that lasted 2-3 hours. There was a pretest phase, a training phase, and a posttest phase. Both the pretest phase and the posttest phase had a sample of multiple choice questions on computer literacy. There were two versions of the multiple choice test (version A vs. B); assignment of version to pretest vs. posttest was counterbalanced across subjects. Each multiple choice had a sample of questions that were classified as **shallow** questions according to Bloom's taxonomy of cognitive difficulty [4] and another sample of **deep** questions. Shallow questions consisted of definitions, properties, and examples of technical components. For example, one shallow question was "Which of the following is only a sequential access mechanism? (a) RAM, (b) optical disk, (c) tape, (d) hard disk." Deep questions tapped causal antecedents and consequences of events, and methods of solving problems. An example deep question is "You buy a new graphics program but it will not run on your computer, so what is the best solution? (a) partition your hard drive, (b) delete unnecessary files, (c) increase RAM, (d) return the program to the store." There were an equal number of multiple choice questions for each of three topics: Hardware, Operating System, and the Internet. Each version (A,B) of each of the three topics had 6 shallow questions and 24 deep questions (180 questions total). There also was a **cloze** task at the posttest phase only. The cloze task presented the ideal answers to all of the questions covered by AutoTutor in the tutoring sessions, but with content words in the ideal answers randomly deleted. The student's task was to to fill in these content words. There were 12 ideal answers each for Hardware, Operating Systems, and Internet. Performance on these three tests (shallow, deep, cloze) was measured as the proportion of correct responses.

During the training phase, one of the three topics was assigned to the **AutoTutor** condition, a second to a **Read-text** condition, and the other to a read-nothing **Control** condition. The assignment of topics (Hardware, OS, Internet) to condition was counterbalanced across students. So every student learned from AutoTutor for one of the topics, read excerpts from a textbook on another topic, and did nothing for the remaining topic. The excerpts from the textbook focused on those paragraphs and pages that were directly relevant to the questions covered by AutoTutor. We did this as an attempt to achieve information equivalence between the textbook content and the AutoTutor content. We of course expected learning to be higher for AutoTutor than the other two training conditions. Orthogonal to this training variable, we manipulated the media of the dialog moves presented by AutoTutor. College students were randomly assigned to one of four media conditions: (1) **Print** (the tutors content for each turn was presented as text in the left window), (2) **Speech** (the tutors content for each turn was spoken with the Microsoft Agent voice, (3) **Talking Head** (TH, the face and speech used in our previous studies of AutoTutor), and (4) Talking Head + Print (THP, which included the face, speech and the print in a bubble above the talking head). Although 96 college students originally participated (24 in each media condition), only 81 completed all phases of the experiment: Print (N=22), Speech (19), TH (21), and THP (19).

3.2 Results

An analysis of the pre-test scores indicated the students in the four media conditions started out on an even playing field before the training sessions. The proportion of correct responses did not significantly differ for shallow questions among the four media conditions, with means of .41, .42, .45, and .43 in the Print, Speech, TH, and THP condition, respectively. The corresponding standard deviations were .21, .22, .21, and .20. Similarly the proportion scores for deep questions did not significantly differ, with means of .40, .41, .38, and .40 and with SDs of .15, .14, .13, and .18, respectively. The cloze task was not administered at pretest.

The results of the posttest were analyzed by an analysis of variance (ANOVA) with a factorial design that had two independent variables: training (AutoTutor, Read-text, Control) and media (Print, Speech, TH, THP). A separate ANOVA was conducted on each of the three dependent measures: shallow questions, deep questions, and cloze.

The results of the shallow test were very simple to interpret. Scores on the shallow posttest were not significantly affected by the training condition, by the media condition, or by a training X media interaction. For example, the main effect means for training condition were .41,.41, and .40 for the AutoTutor, Read-text, and Control conditions, respectively. Previous tests of AutoTutor also showed little or no affects of AutoTutor on the acquisition of shallow knowledge (an effect size sigma of .15 is reported in [23], so this result is not surprising. AutoTutor was designed for helping students construct deep explanations, not to acquire shallow knowledge. These scores were slightly lower for the posttest (.41) than the pretest (.43), which suggests there may be some element of fatigue or decreased motivation over the experimental session. However, the decrease was not statistically significant.

The results of the cloze posttest was also simple to interpret. There was a statistically significant main affect of training, with means of .30, .23, and .23 in the AutoTutor, Read-text, and Control conditions, respectively, $F(2, 154) = 25.69$, $MS_e = .001$, $p < .05$. The corresponding SD's were .19, .16, and .16. Planned comparisons confirmed the following pattern: AutoTutor > Read-test = Control. This result replicates previous tests of AutoTutor [23]. The effect size of AutoTutor compared with the Control was .44, which compares favorably to the .68 sigma reported in [23]. In contrast, there was no significant affect of media on the cloze posttest scores ($F < 1$) and no significant training X media interaction ($F < 1$). The mean cloze scores in the AutoTutor condition were .31, .30, .29, and .31 in the Print, Speech, TH, and THP conditions, respectively. These scores are remarkably constant. This result supports the dialog facilitation hypothesis rather than the media facilitation hypothesis.

The results of the posttest scores for deep questions was the most interesting. Table 1 shows the cell means and standard deviations in the factorial design. There was a significant main effect of training condition, $F(2, 154) = 29.84$, $MS_e = .019$, $p < .05$, with means of .50, .46, and .34 in the AutoTutor, Read-text, and Control conditions, respectively. Planned comparisons showed the following ordering among the means: AutoTutor > Read-text > Control. The effect size of AutoTutor compared with the Control was 1.23 sigma, and with the Read-text was .31. An effect size of .28 has been reported when AutoTutor was compared with their comparison conditions [23]. In contrast to the significant effects of tutoring, the effects of the media were subtle. There was no significant main effect of media, $F(3, 77) = .61$, $MS_e = .049$, $p > .50$, but a significant training X media interaction, $F(6, 154) = 6.71$, $MS_e = .019$, $p < .05$. AutoTutor did appear to be more effective when it had a talking head compared to speech or print alone, but this pattern was subtle and merits replication.

Table 1. Posttest Scores for Deep Questions

Training Conditions	Media Conditions			
	Print	**Speech**	**Talking Head**	**Talking Head +Print**
AutoTutor	.47 (.18)	.47 (.16)	.49 (.17)	.55 (.17)
Read-text	.44 (.18)	.46 (.14)	.42 (.16)	.50 (.18)
Control	.34 (.16)	.38 (.13)	.32 (.11)	.33 (.13)

4. Conclusions

The results of this study clearly support the dialog facilitation hypothesis rather than the media facilitation hypothesis. There is something about the dialog facilities of AutoTutor that facilitate learning, particularly at deeper levels of comprehension. In contrast, the effects of the media are either subtle or nonexistent. Simply put, it is the message that is the message – the media is not the message. Exactly what it is about dialog that facilitates learning remains an open question, but a question that we will continue to explore in the future.

5. Acknowledgements

This research was supported by grants from DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research (N00014-00-1-0600) and the National Science Foundation (SBR 9720314 and REC 0106965). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ONR or NSF.

6 References

[1] Aleven V. & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. Cognitive Science, 26, xx-xx.

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

[3] Baylor, A.L. (2001). Investigating multiple pedagogical perspectives through MIMIC (Multiple Intelligent Mentors Instructing Collaboratively). *Workshop on agents at the AI-ED International Conference*. San Antonio, TX.

[4] Bloom, B.S. (Ed.) (1956) *Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain.* New York ; Toronto: Longmans, Green.

[5] Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4-16.

[6] Cassell, J., and Thorisson, K.R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence, 13*, 519-538.

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

[8] Chi, M. T. H., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science, 25*, 471-533.
[9] Clarebout, G., Elen, J., Johnson, W.L., & Shaw, E. (2002). Animated pedagogical agents: An opportunity to be grasped? *Journal of Educational Multimedia and Hypermedia, 11*, 267-286.
[10] Cohen, M.M., and Massaro, D.W. (1994). Development and experimentation with synthetic visible speech. *Behavior Research Methods, Instruments, and Computers, 26*, 260-265.
[11] 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.
[12] Corbett, A.T. (2001). Cognitive computer tutors: Solving the two-sigma problem. *User Modeling: Proceedings of the Eighth International Conference, UM 2001,* 137-147.
[13] Graesser, A.C., Person, N., Harter, D., & TRG (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education, 12*, 257-279.
[14] Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology, 9*, 1-28.
[15] Graesser, A.C., VanLehn, K., Rose, C., Jordan, P., and Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine, 22*, 39-51.
[16] Graesser, A.C., Ventura, M., Jackson, G.T., Mueller, J., Hu, X., & Person, N. (2003). The impact of conversational navigational guides on the learning, use, and perceptions of users of a web site. Proceedings of the AAAI Spring Symposium 2003 on Agent-mediated Knowledge Management. AAAI Press.
[17] Graesser, A.C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & the TRG (1999). Auto Tutor: A simulation of a human tutor. *Journal of Cognitive Systems Research, 1*, 35-51.
[18] Johnson, W. L., & Rickel, J. W., and Lester, J.C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, *11*, 47-78.
[19] Massaro, D.W., & Cohen, M.M. (1995). Perceiving talking faces. *Psychological Science, 4,* 104-109.
Moreno, K.N., Klettke, B., Nibbaragandla, K., Graesser, A.C., & the Tutoring Research Group (2002). Perceived characteristics and pedagogical efficacy of animated conversational agents. In S. A. Cerri, G. Gouarderes, & F. Paraguacu (Eds.), *Intelligent Tutoring Systems 2002* (pp. 963-971). Berlin, Germany: Springer.
[20] Moreno, R., Mayer, R.E., Spires, H.A., & Lester, J.C. (2002). The case for social agency in computer-based teaching: Do students learn more deeply when the interact with animated pedagogical agents? *Cognition & Instruction, 19*, 177-213.
[21] Olde, B. A., Franceschetti, D.R., Karnavat, Graesser, A. C. & the Tutoring Research Group (Aug., 2002). The right stuff: Do you need to sanitize your corpus when using latent semantic analysis? *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 708-713). Mahwah, NJ: Erlbaum.
[22] Person, N.K., Graesser, A.C., & the Tutoring Research Group (2002). Human or computer?: AutoTutor in a bystander Turing test. In S. A. Cerri, G. Gouarderes, & F. Paraguacu (Eds.), *Intelligent Tutoring Systems 2002* (pp. 821-830). Berlin, Germany: Springer.
[23] Person, N. K., Graesser, A. C., Bautista, L., Mathews, E. C., & the Tutoring Research Group (2001). Evaluating student learning gains in two versions of AutoTutor. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.) *Artificial intelligence in education: AI-ED in the wired and wireless future* (pp. 286-293). Amsterdam, IOS Press.
[24] Person, N.K., Graesser, A.C., Kreuz, R.J., Pomeroy, V., & TRG (2001). Simulating human tutor dialog moves in AutoTutor. *International Journal of Artificial Intelligence in Education. 12*, 23-39.
[25] Picard, R.W. (1997). *Affective Computing*. Cambridge, MA: MIT Press.
[26] Shah, F., Evens, M., Michael, J., & Rovick, A. (2002). Classifying student initiatives and tutor responses in human keyboard-to-keyboard tutoring sessions. *Discourse Processes, 33*, 23-52.
[27] VanLehn, K. & Graesser, A. C. (2002). Why2 Report: Evaluation of Why/Atlas, Why/AutoTutor, and accomplished human tutors on learning gains for qualitative physics problems and explanations. Unpublished report prepared by the University of Pittsburgh CIRCLE group and the University of Memphis Tutoring Research Group.
[28] VanLehn, K., Jones, R. M. & Chi, M. T. H. (1992). A model of the self- explanation effect. *Journal of the Learning Sciences, 2*(1), pp. 1-60.
[29] VanLeh, K., Lynch, C., Taylor, L.,Weinstein, A., Shelby, R., Schulze, K., Treacy, D., & Wintersgill, M. (2002). In S. A. Cerri, G. Gouarderes, & F. Paraguacu (Eds.), *Intelligent Tutoring Systems 2002* (pp. 367-376). Berlin, Germany: Springer.
[30] Whittaker, S. (2003). Mediated communication. In A.C. Graesser, M.A. Gernsbacher, and S.A. Goldman (Eds.), *Handbook of Discourse Processes*. Hillsdale, NJ: Erlbaum.