Fourteen Facts About Human Tutoring: Food for Thought for ITS Developers

Natalie K. Person
Department of Psychology
Rhodes College
2000 North Parkway, Memphis, TN 38112 USA
person@rhodes.edu

Arthur C. Graesser
Department of Psychology and the Institute for Intelligent Systems
University of Memphis
202 Psychology Building, Memphis, TN 38152-3230, USA
a-graesser@memphis.edu

Abstract. This paper contains 14 facts about the dialog that occurs in one-to-one human tutoring sessions. All of the facts are based on the empirical findings of our human tutoring research from the last 14 years. ITS developers are encouraged to consider these facts as they attempt to improve the dialog facilities and pedagogical effectiveness of their systems.

1 Introduction

The Tutoring Research Group (TRG) at the University of Memphis has been studying the tutoring process for the past 15 years. During this time, we have performed in-depth analyses on human-to-human tutoring dialogs (3, 5, 9, 20, 22, 24), developed computer tutors for computer literacy and conceptual physics (6, 11, 23,), and designed and tested computational linguistics modules, simulations, authoring tools, and web facilities that can be used in other intelligent tutoring systems (ITSs) (10, 12, 19, 29). Our work has always been driven by theories in cognitive psychology, education, computer science, and psycho- and computational linguistics. However, we like many other ITS developers have at times become so involved in simply making the modules of our systems work that we have lost sight of our original goal: to build a tutoring system that is as good as or better than highly effective human tutors. The purpose of the paper is to remind ITS developers of what we know about effective human tutors and to stimulate thought on how existing ITS can be improved.

This paper is structured around fourteen facts that we have learned from our research of effective human tutoring over the years. It should be noted that the human tutors that we have studied were not trained to use sophisticated tutoring techniques, but rather were representative of tutors that generally do most of the tutoring in school systems. We have analyzed hundreds of hours of tutoring dialogs on a variety of subjects (e.g., algebra, statistics, and research methods), and the participants in the tutoring sessions have included students in middle school, high school, college, and graduate school.

The focus of our human tutoring research has always been on the tutorial dialog. We have developed elaborate classification schemes that have contributed to our understanding of the questions that tutors and students ask during tutoring sessions (7, 8, 9, 22). We have also developed classification schemes that have furthered our knowledge of the type and timing of tutor-generated examples. And lastly, we have gained considerable knowledge about pedagogical strategies, dialog patterns, and tutor feedback from detailed Turn Transition Matrix (TTM) analyses. The basic TTM framework used in our research is represented in Figure 1. The TTM has elucidated many dialog patterns that frequently occur in human tutoring. By recording each tutor turn as a function of each student turn, we have computed likelihood scores for numerous tutor turn categories. Descriptions of the turn categories that we have studied for both students and tutors are provided in Tables 1 and 2, respectively.

	Tutor Turn N + 1		
Student Turn N	Category 1	Category 2	Category j
Category 1			
Category 2			
Category k			

Figure 1. Structure for Turn Transition Matix (TTM).

Table 1. Descriptions of Student Turn Categories

STUDENT CATEGORY	CATEGORY DESCRIPTION
1. Contribution Quality	
Complete	Student provides complete answer to tutor question.
Partial	Student provides partial answer to tutor question.
Vague	Student provides vague answer to tutor question.
Error-ridden	Student provides error-ridden answer to tutor question.
No Answer	Student fails to provide any answer to tutor question.
2. Asks Question	
Makes request	Student makes request unrelated to the problem/example.
Counter- clarification	Student needs clarification on tutor's previous statement.
Problem-related	Student asks question directly related to the problem/example.
Other	Any question not assigned to one of the other three question categories.
3. Misconception	Student states his or her own misconception.
4. Reminding Example	Student comments on a similar example.
5. Meta-comment	Student comments on own ability or attribute of problem.
6. Acknowledgement	Student acknowledges tutor's contribution (e.g., Uh-huh).
7. Gripes	Student complains.
8. Think aloud	Student thinks aloud.
9. Nonverbal	Student makes a nonverbal response (e.g., laughs).
10. Draw	Student draws on board.
11. Other	Any speech act not assigned to one of the other student categories.

Table 2. Descriptions of Tutor Turn Categories

TUTOR CATEGORY	CATEGORY DESCRIPTION
1. Additional example	
Easier	Tutor provides student an easier example than the previous example.
Difficult	Tutor provides a more difficult example than the previous example.
Equal	Tutor provides an example of equal difficulty with the previous example.
2. Asks question	

Error-repair	Tutor asks question specifically related to student error.
Directed- Activity	Tutor asks question in order to redirect student's activity.
Leading	Tutor asks question to expose student's misconception.
Counter-clarification	Tutor requests clarification of student's previous statement.
Pump	Tutor pumps student for additional information.
Assessment	Tutor assesses student's knowledge about a particular topic.
Global	Tutor globally assesses student's knowledge (e.g., "Do you understand?")
Other	Any question not assigned to one of the other question categories.
3. Feedback	
Positive	Tutor gives positive feedback to student.
Negative	Tutors gives negative feedback to student.
Neutral	Tutor gives neutral feedback to student.
Immediate	Tutor provides immediate feedback for a student error.
Delayed	Tutor provides delayed feedback for a student error.
4. Reminding Example	Tutor comments on a similar example.
5. Specific Component	Tutor focuses on specific component of current problem/example.
6. General Level	Tutor discusses current example in more general terms.
7. Hint	Tutor provides the student with a hint.
8. Splice	Tutor splices in the correct answer.
9. Elaborates	Tutor elaborates current problem/example.
10. Answers	Tutor answers student question.
11. Rearticulates	
Solution	Tutor rearticulates the current problem's solution.
Representation	Tutor rearticulates the problem's representation.
12. Affective	
Own ability	Tutor comments on his or her own ability.
Student ability	Tutor comments on student's ability.
Problem	Tutor comments on the difficulty of the problem/example.
General	Tutor makes general empathetic comment.
13. Gripes	Tutor complains.
14. Directive	Tutor tells the student what to do.
15. Draw	Tutor draws on the board.
16. Nonverbal	Tutor makes some type of nonverbal response (e.g., laughs).
17. Other	Any speech act not assigned to one of the other tutor categories.

1.1 Facts About Real vs. Ideal Tutoring Strategies

Fact 1. Tutors rarely adhere to ideal tutoring models. We have conducted a number of analyses to determine whether human tutors naturally adhere to ideal tutoring models. The ideal models that we have primarily been interested in are those that include strategies that are based on various theories of learning: (1) Vygotsky's developmental theory, (2) building block/prerequisite approaches, (3) feedback-based theories, (4) Socratic tutoring, and (5) case-based reasoning. We have attempted to identify ideal strategies by analyzing tutorial dialog in global and fine-grained ways. In all attempts, we've come up empty-handed. We have seen vestiges of Vygotskian approaches (e.g., modeling-scaffolding-fading techniques); however, we have no evidence that tutors implement prerequisite or cascade approaches when students make errors or employ Socratic tutoring strategies.

Fact 2. Tutors respond to student turns with pumps, hints, elaborations, and splices. We have documented that human tutors rarely use ideal strategies when responding to students contributions. Instead, human tutors rely on a rather small set of dialog moves that are tailored to address the immediately preceding student contribution. Some of the most frequently occurring dialog moves are pump, hint, elaboration, and splice (and, of course, feedback which is discussed below). A pump refers to a tutor pumping a student for

additional information. The likelihood scores from a TTM analysis indicate that tutors use pumps when students provide partial answers (.39), vague answers (.28), error-ridden answers (.11), or no answers (.18) to tutor questions. We know that tutors often provide hints instead of answers. For example, as indicated by likelihood scores, tutors offer hints when students ask questions (.25), provide error-ridden answers (.11), or no answers (.16). We have also documented that tutor hints become progressively more specific until the students can provide answers. The majority of human tutoring dialogs consists of elaborations. Elaborations tend to occur after the following student turn categories, Complete answer (.35), Partial Answer (.27), Vague Answer (.33), Problem-related question (.32), Meta-comments (.53), and Acknowledgements (.62). A Splice is a dialog move in which the tutor automatically supplies the correct answer rather than waiting for the student to do so. We have found that the only conditions under which tutors Splice are when students provide Error-ridden (.31) or Partial answers (.21).

1.2 Facts About Tutor Feedback

Fact 3. Tutors rely on immediate feedback more often than delayed feedback. When should tutors provide feedback? Researchers in the immediate feedback camp argue that such feedback prevents students from floundering and becoming discouraged (1, 13, 16, 30). Those in the delayed feedback camp claim that immediate intervention by the tutor discourages students from diagnosing their own errors and also impedes the acquisition of self-regulated learning strategies (18, 26, 27, 28). Although one type of feedback may be preferable for an optimal learning experience, our feedback analysis indicates that human tutors clearly prefer to respond to student errors with immediate feedback (likelihood score = .59) rather than delayed feedback (likelihood score = .12). Immediate feedback was defined as any corrective utterance (e.g., Error-repair Question, Leading Question, Hint, and Splice) provided by the tutor that occurred immediately after a student's error-ridden contribution. Delayed feedback was defined as any corrective utterance that occurred in subsequent tutor turns of the problem or example discussion.

Fact 4. Tutors provide indiscriminate feedback to student errors. Tutors have the option of providing positive, negative, or neutral feedback when students make errors. Presumably, a rational tutor would provide negative feedback when students provide error-ridden answers, neutral feedback for partial or vague answers, and positive feedback for correct answers. Our analyses of tutor feedback indicate that tutors are not particularly discriminating when responding to student errors. Consider the following excerpt from an algebra tutoring session.

Indiscriminate Feedback

TUTOR: . . . okay, now, what it is, just FOIL. Okay, FOIL, it stands for first, outside, inside, last. Okay, so what you do is you take [the] first one, right? You multiply these two, and you take the outside, the inside and the last. Do you see how that works?

STUDENT: Here's the way I've done it [student's method is incorrect].

TUTOR: Right. Well, see that's one way to do it, but they like this [the FOIL method], this is really the way most people like to do it [elaborates on reasons].

In a feedback analysis of tutorial dialog between college students and graduate student tutors, the likelihood scores indicate that the tutors provide positive (.25) and neutral (.17) feedback in response to student errors just as often as they provide negative (.26) feedback to errors. These findings are somewhat perplexing given that we know that tutors do indeed recognize the errors. We have come to believe that tutors are reluctant to provide excessive negative feedback because of student self-esteem and motivational issues (13, 14). In addition, tutors may provide inappropriate feedback to avoid being impolite (24).

Fact 5. The likelihood of tutors remediating or providing error-specific feedback to student errors is extremely low. We have conducted feedback analyses to determine whether tutors use sophisticated approaches to handle student errors. Specifically, we have been interested in whether tutors employ remediation or error-specific approaches when they detect a student error. Remediation typically involves reteaching the material and additional practice problems for the student. Error-specific feedback generally consists of the tutor pointing out a specific error to the student and then demonstrating the consequences of that particular error. Although others have reported that human tutors engage in both of these approaches (17), we have found virtually no evidence of either approach in our analyses of human tutoring dialog.

Fact 6. Tutors never attribute student errors or misconceptions lack of ability. In a nutshell, tutors never make affective remarks about students' ability levels. Unsurprisingly, we never expected tutors to comment on any student's lack of ability (e.g., "Gee, I don't think you're capable of getting this."). However, it was somewhat surprising that tutors refrain from making positive comments about student ability (e.g., "You are picking this up pretty quickly."). Whenever affective comments are made by tutors, they are attributed to either the tutor's ability (e.g., "I've always struggled with this, too.") or to the difficulty of a particular problem (e.g., "This problem is a real bear.").

1.3 Facts About Question Asking and Answering

Fact 7. Tutors and students ask a lot of questions. Compared to classrooms, tutors and students ask significantly more questions. It has been documented that the mean number of questions generated by a particular student in a classroom is less than .20 questions per hour (2, 5). In one-to-one tutoring settings, students ask approximately 26.0 questions per hour. Tutors ask approximately 1.5 more questions than classroom teachers, 105 versus 70 questions per hour, respectively.

Fact 8. Common ground questions occur most often. Most of the questions asked in tutoring sessions are efforts to negotiate the common ground between the participants (5, 22). Common ground questions comprise 89% of all student-generated questions and 82% of tutor-generated questions. Most common ground questions asked by students are attempts to confirm the validity of their own knowledge (e.g., "Intersecting lines mean

there's an interaction, right?"). The majority of tutor common ground questions also serve to assess students' knowledge (e.g., "Do you know what an interaction is?").

Fact 9. Good students ask better questions rather than more questions. We know that good students do not necessarily ask more questions (5, 22). However, good student do ask more deep-reasoning questions. In past analyses, we have shown that the overall frequency of student questions in tutoring sessions is negatively correlated with achievement measures (e.g., Final Grade in course); however, the proportion of deep-reasoning questions asked by students is positively correlated with achievement measures.

Fact 10. Student answers are the best way to infer understanding. Asking students if they understand the material is pointless (22). We have found that student responses to comprehension-gauging questions (e.g., "Do you understand?") are unreliable for assessing student comprehension. Specifically, students usually assert that they do understand (57%), rarely assert that they don't (7%), and often neglect to answer the question altogether (36%). A more accurate way to infer student understanding is by considering the quality of the students' answers to tutor questions. We know that good students tend to give more complete answers to tutor questions, whereas poor students are more likely to provide error-ridden answers, vague answers, or nothing at all.

1.4 Facts About Tutor-generated examples

Fact 11. Tutors use a lot of examples. We have extensively studied the examples that tutors generate during tutoring sessions. Examples in our analyses were defined as assertions or problems tutors use to illustrate or elaborate a tutoring topic. In an analysis of 847 examples that occurred in 44 tutoring sessions, the mean number of examples generated by the tutors was 19.25. Given that the average tutoring session lasted approximately 45 minutes, the tutors generated nearly 26 examples per hour. The mean number of conversational turns for each example was 18.67 indicating that the examples were somewhat detailed and required input from students. Hence, we know that tutor examples are prevalent and detailed.

Fact 12. Tutor do not use examples to correct student errors but do use them to explain difficult material. Intuitively, one would think that an appropriate time for a tutor to present an example would be after a student error or misconception. In an analysis of the conditions that trigger tutor-generated examples, we found that this was not the case. Only 10% of the tutor-generated examples followed student errors and misconceptions. Further, tutors do not use examples to introduce new topics (1%) and rarely use them to review material (10%). Instead, tutors tend to rely on examples to explain difficult topics (53%) and assess students' understanding (26%).

Fact 13. Tutor-generated examples are usually concrete and difficult. We have also performed several analyses on the abstraction and difficulty levels of tutor-generated examples. In an abstraction analysis, the examples were assigned to one of three categories: (1) Situated, (2) Concrete, and (3) Abstract. Situated examples are defined as examples that are coherent, meaningful, authentic, and directly related to the practices of our culture. Concrete examples include variables which have specific referents; however, the overall

simplistic nature of the examples fall short of being ecologically valid (e.g., "The red train leaves New York at 5:00" kind of problem). Abstract examples have no concrete referents, are vague, and have little relevance in real world problems. Our abstraction analyses indicated that most tutor-generated examples are either Concrete (63%) or Abstract (35%). Situated examples comprised only 2% of the entire sample. A Concrete from a tutoring session on factorial designs and an Abstract example from a tutoring session on variables are provided below.

Concrete Example

TUTOR: Okay, let's use heart attack rates as our DV and our two independent variables are personality type, um, have you heard of those? Like A or B?

STUDENT: Yeah.

TUTOR: And our other independent variable can be gender. Okay, what we're gonna do is graph all the possible outcomes.

Abstract Example

TUTOR: Say you see something that looks like this (the tutor draws an inverted U-shape on the board). When your results looks like this, what that means is that, um, usually, your independent variable increases as a function of your dependent variable, but only up to a point (tutors points to apex of curve), at this point, your independent variable starts to decrease as a function of dependent variable, okay?

In the difficulty analysis, each example was assigned to one of three levels: Easy, Intermediate, and Difficult. Easy examples are assertions in which the tutor supplies all of necessary information needed to understand the example. An Intermediate example requires the student to assume some of the cognitive burden in fleshing out the example, and Difficult examples require the student to either supply the majority of the information necessary to complete the example or to integrate multiple information sources. Dialog excerpts illustrating each of these difficulty levels are provided below.

Easy Example (from a tutoring session on variables)

TUTOR: Okay, an example, gender is a variable. It has two levels or values, male and female.

Intermediate Example (from tutoring session on graphing)

TUTOR: (Draws an x and y axis and puts 6 scores on the board). Here are test scores for 6 students. Your job is to construct a frequency distribution for them. Okay, how many students scored a 75?

Difficult Example (from a tutoring session on factorial designs)

TUTOR: Say you're a researcher interested in handedness, you know, like right or left-handed people. Uh, caffeine intake, and heart attacks. I want you to design a study that would look at those variables. First operationally define all of your variables, then pick a design. I basically just want you to tell me everything you need to do in order to make this study happen.

The results indicated that 18% of the tutor examples were easy, 42% were intermediate, and 40% were difficult. It appears that tutors attempt to provide examples that challenge the students to reason on the frontiers of their own knowledge. It may be the case, however, that the tutors themselves consider their examples to be quite easy.

Fact 14. Tutors rely on canned examples. We know that human tutors use a lot of examples during tutoring so it seems worthwhile to ask about the source of these examples. In a source analysis to determine the origin of tutor examples, 847 examples were assigned to one of six categories. We learned that tutors primarily rely on canned examples from their own predetermined curriculum scripts (51%) or from the course textbook (27%). A curriculum script consists of a set of topics, examples, and questions that is constructed by the tutor prior to the tutoring session (25). Thus, most tutor-generated examples were constructed prior to the tutoring sessions. We found some evidence that tutors generate spontaneous examples (17%) perhaps in an attempt cater to students immediate needs. However, we found no evidence that tutors use examples generated by students (3%), from famous studies or cases (2%), or those from their personal research experience (1%).

2 Conclusions

We have provided 14 facts about human tutoring that leave the reader with a pretty good overall impression of what goes on in these one-to-one sessions. These facts should by no means be considered the end all, be all of good tutoring. Although we know that normal, untrained tutors like those we have studied are effective and that students do learn in their interactions with them, we are the first to admit that the pedagogical choices of these tutors were less than optimal. We have little doubt that expert human tutors would have been more effective. We also know that there are many intelligent tutoring systems that rival untrained human tutors and that some ITSs are steadfastly approaching the performance levels of expert tutors. We have attempted to improve the pedagogical effectiveness of our own system, AutoTutor, in light of what we have learned from human tutoring studies. For example, AutoTutor gives discriminating feedback to student responses and uses sophisticated strategies that force students to actively construct their own knowledge (4, 21, 29). AutoTutor is also designed to interpret and answer student questions. In closing, these 14 facts should not be viewed as the foundations of good tutoring, but rather as a "Things to Consider List" for those attempting to build pedagogically sound tutoring systems.

3 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.

4 References

[1] Burton, R. R., & Brown, J. S. (1982). An investigation of computer coaching for informal learning activities. In D. Sleeman & J. S. Brown (Eds.), *Intelligent tutoring systems* (pp. 79-98). New York: Academic Press.
[2] Dillon, T.J. (1988). *Questioning and Teaching: A Manual of Practice*. New York: Teachers College Press.
[3] Graesser, A. C., Bowers, C. A., Hacker, D. J., & Person, N. K. (1997). An anatomy of naturalistic tutoring. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning: Instruction approaches and issues* (pp. 145-184). Cambridge, MA: Brookline Books.
[4] Graesser, A. C., Hu, X., Susarla, S., Harter, D., Person, N. K., Louwerse, M., Olde, B., & the Tutoring Research Group (2001). AutoTutor: An Intelligent Tutor and Conversational Tutoring Scaffold. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.) *Artificial intelligence in education: AI-ED in the wired and wireless future* (pp.47-49). Amsterdam, IOS Press.
[5] Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal, 31,* 104-137.
[6] 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.
[7] Graesser, A. C., Person, N. K., and Huber, J. D. (1992). Mechanisms that generate questions. In T. Lauer, E. Peacock, & A. C. Graesser (Eds.), *Questions and information systems* (pp. 167-187). Hillsdale, NJ: Lawrence Erlbaum Associates.
[8] Graesser, A. C., Person, N. K., and Huber, J. D. (1993). Question asking during tutoring and in the design of educational software. In M. Rabinowitz (Ed.), *Cognitive science foundations of instructional software* (pp. 149-172). Hillsdale, NJ: Lawrence Erlbaum Associates.
[9] Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring sessions. *Applied Cognitive Psychology, 9,* 1-28.
[10] Graesser, A.C., VanLehn, K., Rose, C., Jordan, P., and Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine, 22*, 39-51.
[11] 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.
[12] Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., and the TRG (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments, 8,* 129-148.
[13] Kulik, J. A., & Kulik, C.-L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*, 79-97.
[14] Lepper, M. R., & Henderlong, J. (In Press). Motivation and instruction. In J. W. Guthrie (Ed.), *Encyclopedia of education.* New York: Macmillan.
[15] Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75-105). Hillsdale, NJ: Erlbaum.
[16] 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: Lawrence Erlbaum Associates, Inc.
[17] McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*, 197-244.
[18] Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences, 2*, 277-305.
[19] Olde, B. A., Franceschetti, D.R., Karnavat, A., 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.

[20] Person, N. K., & Graesser, A. C. (1999). Evolution of discourse in cross-age tutoring. In A. M.O'Donnell and A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 69-86). Mahwah, NJ:Erlbaum.
[21] 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.
[22] Person, N. K., Graesser, A. C., Magliano, J. P., & Kreuz, R. J. (1994). Inferring what the student knows in one-to-one tutoring: The role of student questions and answers. *Learning and Individual Differences, 6*, 205-229.
[23] 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.
[24] Person, N. K., Kreuz, R. J., Zwaan, R., & Graesser, A. C. (1995). Pragmatics and pedagogy: Conversational rules and politeness strategies may inhibit effective tutoring. *Cognition and Instruction, 13*, 161-188.
[25] Putnam, R. T. (1987). Structuring and adjusting content for students: A live and simulated tutoring of addition. *American Educational Research Journal, 24*, 13-48.
[26] Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences, 1*, 37-68.
[27] Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer-supported intentional learning environments. *Journal of Educational Computing Research, 5*, 51-68.
[28] Schank, R. C., & Farrell, R. (1987). Creativity in education: A standard for computer-based teaching. Machine-Mediated Learning, 2, 175-194.
[29] 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.
[30] VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences, 2*, 1-59.