Running head: TUTORING

Instruction Based on Tutoring

Arthur C. Graesser, Sidney D'Mello, and Whitney Cade

University of Memphis

Send correspondence to:

Art Graesser
Department of Psychology & Institute for Intelligent Systems
202 Psychology Building
University of Memphis
Memphis, TN 38152-3230
901-678-4857
901-678-2579 (fax)
a-graesser@memphis.edu

Instruction Based on Tutoring

This chapter reviews research on human tutoring, a form of one-on-one instruction between a tutor and a tutee. In most cases the tutor is knowledgeable about the subject matter and helps the tutee (i.e., the student) improve mastery of the knowledge base or skill. However, sometimes the tutor is a peer of the tutee who plays the role of a tutor, even though the tutor and tutee are at approximately the same level of subject matter mastery. The hope is that the tutorial session is tailored to the needs of the individual student by building on what the student already knows, filling in gaps in knowledge, and correcting conceptual errors. We distinguish between tutors and mentors, although the distinction is not entirely clear-cut. A tutor typically is an expert on a particular subject matter and has a tight control over the tutorial session -- turn by turn and moment by moment. In contrast, a mentor has a broader repertoire of knowledge, skills, and wisdom, with only occasional suggestions to the student as the student proceeds with a more self-regulated agenda.

Tutoring is the typical solution that students, parents, teachers, principles and school systems turn to when the students are not achieving expected grades and educational standards. There are serious worries in the community when a school is not meeting the standards of a high stakes test, and teachers are anxious about the prospects of losing their jobs due to the criteria and policies of *No Child Left Behind.* Schools and families worry when a student runs the risk of losing a scholarship or when an athlete may be cut from a team. Tutors step in to help under these conditions. Wealthier families might end up paying $200 per hour for an accomplished tutor to rescue save a son or daughter. However, these expectations may be rather high, considering that most tutors are same-age peers of the students, slightly older cross-age tutors, citizens in the community, and paraprofessionals who have had little or no training on tutoring pedagogy (Cohen,

Kulik, & Kulik, 1982; Graesser & Person, 1994). Nevertheless, their tutoring can be effective in

helping students learn, as we will document in this chapter.

Although most tutors in school systems have little or no tutoring training, there are many

examples of excellent tutoring programs that are grounded in the science of learning. One notable

example is the *Reciprocal Teaching* method that helps students learn how to read text at deeper

levels (Palincsar & Brown, 1984, 1988). The tutoring method engages the tutor and students in a

dialogue that jointly constructs the meaning of the text. The dialogue is supported with the use

of four strategies: generating questions, summarizing text segments, clarifying unfamiliar words

and underlying global ideas, and predicting what will happen next in the text. These strategies

are applied in a context-sensitive manner rather than mechanically applied in scripted lessons.

Moreover, the tutors systematically change their style of tutoring as the lessons proceed. When

students are initially introduced to Reciprocal Teaching, the tutor models the application of these

strategies by actively bringing meaning to the written word (called content strategies) and also

monitoring one's own thinking and learning from text (called meta-cognitive strategies). Over

the course of time, the students assume increased responsibility for leading the dialogues. That

is, after the modeling phase, the tutor has the students try out the strategies while the tutor gives

feedback and scaffolds strategy improvements. Eventually the students take more and more

control as the tutor fades from the process and occasionally intervenes much like a coach. This

modeling-scaffolding-fading instructional process has a long history in the psychology of

learning (Collins, Brown, & Newman, 1989; Rogoff & Gardner, 1984; Vygotsky, 1978).

The Reciprocal Teaching method has been tested in dozens of studies and has been

shown to improve students' reading skills. Rosenshine and Meister (1994) conducted a meta-

analysis of 16 studies of Reciprocal Teaching that were conducted with students from age seven

to adulthood. The method was compared with traditional basal reading instruction, explicit instruction in reading comprehension, and reading and answering questions. When experimenter-developed comprehension tests were used, the median effect size was 0.88 sigma. A sigma is a measure in standard deviation units that compares a mean in the experimental treatment to the mean in a comparison condition. According to Cohen (1992), effect sizes of 0.20, 0.50, and 0.80 are considered small, medium, and large, respectively, so the .088 effect size would be considered large. When standardized measures were used to assess comprehension, the median effect size favoring Reciprocal Teaching was $d = 0.32$. This effect size is considered small to medium according to Cohen, but it is important to acknowledge that it is much more difficult to obtain large effect sizes on standardizes tests and particularly on the skill of reading (as opposed to mathematics and science). According to Hattie (2009), a meta-analysis of meta-analyses revealed that a 0.4 effect size is routinely reported in educational studies for successful interventions. The Reciprocal Teaching method has also been applied in classroom contexts with trained teachers applying the method in front of a classroom of students or in small groups. Given the promise of this method, it was recently accepted as an effective method to try in *What Works Clearinghouse*, a mechanism funded by the US Department of Education to test promising methods of instruction in a large number of schools throughout the country.

Despite encouraging examples like Reciprocal Teaching, there are practical challenges in relying on humans to supply high-quality one-on-one human tutoring (Conley, Kerner, & Reynolds, 2005; Hock, Schumaker, & Deshler, 1995). For example, it is very costly to train tutors on tutoring strategies. There is a high dropout rate when both skilled and unskilled tutors face the realities of how difficult it is to help students learn. Fortunately, the tutoring enterprise has expanded beyond human tutoring and into the realm of computer tutoring. Computers are available 24/7, do not get

fatigued, do not burn out, and can reliably apply pedagogical strategies. There are now intelligent tutoring systems and other advanced learning environments that implement sophisticated instructional procedures (VanLehn, 2006; VanLehn et al., 2007; Woolf, 2009). Intelligent tutoring systems are able to induce the characteristics of individual learners at a fine-grained level, to assign problems or tasks that are sensitive to the students' profile, and to generate specific tutoring actions that attempt to optimize learning according to scientific principles. Unlike human tutors, ITS provide precise control over the instructional activities, which of course is a methodological virtue. ITS have the capacity to scale up in delivering learning assistance to many more students than can be provided by human tutors.

The Cognitive Tutors developed by the Pittsburgh Science of Learning Center is one noteworthy ITS family (Anderson, Corbett, Koedinger, & Pelletier, 1995; Koedinger, Anderson, Hadley, & Mark, 1997; Ritter, Koedinger, Anderson, & Corbett, 2007). The *Cognitive Tutors* help students learn algebra, geometry, and programming languages by applying learning principles inspired by the ACT-R cognitive model (Anderson, 1990).   There is a textbook and curriculum to provide the content and the context of learning these mathematically intensive subject matters, but the salient contribution of the Cognitive Tutors is to help students solve problems. The tutor scaffolds the students to take steps in solving the problem by prompting them to actively take the steps, by comparing the students' actions to ideal models of correct solutions, by giving students feedback on their actions, and by providing hints and other forms of help. The Cognitive Tutor mechanism incorporates a combination of declarative knowledge (facts) and procedural knowledge. Students are expected to learn through enough practice in varied contexts on problems that are tied to the curriculum.

The Cognitive Tutors are now used in over 2000 school systems throughout the country and are among the methods being assessed among the methods included in the *What Works Clearinghouse*. These tutors have been heavily evaluated over the course of 35 years. The effect sizes on experimenter-developed tests are approximately 1.0 sigma compared to normal classroom teaching (Corbett, 2001). According to Ritter et al. (2007), standardized tests show overall effect sizes of 0.3 sigma, but particularly shine for the subcomponents of problem solving and multiple representations, which show effect sizes of $d = 0.7$ to 1.2. The What Works Clearinghouse investigations show an effect size of 0.4 sigma. The Cognitive Tutors are an excellent example of how scientific principles of learning can be implemented in a technology that not only helps learning but also scales up to widespread use in thousands of school systems.

This chapter reviews the research on human tutoring. We examine the pedagogical theories, conversation patterns, and empirical evidence for the effectiveness of one-on-one human tutoring. Although computer tutors are becoming more prevalent, it is beyond the scope of this chapter to cover intelligent tutoring systems and other advanced computer environments that attempt to adapt to individual students. The final section identifies some future directions for the field to pursue.

Does Human Tutoring Help Learning?

It could be argued that tutoring was the very first form of instruction. Children were trained one-on-one by parents, other relatives, and members of the village who had particular specialized skills. The apprenticeship model reigned for several millennia before we encountered the industrial revolution and classroom education (Collins & Halverson, 2009). Throughout that part of history, the modelling-scaffolding-fading process was probably the most sophisticated form of early tutoring. The alternative of lecturing was probably more prevalent: The master simply lectured to the apprentice, the apprentice nodded (knowingly or unknowingly), and the master undoubtedly grew

frustrated when very few of the ideas and skills were sinking in. Lecturing is ubiquitous in the repertoire of today's unskilled tutors (Graesser, Person, & Magliano, 1995), but there are some other strategies that come naturally, as will be elaborated in this chapter.

Evaluations of one-on-one tutoring have shown that the method is quite effective, even when the tutors are unskilled tutors. Unskilled tutors are defined in this chapter as tutors who are not experts on subject matter knowledge, are not trained systematically on tutoring skills, and are virtually never evaluated on their impact on student learning. Unskilled tutors are paraprofessionals, parents, community citizens, cross-age tutors, or same-age peers. Meta-analyses show learning gains from typical human tutors, the majority being unskilled, of approximately 0.4 sigma when compared to classroom controls and other suitable controls (Cohen, Kulik & Kulik, 1982).

There are many possible explanations of these learning gains from tutors who are unskilled. Perhaps the tutor can detect whether or not the student is generally mastering the subject matter on the basis of the student's verbal responses, from their nonverbal reactions, or from the student's attempts to perform a task. The tutor would then re-plan and make adjustments to help the student move forward. Perhaps the one-on-one attention motivates the student or encourages sufficient mastery to prevent embarrassing performance deficits. Perhaps the nature of conversation encourages a meeting of the minds with sufficient common ground for learning to be built on a solid discourse foundation. The question is still unsettled why one-on-one tutoring is so effective when the tutor is unskilled.

Available evidence suggests that the expertise of the tutor does matter, but the evidence is not strong. Collaborative peer tutoring shows an effect size advantage of 0.2 to 0.9 sigma (Johnson & Johnson, 1992; Mathes & Fuchs, 1994; Slavin, 1990; Topping, 1996), which appears to be slightly lower than older unskilled human tutors. Peer tutoring is a low-cost effective solution because

expert tutors are expensive and hard to find. Unfortunately, there have not been many systematic studies on learning gains from expert tutors because they are expensive, they are difficult to recruit in research projects, and tutors tend to stay in the tutoring profession for a short amount of time (Person, Lehman, & Ozbun, 2007). Certified tutors appear to yield the largest gains in tutoring (Slavin, Karweit, & Madden, 1989), so there is some evidence that training facilitates tutoring quality. Available studies report effect sizes of 0.8 to 2.0 (Bloom, 1984; Chi, Roy, & Hausmann, 2008; VanLehn et al., 2007), which is presumably higher than other forms of tutoring. The question is still unsettled on the impact of tutoring expertise on learning gains.

The impact of tutoring expertise on student learning is complicated by the fact that much of the answer lies in what the student does, not what the tutor does. Constructionist theories of learning have routinely emphasized the importance of getting the student to construct the knowledge, as opposed to an instruction delivery system that transfers the knowledge to the student (Bransford, Brown & Cocking, 2000; Mayer, 2009). Students learn by expressing, doing, explaining, and being responsible for their knowledge construction, as opposed to being passive recipients of exposure to information. There is considerable evidence for the constructivist thesis in general, (Bransford et al., 2000), but this chapter considers the evidence for constructivism in tutoring per se.

One form of evidence is that the tutors in these same-age and cross-age collaborations tend to learn more than the tutees (Cohen et al., 1982; Mathes & Fuchs, 1994; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003). Playing the role of tutor rather than tutee undoubtedly increases study, effort, initiative, and organization, all of which contribute to learning. In peer tutoring, students often are randomly assigned to tutor versus peer, so any advantages of the tutor role can not be explained by prior abilities. Another form of evidence lies in who contributes most to the tutoring

session. Is it the student or tutor? Correlational evidence reveals that students learn more when they contribute a higher percentage of the words and ideas to the tutoring sessions (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; Litman et al., 2006). A good tutor apparently says very little when the student is on a roll and learning. Yet another form of evidence is that it does not help much for the tutor to articulate explanations, solutions, and other critical content in the form of information delivery, without making any attempt to connect with what the learner knows (Chi et al., 2001; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003).   Explanations and other forms of high quality information are of course important when students are maximally receptive, for example after they try to solve a problem and fail (Schwartz & Bransford, 1998).   However, information delivery very often has a limited impact on the student when the content involves complex conceptualizations.

The obvious question that learning scientists have been asking over the years is why tutoring is effective in promoting learning? There are many approaches to answering this question. One approach is to conduct meta-analyses that relate learning gains with characteristics of the subject matter, tutee, tutor, and general structure of the tutoring session. There is evidence, for example, that (a) learning gains tend to be higher for well-structured, precise domains (mathematics, physics) than for ill-structured domains (reading), (b) that learning gains from tutors are more pronounced for tutees who start out with comparatively lower amounts of knowledge and skill, (c) that the quality of tutor training is much more important than the quantity of training, and (d) that a tutoring session shows more benefits when there are particular pedagogical activities (Cohen et al., 1982; Fuchs et al., 1994; King, Staffieri, & Adelgais, 1998; Mathes & Fuchs, 1994; Rohrbeck et al., 2003).

A second approach is to perform a very detailed analysis of the tutoring session structure, tasks, curriculum content, discourse, actions, and cognitive activities manifested in the sessions and to speculate how these might account for the advantages of tutoring (Chi, Roy, & Hausmann, 2008; Chi et al., 2001; Graesser & Person, 1994; Graesser, Person, & Magliano, 1995; Hacker & Graesser, 2007; Lepper, Drake, & O'Donnell-Johnson, 1997; McArthur, Stasz, & Zmuidzinas, 1990; Merrill, Reiser, Merrill, & Landes, 1995; Person & Graesser, 1999; Person, Kreuz, Zwaan, & Graesser, 1995; Shah, Evens, Michael, & Rovick, 2002; VanLehn et al., 2003). This chapter addresses these process factors in more detail.

A third approach is to manipulate the tutoring activities through trained human tutors or computer tutors and to observe the impact of the manipulations on learning gains (Chi et al., 2001, 2008; Graesser, Lu et al., 2004; Litman et al., 2003; VanLehn et al., 2003; VanLehn et al., 2007). Manipulation studies allow us to infer what characteristics of the tutoring directly cause increases in learning gains, barring potential confounding variables.

What Are the Common Tutoring Strategies and Processes?

As discussed, the typical tutors in school systems are unskilled. These tutors are nonetheless effective in helping students learn so it is worthwhile to explore what tutoring strategies and processes they frequently implement. Graesser and Person analyzed the discourse patterns of 13 unskilled tutors in great detail (Graesser, & Person, 1994; Graesser et al., 1995; Person & Graesser, 1999). They videotaped over 100 hours of naturalistic tutoring in a corpus of unskilled tutors who tutored middle school students in mathematics or college students in research methods. The research team transcribed the tutorial dialogues, classified the speech act utterances into discourse categories, and analyzed the rate of particular discourse patterns. We refer to this as the Graesser-Person *unskilled tutor corpus.* Regarding expert tutors, Person et al. (2007)

conducted a literature review of studies with accomplished tutors. Unfortunately, the sample sizes of expert tutors have been extremely small (N <3) in empirical investigations of expert tutoring and often the same expert tutors are used in different research studies; at times the tutors are co-authors on publications. Claims about expert tutoring may therefore be biased by the idiosyncratic characteristics of the small sample of tutors and the tutors' authorship role. Person et al. (2007) recently conducted a study on a sample of 12 tutors who were nominated by teachers in the Memphis community who were truly outstanding. The discourse patterns of these outstanding tutors in Person's *expert tutor corpus* were dissected in great detail.

Unfortunately, neither the unskilled tutor corpus nor the expert tutor corpus had outcome scores. There is a large void in the literature on detailed analyses of human tutorial dialogue that are related to outcome measures and that have a large sample of tutors. Part of the problem lies in logistical problems in obtaining such data. The subject matters of the tutoring sessions are difficult to predict in advance so it is difficult to proactively identify suitable pre-test and posttest measures from normative testbanks. Nevertheless, these tutoring corpora can be analyzed to identify the tutoring processes.

Sophistication of Tutoring Strategies

As one might expect, unskilled human tutors are not prone to implement sophisticated tutoring strategies that have been proposed in the fields of education, the learning sciences, and developers of ITSs (Graesser et al., 1995; Graesser, D'Mello, & Person, 2009; Person et al., 1995). Tutors rarely implement pedagogical techniques such as *bona fide* Socratic tutoring strategies, modeling-scaffolding-fading, Reciprocal Teaching, frontier learning, building on prerequisites, or diagnosis/remediation of deep misconceptions. In Socratic tutoring, the tutor asks learners illuminating questions that lead the learners to discover and correct their own

misconceptions in an active, self-regulated fashion (Collins et al., 1975). Thus, Socratic tutoring

is not merely bombarding the student with a large number of questions, as some practitioners and

researchers erroneously believe. In modeling-scaffolding-fading, the tutor first models a desired

skill, then gets the learners to perform the skill while the tutor provides feedback and explanation,

and finally fades from the process until the learners perform the skill all by themselves (Rogoff

& Gardner, 1984). As discussed, in Reciprocal Teaching, the tutor and learner take turns reading

and thinking aloud with the goal of lacing in question generation, summarization, clarification,

and prediction (Palincsar & Brown, 1984). Tutors who use frontier learning select problems and

give guidance in a fashion that slightly extends the boundaries of what the learner already knows

or has mastered (Sleeman & Brown, 1982). Tutors who build on prerequisites cover the

prerequisite concepts or skills in a session before moving to more complex problems and tasks

that require mastery of the prerequisites (Gagné, 1985). Tutors who diagnose and remediate deep

misconceptions are on the lookout for errors that are manifestation of more global problematic

mental models (Lesgold, Lajoie, Bunzo, & Eggan, 1992). When a deep misconception is

recognized, the tutor attempts to supplant the error-ridden mental model with a correct mental

model.

One would expect tutors to be able to help the students correct their idiosyncratic deficits in

knowledge and skills. Tutors are no doubt sensitive to some of these deficits but available data

suggest there are limitations. Two examples speak to such limitations. First, if a tutor is truly

adaptive to the student's learning profile, then the tutor should initiate some discussion or activity at

the beginning of the session that diagnoses what the student is struggling with. This adaptation is

manifested when the tutor: (a) inspects previous test materials and scores of the student, (b) selects

problems in the tutoring session that are associated with the student's deficits, and (c) asks the tutee

what they are having problems with. A tutor would lack the principle of adaptation if the tutor

immediately presents problems to work on in a scripted fashion for all students. Whereas *a* and *c*

occur with some frequency, tutors are not prone to do *b* (Chi et al., 2008; Graesser et al., 1995).

The second example of the tutor's limited ability to detect student's knowledge deficits

addresses metacognitive knowledge. Tutors frequently ask students comprehension-gauging

questions, such as *Do you understand? Are you following?*,and *Does that make sense?* If the

student's comprehension calibration skills are accurate, then the student should answer YES when

the student understands and NO when there is little or no understanding. One counterintuitive

finding in the tutoring literature is that there sometimes is a positive correlation between a student's

knowledge of the material (based on pre-test scores or post-test scores) and their likelihood of

saying NO rather than YES to the tutors' comprehension-gauging questions (Chi, Bassock, Lewis,

Reimann, & Glaser, 1989; Graesser et al., 1995). Thus, it is the knowledgeable students who tend

to say *No, I don't understand*.   This result suggests that deeper learners have higher standards of

comprehension (Baker, 1985; Otero & Graesser, 2001) and that many students have poor

comprehension calibration skills. The finding that students have disappointing comprehension

calibration is well documented in the metacognitive literature, where meta-analyses have shown

only a 0.27 correlation between comprehension scores on expository texts and the students'

judgments on how well they understand the texts (Dunlosky & Lipko, 2007; Glenberg, Wilkinson,

& Epstein, 1982; Maki, 1998). It is perhaps not surprising that the student's comprehension

calibration is poor because they are low in domain knowledge. From the perspective of the tutor,

many tutors mistakenly believe the students' answers to the comprehension-gauging questions and

that reflects insensitivity to the students' knowledge states. A good tutor would periodically ask

follow-up questions when students say YES, they understand.

The aforementioned examples suggest that human tutors are insensitive to the students' knowledge states, but such a generalization would be too sweeping. Tutors are often adaptive to the students' knowledge and skills at a micro-level, as opposed to the macro-levels in the above two examples. The distinction is what VanLehn (2006) calls the *inner loop* versus the *outer loop*. The inner loop consists of covering individual steps or expectations within a problem whereas the outer loop involves the selection of problems, the judgment of mastery of a problem, and other more global aspects of the tutorial interaction. Available analyses of human tutoring suggest that human tutors are more sensitive to the students' knowledge at the inner loop than the outer loop.

Dialogue Patterns in Tutoring

Graesser and Person's analyses of tutorial dialogue uncovered a number of frequent dialogue structures (Graesser & Person, 1994; Graesser et al., 1995; Graesser, Hu, & McNamara, 2005). Many of these structures were also prominent in the work of other researchers who have conducted fine-grained analyses of tutoring (Chi et al., 2001; Chi et al., 2004; Chi et al., 2008; Evens & Michael, 2005; Litman et al., 2006; Shah et al., 2002). The following three dialogue structures are prominent: (a) the 5-step tutoring frame, (b) expectation and misconception tailored dialogue, and (c) conversational turn management. All of these structures are in the inner loop: c is embedded in b, which in turn is embedded in a. It should be noted that it is the tutor who takes the initiative in implementing these structures, not the student. It is rare to have the student take charge of the tutorial session in a self-regulated manner.

5-step tutoring frame. Once a problem or difficult main question is selected to work on, the 5-step tutoring frame is launched, as specified below.

(1) TUTOR asks a difficult question or presents a problem.

(2) STUDENT gives an initial answer.

(3) TUTOR gives short feedback on the quality of the answer.

(4) TUTOR and STUDENT have a multi-turn dialogue to improve the answer.

(5) TUTOR assesses whether the student understands the correct answer.

This 5-step tutoring frame involves collaborative discussion, joint action, and encouragement for the student to construct knowledge rather than merely receiving knowledge.

The first 3 steps often occur in a classroom context, but the questions are easier short-answer questions. The Initiate-Respond-Evaluate (IRE) sequence in a classroom consists of the teacher initiating a question, the student giving a short-answer response, and the teacher giving a positive or negative evaluation of the response (Sinclair & Coulthart, 1975). This is illustrated in the exchange below on the subject matter of Newtonian physics.

(1) TEACHER: According to Newton's second law, force equals mass times what?

(2) STUDENT: acceleration

(3) TEACHER: Right, mass times acceleration.

Or

(2) STUDENT: velocity

(3) TEACHER: Wrong, it's not velocity, it is acceleration.

Thus, tutoring goes beyond the IRE sequence in the classroom by having more difficult questions and more collaborative interactions during step 4 of the 5-step tutoring frame.

Expectation and misconception tailored dialogue. Human tutors typically have a list of expectations (anticipated good answers, steps in a procedure) and a list of anticipated *misconceptions* associated with each main question. For example, expectations E1 and E2 and misconceptions M1 and M2 are relevant to the example physics problem below.

PHYSICS QUESTION: If a lightweight car and a massive truck have a head-on collision, upon which vehicle is the impact force greater? Which vehicle undergoes the greater change in its motion, and why?

E1. The magnitudes of the forces exerted by A and B on each other are equal.

E2. If A exerts a force on B, then B exerts a force on A in the opposite direction.

M1: A lighter/smaller object exerts no force on a heavier/larger object.

M2: Heavier objects accelerate faster for the same force than lighter objects.

The tutor guides the student in articulating the expectations through a number of dialogue moves: *pumps*, *hints*, and *prompts* for the student to fill in missing words. A pump is a generic expression to get the student to provide more information, such as "What else" or "Tell me more." Hints and prompts are selected by the tutor to get the student to articulate missing content words, phrases, and propositions. A hint tries to get the student to express a complex idea (e.g., proposition, clause, sentence) whereas a prompt is a question that tries to get the student to express a single word or phrase. For example, a hint to get the student to articulate expectation E1 might be "What about the forces exerted by the vehicles on each other?"; this hint would ideally elicit the answer "The magnitudes of the forces are equal." A prompt to get the student to say "equal" would be "What are the magnitudes of the forces of the two vehicles on each other?" As the learner expresses information over many turns, the list of expectations is eventually covered and the main question is scored as answered.

Human tutors are dynamically adaptive to the learner in ways other than prompting them to articulate expectations. There also is the goal of correcting misconceptions that arise in the student's responses. When the student articulates a misconception, the tutor acknowledges the error and corrects it. There is another conversational goal of giving feedback to the student on

their contributions. For example, the tutor gives short *feedback* on the quality of student

contributions. The tutor accommodates a mixed-initiative dialogue by attempting to answer the

student's questions when the student is sufficiently inquisitive to ask questions. However, it is

well documented that students rarely ask questions, even in tutoring environments (Graesser &

Person, 1994; Graesser, McNamara, & VanLehn, 2005), because they have limited self-regulated

learning strategies (Azevedo & Cromley, 2004). Tutors are considered more adaptive to the

student to the extent that they correct student misconceptions, give correct feedback, and answer

student questions.

      Conversational turn management. Human tutors structure their conversational turns

systematically. Nearly every turn of the tutor has three informational components after the main

problem or question has been introduced and the collaboration is flowing in step 4 of the tutoring

frame:

      Tutor Turn → Short Feedback + Dialogue Advancer + Floor Shift

 The first component of most turns is feedback on the quality of the student's last turn.   This

feedback is either positive (*very good, yeah*), neutral (*uh huh, I see*), or negative (*not quite, not

really*). Sometimes the tutor expresses this short feedback through nonverbal paralinguistic cues,

such as intonation, facial expressions, gestures, or body movements. The second dialogue

advancer component moves the tutoring agenda forward with either pumps, hints, prompts,

assertions with correct information, corrections of misconceptions, or answers to student

questions. The third floor shift component is a cue for the conversational floor to shift from the

tutor as the speaker to the student. For example, the human ends each turn with a question or a

gesture to cue the student to do the talking. Questions strongly invite responses from the

conversation partner so the student is expected say something after the tutor asks a question.

Alternatively, the tutor can signal a floor shift through a hand gesture, posture display, or facial expression that invites the student to contribute. These floor shift signals need to be dramatic when the student is reluctant to contribute.

The three conversational structures together present challenging problems or questions to the student, adaptively scaffold good answers through collaborative interactions, provide feedback when students express erroneous information, and answers student questions that infrequently are asked. What is absent is sophisticated pedagogical strategies. This is perhaps unsurprising because these strategies are complex and took centuries to discover by scholars. However, it is a very important finding to document because it is conceivable that deep learning could improve tremendously by training human tutors and programming computer tutors to implement the sophisticated strategies.

The pedagogical strategies of expert tutors are very similar to those of unskilled tutors in most ways (Cade, Copeland, Person, & D'Mello, 2008; Person, Lehman, & Ozbun, 2007). However, Cade et al. (2008) did identify a few notable trends in pedagogy in the expert tutor corpus. The expert tutors did occasionally implement modeling-scaffolding-fading, although the relative frequencies of the dialogue moves for this pedagogical strategy were not impressively high. The tutors did a modest amount of modeling, a large amount of scaffolding, and very little fading. These tutors periodically had just-in-time direct instruction or mini-lectures when the student was struggling with a particular conceptualization. These content-sensitive mini-lectures allegedly were sensitive to what the student was having trouble with rather than being routinely delivered to all students.

The expert tutors also appeared to differ from unskilled tutors on some metacognitive dimensions, as addressed below. However, it is important to qualify these claims about expert

tutors because there was never a systematic comparison of tutors with different expertise in any

given study. Instead, the relative frequencies of tutor strategies and discourse moves were

computed in the expert tutor corpus and compared with the relative frequencies of the same

theoretical categories in published studies with unskilled tutors. One pressing research need is to

systematically compare tutors with varying expertise.

What Is the Role of Metacognition and Meta-communication in Tutoring?

Graesser, D'Mello, and Person (2009) have documented some of the illusions that typical

human tutors have about cognition and communication. These illusions may get in the way of

optimizing learning. Expert tutors also may be less likely to fall prey to these illusions. The five

illusions below were identified.

(1) *Illusion of grounding*. The unwarranted assumption that the speaker and listener have

shared knowledge about a word, referent, or idea being discussed in the tutoring session.

Failure to establish common ground threatens successful communication and the joint

construction of knowledge (Clark, 1996). A good tutor is sufficiently skeptical of the

student's level of understanding that the tutor trouble-shoots potential communication

breakdowns between the tutor and student.

(2) *Illusion of feedback accuracy*. The unwarranted assumption that the feedback that the

other person gives to a speaker's contribution is accurate. For example, tutors

incorrectly believe the students' answers to their comprehension gauging questions (e.g.,

"Do you understand?").

(3) *Illusion of discourse alignment*. The unwarranted assumption that the listener does

understand or is expected to understand the discourse function, intention, and meaning

of the speaker's dialogue contributions. For example, tutors sometimes give hints, but the students do not realize they are hints.

(4) *Illusion of student mastery*. The unwarranted assumption that the student has mastered much more than the student has really mastered. For example, the fact that a student expresses a word or phrase does not mean that the student understands an underlying complex idea.

(5) *Illusion of knowledge transfer*. The speaker's unwarranted assumption that the listener understands whatever the speaker says and thereby knowledge is accurately transferred. For example, the tutor assumes that the student understands whatever the tutor says, when in fact the student absorbs very little.

Both the tutor and student may each have these illusions and thereby compromise the effectiveness of tutoring.

These illusions undermine the tutor's building an accurate and detailed model of the cognitive states of the student, or what is called the *student model*. Indeed, there are reasons for being pessimistic about the quality of the student model that tutors construct. A more realistic picture is that the tutor has only an approximate appraisal of the cognitive states of students and that they formulate responses that do not require fine-tuning of the student model (Chi et al., 2004; Graesser et al., 1995).

There are three sources of evidence for this claim. First, the short feedback to students on the quality of the students' contributions is often incorrect. In particular, the short feedback has a higher likelihood of being positive than negative after student contributions that are vague or error-ridden (Graesser et al., 1995). Tutors have the tendency to be polite or to resist discouraging the student by giving a large amount of negative feedback (Person et al., 1995).

Second, tutors do not have a high likelihood of detecting misconceptions and error-ridden contributions of students (Chi, Siler, & Jeong, 2004; VanLehn et al., 2007). Third, as mentioned earlier, tutors do not select new cases or problems to work on that are sensitive to the abilities and knowledge deficits of students (Chi et al., 2008). One would expect the selection of problems to be tailored to the student's profile according to the zone of proximal development, i.e., not too easy or not too hard, but just right. However, Chi et al. (2008) reported that there was no relation between problem selection and student's profile. Data such as these lead one to conclude that tutors have a modest ability to conduct student modeling.

A good tutor is sufficiently skeptical of the student's level of understanding. The tutor trouble-shoots potential communication breakdowns between the tutor and student. This is illustrated in the simple hypothetical exchange below.

TUTOR: We know from Newton's law that net force equals mass times acceleration.

   This law ….

STUDENT: Yeah, that is Newton's second law.

TUTOR: Do you get this?

STUDENT: Yeah. I know that one.

TUTOR: Okay, let's make sure. Force equals mass times what?

STUDENT: times velocity.

TUTOR: No, it's mass times acceleration.

A good tutor assumes that the student understands very little of what the tutor says and that knowledge transfer approaches zero. Person et al. (2007) has reported that expert tutors are more likely to verify that the student understands what the tutor expresses by asking follow up questions or giving follow-up trouble-shooting problems.

What Is the Role of Emotions During Tutoring?

It is important to consider motivation and emotion in tutoring in addition to the cognitive subject matter. Indeed, connections between complex learning and emotions have received increasing attention in the fields of psychology and education (Deci & Ryan, 2002; Dweck, 2002; Gee, 2003; Lepper & Henderlong, 2000; Linnenbrink & Pintrich, 2002; Meyer & Turner, 2006). Studies that have tracked the emotions during tutoring have identified the predominate emotions, namely confusion, frustration, boredom, anxiety, and flow/engagement, with delight and surprise occurring less frequently (Baker, D'Mello, Rodrigo, Graesser, in press; Craig, Graesser, Sullins, & Gholson, 2004; D'Mello et al., 2008; Graesser, Picard, & Graesser, 2007; Lehman, Matthews, D'Mello, & Person, 2008). These data are informative, but the important question is how these emotions can be coordinated productively with learning.

The central assumption is that it is important for tutors to adopt pedagogical and motivational strategies that are effectively coordinated with the students' emotions. Lepper, Drake, and O'Donnell (1998) proposed an INSPIRE model to promote this integration. This model encourages the tutor to nurture the student by being empathetic and attentive to the student's needs, to assign tasks that are not too easy or difficult, to give indirect feedback on erroneous student contributions rather than harsh feedback, to encourage the student to work hard and face challenges, to empower the student with useful skills, and to pursue topics they are curious about. One of the interesting tutor strategies is to assign an easy problem to the student, but to claim that the problem is difficult and to encourage the student to give it a try anyway. When the student readily solves the problem, the student builds self-confidence and self-efficacy in conquering difficult material (Zimmerman, 2001).

Several theories linking emotions and learning have been proposed. Meyer and Turner (2006) identified three theories that are particularly relevant to understanding the links between emotions and learning: academic risk taking, flow, and goals (Meyer & Turner, 2006). The academic risk theory contrasts (a) the adventuresome learners who want to be challenged with difficult tasks, take risks of failure, and manage negative emotions when they occur and (b) the cautious learners who tackle easier tasks, take fewer risks, and minimize failure and the resulting negative emotions. According to flow theory (Csikszentmihaly, 1990), the learner is in a state of flow when the learner is so deeply engaged in learning the material that time and fatigue disappear. When students are in the flow state, they are at an optimal zone of facing challenges and conquering the challenges by applying their knowledge and skills. Goal theory emphasizes the role of goals in predicting and regulating emotions (Dweck, 2002; Stein & Hernandez, 2007). Outcomes that achieve challenging goals result in positive emotions whereas outcomes that jeopardize goal accomplishment result in negative emotions.

A complementary perspective is to focus on learning impasses and obstacles rather than on flow and goals. Obstacles to goals are particularly diagnostic of both learning and emotions. For example, the affective state of confusion correlates with learning gains perhaps because it is a direct reflection of deep thinking (Craig et al., 2004; D'Mello eta l., 2008; Graesser, Jackson, & McDaniel, 2007). Confusion is diagnostic of *cognitive disequilibrium*, a state that occurs when learners face obstacles to goals, contradictions, incongruities, anomalies, uncertainty, and salient contrasts (Graesser, Lu, Olde, Pye-Cooper, & Whitten, 2005; Otero & Graesser, 2001). Cognitive equilibrium is ideally restored after thought, reflection, problem solving and other effortful deliberations. It is important to differentiate being productively confused, which leads to

learning and ultimately positive emotions, from being hopelessly confused, which has no pedagogical value.

Research is conspicuously absent on how the tutees perceive the causes and consequences of these emotions and what they think they should do to regulate each affect state. The negative emotions are particularly in need of research. When a student is frustrated from being stuck, the student might attribute the frustration either to themselves ("I'm not at all good at physics"), the tutor ("My tutor doesn't understand this either"), or the materials ("This must be a lousy textbook"). Solutions to handle the frustration would presumably depend on these attributions of cause of the frustration. When a student is confused, some students may view this as a positive event to stimulate thinking and show their metal in conquering the challenge; other students will attribute the confusion to their poor ability, an inadequate tutor, or poorly prepared academic materials. When students are bored, they are likely to blame the tutor or material rather than themselves. Tutors of the future will need to manage the tutorial interaction in a fashion that is sensitive to the students' emotions in addition to their cognitive states.

What Are Tutoring Strategies that Influence Deep Learning?

So far this chapter has provided evidence for the effectiveness of human tutoring and has identified various strategies and processes of naturalistic human tutoring. The obvious next question is which of the strategies help learning? Surprisingly, there is not an abundance of research on this question because it is difficult to control what human tutors do in controlled experiments, but we review some relevant research on tutoring strategies in this section.

Chi, Roy, and Hausmann (2008) compared 5 conditions in order to test a hypothesis they were advancing called the *active/constructive/interactive/observing* hypothesis. As the expression indicates, the hypothesis asserts that learning is facilitated from active student learning, knowledge

construction, and collaborative interaction, as we have discussed in this chapter. The other aspect of

the expression refers to observing a tutoring session vicariously. Their ideal experimental condition

involves 4 people: two student participants watching and discussing a tutorial interaction that occurs

between a tutor and another student. According to the hypothesis, the participants would learn a

great deal from this interactive vicarious observation condition because it has all four components

(action, construction, interaction, observation). To test this hypothesis, students trying to learn

physics were randomly assigned to the ideal treatment (condition 1) versus to one-on-one tutoring

(condition 2), vicarious observation (all alone) of the tutoring session (condition 3), collaboratively

interacting with another student without observing the interaction (condition 4), and studying from a

text alone (condition 5). Conditions 1 and 2 were approximately the same in learning gains and

significantly higher than conditions 3-5. Thus, it appears that multiple components are needed for

learning to be optimal.

As discussed earlier, there is evidence that learning from tutorial interactions improves when

the learner constructs explanations and when the student does more of the talking than the tutor

(Litman et al., 2006; Siler & VanLehn, 2009). However, Chi et al. (2001) examined the type of tutor

moves in detail for students learning physics. For deep learners, it was the tutor moves that

encouraged reflection that helped; for shallow learners, the tutor's responses to scaffolding and

explanations were important. Unfortunately, the sample sizes in these studies reported by Chi,

Litman, and Siler are modest and very much in need of replication. The door is clearly open for

discovering the particular dialogue moves of tutors predict learning.

Research on reading tutors have also investigated what aspects of tutoring help reading at

deeper levels of compression (McNamara, 2007). Three notable examples are *Reciprocal Teaching*

(Palincsar & Brown, 1984), *Self Explanation Reading Training, SERT* (McNamara, 2004), and

Questioning the Author (Beck, McKeown, Hamilton, & Kucan, 1997). Reciprocal Teaching was

mentioned at the beginning of this chapter and we reported this method has solid learning gains. The

key strategies were clarifying, questioning, summarizing, and predicting content as students read

text. The SERT method helps students build self-explanations when reading the text, which

includes the strategies of paraphrasing, generating inferences, bridging ideas expressed in the text,

and connecting the text to what the student knows. Questioning the Author encourages the student

to critically evaluate the content of what is writing by asking the author such questions as "what is

the evidence for this claim?" and "Why did the author mention this?" The available evidence,

including meta-analyses and reviews of research (Roscoe & Chi, 2007; Rosenshine & Meister,

1996; Rosenshine, Meister, & Chapman, 1996), is that the scaffolding of explanations and of deep

questions and answers are particularly important components. Explanations involve causal chains

and networks, plans of agents, and logical justifications of claims. Deep questions have been

defined systematically (Graesser & Person, 1994) and include questions stem such as why, how,

what if, what if not, and so what? In contrast, the strategy to predict future content in the text has

little or no impact on improving reading at deeper levels, whereas summarization and clarification

are somewhere in between.

Part of the challenge of conducting experimental research on human tutoring is that it is

difficult to train tutors to adopt particular strategies. They rely on their normal conversational and

pedagogical styles. It is nearly impossible to run repeated measures designs where a tutor adopts a

normal style on some days and an experimental style on other days. The treatments end up

contaminating each other and it is difficult to force the human tutors to adopt changes in their

language and discourse, particularly those levels that are unconscious and involuntary. However,

computers can supply such experimental control. Therefore, computer tutors are expected to play a

more important role both in future scientific investigations and also in increasing tutoring pedagogy to an increasing number of students.

Future Directions

This chapter has made a convincing case that tutoring by humans is a powerful learning environment. It could be argued that tutoring is the most effective learning environment we know of in addition to being the oldest. Tutoring has been around for millennia and has been shown to help learning in several meta-analyses, as we have documented in this chapter. However, there are still a large number of unanswered fundamental questions that need attention in future research.

Rather surprisingly, there needs to be a systematic line of research that investigates the impact of tutoring expertise on learning gains as well as learner emotions and motivation. We had hoped to find a rigorous study that randomly assigns students to human tutors with varying levels of expertise and that collects suitable outcome measures. The fact that we came up empty is remarkable, but it also sets the stage for new research initiatives. To what extent is student learning and motivation facilitated as a function of increased tutor training on pedagogy and/or increased subject matter knowledge? To what extend does tutoring experience matter? How do different schools of tutoring pedagogy compare? Are there interactions between tutor pedagogy, subject matter, and student profiles? How do we best train tutors? Decades of research is needed to answer these questions?

Computer tutors have some promise in providing more control over the tutoring process than human tutors can provide. This opens up the possibility of new programs of research that systematically compare different versions of computer tutors and other advanced learning environments. These systems have multiple modules, such as the knowledge base, the profile of

student ability and mastery on particular topics, decision rules that select problems, scaffolding strategies, help systems, feedback, media on the human-computer interface, and so on. Which of these components are responsible for any learning gains of the computer tutors? It is possible to systematically manipulate the quality or presence of each component in *lesion* studies that systematically remove particular components and then assess the impact of the removal on learning. The number of conditions in manipulation studies of course grows with the number of components. If there are 6 major components, with each level varying in 2 levels of quality, then there would be $2^6 = 64$ conditions in a factorial design. That would require nearly 2000 students in a between-subjects design with 30 students randomly assigned to each of the 64 conditions. It indeed might be realistic to perform such a lesion study to the extent that the computer tutor enterprise scales up and delivers training on the web (see Heffernan, Koedinger, & Razzaq, 2008). The alternative would be to selectively focus on one or two modules at a time.

The same comparisons could be made between alternative human tutors. Studies could be conducted to carefully train human tutors to include versus exclude particular tutoring components. For example, should the human tutor respond to the student emotions or ignore their emotions? Should the tutor give negative feedback or stick with positive feedback? Should the human tutor explain the rationale behind answers, or merely give the correct answers? Once again, there are many variables and combinations to test, so this is a research area that could attract the attention of researchers for years.

One of the provocative tests in the future will pit human versus machine as tutors. Most people place their bets on the human tutors under the assumption that they are more sensitive to the student's profile and are more creatively adaptive in guiding the student. However, the detailed analyses of human tutoring challenge such assumptions in light of the many illusions

that humans have about communication and the modest pedagogical strategies in their repertoire. Computers may do a better job in cracking the illusions of communication, in inducing student knowledge states, and in implementing complex intelligent tutoring strategies. A plausible case could easily be made for betting on the computer over the human tutor. Perhaps the ideal computer tutor emulates humans in some ways and complex non-human computations in other ways. Comparisons between human and computer tutors need to be made in a manner that equilibrates the conditions on content, time on task, and other extraneous variables that are secondary to pedagogy. As data roll in from these needed empirical studies, we need to be open to the prospects of some unpredictable and counterintuitive discoveries.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors:

Lessons learned. *Journal of the Learning Sciences*, *4*, 167-207.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning faciliate

students' learning with hypermedia. *Journal of Educational Psychology, 96,* 523-535.

Baker, L. (1985). Differences in standards used by college students to evaluate their

comprehension of expository prose. *Reading Research Quarterly, 20*, 298-313.

Baker, R.S., D'Mello, S.K., Rodrigo, M.T., Graesser, A.C. (in press). Better to be frustrated than

bored: The incidence, persistence, and impact of learners' affect during interactions with

three different computer-based learning environments. *International Journal of Human-*

Computer Studies.

Beck, I.L., McKeown, M.G., Hamilton, R.L., & Kucan, L. (1997). *Questioning the Author: An*

approach for enhancing student engagement with text. Delaware: International Reading

Association.

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*, 4-16.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How People Learn*.

Washington, D.C.: National Academy Press.

Cade, W., Copeland, J. Person, N., and D'Mello, S. K. (2008). Dialogue modes in expert

tutoring. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the Ninth*

International Conference on Intelligent Tutoring Systems (pp. 470-479). Berlin, Heidelberg:

Springer-Verlag

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How

students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145-

182.

Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008) Observing tutorial dialogues

collaboratively: Insights about human tutoring effectiveness from vicarious learning.

Cognitive Science, *32*(2)*,* 301-341.

Chi, M.T.H., Siler, S.A. & Jeong, H. (2004). Can tutors monitor students' understanding

accurately? *Cognition and Instruction, 22*(3), 363-387.

Chi, M.T.H., Siler, S., Yamauchi, T., Jeong, H. & Hausmann, R. (2001). Learning from human

tutoring. *Cognitive Science*, *25*, 471- 534.

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

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

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.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts

of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and*

instruction: Essays in honor of Robert Glaser (pp. 453-494). Hillsdale, NJ: Lawrence

Erlbaum Associates.

Collins, A., & Halverson, R. (2009). *Rethinking education in the age of technology: The digital*

revolution and schooling in America. New York: Teacher College Press.

Collins, A., Warnock, E. H., Aeillo, N., Miller, M. L. (1975). Reasoning from incomplete

knowledge. In D. G. Bobrow A. Collins (Eds.), *Representation and understanding* (pp. 453-

494). New York: Academic.

Conley, M., Kerner, M., & Reynolds, J. (2005). Not a question of should, but a question of how: Literacy knowledge and practice into secondary teacher preparation through tutoring in urban middle schools. *Action in Teacher Education, 27*, 22-32.

Corbett, A.T. (2001). Cognitive computer tutors: Solving the two-sigma problem. *User Modeling: Proceedings of the Eighth International Conference, UM 2001,* 137-147.

Craig, S.D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning. *Journal of Educational Media, 29,* 241-250.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper-Row.

D'Mello, S. K., Craig, S.D., Witherspoon, A. W., McDaniel, B. T., and Graesser, A. C. (2008). Automatic Detection of Learner's Affect from Conversational Cues. *User Modeling and User-Adapted Interaction*, *18*(1-2), 45-80.

D'Mello, S.K., Picard, R., & Graesser, A.C. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems, 22*, 53-61.

Deci, E. L., & Ryan, R. M. (2002). The paradox of achievement: The harder you push, the worse it gets. In J.Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 61-87). Orlando, FL: Academic Press.

Dunlosky, J., & Lipko, A. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science, 16,* 228-232.

Dweck, C. S. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving academic*

achievement: Impact of psychological factors on education (pp. 61-87). Orlando, FL:
Academic Press.

Evens, M. W., & Michael, J. (2005). *One-on-one tutoring by humans and computers*. New
York: Rutledge/Taylor & Francis.

Fuchs, L., Fuchs, D., Bentz, J., Phillips, N., & Hamlett, C. (1994). The nature of students'
interactions during peer tutoring with and without prior training and experience. *American
Educational Research Journal, 31*, 75-103.

Gagne, R. M. (1985). *The conditions of learning and theory of instruction* (4th ed.). New York:
Holt, Rinehart, & Winston.

Gee, J.P. (2003). *What video games have to teach us about language and literacy*. New York:
Macmillan.

Glenberg, A. M., Wilkinson, A. C., and Epstein, W. (1982). The illusion of knowing: Failure in
the self-assessment of comprehension. *Memory & Cognition*, *10,* 597-602.

Graesser, A. C., D'Mello, S. K., & Person, N., (2009). Meta-knowledge in tutoring. In D. J.
Hacker, J. Dunlosky, & A. C. Graesser (Eds.). *Metacognition in educational theory and
practice*. Mahwah, NJ: Erlbaum.

Graesser, A.C., Jackson, G.T., & McDaniel, B. (2007). AutoTutor holds conversations with
learners that are responsive to their cognitive and emotional states. *Educational Technology,
47*, 19-22.

Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M.M.
(2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research
Methods, Instruments, and Computers, 36*, 180-193.

Graesser, A.C., Lu, S., Olde, B.A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and

eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices

when the devices break down. *Memory and Cognition, 33*, 1235-1247.

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension

strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist, 40*,

225-234.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American

Educational Research Journal, 31*, 104-137.

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.

Hacker, D.J., & Graesser, A.C. (2007). The role of dialogue in reciprocal teaching and

naturalistic tutoring. In R. Horowitz (Ed.), *Talk about text: How speech and writing interact

in school learning*. Mahwah, NJ: Erlbaum.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses related to

achievement*. New York: Routledge/Taylor & Francis.

Heffernan,N.T., Koedinger, K.R., & Razzaq, L.(2008) Expanding the model-tracing architecture:

A 3rd generation intelligent tutor for Algebra symbolization. *The International Journal of

Artificial Intelligence in Education. 18(2).* 153-178

Hock, M., Schumaker, J., & Deshler, D. (1995). Training strategic tutors to enhance learner

independence. *Journal of Developmental Education, 19,* 18-26.

Johnson, D.W., & Johnson, R.T. (1992). Implementing cooperative learning.*Contemporary

Education,* 63(3), 173–180.

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring

 tutorial interaction to scaffold peer learning. *Journal of Educational Psychology*, *90*, 134-

 152.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. (1997). Intelligent tutoring goes

 to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-

 43.

Lehman, B. A., Matthews, M., D'Mello, S. K., and Person, N. (2008). Understanding students'

 affective states during learning. In B. P. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.),

 Intelligent Tutoring Systems: 9th International Conference. Heidelberg, Germany: Springer.

Lepper, M. R., Drake, M., & O'Donnell-Johnson, T. M. (1997). Scaffolding techniques of expert

 human tutors. In K. Hogan & M. Pressley (Eds), *Scaffolding student learning: Instructional

 approaches and issues* (pp. 108-144). New York: Brookline Books.

Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25

 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J.

 M.Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation

 and performance* (pp.257-307). San Diego, CA: Academic Press.

Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the

 study of highly effective tutors. In J. Aronson (Ed.), *Improving academic achievement:

 Impact of psychological factors on education* (pp. 135-158). Orlando, FL: Academic Press.

Lesgold, A., Lajoie, S. P., Bunzo, M., & Eggan, G. (1992). SHER-LOCK: A coached practice

 environment for an electronics trouble-shooting job. In J. H. Larkin & R. W. Chabay (Eds.),

 *Computer assisted instruction and intelligent tutoring systems: Shared goals and complementary

 approaches* (pp. 201–238). Hillsdale, NJ: Erlbaum.

Linnenbrink, E. A. & Pintrich, P. (2002). The role of motivational beliefs in conceptual change.

In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice.* Dordretch, Netherlands: Kluwer Academic Publishers.

Litman, D.J, Rose, C.P., Forbes-Riley, K., VanLehn, K., Bhembe, D., and Silliman, S. (2006).

Spoken versus typed human and computer dialogue tutoring. *International Journal of Artificial Intelligence in Education*, *16*, 145-170.

Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C.

Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117-144).   Mahwah, NJ: Erlbaum.

Mathes, P. G., & Fuchs, L. S. (1994). Peer tutoring in reading for students with mild disabilities:

A best evidence synthesis. *School Psychology Review*, *23*, 59-80.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed). New York: Cambridge University Press.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*, 197 - 244.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 1-30.

McNamara, D.S. (2007)(Ed.), *Theories of text comprehension: The importance of reading strategies to theoretical foundations of reading comprehension*. Mahwah, NJ: Erlbaum.

Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided learning by

doing. *Cognition and Instruction, 13*(3), 315-372.

Meyer, D. K., & Turner, J. C. (2006). Re-conceptualizing emotion and motivation to learn in

classroom contexts. *Educational Psychology Review*, *18*, 377-390.

Otero, J., & Graesser, A.C. (2001). PREG: Elements of a model of question asking. *Cognition &*
   Instruction, 19, 143-175.

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

Palincsar, A. S., & Brown, A. L. (1988). Teaching and practicing thinking skills to promote
   comprehension in the context of group problem solving. *Remedial and Special Education
   (RASE), 9*(1), 53-59.

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.

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.

Person, N., Lehman, B., & Ozbun, R. (2007). Pedagogical and motivational dialogue moves
   used by expert tutors. Presented at the 17th Annual Meeting of the Society for Text and
   Discourse. Glasgow, Scotland.

Ritter, S., Anderson, J. R., Koedinger, K. R., Corbett, A. (2007) Cognitive Tutor: Applied
   research in mathematics education. *Psychonomic Bulletin & Review, 14*, 249-255.

Rogoff, B. & Gardner, W., (1984). Adult guidance of cognitive development. In: Rogoff, B. and
   Lave, J., Editors, 1984. *Everyday cognition: Its development in social context*, Harvard
   University Press, Cambridge, MA, pp. 95–116.

Rohrbeck, C. A., Ginsburg-Block, M., Fantuzzo, J. W., & Miller, T. R. (2003). Peer assisted

   learning Interventions with elementary school students: A Meta-Analytic Review. *Journal of*

   Educational Psychology, 95 (2), 240-257.

Roscoe, R.D., & Chi, M.T.H. (2007). Understanding tutor learning: Knowledge-building and

   knowledge-telling in peer tutors' explanations and questions. *Review of Educational*

   Research, 77, 534-574.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of*

   Educational Research, 64(4), 479-530.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A

   review of the intervention studies. *Review of Educational Research, 66,* 181-221.

Schwartz, D.L., & Bransford, J.D. (1998). A time for telling. *Cognition & Instruction, 16(4),*

   475-522.

Shah, F., Evens, M.W., Michael, J., & Rovick, A. (2002). Classifying student initiatives and

   tutor responses in human keyboard-to keyboard tutoring sessions. *Discourse Processes*, *33*,

   23-52.

Sinclair, J. & Coulthart, M. (1975) *Towards an analysis of discourse: The English used by*

   teachers and pupils. London: Oxford University Press.

Slavin, R.E. (1990). *Cooperative learning: Theory, research, and practice.* New Jersey: Prentice

   Hall.

Slavin, R., Karweit, N., & Madden, N. (1989). *Effective programs for students at risk*. Boston:

   Allyn and Bacon.

Sleeman D. & J. S. Brown. (1982)(Eds.). *Intelligent Tutoring Systems*. Orlando, Florida:

   Academic Press, Inc.

Stein, N. L., & Hernandez, M.W. (2007). Assessing understanding and appraisals during

emotional experience: The development and use of the Narcoder. In J. A. Coan & J. J. Allen

(Eds.), *Handbook of emotion elicitation and assessment* (pp. 298-317). New York: Oxford

University Press.

Topping, K. (1996). The effectiveness of peer tutoring in further and higher education: A

typology and review of the literature. *Higher Education, 32*, 321-345.

VanLehn, K. (2006) The behavior of tutoring systems. *International Journal of Artificial

Intelligence in Education*. 16, 3, 227-265.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (2007). When

are tutorial dialogues more effective than reading? *Cognitive Science*, *31*, 3-62.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W.B. (2003). Why do only some

events cause learning during human tutoring? *Cognition and Instruction*, 21(3), 209-249.

Vygotsky, L.S. 1978. *Mind in society*. Cambridge, MA: Harvard University Press.

Woolf, B.P. (2009). *Building intelligent tutoring systems*. Burlington, MA: Morgan Kaufman.

Zimmerman, B. (2001). Theories of self-regulated learning and academic achievement: An

overview and analysis. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and

academic achievement: Theoretical perspectives* (pp. 1-37). Mahwah, NJ: Erlbaum.

Author Notes