

Running head: EMOTIONS DURING LEARNING WITH AUTOTUTOR

Emotions during learning with AutoTutor

Sidney D'Mello and Art Graesser

Institute for Intelligent Systems, University of Memphis

### **Emotions during learning with AutoTutor**

Deep learning and problem solving are emotionally rich experiences. Students experience boredom when the material does not appeal to them, confusion when they have difficulty comprehending the material and are unsure about how to proceed, frustration when they make mistakes and get stuck, and perhaps even despair and anxiety when their efforts seem to be futile and the big exam is creeping around the corner. This negative picture of the emotional experiences that accompany learning has a complementary positive side. Students experience curiosity when they encounter topics that interest them, eureka moments when insights are unveiled and major discoveries made, delight when challenges are conquered, and perhaps even flow like states (Csikszentmihalyi, 1990) when they are so engaged in learning that time and fatigue disappear.

In general, emotion and cognition are complementary processes in learning environments that require students to generate inferences, demonstrate causal reasoning, diagnose and solve problems, make conceptual comparisons, produce coherent explanations, and show application and transfer of acquired knowledge. Contemporary theories of emotion and cognition assume that cognitive processes such as memory encoding and retrieval, causal reasoning, deliberation, goal appraisal, and planning operate continually throughout the experience of emotion (Barrett, Mesquita, Ochsner, & Gross, 2007; Bower, 1981; Mandler, 1984, 1999; Ortony, Clore, & Collins, 1988; Scherer, Schorr, & Johnstone, 2001; Stein, Hernandez, & Trabasso, 2008; Stein & Levine, 1991). The inextricable link between emotion and cognition is sufficiently compelling that some claim the scientific distinction between emotion and cognition to be artificial, arbitrary, and of limited value (Lazarus, 1991, 2000).

The last decade has witnessed a burst of research investigating the links between emotions and learning from the fields of psychology (Beilock & DeCaro, 2007; Deci & Ryan, 2002; Dweck, 2002; Stein et al., 2008), education (Lepper & Woolverton, 2002; Meyer & Turner, 2006; Pekrun, Elliot, & Maier, 2006), computer science (Arroyo et al., 2009; Conati & Maclaren, 2009), and neuroscience (Damasio, 2003; Immordino-Yang & Damasio, 2007). Some of the research has focused on student emotions in classrooms, where a broad array of affective responses are elicited in a number of contexts. These affect inducing contexts include classroom activities such as lectures and discussions that may evoke curiosity and boredom, examinations and quizzes that may induce anxiety and joy, and peer interactions where some of the social emotions such as pride and shame play a major role (Meyer & Turner, 2006; Schultz & Pekrun, 2007). Other research has focused on a more in-depth analysis of a smaller set of emotions that arise during deep learning in more restricted contexts and over shorter time spans (Baker, D'Mello, Rodrigo, & Graesser, in review; Conati & Maclaren, 2009). Our own research, which we describe in this chapter, aligns with the latter group, and has explored emotion-learning connections in the context of advanced learning environments such as AutoTutor.

AutoTutor is a validated intelligent tutoring system (ITS) that helps students learn topics in Newtonian physics, computer literacy, and critical thinking via a mixed-initiative conversational dialogue between the student and the tutor (Graesser, Chipman, Haynes, & Olney, 2005; Graesser et al., 2004; VanLehn et al., 2007). AutoTutor's dialogues are organized around difficult questions and problems (called main questions) that require reasoning and explanations in the answers. AutoTutor actively monitors learners' knowledge states and engages them in a turn-based dialogue as they attempt to answer these questions. It adaptively manages the tutorial dialogue by providing feedback (e.g. "good job", "not quite"), pumping the learner for more

information (e.g. “What else”), giving hints (e.g. “What about X”), prompts (e.g. “X is a type of what “), identifying and correcting misconceptions, answering questions, and summarizing answers. Using these strategies, AutoTutor adheres to constructivist theories of pedagogy (Biggs, 1995; Chi, Roy, & Hausmann, 2008; Jonassen, Peck, & Wilson, 1999; Moshman, 1982) by allowing students to chart their own course through the tutorial dialogue and to build their own answers to difficult questions. The AutoTutor interface along with a sample dialogue between a college student and AutoTutor is presented in Figure 1.

*Insert Figure 1 about here*

Intelligent tutoring systems (ITSs) such as AutoTutor (Graesser et al., 2004), Andes physics tutor (VanLehn et al., 2005), and Cognitive Tutor (Koedinger & Corbett, 2006) have come a long way towards modeling and responding to learners’ cognitive states. This allows ITSs to implement some of the ideal tutoring strategies such as error identification and correction, building on prerequisites, frontier learning (expanding on what the learner already knows), student modeling (inferring what the student knows and having that information guide tutoring), and building coherent explanations (Alevan & Koedinger, 2002; Anderson, Douglass, & Qin, 2005; Gertner & VanLehn, 2000; Koedinger, Anderson, Hadley, & Mark, 1997; Lesgold, Lajoie, Bunzo, & Eggan, 1992; Sleeman & Brown, 1982).

However, ITSs can be more than mere cognitive machines, and the link between emotions and learning suggests that they should be affective processors as well (Issroff & del Soldato, 1996; Picard, 1997). Affect-sensitivity is important for ITSs that aspire to model human tutors because it has been claimed 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 (Goleman, 1995; Lepper & Woolverton, 2002). An affect-sensitive ITS would

incorporate assessments of the students' cognitive and affective states into its pedagogical and motivational strategies in order to keep students engaged, boost self-confidence, heighten interest, and presumably maximize learning.

Therefore, in addition to investigating links between emotions and learning with AutoTutor, our research also focused on developing a version of AutoTutor that is dynamically responsive to learners' affective states in addition to their cognitive states. Our research program involving emotions and learning with AutoTutor has encompassed:

1. Identifying the emotions that occur during learning with AutoTutor and other learning environments (Baker et al., in review; Craig, Graesser, Sullins, & Gholson, 2004; D'Mello, Craig, Sullins, & Graesser, 2006; Graesser, Chipman, King, McDaniel, & D'Mello, 2007; Graesser et al., 2006; Lehman, D'Mello, & Person, 2008; Lehman, Matthews, D'Mello, & Person, 2008),
2. Investigating relationships between emotions and learning (Craig et al., 2004; D'Mello, Taylor, & Graesser, 2007; D'Mello & Graesser, in review; Graesser, Chipman et al., 2007),
3. Modeling the temporal dynamics of emotions (D'Mello, Taylor et al., 2007; D'Mello & Graesser, in review),
4. Assessing how reliably humans detect emotions (D'Mello, Taylor, Davidson, & Graesser, 2008; Graesser et al., 2006),
5. Identifying cognitive, bodily, and linguistic correlates of emotional expressions and developing systems to automatically detect emotions (D'Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008; D'Mello & Graesser, 2009; D'Mello, Dale, & Graesser, in review; D'Mello, Dowell, & Graesser, 2009; McDaniel et al., 2007), and

6. Developing computer systems that detect, respond to, and synthesize emotions (D'Mello et al., 2005; D'Mello, Jackson et al., 2008; D'Mello, Picard, & Graesser, 2007; D'Mello, Craig, Fike, & Graesser, 2009).

This chapter provides a synopsis of our research on emotions and learning by focusing on students' emotions during learning sessions with AutoTutor. We discuss: (a) theories on emotions and learning, (b) the emotions that learners experience during interactions with AutoTutor and correlational links between emotions and learning gains, (c) the temporal dynamics of the emotional states, (d) how contextual events influence learners' emotions, and (e) new versions of AutoTutor that detect and respond to learners' emotions.

### **Theories of Emotions and Learning**

The major theories of emotion and cognition have primarily focused on affective taxonomies (Izard & Ackerman, 2000; Ortony et al., 1988), valence-arousal frameworks (Barrett et al., 2007; Russell, 2003), cognitive-affective associative networks (Bower, 1981, 1992; Forgas, 1991), attributions (Gotlib & Abramson, 1999; Heider, 1958; Weiner, 1986), appraisals (Lazarus, 1991; Scherer et al., 2001; Smith & Ellsworth, 1985), and physiological and behavioral correlates of emotional experience (Ekman, 1984, 2003; Scherer, 2003). These theories convey general links between cognition and emotions, but they do not directly explain and predict the sort of emotions that occur during complex learning, such as attempts to master physics, biology, or critical thinking skills.

Fortunately, theoretical frameworks that predict systematic relationships between affective and cognitive processes during complex learning are beginning to emerge in fields of psychology (Barrett, 2006; Deci & Ryan, 2002; Dweck, 2002; Russell, 2003), education (Lepper

& Chabay, 1988; Lepper & Woolverton, 2002; Linnenbrink & Pintrich, 2002; Meyer & Turner, 2006; Stein et al., 2008), and even artificial intelligence (Conati, 2002; Dragon et al., 2008; Forbes-Riley, Rotaru, & Litman, 2008; Kort, Reilly, & Picard, 2001). The theories that have emerged highlight the contributions of academic risk taking, motivation, mood states, flow, goals, and cognitive disequilibrium, as we elaborate below.

The academic risk theory and intrinsic motivation literature address how individual differences in risk taking behavior and motivation influence learners' emotional states and behavior choices. The academic risk theory contrasts (a) adventuresome learners who want to be challenged with difficult tasks, take risks of failure, and manage negative emotions when they occur, with (b) cautious learners who tackle easier tasks, take fewer risks, and minimize failure and its resulting negative emotions (Clifford, 1988). Risk takers choose challenging tasks in order to maximize learning; they perceive failure and the resulting negative emotions as necessary steps towards content mastery (Meyer & Turner, 2006). In contrast, risk avoiders settle for emotional well-being at the expense of learning; they select tasks that are easier than their capabilities and that result in positive feedback on their performance (Boekaerts, 1993; Meyer & Turner, 2006).

The intrinsic motivation literature has identified affective states such as curiosity as indicators of motivation level and learning (Harter, 1992; Stipek, 1988). Intrinsically motivated learners derive pleasure from the task itself (e.g., enjoyment from problem solving), while learners with extrinsic motivation rely on external rewards (e.g., receiving a good grade). Learners with more intrinsic motivation display greater levels of pleasure, more active involvement in tasks (Harter, 1992; Tobias, 1994), more task persistence (Miserandino, 1996), lower levels of boredom (Miserandino, 1996), less anxiety, and less anger (Patrick, Skinner, &

Connell, 1993). Since a person's affective state is linked to their motivation level, intrinsically motivated learners who are affectively engaged should demonstrate more active involvement in tasks and greater task persistence. One consequence of this engaged persistence is a deeper understanding of the material (Jonassen et al., 1999).

Whereas theories of academic risk taking and intrinsic motivation address individual differences, mood theories and flow theory are concerned with how mood states impact emotions and performance. Mood theories highlight the important role of baseline mood states (positive, negative, or neutral) on learning, particularly for creative problem solving. In particular, flexibility, creative thinking, and efficient decision-making in problem solving have been linked to experiences of positive affect (Bless & Fielder, 1995; Fielder, 2001; Isen, 2001; Isen, Daubman, & Nowicki, 1987), while negative affect has been associated with a more methodical approach to assessing the problem and finding the solution (Hertel, Neuhof, Theuer, & Kerr, 2000; Schwarz, 2000; Schwarz & Skurnik, 2003). Mood states also influence emotional reactions by performing a threshold reduction function on emotional elicitation (Rosenberg, 1998). For example, repetitive failure is more likely to trigger frustration when the learner is in a negative versus positive mood.

According to flow theory, learners are in a state of flow (Csikszentmihalyi, 1990) when they are so deeply engaged in learning the material that time and fatigue disappear. The zone of flow occurs when the structure of the learning environment matches a learner's zone of proximal development (Brown, Ellery, & Campione, 1998; Vygotsky, 1978), so that the learner is presented with just the right sort of materials, challenges, and problems to the point of being totally absorbed. The state of flow is characterized by a focus on goals, unbridled attention, a virtual disappearance of time and fatigue, and a critical balance between skills and challenge

(Csikszentmihalyi, 1990; Meyer & Turner, 2006); all factors that make flow the optimal state for learning.

Goal theory and cognitive disequilibrium theory specify how particular events predict emotional reactions and are pitched at a finer temporal resolution than theories that highlight individual differences and mood states. Goal theory emphasizes the role of goals in predicting emotions. Consistent with contemporary appraisal theories, the arousal level (intense/weak) of an emotional episode is dependent upon how great the interruption is to the person's goal whereas the valence (positive/negative) depends on the person's evaluation of the interruption (Lazarus, 1991; Mandler, 1984). Hence, outcomes that achieve challenging goals result in positive emotions, whereas outcomes that jeopardize goal accomplishment result in negative emotions (Dweck, 2002; Stein et al., 2008; Stein & Levine, 1991). For example, getting stuck and not being able to move past an obstacle would be interpreted as intensely negative because goal attainment is obstructed (Dweck, 2002). Obstacles to goals are particularly diagnostic of both learning and emotions.

The cognitive disequilibrium theory postulates an important role for impasses (VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003) in comprehension and learning processes. Cognitive disequilibrium is a state that occurs when learners face obstacles to goals, contradictions, incongruities, anomalies, uncertainty, and salient contrasts (Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Graesser & Olde, 2003; Otero & Graesser, 2001; Piaget, 1952; Schwartz & Bransford, 1998). Cognitive equilibrium is restored after thought, reflection, problem solving, and other effortful deliberations. This theory states that the complex interplay between external events that trigger impasses, and the resultant cognitive disequilibrium, are the key to understanding the cognitive-affective processes that underlie deep learning. In particular, the

affective states of confusion and perhaps frustration are likely to occur during cognitive disequilibrium because confusion indicates an uncertainty about what to do next or how to act (Keltner & Shiota, 2003; Rozin & Cohen, 2003).

These theoretical perspectives make a number of predictions about the affective experiences during learning. We have tested some of these predictions in our analysis of emotion-learning connections in the context of advanced learning environments. We begin with a description of studies that attempted to identify the emotions that are prominent in learning sessions with intelligent tutoring systems, problem solving environments, and human tutors.

### **Identifying the Learning-Centered Emotions and Assessing their Relationship with Learning**

Researchers in different fields are familiar with Ekman's pioneering work on the detection of emotions from facial expressions (Ekman, 1984; Keltner & Ekman, 2000). However, the emotions that Ekman intensely investigated (e.g., sadness, happiness, anger, fear, disgust, surprise), though ubiquitous to everyday experience, are not expected to be relevant to learning sessions that span 30-minutes to 2 hours. The theoretical perspectives described above recommend a somewhat different set of emotions during learning. These include boredom, flow/engagement, confusion, frustration, delight and surprise. We refer to these states as *learning-centered emotions*. We conducted four studies in an attempt to identify the major emotions that accompany deep learning with AutoTutor, and presumably similar learning environments. These are briefly described below.

#### *Brief Description of Studies*

In Study 1 (Observational study), 34 participants' affective states were coded by observers every 5 minutes during interactions with AutoTutor (Craig et al., 2004). The affective

states were boredom, flow, confusion, frustration, eureka, and neutral. All coders were given a training session lasting at least 30 minutes to ensure they understood and were comfortable with coding the affective states of interest.

Study 2 (Emote-aloud study) adopted an emote-aloud procedure, a variant of the think-aloud procedure (Ericsson & Simon, 1993), as an online measure of the learners' emotions. Seven college students were asked to state the affective states they were feeling while working on a task, in this case being tutored in computer literacy with AutoTutor. This method allowed for on-line identification of emotions while working on a task with minimal task interference. The affective states in this study were anger, boredom, confusion, contempt, curiosity, disgust, eureka, and frustration. Flow was not included in this study because of the concern that asking people to report on their flow experiences would disrupt those experiences.

While the observational and emote-aloud studies used online methods to monitor learners' emotions, Studies 3 and 4 used an offline retrospective affect judgment protocol for emotion measurement. Study 3 (Multiple-judge study) had 28 college students who interacted with the AutoTutor system for 35 minutes (Graesser et al., 2006). Videos of the learners' faces, their computer screens, and posture patterns were recorded for offline analyses, as shown in Figure 2.

*Insert Figure 2 about here*

Participant's affective states (boredom, flow/engagement, confusion, frustration, delight, surprise, and neutral) were measured in a retrospective affect-judgment procedure that commenced after their AutoTutor session. The judging process was initiated by synchronizing the video streams from the screen and the face (center and right monitors in Figure 2), and displaying them to the judges. The screen capture included the tutor's synthesized speech,

printed text, students' responses, dialogue history, and images, thereby providing the context of the tutorial interaction.

Judges were instructed to make judgments on what affective states were present in 20-second intervals (*fixed* judgments), at which time the video automatically paused. They were also instructed to indicate any affective states that were present in between the 20-second stops (*spontaneous* judgments). Judgments were provided by the learners themselves (self reports), untrained peers, and two researchers (trained judges) with considerable experience interacting with AutoTutor and with the Facial Action Coding System (Ekman & Friesen, 1978). The judges were provided with a list of emotions with definitions. Boredom was defined as being weary or restless through lack of interest. Confusion was defined as a noticeable lack of understanding, whereas flow was a state of interest that results from involvement in an activity. Frustration was defined as dissatisfaction or annoyance. Delight was a high degree of satisfaction. Surprise was wonder or amazement, especially from the unexpected. Neutral was defined as no apparent emotion or feeling.

Study 4 (Speech recognition study) was a replication of the multiple-judge study with the exception that 30 learners spoke their responses to a new speech-enabled version of AutoTutor (D'Mello, King, Entezari, Chipman, & Graesser, 2008). This study also utilized a retrospective affect judgment procedure with judgments provided by the self and peers.

#### *The Emotions that Occur during Learning with AutoTutor*

Figure 3 shows descriptive statistics on the proportional occurrence of the various affective states in the four studies. There are important differences in the reliability by which different judges (self, peer, trained judges) can classify learners' emotions; these are discussed in previous publications (D'Mello, Taylor et al., 2008; Graesser et al., 2006). In Figure 3,

proportional scores for the multiple-judge study and the speech recognition study were computed by averaging across the different judges. The results indicate that boredom, flow/engagement, confusion, and frustration were the major affective states observed during learning sessions with AutoTutor.

*Insert Figure 3 about here*

These learner-centered emotions have also been found in learning environments other than AutoTutor and with populations other than college students. For example, the learning-centered emotions comprised 86% of the observations in a study where 36 adolescents from the Philippines solved logic problems with a simulation environment (Baker et al., in review). They comprised 86% of the observations in a study where 140 Philippine students were tutored in algebra with an intelligent tutoring system (Baker et al., in review). Confusion, boredom, and frustration were also the dominant states in a study where 41 aspiring law-school students solved difficult problems from the analytical reasoning section of the Law School Admissions Test (LSAT) (flow was not included in that study) (D'Mello, Lehman, & Person, in review; Lehman, D'Mello et al., 2008).

In contrast, eureka, curiosity, anger, contempt, disgust, delight, and surprise were comparatively infrequent. Eureka was well reported in the emote-aloud study but there was only one eureka experience identified in the 20 hours of tutoring in the observational study. Hence, we suspect that eureka responses in the emote-aloud study might functionally signify happiness or delight from giving a correct answer rather than a true eureka experience where there is a flash of deep insight. An examination of the videos captured during the tutorial sessions confirmed this suspicion.

Curiosity was also quite rare in the emote-aloud study, presumably because students had no choice of tutoring topics in our experimental environment. If participants had been given a choice of topics, they might have picked one more relevant to their interests and displayed more curiosity (Lepper & Woolverton, 2002). There is some evidence to support this assertion. In particular, curiosity was the dominant emotion when aspiring law-school students solved problems from the LSAT (D'Mello, Lehman, & Person, in review).

It is interesting to note that four of the six low frequency emotions were basic emotions, namely anger, disgust, surprise (Ekman, 1992; Izard, 1971), and contempt (Izard, 1971). Although the studies with AutoTutor did not incorporate the full set of basic emotions (i.e. happiness and sadness were excluded), two studies that compared the full set of basic emotions to the learning-centered emotions indicated that the basic emotions were infrequent in learning sessions (D'Mello, Lehman, & Person, in review; Lehman, Matthews et al., 2008). For example, 67% of students' emotions in 40 tutoring session with human tutors were the learning-centered emotions. The six basic emotions comprised 32% of the observations (Lehman, Matthews et al., 2008), with happiness accounting for 29% of the emotional expressions. Similarly, the basic and learning-centered emotions accounted for 26% and 74% of the observations, respectively, when students solved analytical reasoning problems for the LSAT (D'Mello, Lehman, & Person, in review; Lehman, D'Mello et al., 2008).

Taken together, the results substantiate the claim that the basic emotions, although ubiquitous in everyday experience, may not be particularly relevant to learning, at least for the short learning sessions of these studies. It is possible that they might be more relevant during learning in more extended time spans (such as completing a dissertation) or high stakes tests

(e.g., final exams in courses). However, this hypothesis needs to be substantiated with some empirical evidence.

### *Relationship between Emotions and Learning*

The theoretical perspectives discussed above make a number of predictions regarding relationships between emotions and learning gains. According to the zone of flow theory, the state of flow should also show a positive correlation with learning (Csikszentmihalyi, 1990), while boredom should be negatively correlated with learning (Csikszentmihalyi, 1990; Miserandino, 1996). If constructivist theory and the claims about cognitive disequilibrium are correct, we should observe a positive relationship between confusion and learning gains if the learning environment productively helps the learners regulate their confusion (Graesser & Olde, 2003; Kort et al., 2001). Similarly, a negative correlation is predicted between frustration and learning (Kort et al., 2001; Patrick et al., 1993).

We tested these predictions by correlating the proportional occurrence of boredom, confusion, flow, and frustration with learning measures collected in each study. Correlations were not performed for the emote-aloud study due to the small number of participants ( $N = 7$ ). Learning gains were obtained from knowledge tests administered both before and after the tutorial session (pretest and posttest, respectively). The testing materials were adapted from computer literacy tests used in previous experiments involving AutoTutor (Graesser et al., 2004). These tests had a 4-alternative multiple-choice format and consisted of questions that required inferences and deep reasoning, such as *why*, *how*, *what-if*, *what if not*, *how is X similar to Y?*.

With the exception of frustration, the predictions were supported in the observational study. As predicted, learning gains were positively correlated with confusion and flow and negatively correlated with boredom. There was no correlation between learning gains and

frustration (Craig et al., 2004). The positive correlation between confusion and learning was replicated in the multiple-judge study (D'Mello & Graesser, in review) and in the speech recognition study (Graesser, Chipman et al., 2007). Learning was not correlated with boredom, flow, and frustration in these studies.

It was somewhat of a surprise to discover that boredom and flow were not correlated with learning gains in the multiple-judge and speech recognition studies. It might be the case that these states operate on longer time-scales, so their effects on learning could not be observed in short 30-35 minute learning sessions. Longer learning sessions would be required before the effects of these states can be observed.

These emotions were correlated with learning in the predicted directions in the observational study, with approximately similar training times, so a comparison of methodologies is warranted. The major differences between studies included (a) the version of AutoTutor (improved version in multiple-judge and speech recognition studies), (b) the population of learners (low domain knowledge students in the observational study), (c) the emotion judgment frequency (5 minutes in observational study versus <20 seconds in the multiple-judge and speech recognition studies), and (d) the emotion judges (observers in observational study versus self, peers, and trained judges in the multiple-judge study, and self and peer in speech recognition study). Additional data and analyses would be needed to isolate which of these factors contributed to the discrepant findings between studies.

Another surprising finding was that frustration was not correlated with learning gains in any of the studies. Frustration is a state that occurs when learners fail to resolve an impasse, they get stuck, and important goals are blocked. The apparent lack of a relationship between frustration and learning might be attributed to the fact that AutoTutor does not let a learner

perseverate in an impasse. When AutoTutor tries to get a learner to articulate an idea, it first provides a hint (e.g. “What about X”). The hint is followed by a prompt (e.g. “X is a type of what?”) if the learner’s response to the hint was unsatisfactory. AutoTutor simply asserts the information when the learner cannot answer the prompt. Hence, impasses presumably caused by hints and prompts are eliminated with assertions. Withholding assertions and repeating hint-prompt cycles would presumably increase frustration and possibly impact learning.

There is some evidence to support this claim. For example negative affect (amalgamation of frustration, anxiety, and annoyed) were negatively correlated with posttests scores when the task was to read a passage in physics without any interference from a tutor (Linnenbrink & Pintrich, 2002). Frustration was also negatively linked to performance outcomes when students solved analytical reasoning problems in the absence of a tutor (D’Mello, Lehman, & Person, in review).

Perhaps the most important finding is that the positive correlation between confusion and learning was discovered in all three studies. This relationship is consistent with the model discussed earlier that claims that cognitive disequilibrium is one precursor to deep learning (Graesser, Lu et al., 2005; Graesser & Olde, 2003), with theories that highlight the merits of impasses during learning (Brown & VanLehn, 1980; VanLehn et al., 2003), and with models that help students learn how to overcome failure from getting stuck (Burlison & Picard, 2004). According to these models, confusion naturally occurs in the learning session when learners are confronted with information that is inconsistent with existing knowledge. Learners are in the state of cognitive disequilibrium, heightened physiological arousal, and more intense thought when they attempt to resolve impasses, discard misconceptions, and actively solve problems.

Confusion itself does not cause learning gains, but the cognitive activities that accompany confusion, cognitive disequilibrium, and impasse resolution are presumably linked to learning.

### **Modeling the Temporal Dynamics of Learners' Emotions**

Identification of the affective states that occur during learning is undoubtedly very important, but it could be argued that there is limited utility in merely knowing *what* states occur and their overall impact on learning. What is missing is a specification of *how* these states evolve, morph, interact, and influence learning and engagement. An analysis of mood states during a learning session will not suffice, because states such as confusion, frustration, surprise, and delight arise and decay at much faster timescales (a few seconds) compared to moods (several minutes or a few hours) (Ekman, 1984; Rosenberg, 1998). Simply put, the affective experiences that accompany learning are seldom static and persistent; instead, they are dynamic and highly transient.

This point is exemplified in the affective trajectory of a sample learner presented in Figure 4. The learner settles into flow after initially oscillating between flow and delight. An impasse or perturbation jerks the learner out of the flow state into a state of confusion. Repetitive oscillations between confusion and flow are observed, presumably as problem solving proceeds. Sometimes the learner gets stuck and experiences frustration. Success in problem solving yields delight and extreme novelty triggers surprise. This is the dominant pattern of emotional transitions until boredom kicks in towards the end of the session.

*Insert Figure 4 about here*

This example illustrates two interesting phenomena pertaining to the temporal dynamics of learners' emotions. First, learners tend to persevere in some states, while others are more transitory. For example, one would expect boredom to be more persistent than surprise, which is

undoubtedly a transitory state. It would be difficult to imagine a learner sustaining a state of surprise for more than a few seconds. The second interesting phenomenon is that some emotional transitions are more likely than others. For example, we would not expect flow to transition into boredom, whereas a flow-to-confusion transition is expected when an impasse is detected.

We investigated these questions by analyzing the persistence of individual states and transitions between states (D'Mello & Graesser, in review; D'Mello & Graesser, in review). The data from the multiple-judge and speech recognition studies were used for this analysis because the sampling methodology in the observational and emotive-aloud studies did not have the requisite sampling rate to warrant a temporal analysis.

### *Persistence of Emotions*

Although the scientific literature on the persistence of the learning-centered emotions is sparse, it is possible to theoretically align them on the following temporal scale in increasing order of persistence: (Delight = Surprise) < (Confusion = Frustration) < (Boredom = Engagement/Flow). These predictions can be understood from the perspective of goal-appraisal theories of emotion (Mandler, 1976, 1999; Stein et al., 2008; Stein & Levine, 1991). In general, learners are typically in a *prolonged* state of either (a) engagement/flow as they pursue the superordinate learning goal of mastering the material or (b) disengagement (boredom) when they abandon pursuit of the superordinate learning goal. When they are deeply engaged, they attempt to assimilate new information into existing knowledge schemas. However, when new or discrepant information is detected, attention shifts to the discrepant information, the autonomic nervous systems increases in arousal, and the learner experiences a variety of possible states depending on the context, the amount of change, and whether important goals are blocked. In the case of extreme novelty, the event evokes surprise. When the novelty triggers the

achievement of a goal, the emotion is positive, such as delight or even one of those rare *eureka* experiences (Knoblich, Ohlsson, & Raney, 2001). Previous research on delight and surprise has indicated that these emotions are typically quite brief (Ekman, 1984, 1992). In contrast, confusion and frustration occur when the discrepancy or novelty trigger an impasse that blocks the superordinate learning goal and possibly results in the student getting stuck. The learner initiates a subgoal of resolving the impasse through effortful reasoning and problem solving. Confusion and frustration address a subgoal, so they should be shorter than the states of flow and boredom that address the major goal. However, confusion and frustration are expected to persist longer than the short-lived reactions of delight and surprise.

We developed a set of exponential decay models to capture graded differences in the decay rates of the various emotions (D'Mello & Graesser, in review). The models supported a tripartite classification of learning-centered emotions along a temporal dimension: persistent emotions (boredom, flow, and confusion), transitory emotions (delight and surprise), and an intermediate emotion (frustration). This pattern somewhat confirms the aforementioned predictions stemming from goal-appraisal theories of emotion, with the exception that confusion was categorized as a persistent rather than an intermediate emotion.

### *Transitions between Emotions*

Cognitive disequilibrium theory makes a number of predictions about the transitions between the learning-centered emotions. Learners who are in a flow/engaged state will experience confusion when an impasse is detected. They engage in effortful problem solving activities in order to resolve the impasse and restore equilibrium. Equilibrium is restored when the impasse is resolved and learners revert back into the flow/engaged state. However, confusion transitions into frustration when the impasse cannot be resolved, the student gets stuck, and

important goals are blocked (Burlison & Picard, 2004). Furthermore, persistent frustration may transition into boredom, a crucial point at which the learner disengages from the learning process.

The major hypotheses of the model were tested by performing time-series analyses on the data from the multiple-judge study and the speech recognition study (D'Mello & Graesser, in review; D'Mello, Taylor et al., 2007). The results confirmed the presence of confusion--flow/engagement and boredom--frustration oscillations as well as confusion to frustration transitions (see Figure 5). Hence, students in the state of engagement/flow are continuously being challenged within their zones of optimal learning (Brown et al., 1998; Vygotsky, 1978) and are experiencing two-step episodes alternating between confusion and insight. In contrast to these *beneficial* flow-confusion-flow cycles, there are the *harmful* oscillations between boredom and frustration. As cognitive disequilibrium theory asserts, confusion plays a central role in the learning process because it the gateway to positive (flow) or negative (frustration) emotions.

*Insert Figure 5 about here*

### **Assessing Contextual Influences on Learners' Emotions**

An investigation into the emotions that occur during learning will not be complete without a discussion of the context surrounding the emotional experiences (Aviezer et al., 2008; Barrett, 2006; Russell, 2003; Stemmler, Heldmann, Pauls, & Scherer, 2001). In many cases, examining the context of an emotional expression can lead to a deeper (and sometimes even causal) explanation of the emotional experience. For example, confusion *while* solving a problem can be contrasted with confusion *after* receiving feedback for the solution. The first form of confusion can be causally attributed to being perplexed with the problem itself, while confusion after feedback is more related to the problem solving outcome. Though similar, these two forms

of confusion might have distinct manifestations (i.e., intensity, valence) and different pedagogical virtues.

In general, the context is critical because it helps to disambiguate between various exemplars of an emotion category (Russell, 2003). For example, the two forms of confusion discussed above are different exemplars of the “confusion” category. Examining confusion (i.e., the category) out of context (i.e., without the exemplar) runs the risk of being meaningless.

One advantage of investigating emotions with a dialogue-based intelligent tutoring system like AutoTutor is that the dialogue history provides a rich trace into the contextual underpinnings of learners’ emotional experiences. For example, consider a four-turn segment of the dialogue history, the accompanying facial expressions, and self-reported emotions of a learner from an actual tutorial session (Figure 6). The learner is in the neutral state when the tutor discusses a change of topic (Figure 6A). She is then engaged while she tries to answer the question (Figure 6B). She experiences an impasse when she is uncertain about her answer (Figure 6C). The answer is incorrect and the tutor responds with negative feedback which frustrates the student (Figure 6D).

*Insert Figure 6 about here*

We examined the tutorial dialogue (i.e. the context) over 15 second intervals that culminated in episodes of boredom, flow/engagement, confusion, and frustration that were reported by the self, peers, and trained judges (D’Mello et al., 2006; D’Mello, Craig et al., 2008). An event triggering an emotional event could either be tutor generated (i.e. boredom because the tutor is providing a long-winded explanation), student generated (i.e. boredom because the student has no interest in computer literacy), or a session related event (i.e. boredom because the tutorial session is dragging on).

We discovered a number of interesting patterns that provide some insights into the emotional experiences during learning with AutoTutor (D'Mello & Graesser, in review; D'Mello, Taylor et al., 2007). In particular, boredom occurs later in the session, after multiple attempts to answer a question. Boredom also occurs when AutoTutor gives more direct dialogue moves (i.e., assertions or summaries are more direct than pumps or hints). In contrast, confusion occurs earlier in the session, within the first few attempts to answer a question, with slower and less verbose responses, with poor answers, with frozen expressions (instead of domain related contributions), when the tutor is less direct, and when the tutor provides negative feedback. Flow occurs within the first few attempts to answer a question, with quicker, longer, proficient responses, and is accompanied by positive feedback from the tutor. Frustration was prevalent later in the temporal span of a session, with longer response times, with good answers towards the immediate question, but poor answers towards the broader topic, and after negative tutor feedback. In summary, the relationships between the various dialogue features and the affective states described above are generally intuitive and in the expected directions.

### **Programming AutoTutor to be Responsive to Learners' Emotions**

Now that we have a better understanding of what emotions are relevant to learning and how they arise and decay within the context of a learning session, we turn our attention to the practical problem of developing an ITS that is responsive to learners' affective and cognitive states. We have recently developed two new versions of AutoTutor that detect and respond to learners' affective and cognitive states (D'Mello, Jackson et al., 2008; D'Mello, Craig et al., 2009). These affect-sensitive variants of AutoTutor detect and respond to boredom, confusion, and frustration. Appropriate responses to these states could potentially have a positive impact on engagement and learning outcomes. These affect-sensitive versions of AutoTutor have a set of

production rules that were designed to map dynamic assessments of the student's cognitive and affective states with tutor actions to address the presence of the negative emotions (see Figure 7). Hence, the learner and the tutor are embedded into an affective loop that involves *detecting* the learner's affective states, *responding* to the detected states, and *synthesizing* emotional expressions via animated pedagogical agents. These processes are briefly described below.

*Insert Figure 7 about here*

#### *Detecting Affect and Cognitive States*

The affect detection system monitors conversational cues, gross body language, and facial features to detect boredom, confusion, frustration, and neutral (no affect). Automated affect-detection systems that detect these emotions have been integrated into AutoTutor (see (D'Mello, Craig et al., 2008; D'Mello & Graesser, 2009; D'Mello, Picard et al., 2007)). Each channel independently provides its own diagnosis of the student's affective state. These individual diagnoses are combined with a decision-level fusion algorithm that selects a single affective state (see Current State and Next State in Figure 7) and a confidence value (see Detection Probability in Figure 7).

The tutor's model of learners' cognitive states include a global measure of student ability (dynamically updated throughout the session) and the conceptual quality of the student's immediate response (Global Ability and Local Ability, respectively, in Figure 7). These parameters are computed by performing a syntactic and semantic analysis of the student's past and immediate utterances (Graesser, Penumatsa, Ventura, Cai, & Hu, 2007; Graesser et al., 2000).

### *Responding to Affective and Cognitive States*

AutoTutor provides short feedback to each student response. In addition to articulating the verbal content of the feedback, the affective AutoTutor also modulates its facial expressions and speech prosody. Positive feedback is delivered with an *approval* expression (big smile and big nod). Neutral positive feedback receives a *mild approval* expression (small smile and slight nod). Negative feedback is delivered with a *disapproval* expression (slight frown and head shake), while the tutor makes a *skeptical* face when delivering neutral-negative feedback (see Figure 8). No facial expression accompanies the delivery of neutral feedback.

After delivering the feedback, the affective AutoTutor delivers an emotional statement if it senses that the student is bored, confused, or frustrated. A non-emotional discourse marker (e.g. “Moving on”, “Try this one”) is selected if the student is neutral. AutoTutor’s strategies to respond to boredom, confusion, and frustration are motivated by attribution theory (Batson, Turk, Shaw, & Klein, 1995; Heider, 1958; Weiner, 1986), cognitive disequilibrium during learning (Craig et al., 2004; Festinger, 1957; Graesser & Olde, 2003; Piaget, 1952), and recommendations by pedagogical experts. These perspectives are integrated in two pedagogically distinct variants of the affect-sensitive AutoTutor. These include a Supportive and a Shakeup AutoTutor.

*Supportive AutoTutor.* The supportive AutoTutor responds to the learners’ affective states via empathetic and motivational responses. These responses always attribute the source of the learners’ emotion to the material instead of the learners themselves. So the supportive AutoTutor might respond to mild boredom with “This stuff can be kind of dull sometimes, so I’m gonna try and help you get through it. Let’s go”. A response to confusion would include attributing the source of confusion to the material (“Some of this *material* can be confusing. Just keep going

and I am sure you will get it”) or the tutor itself (“I know I do not always convey things clearly. I am always happy to repeat myself if you need it. Try this one”).

*Shakeup AutoTutor.* The major difference between the shakeup AutoTutor and the supportive AutoTutor lies in the source of emotion attribution. While the supportive AutoTutor attributes the learners’ negative emotions to the material or itself, the shakeup AutoTutor directly attributes the emotions to the learners. For example, possible shakeup responses to confusion are “This material has got *you* confused, but I think you have the right idea. Try this...” and “You are not as confused as you might think. I’m actually kind of impressed. Keep it up”.

Another difference between the two versions lies in the conversational style. While the supportive AutoTutor is subdued and formal, the shakeup tutor is edgier, flaunts social norms, and is witty. For example, a supportive response to boredom would be “Hang in there a bit longer. Things are about to get interesting.”. The shakeup counterpart of this response is “Geez this stuff sucks. I’d be bored too, but I gotta teach what they tell me”.

### *Synthesizing Affective Expressions*

The affect-sensitive versions of AutoTutor synthesize affect with facial expressions and emotionally modulated speech. These affective expressions include: approval, mild approval, disapproval, empathy, skepticism, mild enthusiasm, and high enthusiasm (Figure 8). The supportive and shakeup responses are always paired with the appropriate expression, which can be neutral in some cases. The facial expressions in each display were informed by Ekman’s work on the facial correlates of emotion expression (Ekman & Friesen, 1978).

*Insert Figure 8 about here*

The facial expressions of emotion displayed by AutoTutor are augmented with emotionally expressive speech synthesized by the agent. The emotional expressivity is obtained

by variations in pitch, speech rate, and other prosodic features. Previous research has led us to conceptualize AutoTutor's affective speech on the indices of pitch range, pitch level, and speech rate (Johnstone & Scherer, 2000).

### *Evaluating the Affect-Sensitive AutoTutor*

We have recently conducted an experiment that evaluated the pedagogical effectiveness of the Supportive AutoTutor when compared to the original tutor (D'Mello, Lehman, Sullins et al., in review). This original AutoTutor has a conventional set of fuzzy production rules that are sensitive to cognitive states of the learner, but not to the emotional states of the learner. The Supportive AutoTutor is sensitive to learners' affective and cognitive states. The obvious prediction is that learning gains should be superior for the Supportive AutoTutor.

The experiment utilized a between-subjects design where learners (a) completed a pretest on topics in computer literacy, (b) were tutored on two computer literacy topics with either the affective or the regular AutoTutor, and (c) completed a posttest. The tests and tutorial sessions were pitched at deeper levels of comprehension with questions that required reasoning and inference instead of the recall of shallow facts and definitions. The tutorial session consisted of two 30-minute sessions on different computer literacy topics but with the same version of AutoTutor (i.e. either Supportive or Original).

The results of this experiment indicated that the supportive AutoTutor was more effective than the regular tutor for low-domain knowledge students in the second session ( $d = .713$ ), but not the first session. This suggests that it is inappropriate for the tutor to be supportive to these students before there has been enough context to show there are problems. Simply put, don't be supportive until the students need support. Second, the students with more knowledge never benefited from the supportive AutoTutor. These students don't need the emotional support, but

rather they need to go directly to the content. Third, there are conditions when emotional support is detrimental, if not irritating to the learner. There appears to be a liability to quick support and empathy compared to no affect-sensitivity.

The central message is that there is an appropriate time for affect-sensitivity in the form of supportive dialogues. Just as there is a “time for telling”; there is a “time for emoting.” We could imagine a strategy where low-knowledge students start out with a non-emotional regular tutor until they see there are problems. Then after that they need support, as manifested in the second tutorial session. Regarding high-knowledge students, they are perfectly fine working on content for an hour or more and may get irritated with an AutoTutor showing compassion, empathy, and care. But later on there may be a time when they want a shake-up AutoTutor for stimulation, challenge, and a playful exchange. Or maybe even a supportive AutoTutor. These are all questions to explore in future research.

### **Conclusions**

The idea of having a tutoring system detect, respond to, and synthesize emotions was once a seductive vision (Picard, 1997). This vision is now a reality as affect-sensitive learning environments are coming online. Our research on emotions during learning with AutoTutor represents one out of a handful of related efforts made by researchers who have a vision of enhancing engagement and motivation, boosting self-efficacy, and promoting learning gains by designing intelligent learning environments that optimally coordinate cognition and emotion (Arroyo et al., 2009; Burleson & Picard, 2007; Chaffar, Derbali, & Frasson, 2009; Conati & Maclaren, 2009; D’Mello, Craig et al., 2009; Forbes-Riley & Litman, 2009; Robison, McQuiggan, & Lester, 2009; Woolf, 2009).

Despite the impressive progress on affect-sensitive ITSs, it should be noted that affect detectors and reactors (i.e. they detect and respond to learners' emotions) are not the panacea for the problem of promoting learning gains. Our research suggests that they need to be affect *anticipators*, *forestallers*, and *inducers* as well. Affect anticipators and forestallers would be required to predict and prevent the occurrence of persistent negative affective states like boredom and presumably frustration. Prediction and prevention is necessary to address boredom because boredom begets frustration and even more boredom (D'Mello & Graesser, in review; D'Mello, Taylor et al., 2007). More importantly, tutorial interventions are not very effective in alleviating boredom when learners tend to experience harmful oscillations between boredom and frustration (D'Mello & Graesser, in review), indicating that advanced prediction is very important to regulate boredom.

Proactively responding to boredom would involve engaging the learner in a task that increases interest and cognitive arousal, such as a challenge, an interactive simulation, or a seductive embedded serious game. These difficult tasks have a high likelihood of getting students to reengage with the material. Another strategy is to provide participants with a choice of tasks and topics so they might pick one that is more relevant to their interests. Curiosity and engagement are enhanced by the learner's freedom of choices (Lepper & Woolverton, 2002).

On the other hand, the positive link between confusion and learning suggests that learning environments need to substantially challenge students in order to illicit critical thought and deep inquiry. Therefore, a promising strategy to promote opportunities for deep learning is to develop affect induction interventions that jolt students out of their perennial state of blasé comprehension by presenting challenges with contradictions, incongruities, anomalies, system breakdowns, and difficult decisions (Bjork & Linn, 2006; Festinger, 1957; Graesser & Olde,

2003; Schwartz & Bransford, 1998). Learners experience impasses, cognitive disequilibrium, and confusion in these conditions. Cognitive equilibrium is restored after thought, reflection, problem solving, self explanations, and other effortful cognitive activities that force learners to pause and think.

In summary, the scientific research on affect and learning is an exciting research direction that we believe will be a priority for the next decade or longer. Progress will depend on breakthroughs in the development of affective computing technologies, such as affect detectors, reactors, synthesizers, anticipators, forestallers, and inducers. Future research will hopefully unveil novel insights into the intricate dance among emotions, cognition, motivation, individual differences, pedagogical interventions, and learning gains.

### **Author Notes**

This research was supported by the National Science Foundation (REC 0106965, ITR 0325428, HCC 0834847). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

We thank our research colleagues in the Emotive Computing Group and the Tutoring Research Group (TRG) at the University of Memphis (<http://emotion.autotutor.org>). We gratefully acknowledge our partners in the Affective Computing group at the MIT Media Lab.

The Tutoring Research Group (TRG) is an interdisciplinary research team comprised of researchers from psychology, computer science, physics, linguistics, engineering, and education (visit <http://www.autotutor.org>, <http://emotion.autotutor.org>, <http://fedex.memphis.edu/iis/> ).

Requests for reprints should be sent to Sidney D'Mello, 202 Psychology Building, University of Memphis, Memphis, TN 38152-3230, [sdmello@memphis.edu](mailto:sdmello@memphis.edu).

### References

- Aleven, V., & Koedinger, K. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, 26(2), 147-179.
- Anderson, J., Douglass, S., & Qin, Y. (2005). How should a theory of learning and cognition inform instruction? In A. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 47-58). Washington, DC.: American Psychological Association.
- Arroyo, I., Woolf, B., Cooper, D., Burleson, W., Muldner, K., & Christopherson, R. (2009). Emotion Sensors Go To School. In V. Dimitrova, R. Mizoguchi, B. Du Boulay & A. Graesser (Eds.), *Proceedings of 14th International Conference on Artificial Intelligence In Education*. Amsterdam: IOS Press.
- Aviezer, H., Hassin, R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, 19(7), 724-732.
- Baker, R., D'Mello, S., Rodrigo, M., & Graesser, A. (in review). Better to be frustrated than bored: The incidence and persistence of affect during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*.
- Barrett, L. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science* 1, 28-58.
- Barrett, L., Mesquita, B., Ochsner, K., & Gross, J. (2007). The experience of emotion. *Annual Review of Psychology*, 58, 373-403.

- Batson, C., Turk, C., Shaw, L., & Klein, T. (1995). Information Function of Empathic Emotion - Learning That We Value the Others Welfare. *Journal of Personality and Social Psychology*, 68(2), 300-313.
- Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology-Learning Memory and Cognition*, 33(6), 983-998.
- Biggs, J. (1995, July). *Enhancing teaching through constructive alignment*. Paper presented at the 20th International Conference on Improving University Teaching, Hong Kong, Hong Kong.
- Bjork, R. A., & Linn, M. C. (2006). The Science of Learning and the Learning of Science: Introducing Desirable Difficulties. *American Psychological Society Observer*, 19, 3.
- Bless, H., & Fielder, K. (1995). Affective states and the influence of activated general knowledge. *Personality and Social Psychology Bulletin*, 21, 766-778.
- Boekaerts, M. (1993). Being concerned with well-being and with learning. *Educational Psychologist*, 28(2), 149-167.
- Bower, G. (1981). Mood and memory. *American Psychologist*, 36, 129-148.
- Bower, G. (1992). How Might Emotions Affect Learning. In S. A. Christianson (Ed.), *The Handbook of Emotion and Memory: Research and Theory* (pp. 3-31). Hillsdale, NJ: Erlbaum.
- Brown, A., Ellery, S., & Campione, J. (1998). Creating Zones of Proximal Development Electronically in Thinking Practices in Mathematics and Science Learning. In J. Greeno & S. Goldman (Eds.), (pp. 341-368). Mahawah, NJ: Lawrence Erlbaum.

- Brown, J., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science*, 4(379-426,).
- Burleson, W., & Picard, R. (2004). *Affective agents: Sustaining motivation to learn through failure and a state of "stuck"*. Paper presented at the ITS 2004 Workshop Proceedings on Social and Emotional, Maceio, Brazil.
- Burleson, W., & Picard, R. (2007). Evidence for Gender Specific Approaches to the Development of Emotionally Intelligent Learning Companions. *IEEE Intelligent Systems*, 22(4), 62-69.
- Chaffar, S., Derbali, L., & Frasson, C. (2009). *Inducing Positive Emotional State in Intelligent Tutoring Systems*. Paper presented at the 14th International Conference on Artificial Intelligence in Education.
- Chi, M., Roy, M., & Hausmann, R. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, 32(2), 301-341.
- Clifford, M. (1988). Failure tolerance and academic risk-taking in ten- to twelve-year-old students. *British Journal of Educational Psychology*, 58(15-27).
- Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, 16(7-8), 555-575.
- Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, 19(3), 267-303.
- Craig, S., Graesser, A., 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 and Row.
- D'Mello, S., Craig, S., Gholson, B., Franklin, S., Picard, R., & Graesser, A. (2005). Integrating affect sensors in an intelligent tutoring system. In *The Computer In The Affective Loop Workshop At 2005 International Conference On Intelligent User Interfaces* (pp. 7-13).
- D'Mello, S., Craig, S., Sullins, J., & Graesser, A. (2006). Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education*, 16(1), 3-28.
- D'Mello, S., Craig, S., Witherspoon, A., McDaniel, B., & Graesser, A. (2008). Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction*, 18(1-2), 45-80.
- D'Mello, S., & Graesser, A. (2009). Automatic detection of learners' affect from gross body language. 23(2), 123 - 150.
- D'Mello, S., & Graesser, A. (in review). Dynamics of Cognitive-Affective States during Deep Learning.
- D'Mello, S., Jackson, G., Craig, S., Morgan, B., Chipman, P., White, H., et al. (2008). *Autotutor detects and responds to learners affective and cognitive states*. Paper presented at the Workshop on Emotional and Cognitive issues in ITS held in conjunction with the Ninth International Conference on Intelligent Tutoring Systems, Montreal, Canada.
- D'Mello, S., King, B., Entezari, O., Chipman, P., & Graesser, A. (2008, March). *The impact of automatic speech recognition errors on learning gains with AutoTutor*. Paper presented at the Annual meeting of the American Educational Research Association, New York, New York.

- D'Mello, S., Lehman, B., & Person, N. (in review). Monitoring affect states during effortful problem solving activities. *International Journal of Artificial Intelligence In Education*.
- D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., et al. (in review). A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning.
- D'Mello, S., Picard, R., & Graesser, A. (2007). Towards an affect-sensitive AutoTutor. *Intelligent Systems, IEEE*, 22(4), 53-61.
- D'Mello, S., Taylor, R., Davidson, K., & Graesser, A. (2008). Self versus teacher judgments of learner emotions during a tutoring session with AutoTutor. In B. Woolf, E. Aimeur, R. Nkambou & S. Lajoie (Eds.), *Proceedings of the 9th international conference on Intelligent Tutoring Systems*. Berlin, Heidelberg: Springer-Verlag.
- D'Mello, S., Taylor, R., & Graesser, A. (2007). Monitoring affective trajectories during complex learning. In D. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 203-208). Austin, TX: Cognitive Science Society.
- D'Mello, S., Craig, S., Fike, K., & Graesser, A. (2009). Responding to learners' cognitive-affective states with supportive and shakeup dialogues. In J. Jacko (Ed.), *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction* (pp. 595-604). Berlin/Heidelberg: Springer.
- D'Mello, S., Dale, R., & Graesser, A. (in review). Disequilibrium in the Mind, Disharmony in the Body.
- D'Mello, S., Dowell, N., & Graesser, A. (2009). Cohesion Relationships in Tutorial Dialogue as Predictors of Affective States. In V. Dimitrova, R. Mizoguchi, B. du Boulay & A. Graesser (Eds.), *Proceedings of 14th International Conference on Artificial Intelligence In Education* (pp. 9-16). Amsterdam: IOS Press.

- D'Mello, S., & Graesser, A. (in review). The Half-Life of Emotions.
- Damasio, A. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*: Harcourt Inc.
- Deci, E., & Ryan, R. (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.
- Dragon, T., Arroyo, I., Woolf, B. P., Burleson, W., el Kaliouby, R., & Eydgahi, H. (2008, Jun 23-Jul 27). *Viewing student affect and learning through classroom observation and physical sensors*. Paper presented at the 9th International Conference on Intelligent Tutoring Systems, Montreal, CANADA.
- Dweck, C. (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. (1984). Expression and the nature of emotion. In K. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 319-344). Hillsdale, NJ: Erlbaum.
- Ekman, P. (1992). An Argument for Basic Emotions. *Cognition & Emotion*, 6(3-4), 169-200.
- Ekman, P. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Henry Holt and Company, LLC.
- Ekman, P., & Friesen, W. (1978). *The Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press.
- Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data (Rev. ed.)*. Cambridge, MA: The MIT Press.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

- Fielder, K. (2001). Affective states trigger processes of assimilation and accommodation. In K. Martin & G. Clore (Eds.), *Theories of Mood and Cognition: A User's Guidebook* (pp. 85-98). Mahwah: Erlbaum.
- Forbes-Riley, K., & Litman, D. (2009). Adapting to student uncertainty improves tutoring dialogues. In V. Dimitrova, R. Mizoguchi & B. Du Boulay (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 33-40). Amsterdam: IOS Press.
- Forbes-Riley, K., Rotaru, M., & Litman, D. (2008). The relative impact of student affect on performance models in a spoken dialogue tutoring system. *User Modeling and User-Adapted Interaction*, 18(1-2), 11-43.
- Forgas, J. (1991). Affect and cognition in close relationships. In J. Fletcher & J. Fincham (Eds.), *Cognition in Close Relationships* (pp. 151-174). Hillsdale, NJ: Erlbaum.
- Gertner, A., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In G. Gauthier, C. Frasson & K. VanLehn (Eds.), *Proceedings of the International Conference on Intelligent Tutoring Systems* (pp. 133-142). Berlin / Heidelberg: Springer.
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.
- Gotlib, I., & Abramson, L. (1999). Attributional Theories of Emotion. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. Sussex: John Wiley & Sons Ltd.
- Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612-618.
- Graesser, A., Chipman, P., King, B., McDaniel, B., & D'Mello, S. (2007). Emotions and learning with AutoTutor. In R. Luckin, K. Koedinger & J. Greer (Eds.), *13th International Conference on Artificial Intelligence in Education* (pp. 569-571): IOS Press.

- Graesser, A., Lu, S., Olde, B., 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., Lu, S. L., Jackson, G., Mitchell, H., Ventura, M., Olney, A., et al. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers*, *36*, 180-193.
- Graesser, A., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. (2006). *Detection of emotions during learning with AutoTutor*. Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, Canada.
- Graesser, A., & Olde, B. (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*(3), 524-536.
- Graesser, A., Penumatsa, P., Ventura, M., Cai, Z., & Hu, X. (2007). Using LSA in AutoTutor: Learning through mixed-initiative dialogue in natural language. In T. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 243-262). Mahwah, NJ: Erlbaum.
- Graesser, A., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & TRG. (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments*, *8*(2), 129-147.
- Harter, S. (1992). The relationship between perceived competence, affect, and motivational orientation within the classroom: process and patterns of change. In A. Boggiano & T. Pittman (Eds.), *Achievement and motivation: a social-developmental perspective* (pp. 77-114). New York: Cambridge University Press.

- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: John Wiley & Sons.
- Hertel, G., Neuhof, J., Theuer, T., & Kerr, N. (2000). Mood effect on cooperation in small groups: Does positive mood simply lead to more cooperation? *Cognition and Emotion*, *14*, 441-472.
- Immordino-Yang, M. H., & Damasio, A. R. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain and Education*, *1*(1), 3-10.
- Isen, A. (2001). An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. *Journal of Consumer Psychology*, *11*, 75-85.
- Isen, A., Daubman, K., & Nowicki, G. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*, 1122-1131.
- Issroff, K., & del Soldato, T. (1996). *Incorporating motivation into computer-supported collaborative learning*. Paper presented at the Proceedings of European conference on artificial intelligence in education.
- Izard, C. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Izard, C., & Ackerman, B. (2000). Motivational, organizational, and regulatory functions of discrete emotions. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed.). New York: Guilford Press.
- Johnstone, T., & Scherer, K. (2000). Vocal communication of emotion. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed., pp. 220-235). New York: Guilford Press.
- Jonassen, D., Peck, K., & Wilson, B. (1999). *Learning with technology: A constructivist perspective*. Upper Saddle River, NJ: Prentice Hall.

- Keltner, D., & Ekman, P. (2000). Facial expression of emotion. In R. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (Vol. 2nd ed, pp. 236–264). New York: Guilford.
- Keltner, D., & Shiota, M. (2003). New displays and new emotions: a commentary on Rozin and Cohen (2003). *Emotion*, 3(86-91).
- Knoblich, G., Ohlsson, S., & Raney, G. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, 29(7), 1000-1009.
- Koedinger, K., Anderson, J., Hadley, W., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8(30-43).
- Koedinger, K., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61-78). New York, NY: Cambridge University Press.
- Kort, B., Reilly, R., & Picard, R. (2001). *An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion*. Paper presented at the Advanced Learning Technologies, 2001. Proceedings. IEEE International Conference on.
- Lazarus, R. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Lazarus, R. (2000). The cognition-emotion debate: A bit of history. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed., pp. 1-20). New York: Guilford Press.
- Lehman, B., D’Mello, S., & Person, N. (2008). *All Alone with your Emotions: An Analysis of Student Emotions during Effortful Problem Solving Activities*. Paper presented at the Workshop on Emotional and Cognitive issues in ITS at the Ninth International Conference on Intelligent Tutoring Systems.

- Lehman, B., Matthews, M., D'Mello, S., & Person, N. (2008). What are you feeling? Investigating student affective states during expert human tutoring sessions. In B. Woolf, A. E., N. R. & L. S. (Eds.), *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 50-59).
- Lepper, M., & Chabay, R. (1988). Socializing the intelligent tutor: Bringing empathy to computer tutors. In H. Mandl & A. Lesgold (Eds.), *Learning Issues for Intelligent Tutoring Systems* (pp. 242-257). Hillsdale, NJ: Erlbaum.
- Lepper, M., & 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., Bunzo, M., & Eggen, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. H. Larkin & R. W. Chabay (Eds.), *Computer-assisted instruction and intelligent tutoring systems* (pp. 201-238). Hillsdale, NJ: Erlbaum.
- Linnenbrink, E., & 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* (pp. 115-135). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Mandler, G. (1976). *Mind and emotion*. New York: Wiley.
- Mandler, G. (1984). *Mind and Body: Psychology of Emotion and Stress*. New York: W.W. Norton & Company.
- Mandler, G. (1999). Emotion. In B. M. Bly & D. E. Rumelhart (Eds.), *Cognitive science. Handbook of perception and cognition* (2nd ed.). San Diego, CA: Academic Press.

- McDaniel, B., D'Mello, S., King, B., Chipman, P., Tapp, K., & Graesser, A. (2007). Facial Features for Affective State Detection in Learning Environments. In D. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 467-472). Austin, TX: Cognitive Science Society.
- Meyer, D., & Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review, 18*(4), 377-390.
- Miserandino, M. (1996). Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. *Journal of Educational Psychology, 88*(2), 203-214.
- Moshman, D. (1982). Exogenous, endogenous, and dialectical constructivism. *Developmental Review, 2*(4), 371-384.
- Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of emotions*. New York: Cambridge University Press.
- Otero, J., & Graesser, A. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction, 19*(2), 143-175.
- Patrick, B., Skinner, E., & Connell, J. (1993). What Motivates Childrens Behavior and Emotion - Joint Effects of Perceived Control and Autonomy in the Academic Domain. *Journal of Personality and Social Psychology, 65*(4), 781-791.
- Pekrun, R., Elliot, A., & Maier, M. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*(3), 583-597.
- Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.
- Picard, R. (1997). *Affective Computing*. Cambridge, Mass: MIT Press.

- Robison, J., McQuiggan, S., & Lester, J. (2009). *Evaluating the Consequences of Affective Feedback in Intelligent Tutoring Systems*. Paper presented at the International Conference on Affective Computing & Intelligent Interaction,, Amsterdam.
- Rosenberg, E. (1998). Levels of analysis and the organization of affect. *Review of General Psychology*, 2(3), 247-270.
- Rozin, P., & Cohen, A. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans. *Emotion*, 3(68-75).
- Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145-172.
- Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1-2), 227-256.
- Scherer, K., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. London: London University Press.
- Schultz, P., & Pekrun, R. (Eds.). (2007). *Emotion in Education*. San Diego, CA: Academic Press.
- Schwartz, D., & Bransford, D. (1998). A time for telling. *Cognition and Instruction*, 16(4), 475-522.
- Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, 14, 433-440.
- Schwarz, N., & Skurnik, I. (2003). Feeling and thinking: Implications for problem solving. In J. D. R. S. (Eds.) (Ed.), *The Psychology of Problem Solving* (pp. 263-290). New York: Cambridge University Press.

- Sleeman, D., & Brown, J. (Eds.). (1982). *Intelligent tutoring systems*. New York: Academic Press.
- Smith, C., & Ellsworth, P. (1985). Patterns of Cognitive Appraisal in Emotion. *Journal of Personality and Social Psychology*, 48(4), 813-838.
- Stein, N., Hernandez, M., & Trabasso, T. (2008). Advances in modeling emotions and thought: The importance of developmental, online, and multilevel analysis. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 574-586). New York: Guilford Press.
- Stein, N., & Levine, L. (1991). Making sense out of emotion. In A. O. W. Kessen, & F. Kraik (Eds.) (Ed.), *Memories, thoughts, and emotions: Essays in honor of George Mandler* (pp. 295-322). Hillsdale, NJ: Erlbaum.
- Stemmler, G., Heldmann, M., Pauls, C., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. *Psychophysiology*, 38(2), 275-291.
- Stipek, D. (1988). *Motivation to learn: from theory to practice*. Boston, MA: Allyn and Bacon.
- Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, 64, 37-54.
- VanLehn, K., Graesser, A., Jackson, G., Jordan, P., Olney, A., & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31(1), 3-62.
- VanLehn, K., Lynch, C., Schulze, K., Shapiro, J., Shelby, R., Taylor, L., et al. (2005). The Andes physics tutoring system: five years of evaluations. *International Journal of Artificial Intelligence in Education* 15, 147-204.
- VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, 21(3), 209-249.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*.

Cambridge, MA: Harvard University Press.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-

Verlag.

Woolf, B. (2009). *Building intelligent interactive tutors*. Burlington, MA: Morgan Kaufmann

Publishers.

### Figure Captions

*Figure 1.* (A) AutoTutor interface. The main question stays on the computer screen throughout the conversation. The animated conversational agent speaks the content of AutoTutor's turns via synthesized speech. The image window is either blank or has auxiliary diagrams pertaining to the main question. The dialogue history contains the conversational exchange between the student and the tutor. The student response window displays the student's answers as they are typed in. (B) Sample dialogue from an actual tutorial session with annotated tutor dialogue moves displayed in brackets.

*Figure 2.* Sensors used in the multiple-judge study. The leftmost and rightmost monitors were turned off during the tutoring session.

*Figure 3.* Proportional occurrence of affective states across four studies.

*Figure 4.* Affective trajectory of a student during a learning session with AutoTutor. This time series is from a student in the multiple judge study. The student self-reported their affective states at 20 second intervals as well as in between the 20-second blocks.

*Figure 5.* Observed pattern of transitions between emotions

*Figure 6.* Sequence of affective states annotated with the text of student or tutor dialogue move when the emotions were experienced.

*Figure 7.* Architecture of affect-sensitive AutoTutor. Blue and red boxes indicate cognitive and affective elements, respectively.

*Figure 8.* Synthesized facial expressions by AutoTutor's animated conversational agent.

Figure 1.

**(A) AUTOTUTOR INTERFACE**

The screenshot shows the AutoTutor interface with a purple background. On the left is a 3D avatar of a man labeled "Conversational Agent". The main area contains a "Main Question" in red: "How does information that you type in get passed from the keyboard to the hard disk?". Below the question is an "Auxillary Diagram" (misspelled) showing a computer system with components: Memory, Central Processing Unit, Input Device, and Hard Drive. Arrows indicate data flow from the Input Device to the CPU, and from the CPU to Memory and the Hard Drive. Below the diagram is a "Student Response" in red: "through the CPU". At the bottom left are "Submit" and "Settings..." buttons.

**(B) SAMPLE TUTORIAL DIALOGUE**

**T:** When you turn on the computer, how is the operating system first activated and loaded into RAM? *[main question]*

**S:** When the computer is turned on a file is automatically booted up.

**T:** Anything else? *[pump for info]*

**S:** This file contains the operating system.

**T:** What about storage? *[hint]*

**S:** The operating system is then read into RAM and activated.

**T:** Right. The CPU executes instructions in ROM, which helps the operating system to start. *[after positive feedback, there is an assertion]*

Figure 2.

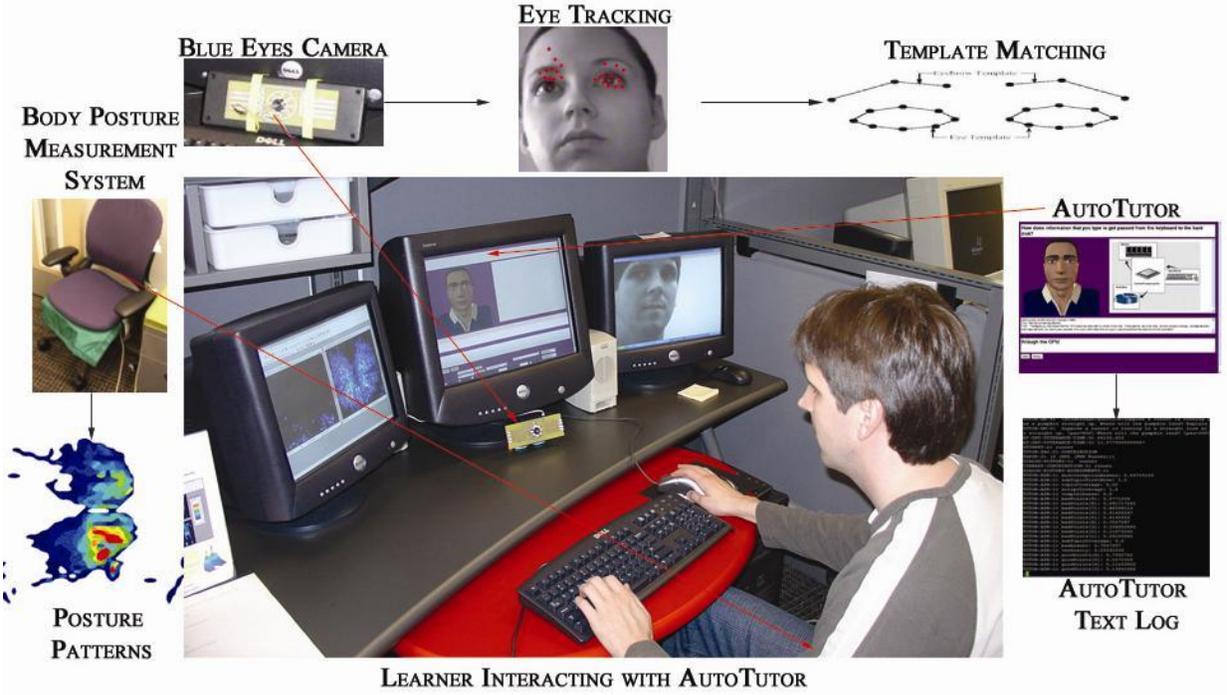


Figure 3.

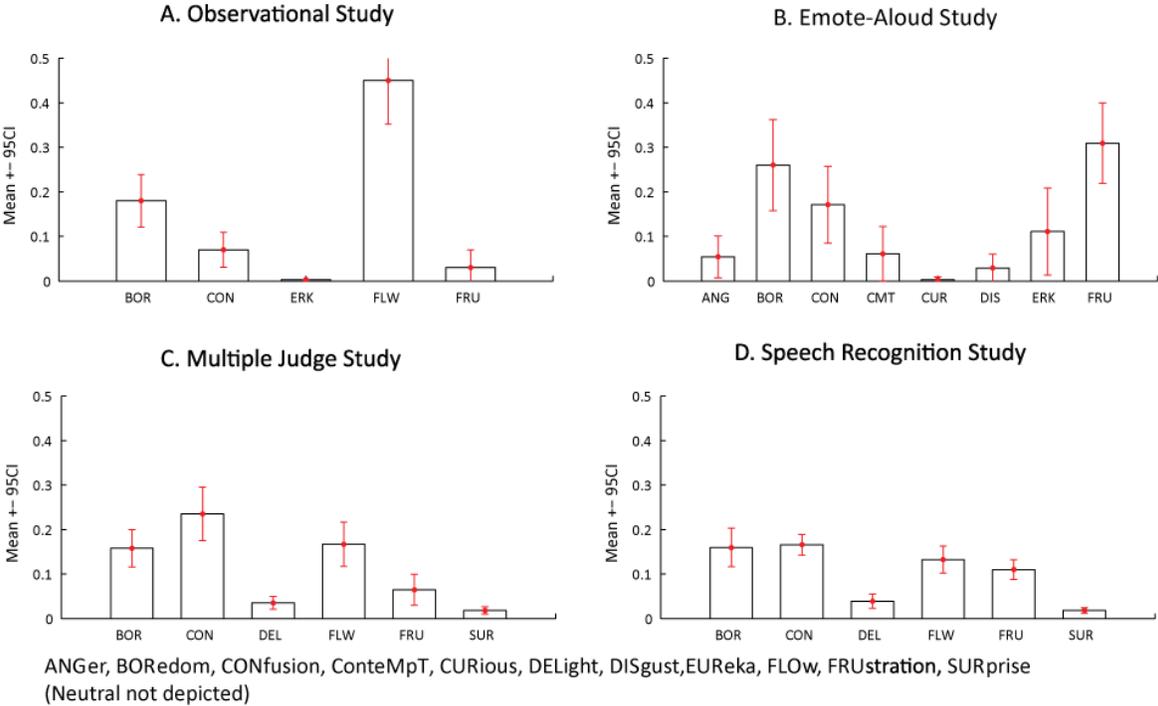


Figure 4.

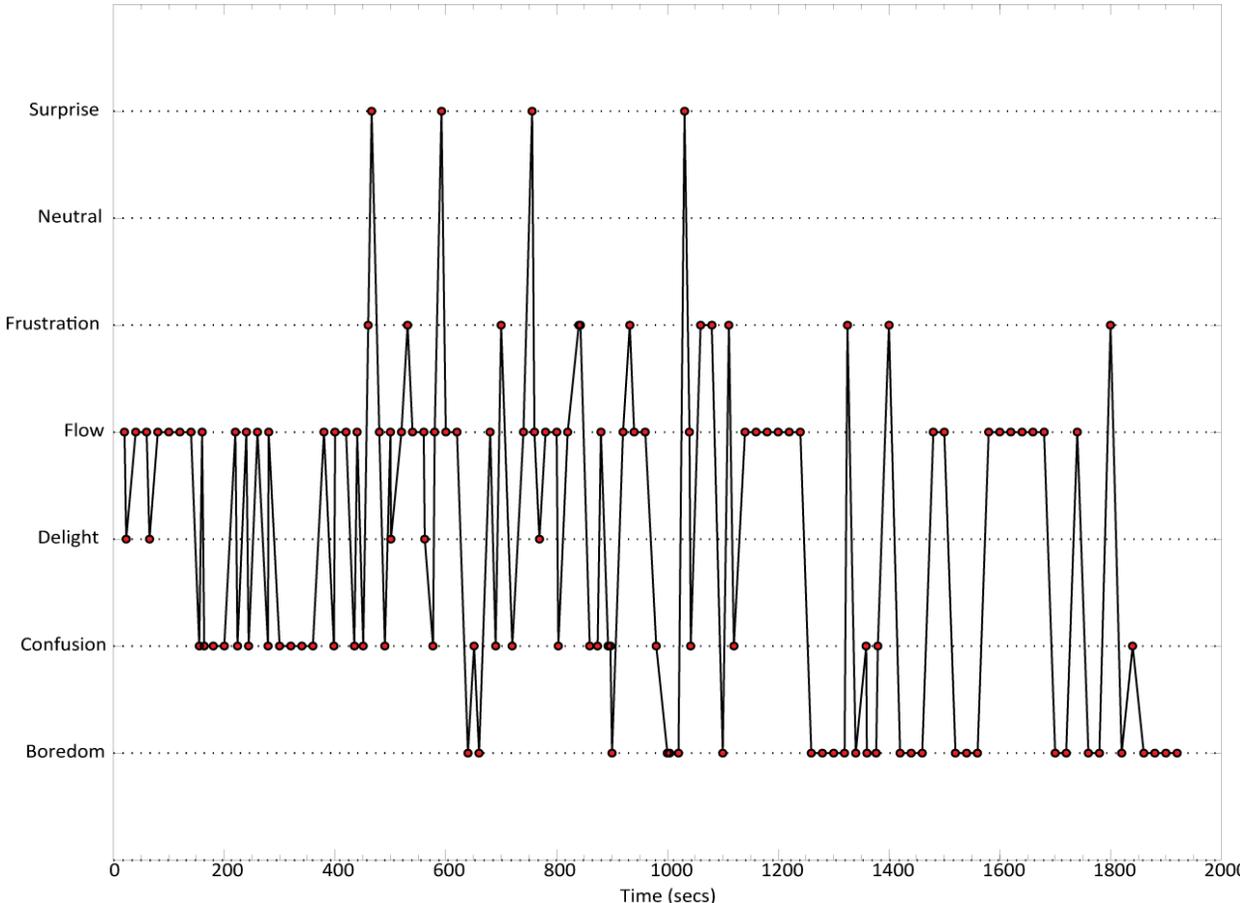


Figure 5.

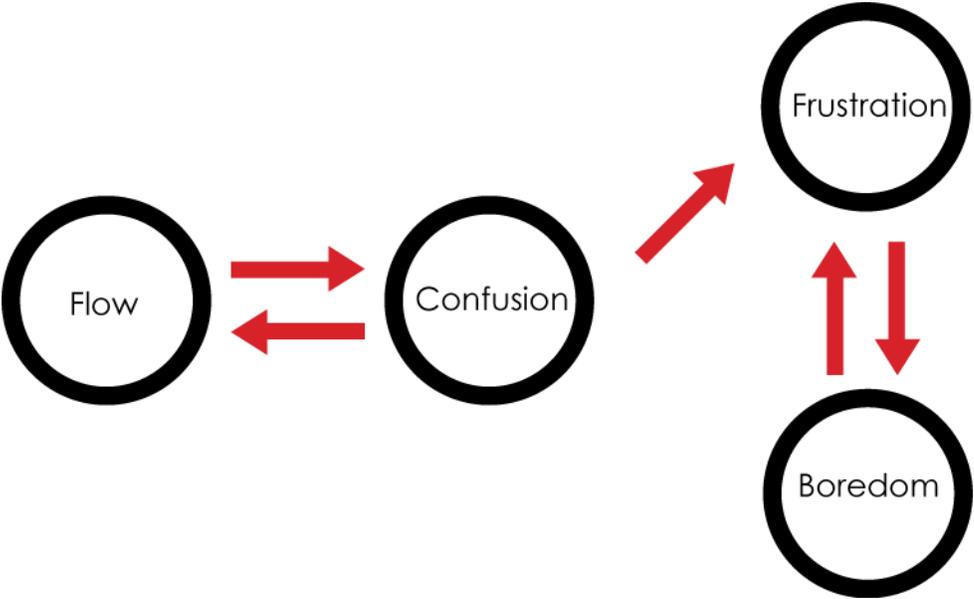


Figure 6.

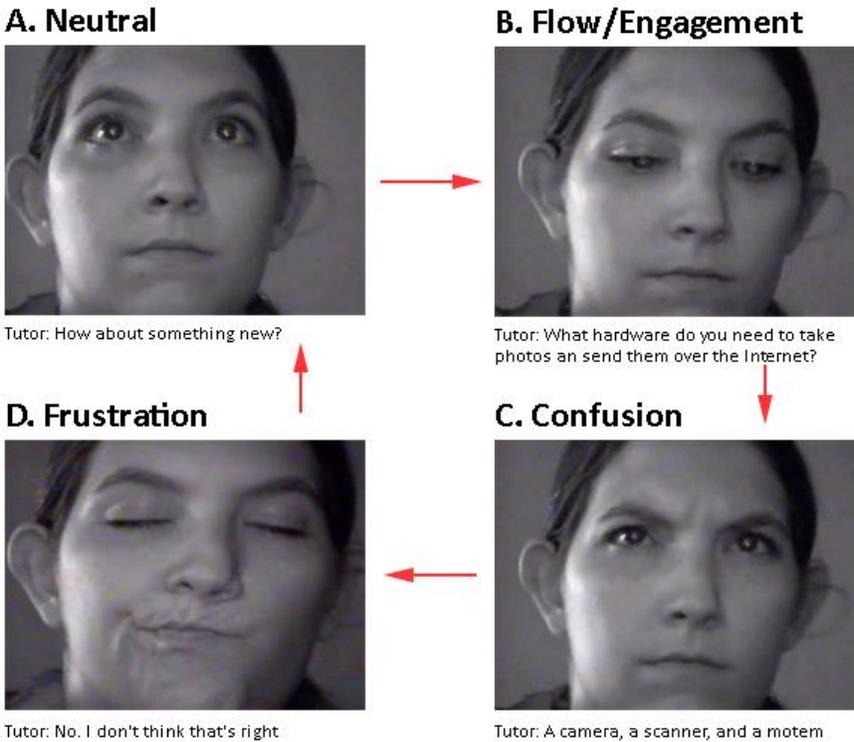


Figure 7.

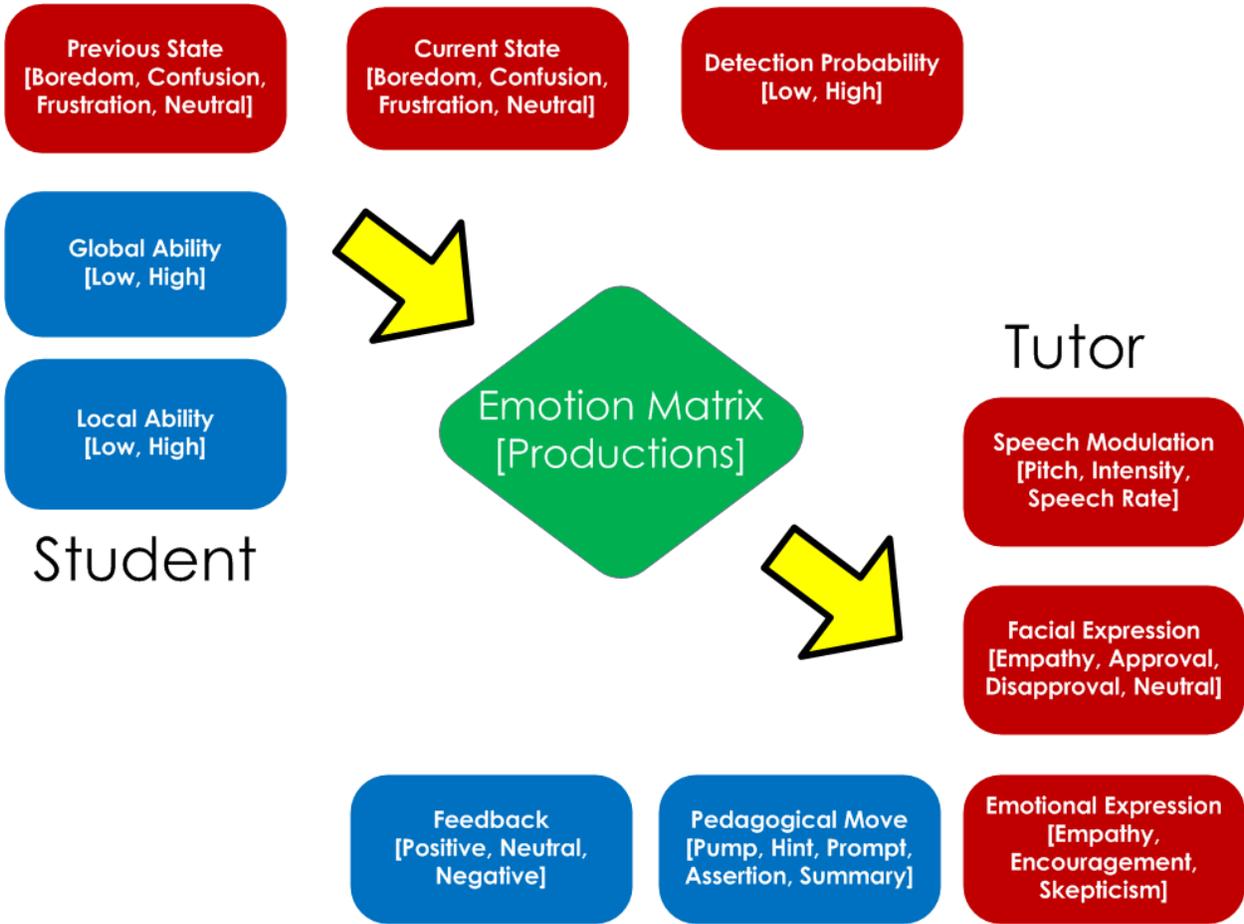


Figure 8.

