Running head: TUTORING

Meta-knowledge in Tutoring

Arthur C. Graesser[1], Sidney D'Mello[1], and Natalie Person[2]

[1] University of Memphis

[2] Rhodes College

Send correspondence to:

Art Graesser
Department of Psychology & Institute for Intelligent Systems
202 Psychology Building
University of Memphis
Memphis, TN 38152-3230
901-678-2146
901-678-2579 (fax)
a-graesser@memphis.edu

Meta-knowledge in Tutoring

It is important to have a chapter in this volume on one-to-one tutoring because it is one of the most effective methods of helping students learn. Tutoring provides the opportunity for an intense "meeting of the minds" between a student and a person with subject matter expertise. Tutoring is the standard panacea when students are failing to meet expected grades and anxiety reverberates among the students, parents, teachers, principles, and school systems. Anxiety turns to panic when a school is not meeting the standards of a high stakes test, when a prize athlete may be cut from a team, and when a student runs the risk of losing a scholarship. A wealthy family might end up paying $200 per hour for an accomplished tutor to save a son or daughter.

It is very easy to justify the use of tutors from the standpoint of learning gains. Meta-analyses show learning gains from typical, non-expert human tutors of approximately 0.4 sigma (effect size in standard deviation units) compared to classroom controls and other suitable controls (Cohen, Kulik & Kulik, 1982). Non-expert tutors are paraprofessionals, cross-aged tutors (i.e., students who are older than the tutee), or same-age peers who have had little or no tutor training and have modest subject-matter expertise. Collaborative peer tutoring shows an effect size advantage of 0.2 to 0.9 sigma (Johnson & Johnson, 1992; Mathes & Fuchs, 1994; Slavin, 1990; Toppin, 1996), and appears to be slightly lower than older non-expert human tutors (Pilkington & Parker-Jones, 1996). 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), presumably because of the added study, effort, and initiative in taking on the role as tutor. Peer tutoring is a low-cost

effective solution because expert tutors are expensive and hard to find. In contrast, 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. However, available studies show effect sizes of

0.8 to 2.0 (Bloom, 1984; Chi, Roy, & Hausmann, in press; Person, Lehman, & Ozbun, 2007;

VanLehn et al., 2007), which is presumably higher than other forms of tutoring.

The obvious question to ask is why is tutoring effective in promoting learning? This

question has been investigated by researchers in education, cognitive science, and discourse

processing in recent years. One approach to answering this question is to conduct a program of

research or to report 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 (1) learning gains tend to be higher for well-structured, precise domains

(mathematics, physics) than for ill-structured domains, (2) that learning gains from tutors are

more pronounced for tutees who start out with comparatively lower amounts of knowledge

and skill, (3) that the quality of tutor training is much more important than the quantity of

training, and that (4) that a tutoring session shows more benefits when there are particular

pedagogical activities (Cohen et al., 1982; Fuchs et al., 1994, in press; 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 et al., in press; Chi, Siler, Jeong,

Yamauchi, & Hausmann, 2001; Fuchs et al., in press; 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; 2003; Person, Kreuz, Zwaan, & Graesser, 1995; Shah, Evens, Michael, & Rovick, 2003; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003). This chapter discusses some of these analyses, particularly from the standpoint of meta-cognition, meta-communication, and meta-affect. 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, in press; Graesser, Lu et al., 2004; 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.

Given that the focus of this Handbook is on metacognition, it is not our goal to provide a comprehensive literature review of the anatomy of tutoring and to identify the factors that best explain the impact of tutoring on learning. There are already several papers, meta-analyses and reviews that address that goal, as we have already cited. Instead, this chapter attempts to reconstruct what meta-cognitive, meta-communicative, and meta-affective knowledge exists in the minds of tutors and tutees. What do tutors know, for example, about cognition, emotions, pedagogy, discourse, and the nature of tutoring? What do the tutees (students) know? Does such "meta" knowledge vary with the expertise of the tutor or tutee?

Most of what we know about the "meta" knowledge of tutoring is inferred from the tutoring process and available outcome measures. Traditional studies of metacognition outside of the tutoring realm have collected ratings or judgments from participants on their "meta" knowledge, such as feeling of knowing (FOK), judgments of learning (JOL),

comprehension calibration, or inventories on the use of strategies (Azevedo & Cromley, 2004; Dunlosky, Rawson, & Middleton, 2005; Hacker, 1998). Such explicit judgments are not routinely collected in tutoring studies. For example, researchers do not collect data at different points in the tutoring session on whether the tutor knows or believes a fact F, the student knows F, the tutor knows the student knows F, or the student knows that the tutor knows that the student knows F. Researchers do not collect data, at different points in the tutoring session, on whether pedagogical strategy S should be used, or not used, and why. Researchers have not collected data on the general expertise that particular tutors have about good tutoring practices, the nature of cognitive processes, and the conditions in which a particular strategy would help learning. Such inventories would be useful to collect, but we were not able to identify a study that reported such data. The data we have available about tutoring is restricted to outcome measures of learning and/or the measures collected on the tutoring activities by trained researchers who analyze sessions recorded on videotape, audiotape, or computer (e.g., transcripts of computer-mediated interaction).

Nevertheless, much can be learned from tutoring process data and outcome measures. Consider a couple of examples that illustrate what we can infer. One of the hallmarks of tutoring is that the tutor should help the tutee correct deficits in his or her knowledge and skills. This is a large part of what makes tutoring adaptive. If a tutor knows this, the tutor should initiate some activity at the beginning of the session that diagnoses what the tutee does or does not know, or what challenges the tutee is struggling with. We can infer that the tutor has this principle of adaptation if the tutor performs any of a number of actions: (a) inspects previous test materials and scores of the tutee, (b) selects problems in the tutoring session that are

associated with the tutee's deficits, or (c) asks the tutee what they are having problems with. We can infer that the tutor lacks the principle of adaptation if the tutor immediately launches into problems from a curriculum script, in the same way for all tutees. This example illustrates that we can infer quite a bit by systematically analyzing the actions that occur during tutoring and comparing these actions to theoretically expected actions.

Consider another example that addresses meta-cognitive knowledge of the tutee. Tutors frequently ask students comprehension-gauging questions, such as *Do you understand?*, *Are you following?*, and *Does that make sense?* If the tutee's comprehension calibration skills are accurate, then the tutee should answer YES when the tutee appears to understand and NO when there is little or no understanding. One provocative finding in the tutoring literature is that there sometimes is a negative correlation between a student's knowledge of the material (based on pre-test scores or post-test scores) and their likelihood of saying YES rather than NO to the tutors' comprehension-gauging questions (Chi et al., 1989; Graesser et al., 1995). So it is the knowledgeable tutees 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 subpar 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). In this example, we can infer something about the tutee's metacognition by comparing a discourse pattern between tutor-tutee and a test score (pretest or posttest).

The remainder of this chapter has two sections. We first discuss what tutees know about meta-cognition, meta-communication, and meta-affect on the basis of activities manifested in tutoring. As a note, we have a broad umbrella of what is considered "meta;" it is defined as their knowledge of pedagogy, cognition, communication, and emotion.   After we cover what is known about the tutee's we will turn to the meta-knowledge of tutors. The tutors will include untrained human tutors, expert tutors, and computer tutors.

<div align="center">Meta Knowledge of Tutees</div>

Consider a sketch of what we would want from ideal learners. They would self regulate their learning by identifying their own knowledge deficits, detecting contradictions, asking good questions that rectify these anomalies, searching knowledge sources for answers, making inferences when answers are not directly available, and actively building knowledge at deep levels of mastery. It is safe to say that the vast majority of tutees do not fit this profile (Chi et al., 2004; Graesser & Person, 1994; Graesser, McNamara, & VanLehn, 2005). It is the tutor who directs the tutoring session, not the tutee, and tutees are rarely inquisitive question askers (Chi et al., 2001; Graesser et al., 1995; Otero, this volume).

The more accurate sketch of the tutee's conception of a tutoring session is that of a passive learner, not too different from a student in a classroom listening to a lecture. They expect the tutor to lecture, explain the content, model procedures in solving problems, and essentially run the session. The tutee also expects to be tested with questions, just as in the classroom. As in a classroom, they expect to be grilled with simple questions that invite short answers and clear-cut feedback on the quality of the answers. 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

(Mehan, 1979; Sinclair & Coulthart, 1975). The analogue in the tutoring session would be the

exchange below on the subject matter of Newtonian physics.

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

(2) STUDENT: acceleration

(3) TUTOR: Right, mass times acceleration.

   Or

STUDENT: velocity

TUTOR: Wrong, it's not velocity, it is acceleration.

It is not surprising that the tutees conception of a tutoring session is not much different from a

classroom because the dynamics of the tutorial session closely mirror the classroom that they

experience every day. Classrooms typically consist of the teacher presenting didactic lessons

aligned with a curriculum, of presenting problems and worked out solutions, and of frequently

grilling the students with IRE sequences. As in the classroom, the tutee expects to be

evaluated and graded on their responses. Obviously, there are more innovative classroom

environments that deviate from this simple sketch, but this does depict most classrooms.

Shortly after the tutee arrives at a tutoring session the tutee learns that the discourse and

pedagogical structure of a tutoring session is somewhat different from the typical classroom.

Although there is a tendency for poor tutors to simply lecture like a teacher, most tutors spend

considerable time presenting problems or asking difficult questions that are answered

collaboratively by the tutor and tutee (Chi et al., 2001; Fuchs et al., in press; Graesser et al.,

1995; Person & Graesser, 1999). According to Graesser and Person (1994), tutors frequently

implement the following 5-step tutoring frame:

    (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.

It is quite apparent that this 5-step tutoring frame involves collaborative discussion, joint

action, and pressure for the tutee to construct knowledge rather than merely receiving

knowledge. The role of the tutor shifts from being a knowledge-teller to a guide for

collaborative knowledge construction. The relevant metaphor shifts from a classroom to

collaborative work.

        Tutors exhibit considerable variation on how they run their tutoring sessions. Later in

this chapter we will discuss what tutors do and don't do, as well as their meta-knowledge.

From the standpoint of the tutee, the focus of the present section, the meta-knowledge appears

to be quite limited when we examine available evidence. The remainder of this section

discusses what we can reconstruct about the tutee's meta-knowledge of cognition,

communication, and emotions.

Metacognition

        Other chapters in this volume discuss at length the fact that children, college students,

and adults are quite limited in their metacognitive knowledge and skills, as well as their ability

to self-regulate their learning. The process of self-regulation of learning theoretically involves

the learners' constructing a plan, monitoring metacognitive activities, implementing learning strategies, and reflecting on their progress and achievements (Azevedo, 2005, this volume; Azevedo & Cromley, 2004; Pintrich, 2000; Winne, 2001; Zimmerman, 2001). Each of these phases can be decomposed further. For example, metacognitive monitoring can be decomposed into judgments of learning, feeling of knowing, content evaluation, monitoring the adequacy of a strategy, and monitoring progress towards goals (Hacker, Dunlosky, & Graesser, 1998). Examples of learning strategies include searching for relevant information in a goal-directed fashion, taking notes, drawing tables or diagrams, re-reading, elaborating the material, making inferences, coordinating information sources such as text and diagrams, and summarizing content. Tutees are limited, just as most learners are limited, in knowing, mastering, and executing most of these components.

Tutors are often stuck by the lack of curiosity and questions of tutees. The lack of student questions can partially be attributed to the fact that the tutor governs the tutoring agenda. However, another reason is that there is a low rate of student questions in classrooms and most other learning environments (Dillon, 1988; Graesser, McNamara, & VanLehn, 2005). According to Graesser and Person's (1994) analysis of student questions, a typical student in a classroom asks approximately 1 question every 6-7 hours in a classroom environment and 1 question every 2 minutes in a tutoring environment. Most of the tutee questions are attempts to confirm or verify their knowledge (e.g., Isn't this Newton's second law?), to request definitions of terms (What is a Newton?), or to ask how to perform a procedure (How do you get this square root?), as opposed to questions that attempt to fill major gaps or contradictions in conceptual knowledge. Students with more domain knowledge ask deeper questions

(Graesser & Olde, 2003; Graesser & Person, 1994; Miyake & Norman, 1979; Otero, this volume; Van der Meij, 1990). As Miyake and Norman (1979) pointed out decades ago in their analyses of tutoring, it takes a considerable amount of knowledge for a student to know what he or she does not know. According to one model of student question asking (Graesser & Olde, 2003; Otero & Graesser, 2001), students need to be put in cognitive disequilibrium before they ask sincere information seeking questions. Cognitive disequilibrium occurs when students are confronted with contradictions, anomalies, obstacles to goals, and difficult decisions among equally attractive options. These conditions rarely occur when students are in a passive learning environment, they have minimal prior knowledge, and no vested interest in the subject matter. From this perspective, it is no surprise that good, deep, sincere information-seeking questions are rarely asked by the tutee.

As will be discussed later, tutors who adopt a student-centered perspective (Chi et al., 2001, in press) attempt to get the tutees to do the talking and doing. The tutors generate pumps, prompts, hints, and questions to stimulate verbal explanations and actions of the tutee. Other tutors adopt a tutor-centered approach that is comparatively insensitive to the student's knowledge and the importance of the student constructing the knowledge; these students have underdeveloped meta-pedagogical knowledge. It is remarkable how few students, tutors, and members of the community fail to appreciate the value of active construction of knowledge. Some students may even feel cheated when they are expected to do the work rather than having the tutor do the knowledge-telling or model the actions. And the same holds for the parents. Parents occasionally criticize same-age peer tutoring because they strongly believe that teachers and tutors should do the teaching rather than having their children wasting their

time teaching other students. These parents do not understand the virtues of active learning

and learning by teaching (Fuchs et al., in press). At this point in tutoring research, it is an open

question precisely what meta-knowledge students (and parents) have about the cognitive and

pedagogical mechanisms of tutoring.

Metacommunication

Metacommunication is knowledge about the process of communication. The focus of

this chapter is on tutoring and the focus of this section is on the tutee, so the relevant question

addresses what knowledge the students have about the communication process during tutoring.

Once again, available relevant data relies on an analysis of discourse patterns in conjunction

with measures of student ability or achievement.

Communication is most successful when there is an alignment at multiple levels of

language and discourse between participants (Pickering & Garrod, 2004). This occurs when

there is high common ground (shared knowledge) and the pragmatic ground-rules are well

understood by both parties (Clark, 1996; Schober & Clark, 1989). Alignment violations

sometimes occur during tutoring, however, because there is a large gap between tutor and

tutee in subject matter knowledge and because the tutee sometimes misunderstands the tutor's

tutoring goals (Person et al., 1995). This is most apparent when we examine the grounding of

referents, feedback, and hints. A good tutor presumably attempts to clarify the conversational

ground rules and to repair misalignments, but this too often is not achieved in normal tutoring

sessions.

Grounding referents. A referent is grounded when both the tutor and the student know

the meaning of the referent and this meaning is shared in the discourse space (i.e., the history

of the multi-turn dialogue and the learning materials that are co-present). Sometimes the meaning consists of the definition of a concept, but more often it is an entity in the tutoring environment being referred to or pointed to (e.g., a sentence in a book, a symbol in a math problem).   Unfortunately, many students do not appreciate the importance of grounding so they let too much go by with a minimal understanding. When the tutor asks whether the tutee understands via a comprehension-gauging question ("Do you understand?"), they usually nod or answer YES, even though many of the referents are ungrounded and their understanding is minimal. As mentioned earlier, it is the more knowledgeable tutee who tends to answer NO (Chi et al., 1989; Graesser et al., 1995). Many tutors believe the student's answer and assume that conversational grounding has been achieved, so there is a serious misalignment in grounding. A good tutor might periodically show scepticism and make the student demonstrate adequate grounding by requesting the student to define a term, to point to an external referent, or to perform a procedure; that would be quite diagnostic of breakdowns in grounding. On the flip side, students often assume that the tutor understands whatever the student expresses. In essence, the student assumes "If I say it, the tutor will understand it." The truth is, however, that tutors often misunderstand the tutee because much of what the student says is vague, underspecified, and error-ridden (Graesser et al., 1995). The tutor's and tutee's *illusion of grounding* is pervasive in tutoring sessions and accounts for many of the misunderstandings and misalignments of communication.

Feedback. Tutors give short feedback after most student turns as an indication of the quality of the student's contribution. The feedback is either positive (*yes, very good*, head nod, smile), negative (*no, not quite*, head shake, pregnant pause, frown), or neutral (*uh huh, I see*).

After the short feedback, the tutor goes on to advance the conversation with questions, hints, elaborations, and other types of dialogue moves. The tutors' short feedback is hopefully accurate because many researchers believe that feedback is an important part of learning, although evidence for the impact of feedback on learning is mixed (Azevedo & Bernard, 1995; Chi et al., in press; Kluger & DiNisi, 1998; Kulhavy & Stock, 1989; Shute, 2007). On the other hand, many tutors tend to be polite conversation partners and resist giving negative feedback to the tutee when the student expresses vague or error-ridden contributions (Person et al., 1995). Graesser and Person (1994) reported that the immediate feedback of unskilled tutors tends have a higher likelihood of being positive than negative after incorrect or vague student contributions (Graesser et al., 1995). One could imagine that the students might clam up when they are barraged with a hefty dose of negative feedback, so there indeed is some virtue in tutors following the politeness principle. If tutor feedback does engender a tradeoff between learning and motivation, then tutors should be particularly sensitive to the use of feedback. At the very least, they should be vigilant of the *illusion of feedback accuracy*.

What do students believe and expect regarding feedback? There is some anecdotal evidence suggesting that students do believe the feedback that they receive from tutors and also that they desire accurate feedback. For example, a computer tutor called AutoTutor helps students learn by holding a conversation with the student in natural language (Graesser, Lu et al., 2004); Graesser, Person, Harter, & TRG, 2001). AutoTutor sometimes does not accurately interpret the tutee's contribution, which leads to incorrect short feedback. Students who detect the incorrect feedback get annoyed, irritated, or angry, sometimes to the point of dismissing the utility of AutoTutor altogether (D'Mello, Craig, Witherspoon, McDaniel, &

Graesser, 2008). One explanation of their negative emotions is that the students expect accurate feedback. Another observation from these sessions with AutoTutor is that students want decisive feedback (*yes* versus *no*) rather than evasive or indecisive feedback (*possibly, uh huh, okay*). A polite or wishy-washy computer tutor does not seem to be as desirable as a decisive one. It appears that the politeness principle may apply to humans but not computers in the case of feedback. These findings suggest that the students' assumptions about pragmatic ground-rules and communication may be very different for human tutors versus computer tutors.

Hints. Good tutors are known to give hints that invite students to do the talking and doing rather than just giving lectures and telling the student the correct information (Chi et al., 2004, in press; DiPaolo, Graesser, White, & Hacker, 2004; Hume et al., 1993). The hints vary from being generic statements or questions (*What about X?, Why?*) to speech acts that more directly lead the student to a particular answer. Hints serve as memory cues, problem solving clues, and directives to shift the burden of conversation from the tutor to the tutee. Hints may be the ideal scaffolding move to promote active student learning, but at the same time directing the student to focus on important relevant material.

Two limitations of hints potentially arise from the standpoint of student perceptions. Both of these limitations can be explained by appealing to an *illusion of discourse alignment*, the unwarranted assumption that the listener does or is expect to understand the discourse function, intention, and meaning of the speaker's dialogue contributions. The first limitation is that the students may not understand the discourse function of hints, particularly if they believe the ground-rules dictate that the tutor should do the telling and showing. Students may

have a tutor-centered epistemology of the tutoring register (i.e., genre) rather and a student-centered or interaction-centered epistemology (Chi et al., 2004, in press). As a consequence, the student may not know how to respond to hints because they regard them as confusing or pragmatically infelicitous. A second limitation occurs when (a) the tutors have an answer in mind when they give a hint, (b) the students respond with an answer they think is good, (c) the students' responses do not match the tutors' expected answers, (d) the tutors give negative feedback, and (e) the students get frustrated or confused from being misunderstood and not adequately credited. D'Mello et al. (2008) have indeed reported that student confusion tends to occur after tutor hints and after negative feedback to students' responses to hints. It is interesting to speculate how tutorial dialogues would evolve if students assumed that many contributions in the conversation have discourse misalignments and that perfectly meshed conversations are very rare.

Meta-affect

Connections between complex learning and emotions have received increasing attention in the fields of psychology (Carver, 2004; Deci & Ryan, 2002; Dweck, 2002) and education (Gee, 2003; Lepper & Henderlong, 2000; Linnenbrink & Pintrich, 2002; Meyer & Turner, 2002, 2006). It is important to understand affect-learning connections in order to adequately train tutors and also to design engaging educational artifacts that range from responsive intelligent tutoring systems (DeVicente & Pain, 2002; Graesser, Person, Lu, Jeon, & McDaniel, 2005; Guhe, Gray, Schoelles, & Ji, 2004; Litman & Silliman, 2004) to entertaining media and games (Conati, 2002; Gee, 2003; Vorderer, 2003). Researchers in many different fields are familiar with Ekman's work on the six universal emotions

(sadness, happiness, anger, fear, disgust, surprise) that are manifested through facial

expressions (Ekman & Friesen, 1978) and paralinguistic features of speech (Scherer, 1986).

However, these are not the emotions that occur during most tutoring sessions. The

pervasive affective states during the tutoring of technical material are confusion, frustration,

boredom, anxiety, and flow/engagement, with delight and surprise occurring less frequently

(Craig, Graesser, Sullins, & Gholson, 2004; D'Mello,Graesser et al., 2006; 2007; 2008;

Lehman, Matthews, D'Mello, & Person, in press).

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 (Clifford, 1991).

According to flow theory, the learner is in a state of flow (Csikszentmihaly, 1990) 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; Murphy et al., in

press; Stein & Hernandez, 2008).   Outcomes that achieve challenging goals result in

positive emotions whereas outcomes that jeopardize goal accomplishment result in negative

emotions.

Obstacles to goals are particularly diagnostic of both learning and emotions. The affective state of confusion correlates with learning gains perhaps because it is a direct reflection of deep thinking (Craig et al., 2004; Graesser et al., 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.

There have been very few studies on meta-affect in tutees. This leaves the door wide open for some informed speculation and some encouragement to researchers to conduct more research in this area. We do know that tutee's identify the following affective states when they are asked to emote-aloud (i.e., articulate their emotions) during learning or when they view videotapes of their tutoring sessions and judge their emotions at different points in time (D'Mello et al., 2006; Graesser et al., 2006): confusion, frustration, boredom, flow/engagement, delight and surprise. However, we do not know how reliably different classes of learners can identify these emotions. We suspect from 150 years of psychological research on emotions that some learners lack sensitivity to their own emotions, that other learners are hypersensitive, and that there is a large continuum of possibilities in between (see Lewis, Haviland-Jones, & Barrett, 2008).

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 tutee is frustrated from being stuck, the student might attribute the frustration to either to themselves ("I'm not at all good at physics"), the tutor ("My tutor doesn't understand this either"), or the materials ("There must be a lousy textbook"). Solutions to handle the frustration would presumably depend on these attributions of cause (Heider, 1958; Weiner, 1995). 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 a student is bored, they are likely to blame the tutor or material rather than themselves. Tutoring environments of the future need to manage the tutorial interaction in a fashion that is sensitive to the tutees' emotions in addition to their cognitive states.

Meta Knowledge of Tutors

What might we expect of the tutors' knowledge about pedagogy, metacognition, meta-communication, and meta-affect? In an ideal world, the tutor would be able to reliably identify the cognitive and emotional states of the tutee. The tutor would be able to respond to these states in a pedagogically justifiable matter. That would include a large repertoire of excellent tutoring and teaching strategies that are tuned to different classes of students. The tutor would understand communication processes and how to formulate discourse moves that

establish common ground and that advance the conversation in a manner that optimizes

pedagogical goals.

The typical tutors in the school system and life long learning are far from ideal, as we

expressed earlier in the chapter. It is rare to have an experienced tutor with adequate training

in the subject matter knowledge, pedagogy, and discourse processes.   Person et al.'s (2007)

review of studies with accomplished tutors makes the points that the sample sizes of expert

tutors are extremely small (N <3) in empirical investigations of expert tutoring, that the same

expert tutors are used in different research studies, and the tutors are frequently co-authors on

publications. Therefore, claims about expert tutoring are frequently biased by the

idiosyncratic characteristics of the small sample and the tutors' authorship role. Person et al.

therefore 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, but not linked to outcome

scores on learning. The discourse patterns of 13 normal unskilled tutors, i.e., cross aged tutors

and paraprofessional, have been analyzed by Graesser and Person (Graesser, & Person, 1994;

Graesser et al., 1995; Person & Graesser, 1999, 2003). This Graesser-Person *Unskilled Tutor*

corpus also was analyzed in great detail at the level of discourse, but not outcome scores on

learning. Future research will need to systematically relate the discourse characteristics of

tutors of varying ability to different outcome measures.

This section will sometimes consider *Computer Tutors* in addition to the Expert and

Unskilled Tutors. The advantage of computer tutors is that they systematically analyze the

students' knowledge and actions and they implement strategies to adaptively respond to the

student. Unlike human tutors, the tutoring strategies are implemented systematically and

follow complex algorithms.

Pedagogy

Unskilled human tutors are not particularly sophisticated from the standpoint of

ideal tutoring strategies that have been proposed in the fields of education and artificial

intelligence (Graesser et al., 1995). Graesser and colleagues videotaped over 100 hours of

naturalistic tutoring in the Unskilled Tutoring corpus, transcribed the data, classified the

speech act utterances into discourse categories, and analyzed the rate of particular discourse

patterns. These analyses revealed that the tutors rarely implement intelligent 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 (Collins, Brown, & Newman, 1989; Palincsar & Brown, 1984;

Sleeman & Brown, 1982). 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). 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). In reciprocal teaching, the

tutor and learner take turns working on problems or performing a skill, as well as giving

feedback to each other along the way (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. 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 (Gagne, 1985).

It is perhaps unsurprising that unskilled human tutors rarely implement sophisticated tutoring strategies such as the ones described above. They were never trained to implement such strategies and it is unlikely these sophisticated strategies with be spontaneously discovered by the tutor. Moreover, there is a large computational overhead to implementing these strategies, far beyond most of the computer technologies of today. Researchers in computational linguistics and artificial intelligence have designed computer tutors to simulate tutors who interact with the students in natural language. These include AutoTutor (Graesser et al., 1999, 2001; 2004, 2006; VanLehn et al., 2007), why-Atlas (Graesser, VanLehn, Rose, Jordon, & Harter, 2001; Van-Lehn et al., 2003), CIRCSIM-Tutor (Shah, Evens, Michael, & Rovick, 2002), DC-Trains (Peters, Bratt, Clark, Pon-Barry, Schultz, 2004), and Mission Rehearsal (Gratch et al., 2002). These different computer tutors vary in the extent to which they simulate human dialogue mechanisms, but all of them attempt to comprehend natural language and to formulate adaptive responses and strategies to help students learn.

The structure of the dialogue in both AutoTutor and human tutoring (Chi et al., 2001, 2004, in press; Graesser, & Hu, & McNamara, 2005; Shah et al., 2002) follows an *Expectation and Misconception Tailored* (EMT) dialogue. EMT dialogue is the primary pedagogical method of scaffolding good student answers. Both AutoTutor and human tutors typically have a list of expectations (anticipated good answers) 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

AutoTutor guides the student in articulating the expectations through a number of dialogue moves: *pumps* (what else?), *hints*, and *prompts* for the student to fill in missing words. Hints and prompts are carefully selected by AutoTutor to produce content in the answers that fill in missing content words, phrases, and propositions. 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. Complete coverage of the answer requires AutoTutor to have a pool of hints and prompts available to extract all of the content words, phrases, and propositions in each expectation. AutoTutor adaptively selects those hints and prompts that fill missing constituents and thereby achieves pattern completion.

Human tutors and AutoTutor are dynamically adaptive to the learner in ways other than coaching them to articulate expectations. There is the conversational 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, as discussed earlier. The tutor accommodates a mixed-initiative dialogue by attempting to answer the student's questions when the student is sufficiently inquisitive to ask questions. The tutor asks counter-clarification questions (e.g., I don't understand your question, so could you ask it in another way?) when the tutor does not understand the student's question. Tutors are considered more adaptive to the student to the extent that they correct student misconceptions, give correct feedback, answer student questions, and ask clarification questions to insure the grounding of content. AutoTutor and other dialogue-based intelligent tutoring systems implement these features of conversational responsiveness.

In addition to engaging in EMT dialogue, unskilled human tutors and AutoTutor have a case-based learning foundation. That is, they present challenging problems for the student to collaboratively reason with the tutor in route to an answer. The hope is that deep learning eventually emerges after the student and tutor collaboratively attempt to solve a large number of problems that vary in scope and sophistication. However, the cases are not necessarily anchored in real life situations (Bransford, Browing, and Cocking, 2000) because those problems require a lengthy process of science and engineering to design.

The general message to be conveyed here is that sophisticated pedagogical strategies are not generated by unskilled tutors, which 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. It is a question for future research to determine whether it is technically feasible for these strategies to be reliably implemented in computers and trained human tutors. It is also an open question whether the strategies significantly improve learning over and above the normal strategies of unskilled human tutors.

The pedagogical strategies of expert tutors are very similar to those of unskilled tutors and computer tutors in most ways. However, Person et al. (2007) identified the a few improvements in the 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 high. One pedagogical strategy evident in the Expert Tutor corpus was just-in-time direct instruction or mini-lectures when the tutee was struggling with a particular conceptualization. These content-sensitive lectures were sensitive to what the tutee was having trouble with rather than being routinely delivered to all tutees. The expert tutors also appeared to differ from unskilled tutors on meta-cognitive and meta-communicative dimensions, to which we now turn.

Meta-cognition

The essential question is whether the tutor has knowledge of the cognitive states of the tutee. The tutor's knowledge of the tutor's own cognitive states is not directly relevant

to this section; presumably the meta-cognitive proficiencies of tutors are equal or occasionally higher than the tutees, particularly in the case of expert tutors. We would hope that the tutor builds an accurate and detailed model of the cognitive states of the tutee, or what is called the *student model* by researchers who develop intelligent tutoring systems. Thus, the central question is how accurate and detailed the student model is that gets constructed and used by the tutor.

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 tutees and that they formulate responses that do not require fine tuning of the student model (Chi et al., 2004; Graesser et al., 1995). What is the evidence for such claims? There are three sources of evidence. First, the short feedback to students on the quality of the students' contributions is often incorrect. For example, 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). 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). Computer tutors also have a similar problem of having difficulties detecting errors in verbal descriptions (Graesser, Hu, & McNamara, 2005). Third, 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., in press). You would expect the selection of problems to be tailored to the tutee's profile according to the Goldilock's principle or zone of proximal development, i.e., not too easy or not too hard, but just right. However, Chi et al. (in press) reported that there was no relation between problem selection

and tutee's profile. Data such as these lead one to conclude that tutors have a modest ability to conduct student modeling. Perhaps the only student parameter that the tutor can reliably handle is the verbosity of the tutee, namely the extent to which the express their ideas, as measured in number of words. However, the depth of the tutee's contributions is apparently difficult to calibrate and the specific knowledge deficits extremely difficult to identify. It may take intelligent tutoring systems to detect such subtleties and execute context-sensitive adaptive responses to promote learning.

Tutors often fall prey to the *illusion of student mastery*, namely the belief that the tutee has mastered much more than the tutee really has mastered. As discussed earlier in this chapter, tutors often believe the students when they ask the comprehension-gauging question "Do you understand?" and most students answer yes. The tutor often believes the tutee has covered an expectation, even when the students' contribution is vague or error ridden. The tutors tend to miss misconceptions expressed by students. Tutors frequently believe that the student has expressed a complete sentence-like expectations, even when the student only articulates a couple of content words in the expectation. All of these findings point in the direction that the tutor gives the student the benefit of the doubt with respect to student modeling. They assume that the tutee has mastered whatever is said or covered in the tutorial session, even though (a) much of what is covered in the dialogue is vague, incomplete, and error-ridden and (b) the tutee does not understand much of what the tutor expresses.

It is conceivable that expert tutors are more successful in detecting and tracking the knowledge of the students (Person et al., 2007). This is manifested in the question asking

analyses that Person et al. performed on the Expert Tutor corpus and compared to the previously published studies on the Unskilled Tutor corpus (Graesser & Person, 1994; Person & Graesser, 1999). Expert tutors and students ask proportionately more low specificity questions (e.g., So?) and more common ground questions (e.g., So, I use the Pythagorean theorem?) than tutors and students in non-expert sessions. We interpret these findings to mean that expert tutors are more attuned to the needs of their students and have established considerable common ground. If this were not the case, low specificity questions (e.g., So?) would result in conversation breakdowns. As another form of evidence, expert tutors are more dynamic in their instruction and do not rely on curriculum scripts. Experts typically begin the tutoring sessions by figuring out the topics/problems that students are having difficulty with and by asking questions about the students' performance on quizzes, homework, and exams. After this data collection phase, the tutor decides where to begin the session and what material will be covered. Expert tutors do not begin a session with any pre-planned teaching agenda, but rather base their instruction on students' particular needs at the time of the tutoring session.

One research question for the future addresses the fidelity of student modeling for unskilled, expert, and computer tutors. Which of these classes of tutors is most accurate in detecting the cognitive profile of the learners? How adaptive are the different tutors in producing discourse moves to facilitate learning in individual tutees? It is likely that the mental models of some students are so error-ridden and distorted that it would not be worth the time for the tutor to unravel the pathological conceptual spaghetti. When this occurs, the best the tutor can do is to model good reasoning and strategies. However, some students

may have mental models that approximate some semblance of accuracy. These are the students who might benefit from accurate student modeling and intelligent tutor responses. Answers to these questions await further research.

Meta-communication

Speakers typically assume that the meaning of their messages is correctly transferred to the minds of the addressees. In the tutoring context, knowledge transfer is successful when the tutee understands whatever the tutor expresses in the discourse space. However, there is abundant evidence, as we have discussed throughout this chapter, that tutees often do not understand the subject matter and the discourse contributions expressed by the tutor. The *illusion of knowledge transfer* occurs when the tutor assumes that the student understands whatever the tutor says. A good tutor is aware of this potential illusion. A good tutor is sufficiently skeptical of the tutee's level understanding that the tutor trouble-shoots potential communication breakdowns between the tutor and tutee. The tutor does this by not trusting the tutee's answer to comprehension gauging questions (Do you understand?) and instead asking follow-up questions that verify the student's understanding. This is illustrated in the simple 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.

Students frequently have low domain knowledge so they do not reliably attend to, understand, or remember what the tutor says. They have difficulty applying abstract principles to concrete applications. Therefore, a good tutor should not fall prey to the illusion of knowledge transfer. 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 tutee understands what the tutor expresses by asking follow up questions or giving follow-up trouble-shooting problems.

Knowledge transfer is facilitated when the tutor gives accurate feedback on the tutee's contributions. As discussed earlier, unskilled tutors frequently give incorrect feedback, as in the case of expressing more positive than negative short feedback after vague or incorrect student contributions (Graesser et al., 1995). In contrast, Person et al (2007) report that the expert tutors tend to give correct short feedback. Accountability of student understanding is most successful when the tutor questions have precise specific answers (as in mathematics) rather than a family of alternative imprecise answers (Person et al., 1995). Therefore, a good tutor should ask some discriminating questions with ideal answers that are sufficiently clear-cut that they can quickly diagnose incorrect student understanding and misconceptions. Without such precise accountability of student knowledge, it is unlikely that much communication is taking place in a tutoring session.

Meta-emotion

As discussed earlier, emotions play a critical role in the learning process. Therefore, it is presumably important for tutors to adopt pedagogical and motivational strategies that are effectively coordinated with the students' emotions (Issroff & del Soldato, 1996; Lepper & Chabay, 1988; Lepper & Woolverton, 2002). Lepper, Drake, and O'Donnell (1998) proposed an INSPIRE model to promote this integration. This model encourages the tutor to nurture the tutee by being empathetic and attentive the to the tutee'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 tutee 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 tutee, but to claim that the problem is difficult and to encourage the student to give it a try anyways. When the tutee readily solves the problem, the student builds self-confidence and self-efficacy in conquering difficult material (Winne, 2001; Zimmerman, 2001).

A tutor that is aware of the students' affective states would be expected to adaptively respond to their emotions during the course of enhancing learning. For example, if the tutee is frustrated, the tutor might give hints to advance the tutee's construction of knowledge (Graesser, Rus, D'Mello, & Jackson, 2008) or make supportive empathetic comments to enhance motivation (Burleson & Picard, 2004; Lepper et al., 1998; Lepper & Woolverton, 2002). If the tutee is bored, a good tutor would present more engaging or challenging problems for the tutee to work on. The tutor would probably want to lay low and stay out the student's way when the tutee is deeply engrossed in a state of flow

(Csíkszentmihályi, 1990). The flow experience is believed to occur when the learning rate is high and the student has achieved a high level of mastery at the *region of proximal learning* (Metcalfe & Kornell, 2005). When the tutee is confused, there is variety of productive paths for the tutor to pursue (D'Mello et al., 2007; Graesser et al., 2008). The tutor could allow the student to continue being confused during the cognitive disequilibrium; the student's self-regulated thoughts might hopefully restore equilibrium. Alternatively, after some period of time waiting for the student to progress, the tutor might give indirect hints to nudge the tutee into more productive trajectories of thought.

Goleman (1995) stated in his book, *Emotional Intelligence*, that expert teachers are able to recognize a student's emotional state and respond in an appropriate manner that has a positive impact on the learning process. Lepper and Woolverton (2002) have claimed that it takes expertise in tutoring before accurate detection of learner emotions can be achieved. This requirement of expertise is apparently quite important because, according to Lepper and Woolverton (2002), roughly half of expert tutors' interactions with the student are focused on affective elements.

These important claims would be seriously limited if tutors are unable to detect the affective states of the learner. Unfortunately, there is some evidence that humans untrained in detecting emotions are not particularly adept at detecting the emotions of a learner. For example, Graesser and colleagues conducted a study that assessed the reliability of emotion judgments of tutees, untrained peers, and accomplished teachers (D'Mello, Taylor, Davidson, & Graesser, 2008; Graesser et al., 2006). There were also two trained judges: research assistants who were trained extensively on tutorial dialogue characteristics and

how to detect facial action units according to Paul Ekman's Facial Action Coding System (Ekman & Friesen, 1978). The study involved 28 college students being tutored on computer literacy with AutoTutor. Judgments were made on the basis of videos of the participants face and screen that were recorded during the interaction.

The trained judges who are experienced in coding facial actions and tutorial dialogue provided affective judgments that were reliable and that matched the learner's self reports better than the affect judgments of untrained peers and accomplished teachers. Training on facial expressions apparently makes judges more mindful of relevant facial features and transient facial movements. However, in many situations the face is not the best indicator of the emotions of the learner. For example, while confusion and delight are accompanied by animated facial expressions, the face is virtually expressionless during expressions of boredom and flow (McDaniel et al., 2007; Craig, D'Mello, Witherspoon, & Graesser, in press). Therefore, experience with tutorial dialogue is also an important requirement for affect detection.

The finding that peers and teachers are not very accurate at classifying the emotions of the learner has considerable implications for peer tutoring and expert tutors. One potential advantage of peer tutoring is that there is no appreciable status difference between peers, compared to when a teacher tutors a student or when an older tutor helps a younger learner (Fuchs et al., 1994; Rogoff, 1990). There are several advantages of expert tutors who construct more accurate models of student knowledge, provide more evaluative and discriminating feedback, adapt their instructional strategies dynamically, and are more direct and task-oriented (Person et al., 2007). However, the results of the studies on

emotion tracking (D'Mello et al., 2008; Graesser et al., 2006) suggest that the inability to detect the affective states of a learner might be a drawback of peer and teacher tutoring. We are uncertain at this point on the importance of meta-affective abilities, status differences, experience with domain knowledge, and pedagogy with respect to learning gains and emotion tracking. Future research is needed to resolve this.

<div align="center">Summary</div>

This chapter has discussed the meta-cognitive, meta-communicative, and meta-affective mechanisms that occur in one-one-one human tutoring. We examined these capabilities from the perspective of both the tutor and the student. The typical unskilled tutors that exist in K-12 and college do not have a sophisticated repertoire of tutoring strategies that monitor the course of learning, such as building on pre-prerequisites, modeling-scaffolding-fading, Socratic tutoring, and misconception diagnosis and repair. These tutors are also limited in detecting the idiosyncratic cognitive and emotional profiles of individual students, generating dialogue moves that adapt to these profiles in a pedagogically sensitive manner, and incorporating most theoretical phases of meta-cognition and meta-affect.   Instead, there is a strong tendency to follow curriculum scripts and dialogue tactics that monitor expected correct information and anticipated errors. Accomplished and trained tutors do exhibit a few enhancements in meta-knowledge, but even these are limited. From the perspective of the student, there are well-documented limitations in their ability to monitor and strategically augment their own knowledge and learning, whereas a handful of studies have examined how well they detect and manage their emotions. Given the limitations in the meta-cognitive, meta-communicative, and

meta-affective capabilities in human tutoring, researchers are encouraged to increase their

efforts in building trainers, intelligent tutoring systems, and other learning environments

that attempt to improve these mechanisms.

References

Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist, 40*, 193-198.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effect of feedback in computer-based instruction. *Journal of Educational Computing Research, 13*, 109-125.

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.

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* (expanded ed.). Washington, D.C.: National Academy Press.

Burleson, W. & Picard, R.W. (2004). Affective Agents: Sustaining Motivation to Learn Through Failure and a State of Stuck. *Social and Emotional Intelligence in Learning Environments Workshop In Conjunction with the 7th International Conference on Intelligent Tutoring Systems*. August 31, 2004. Maceio - Alagoas, Brasil.

Carver, C. S. (2004). Negative affects deriving from the behavioral approach system. *Emotion, 4,* 3-22.

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. (in press). Learning from observing tutoring collaboratively: Insights about tutoring effectiveness from vicarious learning. *Cognitive Science*.

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., & Jeong, H. & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471- 534.

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

Clifford, M. M. (1991). Risk taking: Theoretical, empirical, and educational considerations. *Educational Psychologist, 26,* 263–298.

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., 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.

Conati C. (2002). Probabilistic Assessment Of User's Emotions In Educational Games. *Journal Of Applied Artificial Intelligence, 16, 555-575.

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.

Craig S., D'Mello S., Witherspoon A., & Graesser A (in press). Emote-Aloud During Learning With AutoTutor: Applying The Facial Action Coding System To Cognitive-Affective States During Learning. *Cognition And Emotion.*

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

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.

De Vicente, A. & Pain, H. (2002). Informing the detection of students' motivational state: An empirical study. *Intelligent Tutoring Systems 2002,* 933-943.

Dillon, J. T. (1988). The remedial status of student questioning, *Journal of Curriculum Studies*, 20 (3), 197-210.

DiPaolo, R.E., Graesser, A.C., White, H.A., & Hacker, D.J. (2004). Hints in human and computer tutoring. In M. Rabinowitz (Ed.), *The design of instruction and evaluation* (pp. 155-182). Mahwah, NJ: Erlbaum.

D'Mello, S. K., Craig, S. D., Sullins, J., & Graesser, A. C. (2006). Predicting Affective States Through An Emote-Aloud Procedure From Autotutor's Mixed-Initiative Dialogue. *International Journal Of Artificial Intelligence In Education,* 16, 3-28.

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.

D'Mello, S. K., Taylor, R. S., Davidson, K., & Graesser, A. C. (2008). Establishing Convergent Validity in Emotion Measurement: Contrasting Self versus Teacher Judgments of Learners Emotions.

Dunlosky, J., Rawson, K. A., & Middleton, E. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language, 52*, 551-565.

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.

Ekman, P, & Friesen, W. V. (1978). The facial action coding system: A technique for the measurement of facial movement. Palo Alto: Consulting Psychologists Press.

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.

Goleman, D. (1995). *Emotional intelligence* (New York, Bantam Books).

Graesser, A. C. Chipman, P., King, B., McDaniel, B., and D'Mello, S. (2007).Emotions and learning with AutoTutor. *13th International Conference on Artificial Intelligence in Education (AIED 2007)*. R. Luckin et al. (Eds), (pp 569-571). IOS Press.

Graesser, A.C., D'Mello, S.K., Craig, S.D., Witherspoon, A., Sullins, J., McDaniel, B., & Gholson, B. (2008). The relationship between affect states and dialogue patterns during interactions with AutoTutor. *Journal of Interactive Learning Research,* 19, 293-312.

Graesser, A.C., Hu, X., & McNamara,D.S. (2005). Computerized learning environments that incoporate research in discourse psychology, cognitive science, and computational linguistics. In A.F. Healy (Ed.), Experimental Cognitive Psychology and its Applications: Festschrift in Honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer. Washington, D.C.: American Psychological Association.

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., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. (2006). Detection Of Emotions During Learning With Autotutor. *Proceedings of the 28th Annual Conference of the Cognitive Science Society,* Mahwah, NJ: Erlbaum, 285-290.

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., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology,* 95, 524-536.

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

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.

Graesser, A. C., Person, N., Lu, Z., Jeon, M., & McDaniel, B. (2005). Learning While Holding A Conversation With A Computer. In L. PytlikZillig, M. Bodvarsson & R. Bruning (Eds.), *Technology-based education: Bringing researchers and practitioners together.* Greenwich, CT: Information Age Publishing.

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.

Graesser, A.C., Rus., V., D'Mello, S., & Jackson, G.T. (2008). AutoTutor: Learning through natural language dialogue that adapts to the cognitive and affective states of the learner. In D.H. Robinson and G. Schraw (Eds.), Current perspectives on cognition, learning, and instruction.

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.

Graesser, A. C., VanLehn,K., Rose, C., & Jordan, R., Harter, D. (1998). Intelligent Tutoring Systems with Conversational Dialogue. *AI Magazine*, Winter 2001, 39-51.

Gratch, J., Rickel, J., Andre, E., Cassell, J., Petajan, E., & Badler, N. (2002). Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems, 17*, 54-63.

Guhe, M., Gray, W. D., Schoelles, M. J., & Ji, Q. (2004, July). *Towards an affective cognitive architecture.* Poster session presented at the Cognitive Science Conference, Chicago, IL.

Hacker, DJ (1998). Definitions and empirical foundations. In DJ Hacker, J. Dunlosky, & AC Graesser (Eds.), *Metacognition in educational theory and practice (*pp. 1-24). Mahwah, NJ: Erlbaum.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (1998). *Metacognition in educational theory and practice.* Mahwah, NJ: Erlbaum.

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.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley & Sons.

Hume, G., Michael, J., Rovick, A., & Evens, M. (1993). The use of hints as a tutorial tactic. *Proceedings of the 15th Annual Conference of the Cognitive Science Society* (pp. 563-568). Boulder, CO.

Issroff, K. & del Soldato, T. (1996). Incorporating motivation into computer-supported collaborative learning. In Brna, P., et al. (Eds.), *Proceedings of the European Conference on Artificial Intelligence in Education.*

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.

Kluger, A. N., & DiNisi, A. (1996) Feedback Interventions: Toward the understanding of a two-edged sword. *Current Directions in Psychological Science*, 7(3), 67-72.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1, 279 - 308.

Lehman, B. A., Matthews, M., D'Mello, S. K., and Person, N. (in press). Understanding Students' Affective States During Learning. *Ninth International Conference on Intelligent Tutoring Systems (ITS'08).*

Lepper. M.R. and R.W. Chabay (1988). Socializing the intelligent tutor: Bringing empathy to computer tutors. In Heinz Mandl and Alan Lesgold (Eds), *Learning Issues for Intelligent Tutoring Systems* (pp. 242-257). Hillsdale, NJ: Erlbaum.

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.

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (Eds.). (2008). *The handbook of emotion, 3rd Edition*. New York: Guilford.

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., & Silliman, S. (2004). ITSPOKE: An intelligent tutoring spoken dialogue system. In Proceedings of the human language technology conference: 3rd meeting of the North American chapter of the association of computational linguistics (pp. 52-54). Edmonton, Canada: ACL.

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.

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

McDaniel, B. T., D'Mello, S. K., King, B. G., Chipman, P., Tapp, K., & Graesser, A. C. (2007). Facial Features for Affective State Detection in Learning Environments. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 467-472). Austin, TX: Cognitive Science Society.

Mehan, H. (1979). *Learning lessons: social organization in the classroom.* Cambridge, MA: Harvard University Press.

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, 52, 463-477.

Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided learning by doing. *Cognition and Instruction,* 13(3), 315-372.

Metcalfe, J., & Kornell, N. (2005). A region or proximal of learning model of study time allocation. *Journal of Memory and Language,* 52, 463-477.

Meyer, D. K., & Turner, J. C. (2002). Discovering emotion in classroom motivation research. *Educational Psychologist*, 37(2), 107-114.

Meyer, D. K., & Turner, J. C. (2006). Re-conceptualizing Emotion And Motivation To Learn In Classroom Contexts. *Educational Psychology Review*, 18 (4), 377-390.

Miyake, N., & Norman, D. A. (1979). To Ask a Question One Must Know What is Not Known. *Journal of verbal learning and verbal behavior, 18,* 357 - 364.

Murphy

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.

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., & Graesser, A. C. (2003). Pedagogical agents and tutors. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed. Vol. 2 586-589). New York: Macmillan.

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.

Peters, S., Bratt, E. O., Clark, B., Pon-Barry, H., Schultz, K. (2004). Intelligent Systems for Training Damage Control Assistants, in Proceedings of I/ITSEC 2004, Orlando FL, USA

Pickering, M.J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Brain and Behavioral Sciences, 27*, 169-190.

Pilkington, R. & Parker-Jones, C. (1996) Interacting with computer based simulation: the role of dialogue. *Computers and Education*, 27, 1, 1-14.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). New York: Academic Press.

Rogoff, B. (1990). *Apprenticeship in thinking.* New York: Oxford University Press.

Rogoff, B. and 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.

 Scherer, K. R. (1986). Vocal affect expression: A review and a model forn future research. *Psychological Bulletin,* 99(2), 143–165.

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

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.

Shute, V. J. (2007). *Focus on formative feedback* (ETS Research Rep. No. RR-07-11). Princeton,NJ: ETS.

Sinclair, J.M., & Coulthard, R.M. (1975). Towards an analysis of discourse: The English used by teachers and their pupils. London: Oxford University Press.

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

Sleeman D. and J. S. Brown. (1982). Introduction: Intelligent Tutoring Systems. In *Intelligent Tutoring Systems*. Orlando, Florida: Academic Press, Inc. 1-10.

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.

Van der Meij, H. (1990). Question Asking: To Know That You Do Not Know Is Not Enough. *Journal of Educational Psychology,* 82, 505-512.

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.

Vorderer, P. (2003). Entertainment theory. In B. Jennings, D. Roskos-Ewoldsen, & J. Cantor (Eds.), *Communication and emotion: Essays in honor of Dolf Zillmann* (pp. 131-153). Mahwah, NJ: Erlbaum.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct.* New York: Guilford.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 153-189). Mahwah, NJ: Erlbaum.

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