Feeling, Thinking, and Computing with Affect-Aware Learning Technologies

Sidney D'Mello[1] and Art Graesser[2]

[1]University of Notre Dame

[2]University of Memphis

Abstract

This chapter discusses some of the exciting research in the nascent field of affect-aware learning technologies (AALTs) – educational technologies that compute affect in addition to cognition, metacognition, and motivation. We begin by positioning AALTs in the complex ecology comprised of diverse phenomena and technologies in the cognitive, affective, learning, and computing sciences. This is followed by an overview of the major findings of a recent meta-analysis aimed at identifying a subset of learning-centered affective states that generalize across students, learning technologies, learning tasks, and experimental methodologies. We then turn our attention to the two major types of AALTs: reactive systems that respond to affect once it occurs and proactive systems that aim to induce or impede certain affective states. Affective AutoTutor, GazeTutor, and UNC-ITSpoke are presented as examples of reactive AALTs, while Crystal Island and ConfusionTutor are included as examples of proactive AALTs. Some of the open issues in the area of AALTs are discussed. These include scalability for real-world contexts, good-enough classification accuracy, and levels of analysis for affective responding. The chapter concludes by outlining two broad avenues of research for the field.

Keywords

affect-aware learning technology, proactive systems, reactive systems, Affective AutoTutor, GazeTutor, UNC-ITSpoke, Crystal Island, ConfusionTutor.

Glossary

Affect-aware learning technology – an intelligent learning technology that considers a learner's affective and cognitive states in its pedagogical decision making.

Proactive systems – an affect-aware learning technology that aims to induce or impede certain affective states.

Reactive systems – an affect-aware learning technology that senses and responds to affective states.

Affective AutoTutor - a natural language intelligent tutoring system that automatically senses and responds to a learner's confusion, boredom, and frustration by monitoring facial features, body movements, and conversational cues.

GazeTutor - a learning environment that senses and responds to patterns of disengagement by monitoring eye gaze.

UNC-ITSpoke - a speech-enabled intelligent tutoring system that automatically senses and responds to a learner's uncertainty and response accuracy.

Crystal Island - a learning technology that embeds the learning content in a narrative-centered game that supports narrativity, realism, and immersion.

ConfusionTutor – a learning environment that aims to promote deeper comprehension by strategically induces confusion in the minds of learners.

1. Introduction

About 15 years ago, the idea of an “affect-aware learning technology”, or the fusion of “affect” and “learning”, was a misnomer at best, and was considered to be as odd of a pairing as “affect” and “computing.” “What on earth does affect have to do with learning?” would have been a pertinent question at that time because learning, characterized as the acquisition of knowledge and the construction of meaning, was considered to be in the exclusive realm of cognition. Affect, at best, was considered to play a negligible role in learning. At worst, it was considered to be harmful to learning, as documented by the approximately 1,000 studies on test anxiety that emerged over the last few decades (Zeidner, 2007).

Much has changed over the last decade, of course. Just as we know that affect and computing have a natural symbiosis that is played out in the field of affect computing, so do affect and learning. This recent emphasis on affect in the learning sciences is paralleled by a similar renaissance in the cognitive sciences, where it is now widely believed that affect and cognition are inextricably coupled and that affect is both served-by and services cognition (Clore & Huntsinger, 2007; Mandler, 1999; Ortony et al., 1988; Schwarz, 2012).

It is easy to make a case for the relevance of affect to cognition in general and specifically to learning. Affect performs *signaling functions* (Schwarz, 2012) by pointing out problems with knowledge (confusion), problems with motivation (boredom), concerns with impending performance (anxiety), and challenges that cannot be easily surpassed (frustration). It performs *evaluative functions* by appraising events in terms of their value, goal relevance, and goal congruence (Izard, 2010). Affect also perform *modulation functions* by constraining or expanding cognitive focus as is the case when negative emotions engender narrow, bottom-up, and focused modes of processing (constrained focus) (Barth & Funke, 2010; Schwarz, 2012) compared to positive emotions that facilitate broader, top-down, generative processing (expanded focus) (Fredrickson & Branigan, 2005; Isen, 2008). Indeed, affect pervades cognition as evident by its effects on memory, problem solving, decision making, and other facets of cognition (see Clore & Huntsinger, 2007 for a review). As such, affect, along with motivation and cognition, has finally claimed its rightful place as a core component of learning (Snow et al., 1996).

The last 15 years has also witnessed a wider-scale adoption of *advanced learning technologies* (ALTs) that serve educational goals (Koedinger et al., 1997; Woolf, 2009). Quite different from the computer-based training systems (CBTs) of the 70s and 80s, 21st century ALTs are an order of magnitude more sophisticated in how they interact with students, how they model student knowledge, how they implement macro- and micro-adaptive strategies to dynamically tailor the instruction to individual students, and how they support motivation, metacognition, and self-regulated learning. Some broad categories of these ALTs include intelligent tutoring systems (VanLehn, 2011; Woolf, 2009), educational games (Dickey, 2005; Gee, 2003), and simulations, animations, multimedia and hypermedia environments (Ainsworth, 2008; Johnson et al., 2000; Mayer, 2005).

ALTSs have come a long way towards modeling students' knowledge levels and cognitive states. However, the inextricable link between affect and learning suggests that next-generation ALTs should be more than mere cognitive machines. They should be affective processors as well. In line with this, the last decade has witnessed the emergence of a few ALTs that incorporate some form of affective modeling in addition to the traditional cognitive modeling performed by these systems (Conati & Maclaren, 2009; D'Mello et al., 2010; D'Mello et al., 2012; Forbes-Riley & Litman, 2011b; Sabourin et al., 2011; Woolf et al., 2009). We refer to these systems as *affect-aware advanced learning technologies* (AALTs). Although AALTs can take on several forms in terms of focus and scope, for the purpose of simplicity, we broadly categorize them as: (a) systems that either induce or impede particular affective states (proactive systems) and (b) systems that respond to specific affective states as they arise (reactive systems).

The purpose of this chapter is to discuss some of the core themes, advances, and open issues in the emerging field of AALTs. We set the stage by positioning AALTs within the complex landscape of affect, learning, and technology. Next, we briefly review the results of a recent meta-analysis aimed at uncovering the affective states that occur during learning with technology. The next two sections provide examples of reactive and AALTS, respectively. Important challenges and open issues in this field are discussed followed by some broad avenues of research that are ripe for exploration.

2. Finding a Niche for AALTs within Affect, Learning, and Technology

Positioning AALTs within the landscape of affect, learning, and technology is no simple task. Figure 1 provides a simplified organization of the constructs, systems, and research in the relevant fields of the affective, cognitive, learning, and computing science as one attempt to find a niche for AALTs in this complex landscape. At the first level, one encounters different forms of learning: shallow, complex, and procedural learning. *Shallow learning* consists of basic encoding of information into memory, such as memorizing a list of words or a definition of a concept. *Procedural learning*, sometimes called implicit learning, concerns the acquisition of perceptual-motor (e.g., playing a musical instrument) and cognitive (e.g., mathematical operations) skills through repeated and/or deliberate practice (Ericsson et al., 1993). Finally, *complex learning* (aka deep or conceptual learning) involves comprehending conceptual information, such as a mathematics proof, a legal document, or a difficult physics principle, or acquiring a complex skill, such as learning computer programming or argumentative writing. Complex learning can be distinguished from shallow learning and procedural learning in that it requires more complex cognitive processes, such as inference generation, problem solving, conceptual caparisons, and explanation generation, in addition to more basic cognitive processes (memory encoding, attentional orientation, rehearsal, tuning of perceptual-motor systems).

Affect plays a role in all three types of learning activities. It influences shallow learning activities by biasing memory encoding and retrieval as documented by the mood-congruent memory encoding (e.g., happy people are more likely to remember happy words than sad words) and retrieval (e.g., happy words are more likely to be retrieved when in a happy mood compared to a sad mood) effects (Blaney, 1986; Bower, 1981). Affective arousal has also been shown to play a facilitative role in some procedural learning tasks (Steidl et al., 2011). However, in our view, affect is most influential in complex learning activities which require considerable concentration, effortful problem solving, and conceptual meaning making, all activities that inevitably give rise to a range of complex affective reactions. For example, cognitive processes that underlie complex learning tasks (e.g., reading Shakespeare or learning about dynamic programming) are accompanied by affective processes such as confusion, frustration, irritation, anger, and sometimes rage when the learner makes mistakes, struggles with troublesome impasses, and experiences failure. On the other hand, positive affective states such as flow, delight, excitement, and eureka are experienced when tasks are completed, challenges are conquered, insights are unveiled, and major discoveries are made.

In addition to affect, complex learning requires thinking and reasoning (cognition), thinking about thinking (metacognition), and motivation (Snow et al., 1996). Although affective processes interact with cognitive, motivational, and metacognitive processes in meaningful and interesting ways, our lens is currently on affect, knowing fully well that affect must ultimately be studied within the context of these other processes.

Affect is a quite a general term that encompasses multiple systems at multiple levels at multiple time scales (Izard, 2010). One possible organization involves distinguishing between affective traits, background moods, and affective states (Rosenberg, 1998). Affective traits are relatively stable, mostly unconscious, predispositions towards particular affective experiences. They operate by lowering the threshold for experiencing certain emotions (i.e. hostile people have a lower threshold for experiencing anger but not necessarily other negative emotions). Moods also perform a threshold reduction function on emotion activation, but are considered to be more transitory and have a background influence on consciousness. In contrast to affective traits and moods, affective states (or emotions) are brief, intense, reactions that occupy the forefront of consciousness, have significant physiological and behavioral manifestations, and rapidly prepare the bodily systems for action. The present emphasis is on affective states, which also have variable durations when analyzed at a finer-grained time scale (D'Mello & Graesser, 2011) States, such as engagement and boredom, can last for tens of seconds to a few minutes, while others (e.g., delight, surprise) are more ephemeral. Nevertheless, affective states are more transitory than moods, which can last for several hours to days (Rosenberg, 1998; Watson & Clark, 1994).

Researchers in the affective sciences have proposed a number of taxonomies to categorize the affective states that occur in everyday experiences (Ekman, 1992; Ortony et al., 1988; Plutchik, 2001). Broadly, the emotions can be divided into *basic* and *non-basic* emotions. According to Ekman (1992) basic emotions are considered to be (a) universal, (b) innate, (c) cross-cultural, (d) have unique elicitation contexts, (e) have distinguishable neural, physiological, and behavioral correlations, (f) show coherence among different response systems, (g) have a quick onset and brief duration, and (h) automatically occur . Emotions such as anger, surprise, happiness, disgust, sadness, and fear typically make the list of basic emotions (Ekman, 1992). Affective states such as boredom, confusion, frustration, engagement, and curiosity share some, but not all, of the features commonly attributed to basic emotions. Consequently, these emotions are labeled as *non-basic* states. The term "affective states" is being used broadly in this chapter to include both bona fide emotions (e.g., anger) and cognitive-affective blends such as engagement and confusion. As discussed in the next section, non-basic affective states appear to be more relevant during interactions with ALTs, so most AALTs focus on modeling these non-basic states.

AALTs that model non-basic affect can adopt many forms based on the strategies used for intelligent handling of learner affect. One simple categorization broadly distinguishes between *proactive* and *reactive* systems. At the most basic level, proactive systems aspire to increase the likelihood that the learner will experience affective states deemed beneficial to learning (e.g., interest, curiosity, engagement), while simultaneously decreasing the likelihood of states believed to negatively influence the process and product of learning (e.g., boredom, frustration). Reactive systems make no notable *a priori* attempt to up-regulate or down-regulate positive and negative affect, respectively. Instead, they simply detect and respond to affective states as they arise. Reactive systems typically focus on identifying and responding to negative affective states such as frustration and boredom.

3. Which Affective States are Relevant to Learning with Technology?

An AALT can never model learner affect if it does not know which affective states to model. Hence, there is the foundational question of identifying the affective states that learners experience during interactions with ALTs? D'Mello (in press)recently completed a meta-analysis of 24 studies that have systematically tracked the affective states that naturally occur during interactions with learning technologies. The results of this analysis are summarized in this section.

A total of 1740 students in middle school, high school, college, and adult students from the United States, Canada, the United Kingdom, Philippines, and Australia were represented in the 24 studies. On average, each student spent 45 minutes interacting with a learning technology, so the total aggregate training time was approximately 76,000 minutes. There was considerable variance in learning contexts, topics, and technologies. Learning contexts included classroom (i.e., computer lab in a school), research lab, and online studies. Learning topics (subject domains) included algebra, analytical reasoning, argumentative writing, chemistry, computer, ecology, genetics, geography, graphing, logic puzzles, microbiology, pre-algebra, and social studies. Learning technologies included intelligent tutoring systems, serious games and simulations environments, virtual labs, and interfaces for problem solving, reading comprehension, and essay writing. In all, a total of 17 affective states, consisting of achievement, epistemic, and topic emotions, and encompassing both basic and non-basic emotions, were tracked with a number of methodologies. These methods included online self-reports, emote-aloud protocols, online observations, and retrospective coding of video after a learning session by the students themselves, or by peers, observers, or trained judges.

The analyzes focused on the magnitude and significance of affect incidence across studies (formally quantified with weighted mean effect sizes – see (D'Mello, in press)). Engagement/flow was the most frequent state, while contempt, anger, disgust, sadness, anxiety, delight, fear, and surprise were relatively infrequent. There was considerably between-study variability in the occurrence of boredom, confusion, curiosity, happiness, and frustration. However, these states, particularly boredom and confusion, were more likely to occur than the eight infrequent states listed above. The meta-analysis also revealed that the incidence of the affective states was moderated by the source of the affect judgments (self vs. observers) and the authenticity of the learning context (classroom vs. laboratory).

The take-home message was that engagement/flow, boredom, confusion, curiosity, happiness, and frustration were the major affective states exhibited by learners in these 24 studies. These five states are good candidates for modeling in AALTs. Furthermore, with the exception of happiness, the basic emotions had a relatively minor role to play. This is not entirely surprising as they is inadequate theoretical justification to expect notable levels of some of the basic emotions, such as disgust and fear, during short one-on-one learning sessions with computers.

4. Reactive Affect-Aware Learning Technologies

Reactive AALTs focus on automatically *detecting* student affect and *responding* to the sensed affect in order to up-regulate positive affective states (e.g., curiosity, engagement) and down-regulate negative affective states (e.g., frustration). Fully-automated affect detection can use predictive models which infer student affect by analyzing the context of the interaction and other relevant cues (Baker et al., 2012; Conati & Maclaren, 2009; Sabourin et al., 2011) and/or diagnostic models which sense affect from communicative channels like facial features, speech, postures, gestures, central and peripheral physiology, and textual responses (Calvo & D’Mello, 2010; Chaouachi & Frasson, 2010; Pour et al., 2010). Once the learner’s affective state is detected with reasonable accuracy, a reactive AALT must dynamically alter its pedagogical strategies in response to the detected state. An AALT has a number of paths to pursue once it has detected learner affect. It could do nothing if the learner is engaged and is on a positive learning trajectory. Hints and just-in-time explanations can be provided when confusion or frustration is detected. The system could provide choice, encourage breaks, or adjust levels of challenge when it detects that a student is bored. Some of the implemented responses to student affect include affect mirroring (Burleson & Picard, 2007), empathetic responses (Woolf et al., 2010), and a combination of politeness, empathy, and encouragement (D'Mello et al., 2010).

The ideal affect-detection and affect-response strategies are ultimately tied to aspects of the global and local situational context. The global context refers to stable, non-malleable factors, such as the specific ALT, domain, student, and interaction context. The local context pertains to unstable factors that can be manipulated to regulate student affect. These include the current topic or question, system messages, system feedback, etc. In this section, we briefly discuss three systems as examples of contemporary reactive AALTs. These include the Affective AutoTutor, GazeTutor, and UNC-ITSpoke.

4.1. Affective AutoTutor: Responding to Boredom, Confusion, and Frustration

The Affective Tutor was perhaps the first reactive AALT that was developed from 2004 to 2010. AutoTutor is a conversational intelligent tutoring system that helps students develop mastery on difficult topics in Newtonian physics, computer literacy, and scientific reasoning by holding a mixed-initiative dialog in natural language (Graesser et al., 2005; Graesser et al., 2012). AutoTutor has a set of fuzzy production rules that are sensitive to the cognitive, but not to the affective states of the learner. The Affective AutoTutor augments these rules with the ability to map dynamic assessments of learners' affective and cognitive states with tutor actions to address the presence of boredom, confusion, and frustration (D'Mello et al., 2008b; D'Mello et al., 2010; D'Mello & Graesser, in press).

The Affective AutoTutor automatically detects boredom, confusion, frustration, and neutral by monitoring conversational cues and other discourse features (predictive modeling), along with gross body language, and facial features (diagnostic modeling – see Figure 2A) (D'Mello & Graesser, 2010; D'Mello & Graesser, in press). Each channel independently provides its own evaluation of the learner's affective state. These individual diagnoses are combined with a decision-level fusion algorithm that selects a single affective state and a confidence value of the detection (see Figure 2B). The algorithm relies on a voting rule enhanced with a few simple heuristics.

The Affective AutoTutor's production rules address the presence of boredom, confusion, and frustration by incorporating perspectives from a number of psychological theories, including attribution theory (Weiner, 1986), cognitive disequilibrium during learning (Piaget, 1952), politeness (Brown & Levinson, 1987), and empathy (Lepper & Chabay, 1988), along with recommendations of expert tutor's. The tutor responds with empathetic, encouraging, and motivational dialog-moves and emotional displays. For example, the tutor 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". These affective responses are accompanied by an appropriate emotional facial expression and emotionally modulated speech (see Figure 2C).

We tested the effectiveness of the Affective Tutor in improving learning over the non-affective AutoTutor in a controlled experiment where 84 learners completed two 30-minute training sessions with either tutor (D'Mello et al., 2010). The results indicated that the Affective tutor helped learning for low-domain knowledge learners during the second 30-minute learning session. The Affective tutor was less effective at promoting learning for high-domain knowledge learners during the first 30-minute session. Importantly, learning gains increased from Session 1 to Session 2 with the Affective tutor whereas they plateaued with the non-affective tutor (see Figure 2D). Learners who interacted with the Affective Tutor also demonstrated higher performance on subsequent transfer tests. A follow-up analysis into learners' perceptions of both tutors indicated that their perceptions of how closely the computer tutors resembled human tutors increased across learning sessions, was related to the quality of tutor feedback, and was a powerful predictor of learning (D'Mello & Graesser, 2012). Interestingly, the positive change in perceptions was greater for the Affective tutor.

4.2. GazeTutor: Responding to Boredom and Disengagement

It is generally acknowledged that engagement is an important requirement for learning. Attention to task-related thoughts at the expense of task-irrelevant thoughts is one critical precursor of engagement in a learning activity. Therefore, developing interventions that monitor periods of waning attention and attempt to encourage more productive use of *attentional resources* might be one promising way to increase engagement and promote learning.

We tested this claim with a multimedia interface consisting of an animated conversational agent that provided explanations on biology concepts with synthesized speech that was synchronized with annotated images (see Figure 3A) (D'Mello et al., 2012). The system used a Tobii T60 eye tracker in order to identify when learners were not attending to the important parts of the interface (i.e., the tutor or image). The interface did not directly track off-task thought; it simply assumed that learners were disengaged when their gaze was not on the tutor or image for at least five consecutive seconds. It attempted to reengage learners with statements that directed them to reorient their attention towards the animated pedagogical agent or the image (e.g., “Please pay attention”).

We evaluated the efficacy of the gaze-reactive tutor in promoting learning, motivation, and engagement in an experiment where 48 learners were tutored on four biology topics: two with the gaze-reactive component enabled (experimental condition) and two with the gaze-reactive component disabled (control condition). Learners completed a posttest on all four topics after interacting with the system. The results indicated that the gaze-sensitive intervention was successful in dynamically reorienting learners’ attention patterns to the important areas of the interface (see Figure 3B). Prior to the intervention, the probability that students were looking away from the screen steadily increased while there was a corresponding decrease in focus on the tutor and the image. A reverse pattern was discovered after receiving an intervention message; off-screen gaze behaviors rapidly decreased, while focus on the tutor steadily increased. Importantly, posttest scores for deep reasoning questions were higher when learners interacted with the gaze-sensitive interface compared to its non-reactive counterpart. Interestingly, individual differences in scholastic aptitude moderated the impact of gaze-reactivity on learning gains. Gaze-reactivity was associated with a small improvement in overall learning for learners with average aptitude, but learning gains were substantially higher for learners with high aptitude and somewhat lower for their low-aptitude counterparts.

In summary, the results of this preliminary study suggest that gaze-sensitive statements can reorient attention, and thereby improve comprehension and learning. Future enhancements of the system include replacing the expensive eye-tracker with web-cameras, more fine-grained modeling of disengagement behaviors, a larger repertoire of context-specific gaze-reactive dialogs, and incorporating individual differences in selecting appropriate gaze-sensitive responses.

4.3. UNC-ITSpoke

Forbes-Riley and Litman have developed an AALT called UNC-ITSPOKE to examine whether automatic responses to learner uncertainty could improve learning outcomes (Forbes-Riley & Litman, 2007, 2009; Forbes-Riley & Litman, 2011b). Uncertainty is a state that is similar to confusion and plays an important role on the process and products of learning. ITSPOKE is a speech-enabled ALT that teaches learners about various physics topics with spoken dialogs that are automatically recognized with the Sphinx 2 Speech Recognizer (Litman et al., 2006). UNC-ITSPOKE extends the basic functionality of ITSPOKE with the capability to automatically detect and respond to learners' certainty/uncertainty in addition to correctness/incorrectness. Uncertainty detection is performed by extracting and analyzing the acoustic-prosodic features in the learners' spoken responses in conjunction with lexical features and dialog-based features (similar to contextual cues in Affective AutoTutor).

It is beyond the scope of this chapter to delve deeply into the adaptive strategies to resolve uncertainty except to point out that this involved launching explanation-based sub dialogs when the student was correct in their response but uncertain about the response. This was taken to signal an impasse (see next Section for more details) because the students is unsure about the state of their knowledge despite being correct. The efficacy of the adaptive condition was first verified in a pilot Wizard-of-Oz study where an unseen human (the wizard) performed all of the difficult natural language processing tasks (speech recognition, response understanding, and uncertainty annotation) (Forbes-Riley & Litman, 2010).

In a recent study, Forbes-Riley and Litman (2011b) compared learning outcomes between learners who received adaptive responses to uncertainty (adaptive condition), random responses to uncertainty (random condition), or no responses (control condition). The adaptive condition achieved slightly (but not significantly) higher learning outcomes than the random and control conditions. The findings revealed that it was perhaps not the presence or absence of adaptive responses to uncertainty, but *how many* adaptive responses were given that correlated with learning performance. Unfortunately, the biggest challenge was caused by errors in automatic uncertainty annotation, which reduced the number of opportunities for adaptive responses to uncertainty. Thus, although the findings were somewhat mixed, Forbes-Riley and Litman (2011b) conclude that there is merit in offering adaptive feedback to uncertainty, and that such feedback can improve learning outcomes.

5. Proactive Affect-Aware Learning Technologies

In contrast to reactive AALTs that primarily act once particular affective states have been detected, proactive AALTs have inherent strategies that are specifically designed to engender particular affective responses. One common approach is to implement some of the competitive and motivational features of games in educational environments (Halpern et al., 2012; Jackson et al., in press). These educational games are hypothesized to be the ideal learning environments that are capable of turning work into play by minimizing boredom, optimizing engagement/flow, presenting challenges that reside within the optimal zone of proximal development, preventing persistent frustration, and engineering delight and pleasant surprises (Lepper & Henderlong, 2000; Ritterfeld et al., 2009). One system, discussed in this chapter is called Crystal Island (Rowe et al., 2009) and is an educational game that aims to promote positive affect by carefully embedding the learning content in a game that support narrativity, realism, and immersion (Dede, 2009; Spires et al., 2011).

Taking a very different approach, the second proactive AALT we describe intentionally *confuses* learners in an attempt to promote deeper modes of inquiry (D'Mello et al., in press). Though this may sound somewhat odd at face value, it is important to note that confusion itself has a rather counterintuitive relationship with learning in that it is positively correlated with learning in some ALTs (Craig et al., 2004; D’Mello & Graesser, 2011; Graesser et al., 2007). This suggests that there might be some merits to strategically inducing confusion at key moments in a learning session.

5.1. Crystal Island

Designing educational games can be quite challenging because game designers must balance a trade-off between game environments that are engaging but tangential to learning, and environments that promote deep learning but fail to foster engagement (Johnson & Mayer, 2010). This balance is nicely achieved in Crystal Island (Rowe et al., 2009; Spires et al., 2011), an immersive educational game that capitalizes on the principle of narrativity. The principle of narrativity posits that learners may benefit most from educational games that weave a narrative theme into the context of the learning environment. Ostensibly, immersive narratives motivate learners to initiate and persist in game play, and to remain engaged in the learning activity, which might contribute to increased learning.

In Crystal Island, the learner takes on the role of a protagonist named Alex who arrives on the island. Alex discovers that members of a research team are falling ill and is charged with identifying the source of an infectious disease. Alex proceeds with the investigation by visiting various areas of the island (dining hall, lab, inferminary, dorm), by interviewing other islanders, and by manipulating objects in the world. Through the generation of questions and hypotheses, and the collection and analysis of data, learners make gradual steps toward diagnosing the cause of the disease.

Evidence from learners' interactions with Crystal Island indicate that it is highly engaging and motivating and can improve learning outcomes (McQuiggan et al., 2008). There is also some evidence to suggest that Crystal Island meets one of its intended goals of increasing positive affect, at least when compared to more traditional ALTs like intelligent tutoring systems (Lester et al., 2011). For example, Sabourin et al. (2011) conducted a study in which 450 middle school students interacted with Crystal Island for 55 minutes. Students self-reported one of seven affective states (anxious, bored, confused, curious, excited, focused, and frustrated) at five-minute intervals during the learning session. Positive affective states, such as focused, curious, and excited were found to be more frequent (55%) than negative affective states like confused, frustrated, anxious, and bored (46%). Importantly , recent findings have suggested that engagement and learning, in the context of Crystal Island, share a positive relationship, so his syetms both helps learning and engagement (Rowe et al., 2010).

5.2. ConfusionTutor - Inducing Confusion via Trialogs with Animated Agents

Though most would consider confusion to be a negative affective state, both in terms of its subjective experience (i.e., most people do not like being confused) and its assumed impact on learning (i.e., intuition suggests that confusion is harmful to learning), there is some correlational evidence that suggests a positive relationship between confusion and learning gains (Craig et al., 2004; D'Mello & Graesser, 2011; Graesser et al., 2007). The question arises whether there is a causal relationship between confusion and learning. To test this question, we developed a proactive AALT (ConfusionTutor) that experimentally induced confusion and measured the consequences on the induced confusion and learning (D'Mello et al., in press).

This research was conducted in the context of *trialogs* (conversations between three agents) very similar to those developed in a serious game called *Operation ARA* (Halpern et al., 2012). ARA teaches scientific research methods and critical thinking skills through a series of game modules, including those with two or more animated pedagogical agents. In the trialogs, a 3-way conversation transpired between the human student, a tutor agent, and a student agent. The tutor-agent was an expert on scientific inquiry whereas the student-agent was a peer of the human learner. A series of research case studies with experiments that have a crucial design flaw with respect to proper scientific methodology was presented by the tutor agent.. For example, one case study described an experiment that tested a new pill that purportedly helps people lose weight, but there was no control group. The goal of the human learner was to identify the flaws and express them in natural language and the tutor agent helps scaffold this goal by guiding the human learner and the peer student agent.

ConfusionTutor attempted to induce confusion by manipulating whether or not the tutor agent and the student agent contradicted each other during the trialog. This was accomplished by expressing points that were incorrect and asking the human learners to intervene, decided which opinion had more scientific merit, and provide explanations for their justifications. The tutor agent expressed a correct assertion and the student agent agreed with the tutor in the *true-true* control condition. In the *true-false* condition, the tutor expressed a correct assertion but the student agent disagreed by expressing an incorrect assertion. In the *false-true* condition it was the student agent who provided the correct assertion and the tutor agent who disagreed. In the *false-false* condition, the tutor agent provided an incorrect assertion and the student agent agreed, so both agents were incorrect in this condition. It should be noted that all misleading information was corrected over the course of the learning session and there was no evidence to suggest that the contradictions themselves had a negative impact on learning (see below).

Confusion was operationally defined as occurring if both (a) the human learners manifested uncertainty/incorrectness in their decisions when asked by the agents (e.g. long pauses, changing decisions), (b) the learner reported being confused when probed, and (c) there were visible displays of confusion on video recordings of the learner's face. The data always satisfied criterion (a) and sometimes satisfied criterion (b). Preliminary analyses have indicated that it also satisfies (c), although we are only beginning to analyze the facial videos.

Interestingly, there was evidence that the induced confusion caused more learning at deeper levels of mastery, as reflected in a delayed test on scientific reasoning. The results indicated that contradictions in the experimental conditions produced higher performance on multiple choice questions that tapped deeper levels of comprehension than performance in the no-contradiction true-true condition, but *only if learners were confused* by the contradictions. Similarly, learners who were confused by the contradictions were more likely to correctly identify flaws in subsequent case studies that differed from the case studies discussed in the trialogs on surface-level features (near-transfer) or both surface and structural features (far-transfer). These data suggest there may be a causal relationship between confusion and deep learning, with confusion playing moderating role on the effect of the contradictions on learning.

6. Open Issues

In this section, we discuss some of the critical issues, challenges, and open problems in the field of AALTs. In the interest of brevity, we focus on three core problems, knowing fully well that there are several more issues that warrant further research and development.

6.1. Sensors, Sensorless, or Sensor-lite? Issues of Scalability

Many schools in developed nations have a computer lab that can be used for data collection and system deployment. However, it is unlikely that any school has a computer lab that is specially equipped with affective sensors (e.g., physiological sensors, eye trackers, and posture sensors) because several sensors require customized, non-portable, and expensive hardware and software. Hence, scalability concerns become paramount when one wants to take a program of research on AALTs into the real world. Some of these issues are expected to be mitigated as sensor technologies develop to the point that non-intrusive wireless sensors can be deployed at little cost in authentic educational contexts. This is a future that is likely, but not guaranteed, to occur, so it might be worth to consider alternate solutions.

One possible solution is to take the sensors to the classroom, an approach that was pioneered by Arroyo et al. (2009). In this study, approximately 70 students were monitored for 4-5 days with four affect sensors (a pressure-sensitive mouse, pressure-sensitive mats for a chair, a video camera, and a bracelet to sense skin-conductance) while they completed their normal classroom activities with a mathematics ALT. This *sensor approach* is certainly a viable approach for short-term studies in school contexts, but is unlikely to be sustainable in the longer term. The state of the art in affect detection using sensors is described in several chapters in this volume (see section on Affect Detection).

In contrast to the sensor approach, some researchers have taken a *sensorless* approach, in which affective states are detected by monitoring interaction patterns (e.g., clicks, responses) and other contextual cues from the ALT (e.g., type of feedback being provided, difficulty of problem). No sensors are used in this approach. We first demonstrated the possibility of this approach in a 2006 study (D'Mello et al., 2006) and subsequently expanded its capabilities in a 2008 study (D'Mello et al., 2008a). These studies were conducted in a laboratory setting because of the nature of our research questions. However, Baker and colleagues have since conducted an impressive program of research that focuses on developing sensorless affect detectors in real-world environments (see Chapter X by Baker and Ocumpaugh). Some of the more recent attempts have focused on tracking engagement during writing activities (Bixler & D'Mello, 2013; Liu et al., in press). Although this sensorless approach certainly is a viable solution to the scalability problem, there is the question of quantifying the extent to which affect detection accuracies are impacted by a lack of sensors.

The approach is a *sensor-lite* approach. This involves including scalable sensors when feasible (cameras and microphones) and replacing non-scalable sensors with scalable proxies, or “soft sensors” or “virtual sensors.” Cameras seem to be the ideal sensor for this purpose as web-cams are already integrated in most laptops and can be purchased at low-cost if needed. For example, we have successfully applied motion tracking techniques to video data, thereby effectively replacing a cost-prohibitive posture sensor with simple web-cams (Kory & D'Mello, in review). Some initial work has also suggested that cameras can be used to monitor heart rate (Poh et al., 2010) and eye gaze (Sewell & Komogortsev, 2010). Hence, in our view, camera-based proxy sensing coupled with interactional and contextual features appears to be the most promising way to solve the scalability problem.

6.2 How Good is Good Enough? Issues of Detection Accuracy

A reactive AALT obviously needs to detect learner affect before it can adaptively respond. Detection of naturalistic affective experiences in non-controlled settings has come a long way in the last decade as evidenced by recent surveys (Calvo & D’Mello, 2010; Zeng et al., 2009) and several chapters on affect detection in this volume. Although numerous advances have been made, the field is still grappling with persistent problems. There is the previously discussed problem of intrusive, expensive, and noisy sensors that are largely unscalable. There are the technical challenges associated with detecting latent psychological constructs (i.e., affect) from weak signals embedded in noisy channels. There are difficulties associated with collecting adequate and realistic training data for machine learning models. There are the challenges of incorporating top-down models of context and appraisals with bottom-up body- and physiological-based sensing. There is the lack of clarity of the affective phenomenon being modeled (e.g., moods vs. emotions, categorical vs. dimensional representations). Finally, there are issues pertaining to generalizability across contexts, time, individual differences, and cultural differences.

In general, affect detection is an extremely difficult problem and it is unlikely that *perfect* affect detectors that generalize to new individuals and interactional contexts and that can operate under messy real-world conditions can ever be developed. One can spend a long time waiting for perfect or almost perfect affect detection accuracies before closing the loop by developing affect-aware interventions. In our view, a moderate degree of recognition accuracy is sufficient provided the affect-aware interventions are fail-soft in that they do no harm if delivered incorrectly. Possible fail-soft interventions include never directly acknowledging a learners' affective state, adjusting parameters of the task that are beyond the radar of the learners (e.g., decreasing problem difficulty), and providing cognitive scaffolds in the form of hints and explanations. The severity of the interventions can also be calibrated to the system's confidence in detection accuracy.

6.3 How Adaptive is Adaptive Enough? Issues of Levels of Analysis

Assuming that an AALT can detect a learner's affective state with moderate accuracy, the next pertinent question pertains to how to adaptively respond to that state. This is where there is perhaps the most significant paucity of guiding theory and empirical research. We have discussed some of the strategies that have been implemented in the section on reactive AALTs, but these only scratch the surface of possible affect-sensitive responses. Although there are several critical issues that must be considered in designing an affective response strategy, here we focus on one such issue, namely *levels of analysis*.

The expression "levels of analysis" pertains to the specificity by which an affective state needs to be modeled for a meaningful affect-sensitive response. Current approaches primarily model affect in two ways. Some model arousal and valence independently, jointly, and in conjunction with other relevant dimensions. Others take a categorical view and model discrete affective states, sometimes with their intensity. It is easy to see that simply modeling just the presence of an affective state (either dimensionally or categorically) is unlikely to suffice. Take, for example, the case of boredom. According to the control-value theory of achievement emotions, subjective appraisals of control and value of a learning activity are the critical predictors of boredom (Pekrun, 2010). Subjective control pertains to the perceived influence that a learner has over the activity, while subjective value represents the perceived value of the outcomes of the activity. Boredom is expected to be heightened when learners perceive low value in the outcome of the activity, whereas learners are more engaged when the perceived value is high (Pekrun et al., 2010). The situation is more complex for perceived control because there appears to be a curvilinear relationship between perceived control and boredom. Boredom is frequent when control is too high (i.e., skill outweighs challenges) (Csikszentmihalyi, 1990) but also when control is too low (challenges outweigh skill) (Pekrun et al., 2010). This counterintuitive finding can be interpreted from emerging theoretical perspectives that posit that boredom is not unitary; there are multiple types of boredom which occur in different situations and which impact performance in unique ways (Acee et al., 2010; Forbes-Riley & Litman, 2011a; Vodanovich et al., 2005). Similar arguments can be made for several other affective states. This suggests that affective response strategies must be differentially sensitive to different types of affect, which requires modeling of the immediate situational context surrounding the experience of an affective state in addition to detecting the affect state itself.

7. Some Suggested Broad Steps Forward

This chapter surveyed some of the burgeoning research in the field of affect-aware learning technologies. Although affect has claimed its rightful place as a critical component to learning, there is still much more to be done. In the remainder of this chapter, we highlight two possible avenues of work, with an emphasis on more big-picture advances of this field.

First, there is the pressing need for the development and advancement of a science of non-basic emotions. Researchers in the field of affective sciences are currently engaged in a vigorous debate with respect to the existence of basic emotions and the value of a crisp basic vs. non-basic distinction (Lench et al., 2013; Lindquist et al., 2013). Despite the outcome of this debate, if there ever will be an outcome, it is quite clear that much of the emotion research has still emphasized the basic emotions at the expense of overlooking other non-basic emotions (Rozin & Cohen, 2003). This is particularly unfortunate for the field of AALTs, because available data indicates that the basic emotions are mostly muted during interactions with ALTs. Instead, it is the non-basic affective states, such as confusion, frustration, and boredom, that play a more critical role during interactions with technology. Unfortunately, we know very little about these non-basic states and they are of little interest to the mainstream affect sciences community. For example, efforts to promote interest in studying confusion in the affective sciences (Rozin & Cohen, 2003) were met by prompt resistance from some in the affective science community (Ellsworth, 2003; Hess, 2003; Keltner & Shiota, 2003). The success of the field of AALTs relies on a strong foundation of theory and empirical data, so there is a pressing need for a science of non-basic affect (D'Mello & Calvo, in press).

Second, there is a need to extend the scope of our theories, experiments, and models to be sensitive to the situational contexts in which learning unfolds. This is because the affect-cognition- technology relationship does not exist in a vacuum, but is situated in a complex ecology with multiple layers of influences and interactions. At the individual level, motivation and metacognition interact with affect and cognition in any meaningful learning context. For example, a learner with mastery-oriented motivational tendencies (Daniels et al., 2009), who is an academic risk taker (Clifford, 1988; Meyer & Turner, 2006) and who has a high degree of conscientiousness or persistence or grit (Duckworth et al., 2007) would be able to handle difficulty task, failure, and the resultant negative emotions than one who is performance-oriented, a cautious learner, and easily gives up. Beyond the processes that unfold within the mind of an individual learner, learning itself is situated in an external learning context involving a learner completing a learning task with a learning technology. Different affective profiles are expected to emerge as a function of what the learner brings to the task and what the task and technology afford the learner. For example, a learning environment designed to teach problem solving strategies for an upcoming high stakes standardized test is expected to induce a different affective profile than Crystal Island or AutoTutor. The social environment in which the learner, learning technology, and learning task are situated is also expected to influence learners' affect. For example, social emotions like pride and jealousy are likely to be experienced more profoundly in group learning or other socially-relevant learning contexts, while they are somewhat more muted during one-on-one human-computer learning interactions.

This discussion has only sketched out some of the influences that are likely to influence a learner's affective states. There are numerous other factors that play a role so there is much more theory development and empirical testing that needs to be done. Indeed, a program of research aimed at uncovering the factors that influence affect so as to inform the design of affect-aware technologies can be sustainable and rewarding for several decades.

Acknowledgements

This research was supported by the National Science Foundation (NSF) (ITR 0325428, HCC 0834847, DRL 1235958). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF.

References

Acee, T. W., Kim, H., Kim, H. J., Kim, J. I., Chu, H. N. R., Kim, M., Cho, Y. J., & Wicker, F. W. (2010). Academic boredom in under-and over-challenging situations. *Contemporary Educational Psychology, 35*(1), 17-27.

Ainsworth, S. (2008). How do animations influence learning? In D. Robinson & G. Schraw (Eds.), *Current Perspectives on Cognition, Learning, and Instruction: Recent Innovations in Educational Technology that Facilitate Student Learning* (pp. 37-67): Information Age Publishing.

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 the 14th International Conference on Artificial Intelligence In Education* (pp. 17-24). Amsterdam: IOS Press.

Baker, R. S. J., Kalka, J., Aleven, V., Rossi, L., Gowda, S. M., Wagner, A. Z., Kusbit, G. W., Wixon, M., Salvi, A., & Ocumpaugh, J. (2012). Towards sensor-free affect detection in cognitive tutor algebra. In K. Yacef, O. Zaïane, H. Hershkovitz, M. Yudelson & J. Stamper (Eds.), *Proceedings of the 5th International Conference on Educational Data Mining* (pp. 126-133): International Educational Data Mining Society.

Barth, C. M., & Funke, J. (2010). Negative affective environments improve complex solving performance. *Cognition and Emotion, 24*(7), 1259-1268. doi: 10.1080/02699930903223766

Bixler, R., & D'Mello, S. (2013). Detecting engagement and boredom during writing with keystroke analysis, task appraisals, and stable traits *Proceedings of the 2013 International Conference on Intelligent User Interfaces (IUI 2013)* (pp. 225-234). New York, NY: ACM.

Blaney, P. H. (1986). Affect and memory: a review. *Psychological Bulletin, 99*(2), 229.

Bower, G. (1981). Mood and memory. *American Psychologist, 36*, 129-148.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

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.

Calvo, R. A., & D’Mello, S. K. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1*(1), 18-37. doi: 10.1109/T-AFFC.2010.1

Chaouachi, M., & Frasson, C. (2010). Exploring the relationship between learners EEG mental engagement and affect. In J. Kay & V. Aleven (Eds.), *Proceedings of 10th International Conference on Intelligent Tutoring Systems* (pp. 291-293). Berlin / Heidelberg: Springer.

Clifford, M. (1988). Failure tolerance and academic risk-taking in ten- to twelve-year-old students. *British Journal of Educational Psychology, 58*(15-27). doi: 10.1111/j.2044-8279.1988.tb00875.x

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393-399. doi: 10.1016/j.tics.2007.08.005

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. doi: 10.1080/1358165042000283101

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

D'Mello, S. (in press). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*.

D'Mello, S., & Calvo, R. (in press). Beyond the Basic Emotions: What Should Affective Computing Compute? *Extended Abstracts of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2013)*. New York, NY: ACM.

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. (2008a). 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. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-adapted Interaction, 20*(2), 147-187.

D'Mello, S., & Graesser, A. (2012). Malleability of students' perceptions of an affect-sensitive tutor and its influence on learning. In G. Youngblood & P. Mccarthy (Eds.), *Proceedings of 25th Florida Artificial Intelligence Research Society Conference* (pp. 432-437). Menlo Park, CA: AAAI Press.

D'Mello, S., Jackson, G., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., El Kaliouby, R., Picard, R., & Graesser, A. (2008b). *AutoTutor detects and responds to learners affective and cognitive states.* Paper presented at the Proceedings of 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., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., Perkins, L., & Graesser, A. (2010). A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning. In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 245-254). Berlin / Heidelberg: Springer.

D'Mello, S., Lehman, S., Pekrun, R., & Graesser, A. (in press). Confusion can be beneficial for learning. *Learning and Instruction*.

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of human-computer studies, 70*(5), 377-398.

D'Mello, S. K., & Graesser, A. C. (in press). AutoTutor and Affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*.

D'Mello, S., & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion, 25*(7), 1299-1308.

Daniels, L. M., Pekrun, R., Stupnisky, R. H., Haynes, T. L., Perry, R. P., & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology, 101*(4), 948-963. doi: 10.1037/a0016096

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(5910), 66-69.

Dickey, M. D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development, 53*, 67-93.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology, 92*(6), 1087.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*(3-4), 169-200.

Ellsworth, P. C. (2003). Confusion, concentration, and other emotions of interest: Commentary on Rozin and Cohen (2003). *2003, 3*(1), 81-85.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363.

Forbes-Riley, K., & Litman, D. (2007). *Investigating Human Tutor Responses to Student Uncertainty for Adaptive System Development*. Paper presented at the Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction, Lisbon, Portugal.

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., & Litman, D. (2010). Designing and evaluating a wizarded uncertainty-adaptive spoken dialogue tutoring system. *Computer Speech and Language, 25*(1), 105-126. doi: http://dx.doi.org/10.1016/j.csl.2009.12.002

Forbes-Riley, K., & Litman, D. (2011a). When does disengagement correlate with learning in spoken dialog computer tutoring?,. In S. Bull & G. Biswas (Eds.), *Proceedings of the*

15th International Conference on Artificial Intelligence in Education (pp. 81-89). Berlin/Heidelberg: Springer.

Forbes-Riley, K., & Litman, D. J. (2011b). Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor. *Speech Communication, 53*(9-10), 1115-1136. doi: 10.1016/j.specom.2011.02.006

Fredrickson, B., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion, 19*(3), 313-332. doi: 10.1080/02699930441000238

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

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. doi: 10.1109/TE.2005.856149

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.), *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 569-571). Amsterdam: IOS Press.

Graesser, A. C., D’Mello, S. K., X., H., Cai, Z., Olney, A., & Morgan, B. (2012). AutoTutor. In P. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing: Identification, investigation, and resolution* (pp. 169-187). Hershey, PA: IGI Global.

Halpern, D. F., Millis, K., Graesser, A., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity, 7*(93-100).

Hess, U. (2003). Now you see it, now you don’t - the confusing case of confusion as an emotion: Commentary on Rozin and Cohen (2003). *Emotion, 3*(1), 76-80.

Isen, A. (2008). Some ways in which positive affect influences decision making and problem solving. In M. Lewis, J. Haviland-Jones & L. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 548-573). New York, NY: Guilford.

Izard, C. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review*, *2*(4), 363-370. doi: 10.1177/1754073910374661

Jackson, G. T., Dempsey, K. B., & McNamara, D. S. (in press). Game-based practice in reading strategy tutoring system: Showdown in iSTART-ME. In H. Reinders (Ed.), *Computer games* (pp. 59-68). Bristol, UK: Multilingual Matters.

Johnson, C. I., & Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human behavior, 26*(6), 1246-1252.

Johnson, W., Rickel, J., & Lester, J. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 11*, 47-78.

Keltner, D., & Shiota, M. (2003). New displays and new emotions: A commentary on Rozin and Cohen (2003). *Emotion, 3*, 86-91. doi: 10.1037/1528-3542.3.1.86

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.

Kory, J., & D'Mello, S. (in review). Motion Tracker: Cost-effective, non-intrusive, fully-automated monitoring of bodily movements using motion silhouettes.

Lench, H. C., Bench, S. W., & Flores, S. A. (2013). Searching for evidence, not a war: Reply to Lindquist, Siegel, Quigley, and Barrett (2013). *Psychological Bulletin*, *113*(1), 264-268.

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. 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.

Lester, J. C., McQuiggan, S. W., & Sabourin, J. L. (2011). Affect recognition and expression in narrative-centered learning environments. In R. Calvo & S. D'Mello (Eds.), *New Perspectives on Affect and Learning Technologies* (pp. 85-96). New York: Springer.

Lindquist, K. A., Siegel, E. H., Quigley, K. S., & Barrett, L. F. (2013). The Hundred-Year Emotion War: Are Emotions Natural Kinds or Psychological Constructions? Comment on Lench, Flores, and Bench (2011). *Psychological Bulletin, 139*(1), 264-268.

Litman, D., Rose, C., Forbes-Riley, K., VanLehn, K., Bhembe, D., & Silliman, S. (2006). Spoken versus typed human and computer dialogue tutoring. *International Journal of Artificial Intelligence In Education, 16*(2), 145-170.

Liu, M., Calvo, R., & Pardo, A. (in press). *Tracer: A tool to measure and visualize student engagment in writing activities.* Paper presented at the Proceedings of the13th IEEE International Conference on Advanced Learning Technologies, Beijing, China

Mandler, G. (1999). Emotion. In B. M. Bly & D. E. Rumelhart (Eds.), *Cognitive science. Handbook of perception and cognition* (2nd ed., pp. 367-382). San Diego, CA: Academic Press.

Mayer, R. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Meyer, D., & Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review, 18*(4), 377-390. doi: 10.1007/s10648-006-9032-1

Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of emotions*. New York: Cambridge University Press.

Pekrun, R. (2010). Academic emotions. In T. Urdan (Ed.), *APA educational psychology handbook* (Vol. 2). Washington, DC: American Psychological Association.

Pekrun, R., Goetz, T., Daniels, L., Stupnisky, R. H., & Perry, R. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology, 102*(3), 531-549. doi: 10.1037/a0019243

Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.

Plutchik, R. (2001). The nature of emotions. *American Scientist, 89*(4), 344-350.

Poh, M. Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express, 18*(10), 10762-10774.

Pour, P. A., Hussein, S., AlZoubi, O., D'Mello, S., & Calvo, R. (2010). The impact of system feedback on learners' affective and physiological states. In J. Kay & V. Aleven (Eds.), *Proceedings of 10th International Conference on Intelligent Tutoring Systems* (pp. 264-273). Berlin / Heidelberg: Springer-Verlag.

Ritterfeld, U., Cody, M., & Vorderer, P. (Eds.). (2009). *Serious games: Mechanisms and effects*. New York and London: Routledge.

Rosenberg, E. (1998). Levels of analysis and the organization of affect. *Review of General Psychology, 2*(3), 247-270. doi: 10.1037//1089-2680.2.3.247

Rowe, J., Mott, B., McQuiggan, S., Robison, J., Lee, S., & Lester, J. (2009). *Crystal island: A narrative-centered learning environment for eighth grade microbiology.* Paper presented at the Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education, Brighton, UK.

Rowe, J., Shores, L., Mott, B., & Lester, J. (2010). Integrating learning and engagement in narrative-centered learning environments. In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 166-177). Berlin / Heidelberg: Springer.

Rozin, P., & Cohen, A. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of maturally occurring facial expressions of Americans. *Emotion, 3*, 68-75.

Sabourin, J., Mott, B., & Lester, J. (2011). Modeling learner affect with theoretically grounded dynamic bayesian networks In S. D'Mello, A. Graesser, B. Schuller & J. Martin (Eds.), *Proceedings of the Fourth International Conference on Affective Computing and Intelligent Interaction* (pp. 286-295). Berlin Heidelberg: Springer-Verlag.

Schwarz, N. (2012). Feelings-as-Information Theory. In P. Van Lange, A. Kruglanski & T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 289-308). Thousand Oaks, CA: Sage.

Sewell, W., & Komogortsev, O. (2010). *Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network.* Paper presented at the Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, Austin, TX.

Snow, R. E., Corno, L., & Jackson, D. (1996). Individual differences in affective and conative functions. In D. Berliner & R. Caldee (Eds.), *Handbook of educational psychology* (pp. 243–310). New York: Macmillan.

Spires, H. A., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). Problem solving and game-based learning: Effects of middle grade students' hypothesis testing strategies on learning outcomes. *Journal of Educational Computing Research*, *44*(4), 453-472.

Steidl, S., Razik, F., & Anderson, A. K. (2011). Emotion enhanced retention of cognitive skill learning. *Emotion*, *11*(1), 12.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197-221.

Vodanovich, S. J., Wallace, J. C., & Kass, S. J. (2005). A confirmatory approach to the factor structure of the Boredom Proneness Scale: Evidence for a two-factor short form. *Journal of Personality Assessment*, *85*(3), 295-303.

Watson, D., & Clark, L. A. (1994). Emotions, moods, traits, and temperaments: Conceptual distinctions and empirical findings. In P. Ekman & J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 89-93). New York: Oxford 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.

Woolf, B., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., & Christopherson, R. (2010). The effect of motivational learning companions on low achieving students and students with disabilities In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 327-337). Berlin / Heidelberg: Springer.

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: Recognizing and responding to student affect. *International Journal of Learning Technology, 4*(3/4), 129-163.

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In P. Schutz & R. Pekrun (Eds.), *Emotions in Education* (pp. 165-184). San Diego, CA: Academic Press.

Zeng, Z., Pantic, M., Roisman, G., & Huang, T. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31*(1), 39-58.

Figure Captions

Figure 1. Niche for AALTs in ecology of learning, affect, and technology

Figure 2. Stages in the design and validation of the Affective AutoTutor

Figure 3. Gaze-Tutor interface and gaze-reorienting results

Figure 1.

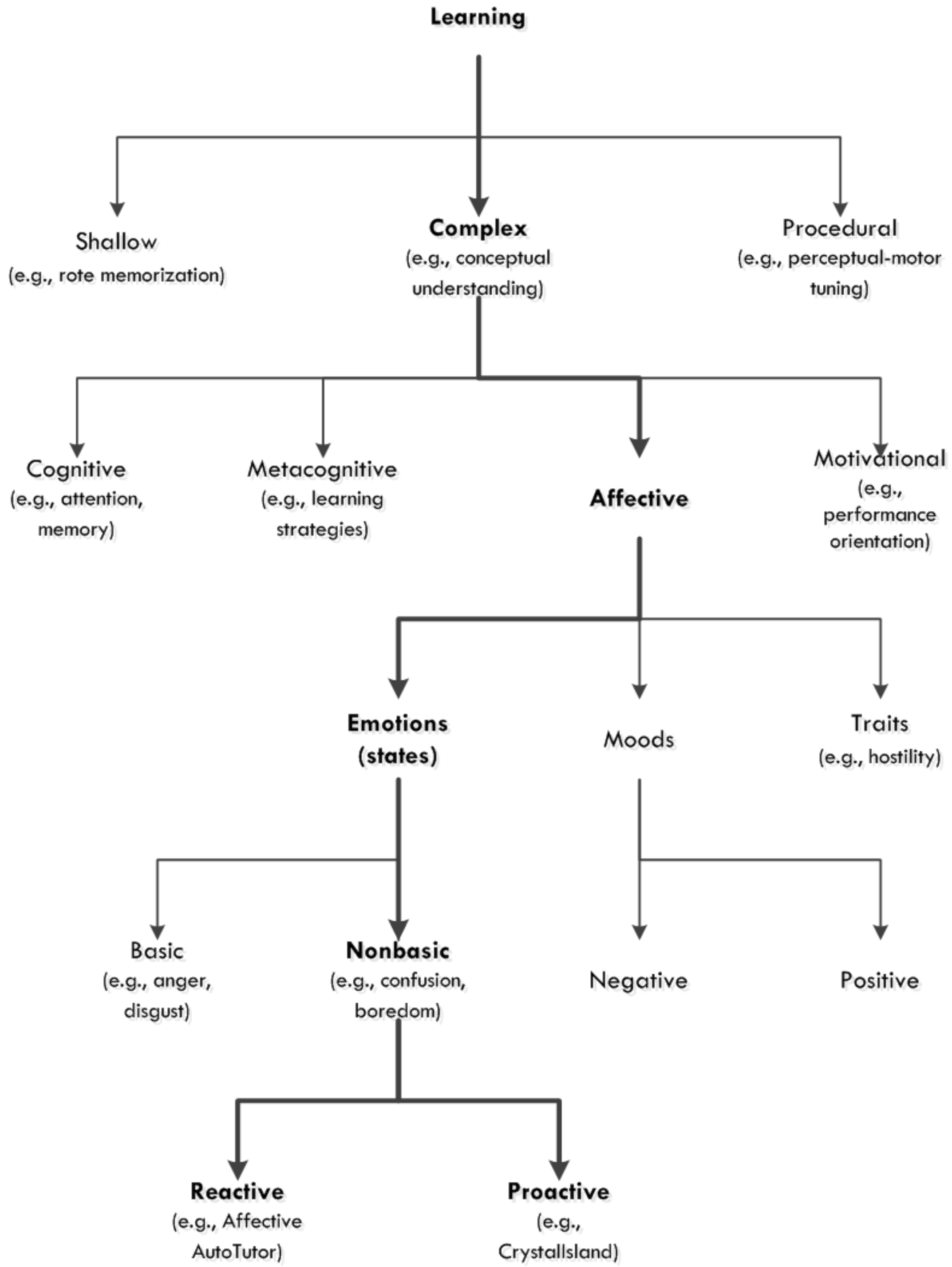
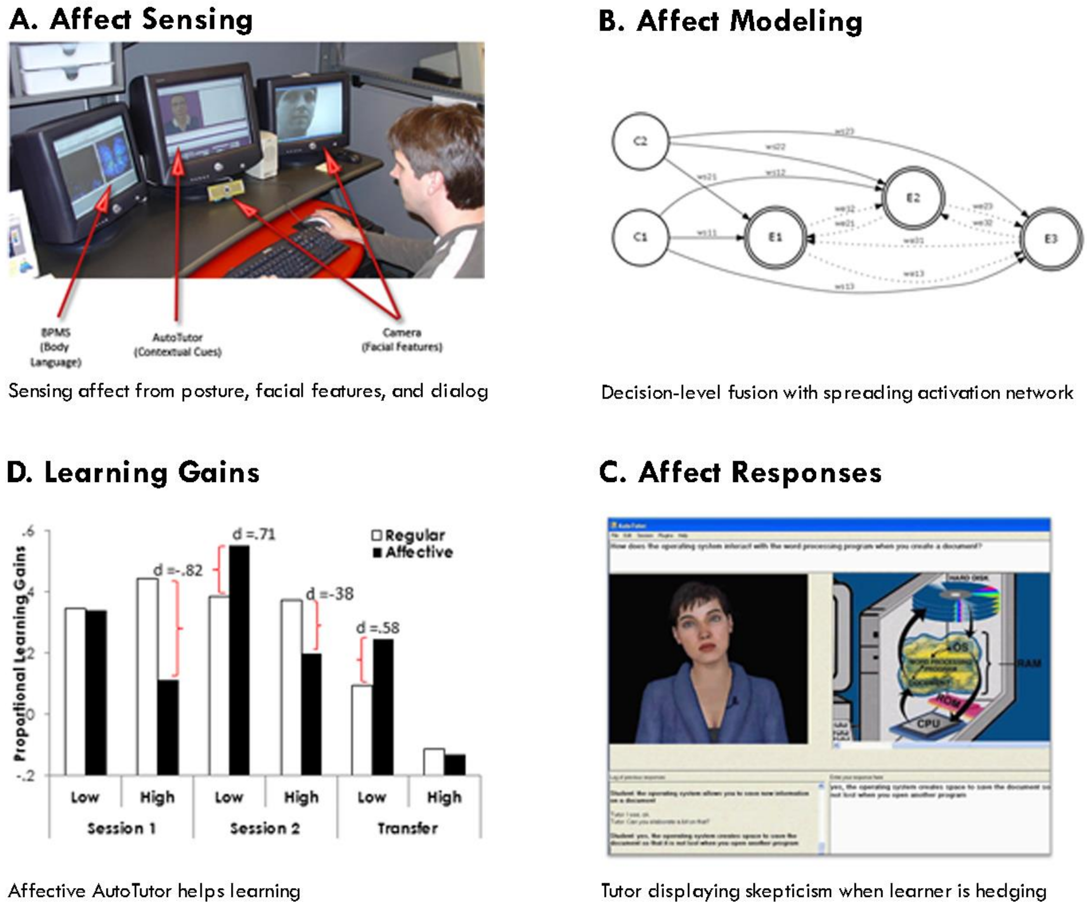
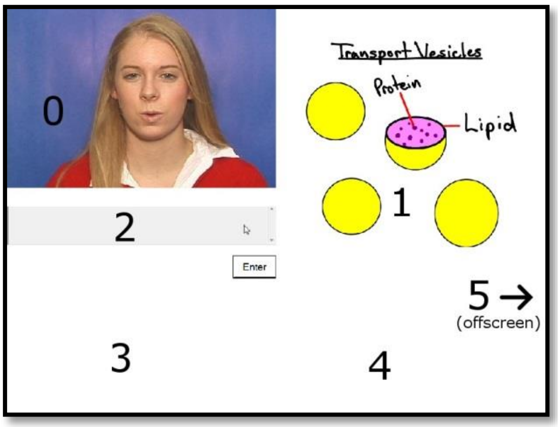

Learning
Shallow
(e.g., rote memorization)
Complex
(e.g., conceptual understanding)
Procedural
(e.g., perceptual-motor tuning)
Cognitive
(e.g., attention, memory)
Metacognitive
(e.g., learning strategies)
Affective
Motivational
(e.g., performance orientation)
Emotions (states)
Moods
Traits
(e.g., hostility)
Basic
(e.g., anger, disgust)
Nonbasic
(e.g., confusion, boredom)
Negative
Positive
Reactive
(e.g., Affective AutoTutor)
Proactive
(e.g., CrystalIsland)

Figure 2

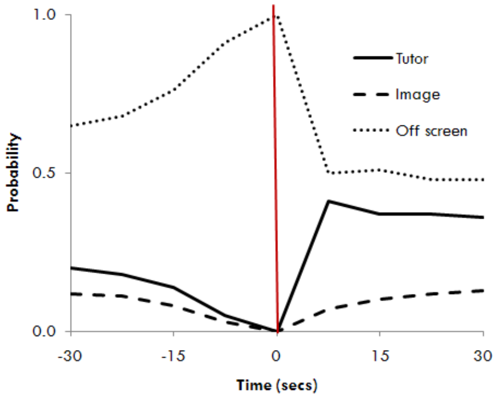


A. Affect Sensing
BPMS (Body Language)
AutoTutor (Contextual Cues)
Camera (Facial Features)
Sensing affect from posture, facial features, and dialog
B. Affect Modeling
C2
C1
E1
E2
E3
Decision-level fusion with spreading activation network
D. Learning Gains
Proportional Learning Gains
.6
.4
.2
0
-.2
d =-.82
d =.71
d =-38
d =.58
Regular
Affective
Low
High
Session 1
Session 2
Transfer
Affective AutoTutor helps learning
C. Affect Responses
HARD DISK
OS
RAM
CPU
Tutor displaying skepticism when learner is hedging

Figure 3



A. Screen shot of interface



B. Gaze before and after intervention