

# Affect-Aware Tutors: Recognizing and Responding to Student Affect

Beverly Woolf,<sup>1</sup> Winslow Burleson,<sup>2</sup> Ivon Arroyo,<sup>1</sup>  
Toby Dragon,<sup>1</sup> David Cooper,<sup>1</sup> Rosalind Picard<sup>3</sup>

<sup>1</sup>Department of Computer Science, University of Massachusetts Amherst

<sup>2</sup>School of Computer Science and Informatics / Arts, Media and Engineering  
Arizona State University

<sup>3</sup>M.I.T. Media Laboratory,

## Abstract

Theories and technologies are needed to understand and integrate knowledge of student affect (e.g., frustration, motivation, and self-confidence) into models of learning. One goal of this research is to integrate tools that model student affect into intelligent tutors and enable tutors to elicit, sense, communicate, measure and respond to student affect. This article presents multiple solutions towards this goal and discusses systems that begin to redress the cognitive vs. affective imbalance in teaching. We describe our broad approach, how we will evaluate the impact of affect intervention on student learning and three main development objectives: i) affect recognition, ii) interventions in response to student affect, and iii) emotionally animated agents.

## I Vision and challenge

Intelligent tutors provide individualized teaching in multiple domains and demonstrate learning gains similar to or greater than those provided by human tutors (Fletcher, 1996; Koedinger et al., 1997; Shute and Psothka, 1995). However, much previous research has tended to privilege the cognitive over the affective in which theories of learning view thinking and learning as information processing, marginalizing or ignoring affect (Picard et al., 2004). If computers are to interact naturally with humans, they must recognize affect and express social competencies. However, the role of affect in instruction is at best in its infancy (Picard et al., 2004). One obvious next frontier in computational instruction is to systematically examine the relationship(s) between student affective state and learning outcomes (Shute, 2008).

This research looks at the role of student affect in learning and the role that new technology plays in recognizing and responding to affect. We describe research to measure, model, study, and support the affective dimension of learning in ways that were not previously possible. Our systems are used in the classroom with the goal of improving learning and perseverance. Affective interventions encourage learning, lessen student humiliation and provide support and motivation that outweighs or distracts from the unpleasant aspects of failure. This research is based on efforts at the University of Massachusetts, Arizona State University and the MIT Media Lab.

The next section describes theories of affect, learning and human emotion. We look at the constellation of student behaviors that we label as emotion and provide a brief overview of how computers can recognize and respond to student affect. The next three sections describe three approaches to affect recognition (human observation, hardware sensors and machine learning

techniques), several interventions that respond to a student's cognitive-affective state and two sets of emotional embodied pedagogical agents. The final two sections provide a discussion and a view of future work.

### **1.1 Theories of affect, learning and human emotion**

Modeling student emotion has become increasingly important for computational teaching systems and emotion has been named as one of the twelve major challenges for cognitive science (Norman, 1981). Human emotion is often defined as an intuitive feeling derived from one's circumstance, mood or relation with others. Teachers have long recognized the central role of emotion in learning and the extent to which emotional upsets can interfere with mental life. Student interest and active participation are important factors in the learning process (e.g. Bransford et al., 2000). Students learn less well if they are anxious, angry, or depressed; students who are caught in these states do not take in information efficiently or deal with it well (Burlison and Picard, 2004; Picard et al., 2004; Goleman, 1995). Teachers often devote as much time to the achievement of students' motivational goals as to their cognitive and informational goals in one-to-one human tutoring situations (Lepper and Hodell, 1989). Emotions can paralyze a student's ability to retain information about a task (Baddeley, 1986). Several studies have addressed emotions involved in learning (e.g. Lepper and Chabay, 1988; Mandler, 1984; Kort et al, 2001).

Human emotion is completely intertwined with cognition in guiding rational behavior, including memory and decision-making. The human brain is a system in which emotion and cognitive functions are inextricably integrated (Cytowic, 1989). Emotional skills have been shown to be more influential than cognitive abilities for personal, career and scholastic success (Goleman, 1996). For instance, in the comparison of impulsivity and verbal IQ as predictors of future delinquent behavior, impulsivity was twice as powerful a predictor (Block, 1995). Recent findings suggest that when basic mechanisms of emotion are missing, intelligent functioning is hindered.

Nearly a hundred definitions of emotion have been categorized as of 1981 (Kleinginna and Kleinginna, 1981). Yet no comprehensive, validated, theory of emotion exists that addresses learning, explains which emotions are most important in learning, or identifies how emotion influences learning (Picard et al., 2004). Most studies of emotion do not include the phenomena observed in natural learning situations, such as interest, boredom, or surprise. Rather, emotion definitions emphasize cognitive and information processing aspects and encode them into machine-based rules used in learning interaction, e.g., OCC model of emotion (Ortony et al., 1988).

Acceptance of ideas about emotion in learning is based largely on intuition and generalized references to constructivist theorists (Piaget and Inhelder 1969; Vygotsky, 1962, 1978). These theories discuss how to motivate, engage, and assist students in a general way. Yet, they do not provide descriptions at the level of individual human-to-human interactions and clearly do not provide methods suitable for implementation in intelligent tutors.

Motivation is one emotion strongly linked to learning and has been defined as a person's direction, intensity and persistence in an activity. Students with high intrinsic motivation often outperform students with low intrinsic motivation. A slight positive approach by a student is often accompanied by a tendency toward greater creativity and flexibility in problem solving, as well as more efficiency and thoroughness in decision-making (Isen, 2000). If student motivation is sustained throughout periods of disengagement, students might persevere to a greater extent through frustration (Burlison and Picard, 2004).

Studies of motivation in learning consider the role of intrinsic versus extrinsic influences, self-efficacy, students' beliefs about their efficacy, the influence of pleasurable past learning experiences, feelings of contributing to something that matters and the importance of having an

audience that cares, among other factors (Vroom, 1964; Keller, 1983; 1987; Ames, 1992; Vail, 1994; Bandura, 1977; Pajares, 1996; Schunk, 1989; Zimmerman, 2000). Theories of motivation are often built around affective and cognitive components of goal directed behavior (e.g. Dweck, 1986, 1999; Dweck and Leggett, 1988).

Flow, or optimal experience is often defined as a feeling of being in control, concentrated and highly focused, enjoying an activity for its own sake, or a match between the challenge at hand and one's skills (Csikszentmihalyi, 1990). In direct contrast Stuck, or a state of non-optimal experience, is characterized by elements of negative affect and includes a feeling of being out of control, a lack of concentration, inability to maintain focused attention, mental fatigue and distress (Burlison and Picard, 2004). The phenomenon of “negative asymmetry” or the staying power of negative affect, which tends to outweigh the more transient experience of positive affect, is also an important component of learning and motivation (Giuseppe & Brass, 2003).

The concept of affect is often distinguished from that of emotion. Affect refers to an observed emotional state or biological response to an external stimuli. Incorporating affect within human-human interactions is very powerful. In their research on “thin slices,” Ambady and Rosenthal demonstrated that when participants were shown a short segment of video, as little as six seconds of a teacher’s first interactions with a student, they could predict that teacher’s effectiveness and student end-of-term grades (Ambady and Rosenthal, 1992). Wentzel has shown that caring bonds between middle school children and their teachers are predictive of learners’ performance (1997).

Some researchers have raised the concern that one cannot begin to measure or respond to emotion until a clear theory of emotion is articulated. However, even without a fully-fledged theory of emotion, computers can be given some ability to recognize and respond to affect (Picard et al., 2004). In fact, research shows that efforts to build models of a less understood phenomenon will aid in improving the understanding of that very phenomenon (Picard et al., 2004). Thus we simultaneously engage in both the practice and the theory directly related to developing affect-aware tutors in an attempt to advance both.

Ekman’s Categorization	Cognitive-Affective Term	Emotion Scale
Joy	High pleasure	“I am enjoying this.”
	Low pleasure	“This is not fun.”
Anger	Frustration	“I am very frustrated.”
	Low-frustration	“I am not frustrated at all.”
Surprise	Novelty	“I am very hooked.”
	Boredom	“I am bored.”
Fear	Anxiety	“I feel anxious”
	Confidence	“I feel very confident”

Table 1. Cognitive-affective terms based on human face studies (Ekman et al., 1972; Ekman 1999)

## 1.2 Categorization of emotion

We identify a subset of emotions that we intend to recognize in student behavior and for which the tutor will provide interventions during learning. This selection of emotion is based on both cognitive and affective analyses. We begin with Paul Ekman's categorization of emotions based on analyses of facial expressions that includes joy, anger, surprise, fear, disgust/contempt, and surprise (Ekman et al., 1972; Ekman, 1999). However, we realize that these emotions are appropriate for general-purpose description and are not specific to learning. Emotions referred to by students and teachers in a learning environment tend take on a slightly different flavor.

To address this, we added a cognitive component to Ekman’s categorization that is present in educational settings, thus initiating what we call “cognitive-affective” terms.

For each of Ekman’s emotion we created a scale, resulting in four orthogonal bipolar axes of cognitive-affect, Table 1. For example, given Ekman's fear category, the proposed scale is: “I feel anxious . . . I feel very confident.” Note that some of these emotions express a similar essence only at opposite ends of the spectrum (such as joy and surprise --the essence is to be low/high in spirits). Since disgust/contempt do not arise frequently in everyday learning settings, we decided not to use these categories.

### 1.3 Recognizing and responding to student affect through computers

The approach of this research is to evaluate learning in classrooms while students work with intelligent tutors and to develop models of student affect along with tools that recognize affect and generate pedagogical interventions. Students are often faced with difficult tasks within computer tutoring situations, tasks which might at times accelerate failure or increase the fear of failure. Recognition of student affect in these situations helps researchers tease apart the learner’s cognitive and affective states and improve tutor intervention. One long-term goal is to help students develop *meta-cognitive* and *meta-affective* skills, such as self-awareness and self-regulations for dealing with failure and frustration (Azevedo and Cromley, 2004; Burleson and Picard, 2004; Dweck, 1999).

Cognitive Clues	Affective Clues	Tutor intervention based on inference about a student’s state
Student makes an error	Student appears curious and focused	<i>No intervention needed;</i> Student is engaged in learning and exploration (Flow)
	Student is frowning, fidgeting, and looking around	<i>Alternate actions are needed;</i> Student is confused (Stuck)
Student has not made much progress	Evidence of stress, fidgeting, high valence and arousal	<i>Alternate actions are needed;</i> Student is under stress (Stuck)
	Evidence of boredom and confusion	<i>Interventions using off-task activities are needed;</i> Student is not engaged (Stuck)
	Student is not frustrated	<i>No intervention needed;</i> Student is curious and involved in exploration (Flow)
Student is solving problems correctly	Student is not frustrated	<i>No intervention needed;</i> Student is in control, concentrated and focused (Flow)

**Table 2: Case studies of students’ cognitive-affective mechanisms**

Prior research shows that student affect (e.g., frustration or boredom) can be detected within intelligent tutoring systems (McQuiggan and Lester, 2006; Graesser et al., 2007; D’Mello et al., 2007). Our research extends this state-of-the-art by dynamically collecting cognitive and affective information within classrooms, detecting a need for interventions and determining which interventions are most successful for individual students and contexts (problem, affective state).

The tutoring system we use responds to students’ cognitive and affective states, see Table 2. If, for example, a student has not exhibited progress in terms of the task, yet sensors indicate that curiosity and exploration (elements of Flow) are at play and related elements of Stuck are not present, the tutor will not intervene; rather will allow the student to further explore the task.

One central focus of this research is to generate a framework for long-term pedagogical decision-making. Affect recognition can significantly improve a tutor’s long-term planning, e.g., when the tutor allows a student to remain frustrated in the short term. Observing a learner continuously, as a skilled mentor or tutor might do, requires that the computer have affect perception and use that knowledge, along with knowledge about cognitive progress, to reason about a series of student

actions and interventions, not simply a single-shot action or interaction, but as an ongoing and evolving relationship (Picard et al., 2004; Bickmore and Picard, 2004).

In this research, we pay particular attention to understanding learners' progress from one emotion to another and use dynamic sensor information to interpret objective measures of student progress.

Research questions include:

- How is affect expressed in student behavior?
- How accurate are different machine learning methods (e.g., Bayesian Networks, hidden Markov models) at predicting affect from student behaviors?
- How effective are interventions at changing negative affect or changing a state of Stuck into a state of Flow? Can machine learning technology learn reasonable policies for improving student attitude and learning?
- How does affect (student emotion and/or computer understanding of it) predict learning?

This article discusses a variety of ways that these research questions are addressed, divided into three general areas: affect recognition, interventions that respond to students and development of emotional embodied pedagogical agents.

## **2 Affect recognition**

The first area of this research is affect recognition, or use of techniques to detect and evaluate student affect. This research area is fairly new and uses exploratory methods and tools that are likely different from techniques used once the field has matured and reached its steady state. Normal conditions for affect recognition might use non-invasive sensors and machine learning techniques to measure student affect on-line. However, at this early research stage, we use a variety of invasive techniques until we can efficiently predict affect with automatic techniques alone.

In one technique described below we invited trained human observers to label students' affect. Although this technique is labor and time-intensive, it provides several advantages, such as identifying high-level student learning behaviors and suggesting how emotion impacts learning. We induce both *static* (based on demographics and emotion instruments) and *dynamic* (based on real-time sensor data as well as inferred hidden variables) student models (McQuiggan & Lester, 2006). Before and after completing the tutoring session (a matter of several days) students are presented with emotion instruments to measure long-term changes in their motivation, self-confidence and boredom. Well-established instruments are employed to measure student emotion before and after interacting with the tutor, Table 3. We use these older instruments because they are verified and used by hundreds of people. These instruments do overlap our cognitive-affective framework; however, we don't yet know how.

Another invasive technique is student self-report, which typically requires interrupting the student during the learning experience and can be unreliable (Picard et al., 2004). Another type of self-report involves asking students to recall their feelings afterward (Graesser et al., 2007); these techniques are less interruptive but very time consuming and still have high variance in reliability. We explore innovative ways to measure affective states, such as 'gaming' or moving rapidly through problems without reading them, or rushing through hints in the hope of being given the answer. It has been estimated that students who game the system learn two thirds of what students who do not game the system learn (Baker et al., 2004). This could be because of frustration, something especially important to detect for students with special needs (Murray et al., 2007). Another possibility is that gaming is a behavior related to poor self-monitoring and/or poor use of meta-cognitive resources.

<b>Dependent variables</b> ⇒⇒⇒	<b>Frustration</b>	<b>Motivation /Flow</b>	<b>Confidence</b>	<b>Boredom</b>	<b>Fatigue</b>
Instruments to measure dependent variables (to be predicted)	Frustration Button (Burleson, 2006); AMAS, reduced mathematics anxiety scale)	Harter's Scale	Fennema-Sherman Scale; Eccles scale	Boredom Proneness Scale; (Are you bored?)	Mental fatigue Scale (Are you tired?)
<b>Behavioral Variables</b>	Sensing data (camera, pressure sensitive chair, skin conductance glove, sensitive mouse)				
Behaviors that help predict the dependent variables	High state of arousal, high gaming; high effort; Gaussian Classification.	Record student effort exerted; dependence on help	Persistence at problem solving after incorrect attempts; dependence on encouragement messages	Low state of arousal combined with low effort and gaming	Increased problem solving time and increased error rate after some time in the tutoring session

**Table 3. Independent behavioral variables and dependent variables to measure student affect**

We triangulate among four techniques (human observations, sensors readings, machine learning techniques and student reports) in an attempt to resolve toward agreement with the realization that we may be far away from realizing any consensus. We intend to empirically identify which methods are more successful in the recognition of human affect level in specific contexts. We also analyze the dependency of specific behavioral variables and use a small subset of these variables to build prototype models where we draw on the relations between emotional state and actions. This section describes three methods to recognize student affect: human observations, a platform of hardware sensors and machine learning techniques.

## 2.1 Human observation to recognize affect

The first experiment described in detail here involved researchers who observed students in the classroom and labelled student emotion. Observations by multiple observers using similar methods have had high inter-rater reliability and report relatively low impact on student behavior once students are used to the observer's presence (Baker et al., 2004).

We trained researchers to conduct unobtrusive quantitative field observations and note students' behavior while using intelligent tutors. Observers identified variables that represent emotions and desirable/undesirable states linked to student learning and physical behaviors linked to affect states. Human observation are a useful exploratory strategy since observers can intuitively discern high-level behaviors and make appropriate judgments on limited information that may be difficult to automatically decide from raw sensor data. Human observers also provide some evidence for understanding the impact of student emotion in learning. They identify behaviors that are worth observing and then sensors are used to gather this behavioral data in bulk, see Section 2.3. These observations help develop a theoretical basis for affect recognition, approximate the type of information the sensors will collect and corroborate what sensor information indicates about perceived student emotional state. Only human observers were used in this experiment; face recognition and skin conductance, as described in Section 2.2, were not used here.

### 2.1.1 Experimental design

The human observation experiment included 34 students in a public school in urban Holyoke, MA, divided into three different classes. Students took a pretest survey to evaluate their attitudes towards mathematics (self-concept and value) and goal (learning vs. performance) orientation (Dweck, 1999), as well as a mathematics pretest with multiple problems to evaluate diverse mathematics concepts. Students used the tutoring software during a period of three weeks while three researchers

coded behavioral and subjective variables. Prior to the experiment, observers studied videos of students using Wayang, an intelligent tutor, see Figure 5, to learn how to code student affect. During the experiment, observers rotated around the classroom, coding one student at a time. Observation periods lasted for approximately 15-20 seconds per student, with an additional 15 seconds used to confirm the observation before they moved on to the next student. Because students may experience several behaviors/emotions during one time period (e.g., the student was seen forward and then back on the chair), we coded the first state seen, but the second one was coded and taken into account during subsequent analysis. More than 200 observations of each behavior were observed as shown in Table 6.

**Behavioral and Task-Based Variables.** Researchers looked for expressed affect and recorded facial expressions (smile, frown, scratch the head, nod), physical expression (relaxed pose, hitting the table, fidgeting), and verbal behavior (loud exclamations, talking with others). They also coded whether students appeared to be on- or off-task, obviously a subjective and noisy variable as students may seem off-task when they are not. Students were marked as being off-task when they were not using the software appropriately (using other programs on the computer) or conversing with peers about other subject matter (Baker et al., 2007). On-task students might be reading/thinking about the problem, talking to a friend about the problem, or writing a solution on paper.

**Table 4. Desirable State Variables and Possible Emotion Models**

Valence	Arousal	On/Off task	Example Student Behavior	Desirability Value	
+	+	On	Aha moment, yes! That's it!	2	Highly Desirable
+	—	On	Concentrated on problem-solving	2	Highly Desirable
—	+	On	Frustrated with tutoring software,	1	Maybe desirable
—	—	On	Yawning, zoned out within software	0	Not desirable
+	+	Off	Laughing with friend	0	Not desirable
+	—	Off	Very focused but on other software	0	Not desirable
—	+	Off	Angry quarrel with friend	0	Not desirable
—	—	Off	Zoned out, or sleeping	0	Not desirable

**Emotional Indicators.** Distinguishing one emotion from another is very difficult, especially using only facial expressions and body movement. For example, neither frustration nor boredom is clearly distinguished from neutral using only a camera and facial action units (McDaniel et al., 2007).

Because of this, we limited the conventional emotional terms (e.g., anxiety or frustration) to emotions that result from the combination of two indicators: *valence* (positive or negative nature of the emotion/energy the student seemed to be expressing) and *arousal* (we analyze physical activity as an expression of arousal). These emotion indicators are used to express the basic emotions in Table 1 and are consistent with early research on emotions (Wundt, 1902). For example:

- positive valence and arousal is related to being excited and joyful,
- positive valence and negative arousal is related to being concentrated or satisfied,
- negative valence and positive arousal is related to being frustrated or angry,
- negative valence and arousal is related to being bored and tired

However, our concern was that these emotional state variables might not be correlated to learning without also considering on-task or off-task behavior. It is highly desirable for a student to experience a state of joy/excitement when she is on-task, but if the student tends to be joyful while off-task, the emotion variable will not correlate strongly with optimal learning. Thus created another variable, *Desirability Value*, which is both task- and emotion-dependent (on/off-task, valence and arousal), see Table 4. Labeling emotional states as desirable or undesirable is problematic as often an undesirable state of confusion precedes learning gains, thus making it a desirable state pedagogically (Graesser et al., 2007). We include frustration as a desirable state while being on-task since learning episodes often have productive moments of frustration. Highly desirable states include states of positive valence while being on-task, whether accompanied by high

arousal or by low levels of arousal where students experience high mental activity without expressing significant observable emotion. Also, while laughing with a friend is desirable in general, this can change to be undesirable when it pulls the learner away from the learning task, changing them to be off-task. Undesirable states include being tired/bored (negative valence, negative arousal) while being on-task, as a student might give up. We could include some negative values for the *Desirability Value* since some states are more undesirable than others. Desirability might also be a function of time spent on the task; sometimes breaks are important to sustain learning.

### 2.1.2 Results of classroom observations

We computed correlations between emotion indicators and intermediate emotion/task-based state variables and analyzed the correlation between these state-based variables and student behaviors. Students were detected to be on-task 76% of the time, slightly lower than previous findings regarding off/on-task behavior with software learning environments (Baker, 2007).

**Table 5. Frequency of Emotion Indicators and Desirable Learning States**

Emotion indicators: Valence & Arousal	Frequency	Percent
+ valence & --arousal (concentrated, satisfied)	148	58%
+ valence & + arousal (excited, joyful, actively engaged)	85	34%
- valence & +arousal (frustrated, angry)	16	6%
- valence & --arousal (bored, tired)	5	2%
Total	254	100%

Desirable State	Frequency	Percent
Highly desirable	181	73%
Not desirable	61	24%
Medium Desirable	7	3%

Table 5 shows the frequencies of different emotional states. Note that negative valence emotions were observed only 8% of the time. This could be largely due to the fact that a neutral or indiscernible valence was coded as positive. Table 5 shows that 73% highly desirable states were observed, 3% medium desirable states, and 24% non-desirable states.

**Correlation Between Emotion Indicators and Learning/Attitudes.** We analyzed whether we can use emotional indicators and other state variables to predict learning and motivation, the variables we want to optimize.

*Valence.* Valence (or student energy) was significantly correlated to pretest mathematics score (N=34, R=.499, p=.003). This suggests that students who are good in mathematics to begin with, also have substantially more positive emotions while using the software, or at least less unpleasant emotions (e.g. boredom, frustration). Valence was also positively correlated to posttest learning orientation (N=34 R=.499, p<.01), but not to pretest learning orientation, suggesting that having positive valence during the tutoring session may instill higher *learning orientation* goals at posttest time. A similar effect happened for posttest self-concept and valence (R=.48, p<0.01) where students who had higher valence emotions had higher posttest self-concept scores. Thus, the presence of *positive* or *negative* emotions can help predict more general attitudes towards mathematics at posttest time.

*Arousal.* Arousal (expressed as physical activity) was negatively correlated with pre-tutor learning orientation (N=34, R=-.373, p<0.05), suggesting that students who are *performance-oriented* (characterized by a desire to be positively evaluated by others) are more likely to be physically active or ‘aroused,’ as opposed to those who are *learning oriented*, who tend to express less physical activity.

*Emotion (Valence + Arousal)*. Our emotional scale was correlated with pretest self-concept (N=34) (R=.385, p<0.05) and posttest learning orientation (R=.463, p<.05), suggesting that the presence of four types of emotion indicators (determined by combinations of valence and arousal) can help predict more general attitudes towards learning math.

**Table 6. Correlations between student behavior and emotion states.**

Pearson correlations among student behavior (chair and head positions), emotion indicators (valence and arousal), the desirability value and student talk.

Note: In this table, N (number) refers to the number of student behaviors recorded by observers.

	VALENCE	AROUSAL	ON TASK?	Desirability Value	TALK
<b>Chair Movement</b> N =	-.467 (.046*) 252	.420 (.000***) 252	-.140 (.027*) 249	-.154 (.015*) 247	----
<b>CHAIR Middle</b> N =	.148 (.018*) 252	.107 (.090) 252	-.002 (.974) 249	-.003 (.967) 247	----
<b>HEAD MOVE</b> N = ?	-.224 (.000***) 249	.345 (.000***) 249	-.417 (.000***) 246	-.435 (.000***) 244	
<b>HEAD SIDE</b> N = ?	-.195 (.002**) 254	.247 (.000***) 254	-.325 (.000***) 251	-.337 (.000***) 249	----
<b>HEAD MOVE SIDE</b> N =	-.270 (.000***) 249	.230 (.000***) 249	-.422 (.000***) 246	-.443 (.000***) 244	----
<b>HEAD MIDDLE</b> N =	.202 (.000***) 254	-.186 (.000***) 254	.427 (.000***) 251	.436 (.000***) 249	----
<b>HEAD UP</b> N =	-.097 (.123) 254	.062 (.326) 254	-.214 (.001**) 251	-.235 (.000***) 249	----
<b>TALK</b> N =	-.117 (.064) 251	.304 (.000***) 251	-.644 (.000***) 251	-.628 (.000***) 249	----
<b>SOUND</b> N =	-.075 (.248) 242	.370 (.000***) 242	-.388 (.000***) 241	-.379 (.000***) 239	----
<b>SMILE</b> N =	-.086 (.185) 240	.313 (.000***) 240	-.430 (.000***) 237	-.420 (.000***) 235	.485 (.000***) 237
<b>NEUTRAL</b> N =	.142 (.028*) 240	-.238 (.000***) 240	.395 (.000***) 237	.409 (.000***) 235	-.285 (.000***) 237

\*\*\* Correlation is significant at the 0.001 level (2-tailed); \*\* Correlation is significant at the 0.01 level (2-tailed); \* Correlation is significant at the 0.05 level (2-tailed).

*On/Off task*. Being on-task is significantly correlated to posttest self-concept in mathematics (N=34, R=.442, p=.02), but not to pretest self-concept in math, suggesting that being on-task is not a result of an incoming high self-concept in math. However, it indicates that being on-task may generate better self-concept after using the tutor. There is a significant correlation between mathematics posttest performance and being on-task (R=.640, p<.018). Again, being on-task is not correlated with mathematics pretest performance, meaning that prior mathematics knowledge will not predict students' tendencies towards on- or off-task behavior. Instead, being on-task seems to lead to higher posttest scores, again implying that being engaged with the tutoring system is part of the reason for achieving higher posttest scores. This is consistent with past research results on on/off-task behavior (Baker, 2007). If we can encourage students to be on-task, we will foster better attitudes for mathematics and higher posttest scores.

*Desirable Learning State*. Similar significant correlations were found for this variable (i.e., it predicted posttest scores and posttest self-concept in mathematics to a similar extent as did on/off-task behavior). If we can encourage students to be in our defined desirable learning states (Table 6), we will also foster better attitudes for mathematics and higher posttest scores.

**Correlations Between Student Behavior and Emotion States.** Several correlations were discovered among student behavior (chair and head position), emotion indicators (valence and arousal) and the desirability value, see Table 6. Clearly, a high positive correlation exists for arousal and chair movement since we defined arousal as being expressed by physical activity. Meanwhile, valence is not linked to chair movement, meaning that students do not express their

positive or negative emotions with chair movement. A negative correlation exists for desirable state and being on-task, meaning that students are in a more desirable learning state (and more on-task) when they don't move so much in the chair.

Other interesting findings (some not shown) are that students with positive valence emotions tend to sit in the middle of the chair, instead of being towards the side, the front or the back of the chair. Last, students leaning on their hands correlated negatively with arousal –as leaning is a fairly inactive posture. It is not obvious that students in a state of positive valence also tend to lean on their hands.

Head movement was correlated with negative valence, high arousal, off-task behavior and non-desirable states. This implies that students move their heads when they feel negative emotions, when being off-task and in a non-desirable learning state. When students are in such unproductive learning states, and are off-task, they tend to move their heads to the side. Also, students tend to move their head to the side when they have negative feelings. It is possible that students avoid the computer screen when they don't feel good about the software or the learning situation. At the same time, having their head in the middle had the opposite effect: it was correlated with positive valence, low arousal, on-task behavior, and desirable state for learning.

Students indicate off-task behavior with their head movement; holding one's head up looking over the top of their screen is correlated with an undesirable state for learning, while holding one's head down is not related to an undesirable state for learning (possibly because many students tend to work on paper on their desk). Again, head up could be an indication of screen avoidance. It seems obvious that frowning is related to having a negative valence emotion. However, frowning doesn't appear to be a good predictor of being on-task or being in a desirable learning state (not shown). A smile on the face does predict off-task behavior ( $R=-.430$  with on-task) and undesirable state for learning ( $R=-.420$ ), Table 6. Surprisingly, smiling was not linked to valence, but it is positively correlated with arousal and talk (students probably moved and talked with friends while they smiled). The opposite effect happened for a neutral face: it was positively correlated to desirable learning state and on-task behavior. A neutral face was linked to positive valence, most likely because we coded seeing a neutral emotion as positive valence. A neutral face was an indicator that the student was not moving (negative arousal) and not talking.

## **2.2 Hardware sensors to recognize affect**

The second method used for affect recognition is a research platform of non-invasive hardware sensors. The computer assesses a constellation of patterns from sensors and relates them to students' affective state. Clearly sensors can't really see the student's feelings, rather they record a pattern of external changes (on the face, in the posture, on the skin) associated with feelings. Sensors record patterns of student behavior (cameras or pressure sensors) applied to objects the student is in contact with (mouse, chair, keyboard) and the computer associates these patterns with probable affective state information. In the research described below a camera and computer, equipped with pattern recognition software, are used to recognize facial muscle movements associated with a smile, and the smile-detection might then be used to help reason about the probability the person is actually happy.

Recent research has focused on recognizing specific muscle movements known as 'facial actions' (Ekman et al., 1972; Ekman 1999) that can be used to construct any facial expression (el Kaliouby 2005, McDaniel et al., 2007; Kapoor and Picard, 2002, Bartlett et al., 2003). Under certain restricted conditions the automated recognizers have been shown to perform comparably to humans trained in recognizing facial actions (Cohn et al., 1999). Combining visual information with other modalities can give improved results (Chen et al., 1998; Kapoor et al., 2004).

Our platform includes four sensors (Figures 1-4): a facial expression system, posture analysis seat, pressure mouse, and wireless skin conductance sensor.

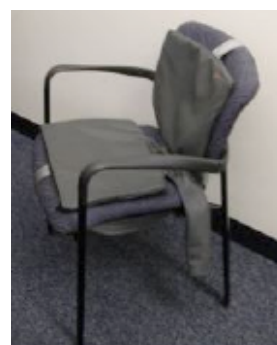
This hardware platform (with the exception of the camera) was manufactured at Arizona State University from validated instruments first developed by the Affective Computing group at MIT. Pre-production prototypes of each sensor were developed and twenty-five sets manufactured for simultaneous use in classrooms in Fall 2008. They were developed at an order of magnitude reduction in the overall cost of the MIT sensors and then integrated into the Wayang Intelligent Tutor, see Section 3.1. Sensors collect constant streams of data in parallel, allowing for much more consistent observation than a human ever could accomplish.



**Figure 1. Facial Expression Sensor**  
Detection, tracking, and recognition from video. Logitech QuickCam STX used with MIT algorithms.

*Facial Expression Camera.* A person's mental state is typically inferred from a range of non-verbal cues including facial expressions. The facial expression recognition system incorporates a computational framework that aims to infer a user's state of mind (el Kaliouby, 2005; el Kaliouby et al., 2006; el Kaliouby and Robinson, 2005). This facial action analysis is based on a combination of bottom-up vision-based processing of the face (e.g. head nod or smile) with top-down predictions of mental state models (e.g. interest and confusion) to interpret the meaning underlying head and facial signals over time (el Kaliouby and Robinson, 2005).

A multilevel, probabilistic architecture (using dynamic Bayesian networks) mimics the hierarchical manner with which people perceive facial and other human behavior (Zacks et al., 2001) and handles the uncertainty inherent in the process of attributing mental states to others. The output probabilities represent a rich modality that technology can use to represent a person's state and respond accordingly. The resulting visual system infers mental states from head gestures and facial expressions in a video stream in real-time. At 30 fps, the inference system locates and tracks 24 feature points on the face and uses motion, shape and color deformations of these features to identify 20 facial and head movements (e.g., head pitch, lip corner pull) and 11 communicative gestures (e.g., head nod, smile, eyebrow flash) (Zacks et al., 2001). Dynamic Bayesian networks model these head and facial movements over time, and infer the student's "hidden" affective-cognitive state.



**Figure 2. Posture Analysis Seat**  
Manufactured at Arizona State University and based on a seat that resulted in accurate recognition of high/low interest and break-taking (Burleson,

*Posture Analysis Seat.* We have manufactured a low-cost/low resolution pressure sensitive seat cushion and back pad with an incorporated accelerometer to measure elements of a student's posture and activity, Figure 2. We have also developed algorithms based on analyzing movement from this posture analysis chair. The learning companion system discussed in Section 4.1 used an extremely expensive Posture Analysis Seat, developed for medical and automotive applications (Burleson, 2006; Tekscan, 1997). This earlier system used pattern recognition techniques while watching a student's natural behaviors to *learn* which behaviors tended to accompany states such as interest and boredom.

*Pressure mouse.* A pressure mouse is used to detect the increasing amounts of pressure that students place on their mice related to increased levels of frustration. The pressure mouse was developed at Arizona State University based on an MIT system (Reynolds and Picard 2004). It has six force sensitive resistor sensors and an embedded microprocessor, Figure 3 and measures the overall pressure of the student's hand across the surface of the mouse. It uses the standard communication channel of a USB mouse for pointing and clicking functions and then in parallel uses a second channel, a serial communications port, to provide pressure data at 20ms intervals from each of the six sensors.



**Figure 3. Pressure mouse sensor.**  
Developed at Arizona State University based on an MIT svstem.

*Wireless skin conductance.* A wireless conductance bracelet, see Figure 4, was developed based on an earlier glove that sensed skin conductance, developed by Carson Reynolds and Marc Strauss at the MIT Media Lab, in collaboration with Gary McDarby, at Media Lab Europe (Strauss et al., 2005). While the skin conductance signal is not valenced (i.e. does not describe how positive or negative the affective state is) it is strongly related to arousal. A certain amount of arousal is a motivator toward learning and tends to accompany significant, new, or attention-getting events (Boucsein, 1992).

Information from these four sensors is analyzed along with cognitive activities from the tutor (time in each problem, number of hints requested, correct solutions, etc), stored in an episodic database. To coordinate the four sensors, two client programs were developed, one on each student computer for the video, chair and mouse data and an additional one located on a separate computer to processes and relay skin conductance data. The client program on each student's computer is initialized with the wrist sensor ID to coordinate the four sources of sensor data. One integrated log file is produced from the two sets of server software.

All sensor data is time stamped and sent to a Java remote method invocation (RMI) Server, an interface for performing remote procedure calls, in which methods are invoked from other Java virtual machines, possibly on different hosts. The server processes a second of data at a time and sorts and aggregates the sensor data into a string with the time stamp and the latest values from each sensor. In addition, the Wayang Tutor has the wrist sensor ID as part of its login so cognitive activities are correlated with the sensor data during the time period that the student is connected.

To synchronize the wrist sensor with the other data, software is used to time stamp and relay data from each wrist sensor to the sensor server. While students interact with Wayang, episodic data is written to local files and sent to a server. Ultimately, episodic data from the tutor will be sent directly to the Sensor Server so that the sensor and tutor data can be aggregated in real time. Sensor data (comma delimited) is written out each time an update is received. To address delays with respect to sensor data, a one second buffer is used to handle any timing mix-ups in transit. As sensor data comes in from the four sources, they are aggregated based on the wrist ID number, then printed to standard output and logged to a database. Two connections manage the information between the Sensor and Wayang Server: an episodic data connection (data sent to the Sensor Server from the Wayang Server) and a sensor state connection (data requested by the Wayang Server).



**Figure 4. Conductance Bracelet**  
A wireless skin conductance sensor manufactured through collaboration between ASU and MIT (Reynolds and Picard 2004).

### **2.3 Machine learning techniques to recognize affect**

The third and final method for achieving affect recognition is through the use of machine learning techniques. These techniques are very versatile and have been used with intelligent tutors to answer a variety of questions (Woolf, 2008): Is the student engaged? Is the student motivated? What should the student learn next? What type of intervention should be tried next? How is learning progress measured? How should student errors be corrected? How can student success be recognized? How and when should help be provided?

Pattern recognition machine learning techniques are used with sensors to learn a mapping from a set of sensor input features to an output label (e.g., appears to be frustrated). The input features might be associated with sensor readings from the camera (movement of the mouth or head) or the skin conductance bracelet (high arousal). Machine learning techniques typically learn the mapping through a statistical analysis of hundreds or thousands of training examples chosen by an external supervisor, in the case of supervised learning techniques, where an example contains both the input features and the desired output label.

Additionally, machine learning techniques are often used independent of hardware sensors to infer student affect (Conati et al., 2002; Murray & VanLehn, 2000; Baffes & Mooney, 1996; Mayo & Mitrovic, 2001). Marsella and Johnson used affective tutors to alter student affective states through changes in the tutor's perspective rather than in the task (Marsella and Johnson 2003). Machine learning also provides useful information for detecting inappropriate task strategies, procedural errors, or misconceptions.

We used Bayesian networks to infer affect based on student's observed problem-solving behavior and estimations from surveys filled out by prior students (Arroyo & Woolf, 2005). Networks were used to discover links between affect (revealed in a post-survey) and observable behavior (time spent on hints, number of hints selected, etc.) (Arroyo et al., 2004). The probability of being correct about a student's affective state (e.g., predicting a student's response about motivation as shown in the post-survey) was measured within a window of 80-90%. We correlated observable student activities and survey responses, converted this into a Bayesian network and then tested the predictions on the log data of new students. Hidden affective variables were integrated into the student model, enabling the tutor to refine its inference of student frustration, engagement and confidence. Links between students' behaviors, attitudes and perceptions exist and correlations between help requests and learning have been shown to be consistent with other authors' findings (Wood & Wood, 1999; Renkl, 2002).

In another study, machine learning techniques were used to show that disengagement negatively correlates with performance gain (Johns & Woolf, 2006). Hidden Markov models were integrated with an Item Response Theory dynamic mixture model to simultaneously estimate a student's changing motivation level and proficiency (Johns et al., 2006). This tutor predicted the probability of a correct student response with up to 75% accuracy. It was tested dynamically with high school students using the Wayang tutor, described in Section 3.1. By accounting for a student's motivation, the dynamic mixture model accurately estimated proficiency and the probability of a correct response. Motivation was modeled as a dynamic, discrete variable and proficiency as a static, continuous variable. These assumptions are based on a student's tendency to exhibit different behavioral patterns over the course of a tutoring session.

### **3 Interventions that respond to student cognitive-affective state**

The second area of this research is to use interventions to respond to students' cognitive-affective state. Evidence from the literature shows the value of interventions that are adapted to both a student's cognition and affect. Sweller et al. (1998) showed that student become overwhelmed when they can not solve mathematics problems. Presentation of worked examples reduces the cognitive

load for low-ability, novice or struggling students. One general recommendation is that immediate feedback for students with low achievement levels in the context of either simple (lower-level) or complex (higher-level) tasks is superior to delayed feedback; while delayed feedback is suggested for students with high achievement levels, especially for complex tasks (Shute, 2008). Appropriate feedback does improve learning; it can reduce uncertainty about how well (or poorly) students are performing and motivate strategies aimed at reducing that uncertainty (Ashford et al., 2003).

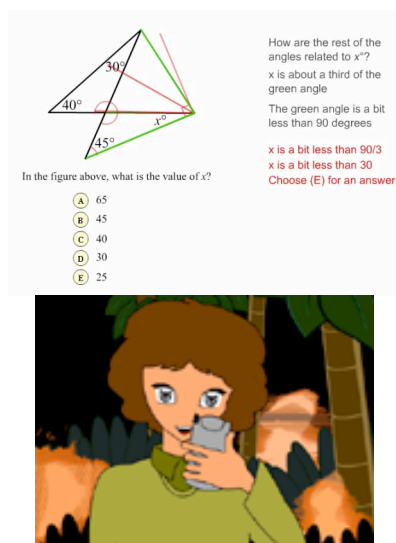
Computational instruction provides an opportunity to vary interventions for every student and every context (e.g. topic, emotional state). Instructional feedback can be varied according to type (explanation, hints, worked examples) and timing (immediately following an answer, after some elapsed time) (Shute, 2008). Complex interventions can be applied to bring learners back on track, or into a state of Flow, increasing the probability that the student will actually learn. We measure interventions in relation to their impact on student affect, behavior and learning. We also measure how intervention variables interact to promote learning in context (characteristics of the learner, aspects of the task). One goal is to specify in detail which behavioral variables and which interactions between variables impact student behavior and learning. This section first outlines the tutor into which the recognition and intervention mechanisms described above have been embedded and then describes how we identify the timing and type of intervention to use in the tutor.

### 3.1 Intelligent tutors

The affective sensors described above have been integrated into Wayang Outpost, an intelligent tutor that infers a student's cognitive skills and reasons about which type of hints are best to present in each context. The tutor teaches mathematics (geometry, statistics) and prepares students for standardized state exams<sup>1</sup> (Arroyo et al, 2004, 2005, 2007). The theme and setting of the tutor is a research station on the island of Borneo and features storylines, animated characters and problem solving hints that foster student engagement with mathematical thinking. Meanwhile it embeds mathematics problems into investigations of the ecology and biology of tropical rain forests, see Figure 5.

Wayang has been used with thousands of students and has demonstrated improved learning gains (an average 12% improvement from pretest to posttest) after only 2 class periods. Students passed the state standard exam at a higher rate (92%) as compared with students not using the tutor (76%).

The tutor provides a complex learning environment that can be explored at length by students or teams without supervision. Like a human tutor, it provides the sustained engagement and structured practice needed for students to become better learners and test takers. Multimedia (e.g., animation and audio) is provided with help and hints to support problem solving. Exercises support literacy while engaging students in role-playing around case studies (e.g., endangered species -- orangutans). The tutor incorporates knowledge of student group characteristics (e.g., profile of cognitive skills, gender) to guide instruction and customizes the choice of hint type for individual students based on their cognitive profile, gender, spatial ability, and mathematics fact retrieval speed.

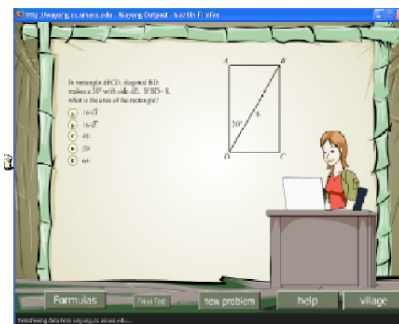


**Figure 5. Wayang Outpost.**

The tutor animates geometry problems when help is requested (top). Scientist Anne leads the student to save three orangutans after a fire breaks out in the forest. "We need to find the shortest route."(bottom).

<sup>1</sup> See <http://althea.cs.umass.edu/wayang/wayangindex.html> and <http://wayang.cs.umass.edu/Wayang/flash/>

The tutor facilitates the task of logging, pre/post-testing and data collection. An adaptive module tailors the sequencing of problems depending on students' performance in past problems. Decision-making is performed with a database-backed Java servlet with a Flash interface. Students are presented with questions by an adaptive module according to the tutor's inference about student cognitive and affective state. Students can request help at any time and receive multimedia support specific to the problem at hand.



**Figure 6. Affect-aware agents are integrated in the Wayang Geometry Tutor.**

We implemented and are evaluating two animated affective agents that work with students as learning companions, Figures 6 and 7. “Jake” and “Jane” are amusing and friendly animated study partners who offer advice and encouragement while reflecting on the range of the student's own emotions. They are integrated into the Wayang mathematics tutoring software both for sixth grade and tenth grade. Gendered characters are useful as they act out their emotion and talk with the student. They express full sentences of cognitive, meta-cognitive and emotional feedback, as outlined in Table 8. Both agents will be extended to multiple ethnicity by modifying their face module, hair texture and skin color.

Students frequently bring baggage of negative attributions in their self-perception of their mathematics ability. Thus if a mathematics tutor can recognize that a student is frustrated and support her or him, the student may persist longer and move beyond frustration. When animated agents mirror and animate student emotion (acting frustrated, bored, or confused) students might believe the agents are part of the learning experience and feel that they have a companion to work with them. The characters may seem like a mentor, someone who is together with the student against the computer, but who is more knowledgeable most of the time (not always) both cognitively and emotionally.



**Figure 7. Emotional pedagogical agents show a range of emotion. They act out their emotion and talk with the student expressing full sentences of cognitive, meta-cognitive and emotional feedback, see Table 8.**

We specified the messages and forms of interaction the agents have with students, see Table 8. We outlined not only cognitive and meta-

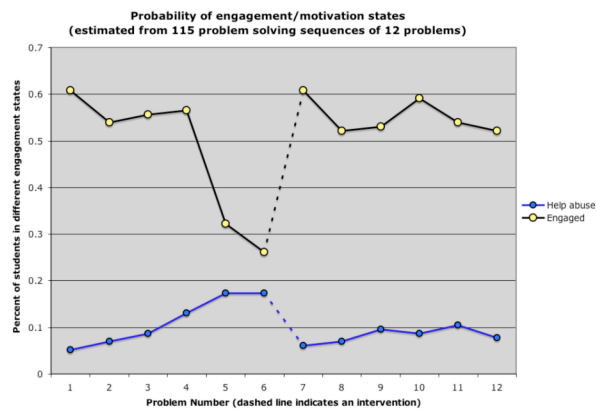
cognitive feedback, but also emotional feedback in the form of messages that allow agents to attribute failure to something different than lack of inherent ability, and to empathize with students to help them cope with frustration and anxiety. This allows us to study the benefits of feedback at key moments of student disengagement and frustration. These agents are being evaluated in the classroom along with the sensors described in Section 2.2.

### 3.2 Interventions based on a student's affective state

A variety of heuristic policies (providing text messages, mirroring student actions) have been used with the Wayang intelligent tutor to respond to student affect. For example, policies such as agent responses listed in Table 7 have been applied when a specific student emotional state is detected. Machine learning optimization algorithms have been used to search for policies for individual students in different affective and cognitive states, with the goal of achieving high learning and positive attitudes towards the subject, compared to pre-defined heuristic policies.

However, the interventions listed in Table 7 need to be evaluated with numerous students in a variety of contexts. For example, mirroring student emotion (see Section 4.1) can be good for increasing self-awareness and building rapport, e.g., mirroring sadness shows understanding and mirroring joy can amplify that joy. However, mirroring is not the right response for all emotions. An increasingly frustrated student might be moved to anger if the tutor mirrored back his or her increasing frustration. Rather, a look of concern, or appearing subdued in response, and certainly not smiling, may be an appropriate response to frustration.

We use off-line unsupervised learning to help the system learn optimal policies from student data from previous years. The following are a few examples of our prior studies to identify the optimal intervention based on context. In one study, we measured student reaction to interventions in Wayang based on recognition of student engagement (Arroyo et al., 2007). The tutor intervened when unmotivated behavior was recognized after the 6th problem, see Figure 8, top graph. Interventions included either performance graphs with accompanying messages or tips that suggested more productive learning behavior. The tutor provided two kinds of tips, one encouraged students to read the problem and hints more carefully and to slow down, and the second hint encouraged students to think about the problem, make a guess and, if the guess was wrong, to ask for hints. Evidence gleaned from 115 problem-solving sequences showed that students do change their behavior based on digital intervention Figure 8. Once interventions were presented on-target engaged student behavior returned (top line) and hint abuse (quickly asking for hints) subsided.



**Figure 8. Students became motivated after intervention.** Student engagement declined after the 6<sup>th</sup> problem (top). After an intervention was presented on-target student performance returned. Help abuse mirrored student engagement (bottom).

<p><b><u>High frustration</u></b> Agent looks concerned and provides an empathetic responses: “That was frustrating. Let’s move to something easier.” It give student control: “Would you like to choose the next problem? What kind of problem would you like?”</p>	<p><b><u>Low motivation</u></b> Agent mirrors low motivation and changes its voice, motion and gestures; it may present graph, hints, adventures.</p>	<p><b><u>Low confidence</u></b> Agent provide encouragement; It links performance to student effort; It attributes failure to external issue (hard problem) and success to internal issues (you doing great).</p>
<p><b><u>Boredom because student can’t do the work</u></b> Agent moves to an easier topic and Identifies material that the student can Accomplish.</p>	<p><b><u>Boredom because work is too easy</u></b> Agent mirror boredom and increases the challenge level of the activity; it provide empathy messages: “Maybe this is boring? Would you like to move something more challenging?”</p>	<p><b><u>Fatigue</u></b> Agent mirrors fatigue and presents an empathetic message: “Is this getting tiring? Shall we switch to something more fun?” It changes the activity, e.g., moves to adventures, animation or game</p>

**Table 7. Interventions generated by a pedagogical agent based on student affect.**

We also generated interventions based on a probabilistic model of student proficiency using Item Response Theory. The model consisted of four variables: student proficiency, motivation, evidence of motivation and a student’s response to a problem (Johns & Woolf, 2006). Motivation was represented as a dynamic variable that changed during a session as students became more or less engaged with the material. Latent variables in the student model correspond to proficiency and motivation. Proficiency is a static variable that does not change over time and we investigated three types of motivation: motivated, unmotivated (abusing hints and quickly guessing).

In the current research, we measure the impact of a variety of different interventions on student emotion. We have two sets of dependent variables, those that track student engagement (or Flow) and those that track negative affect detrimental to learning (being stuck). Once a particular emotion is teased out, this information is synchronized with the learning tools to train classifier algorithms. We are investigating interventions for students who exhibit self-confidence, frustration, boredom and self-concept.

Interventions for *low self-confidence*. We measured the impact of interventions on student self-confidence or belief in one's own powers, abilities, or capacities. For example, in one version of the Wayang Tutor, the tutor selected new problems based on student proficiency prediction using machine learning and a hidden Markov model, and a second version provided friendly comments (graphs, tips, offering help) (Arroyo et al., 2004). Responses to questions such as “How will you do in mathematics next year?” have shown significant differences in the two intervention groups. Students in the motivational version showed improvement in attitude and had better feelings towards the system (“The tutor is friendly/smart”). Both groups learned more and perceived that they learned more (“How much did you learn?”).

Agent’s Goal	Example Agent Intervention
Mirror student emotional state visually, as a way to <i>empathize</i> with the student. Mirror the last reported feeling of the student if appropriate.	If the student is sad/delighted, the agent might look sad/pleased.
Implement Carol Dweck’s messages praising effort rather than correctness of response.	<i>Agent says</i> “You seem to know this pretty well so let’s move onto something more challenging that you can learn from.” “Congratulations! Your effort paid off!”
Request emotional information from the student.	<i>Agent says</i> “Students sometimes get bored with this problem. Are you bored?”
Acknowledge student emotion if it is negative. Provide a helpful hint.	<i>Agent says</i> “Some students are frustrated by this problem.” Let’s look at some similar problems already worked out.”
Meta-cognitive response about students’ progress and about good learning habits.	<i>Agent says</i> “Congratulations! You are getting more questions right than before. Do you see that from the chart?”

**Table 8: Responses of emotional animated agents.**

Interventions for *frustrated* and *bored* students. We are using the hardware/software research platform described above to measure the impact of interventions on students who appear frustrated (feelings, thoughts, and behaviors associated with not achieving a particular goal), and bored (restlessness, or irritability resulting from a lack of stimulation). First the platform is used to distinguish between bored and frustrated students. Stress sensors (mouse and chair) help tease apart behavior that could be frustration (arousal and hyperactive behavior) or boredom (gaming but with low arousal). In conjunction with activity behavior (hint requests, ‘gaming the system’) pattern matching methods help infer these states in real time. For cases of frustration, we provide motivational and empathetic feedback to support students to understand failure and use it to move the student forward. For cases of boredom, we provide alternative activities (animation and exploratory modules) or more challenging projects.

Interventions for *self-concept*. We are investigating student self-concept (students’ assessment of their own performance in a discipline). This is related to academic outcomes and motivation (Narciss, 2004). Students differ in their task specific self-concept and tend to explain their success or failure based on internal (their original talents) or external (originating in our environment) factors. Sadly, people with low self-concept attribute their failures to themselves and the reverse happens for people with high self-concept. We will use external responses (“That problem was really hard”) when students of low self-concept fail, and use internal responses (“Congratulations, you did an amazing job with that!”) when they succeed, hopefully reversing their negative beliefs.

#### 4 Emotional embodied pedagogical agents

The third area of this research is the design, implementation and evaluation of emotional embodied animated pedagogical agents. If computers are to tailor themselves to individual learner needs and capabilities, the software needs to provide a flexible and protean environment. Animated pedagogical characters help do this by engaging students and tailoring a curriculum for the individual. Many research issues remain to be addressed. Do human-like learning companions (that actually help the student in the learning process) improve student’s self-concept and attitudes towards mathematics? Does the presence of learning companions affect students’ learning? Are learning companions that resemble a student’s gender/ethnicity more effective? This section discusses pedagogical agents, their potential impact on learning and two agents that we are using.

Learning is enhanced when human empathy or support is present (Graham and Weiner, 1996; Zimmerman, 2000). The presence of someone who cares, or at least appears to care, can be motivating (Wentzel, 1997). Various studies have linked interpersonal relationships between teachers and students to motivational outcomes over the long term (Picard et al., 2004; Pianta, 1992; Wentzel and Asher, 1995). Can this noted

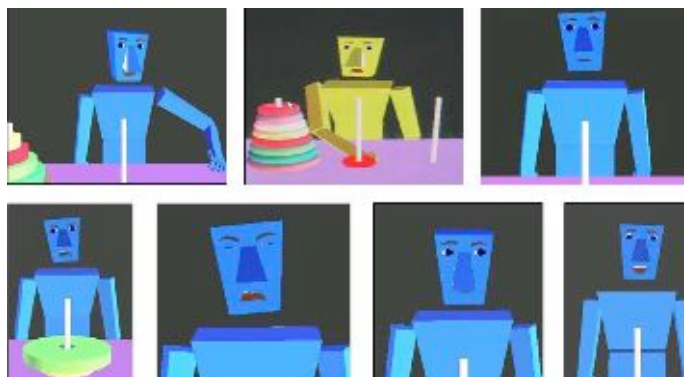


Figure 9. Affective Learning Companion

human relationship be reproduced, in part, by assistance and apparent empathy from a computer character? Apparently the answer is yes (Bickmore and Picard, 2004). Research shows that people relate to computers in the same way they relate to other humans and some relationships are identical to real social relationships (Reeves and Nass, 1998). One reason to use pedagogical agents is to further enhance this “personal” relationship between computers (whose logic is quantitative and

precise) and students (whose reasoning is more fuzzy and qualitative). For example, students continue to engage in frustrating tasks on a computer significantly longer after an empathetic computational response (Klein et al., 2002); users have immediately lowered stress level (via skin conductance) after empathy and after apology (Prendinger et al., 2003; Prendinger & Ishizuka, 2005) and relational skills improve long-term ratings of caring, trust, respect, desire to keep working (Bickmore & Picard, 2004).

Interactive animated pedagogical agents offer a low-pressure learning environment that allows students to gain knowledge at their own pace (Slater, 2000). Agents become excited when learners do well, yet students don't feel embarrassed if they ask the same question over and over again. Creating lifelike and emotive agents potentially provides important educational benefits based on generating human-like features (Lester et al., 1997). Agents can act like companions and appear to care about a learner's progress, which conveys to the learner that they are "in this thing together." Agents encourage students to care about progress made, react in a sensitive way to learner progress and intervene when students becomes frustrated or begin to lose interest. They convey enthusiasm for the subject matter and foster similar levels of enthusiasm in the learner. A learner who enjoys interacting with a pedagogical agent may have a more positive perception of the overall learning experience and may spend more time in the learning environment.

An animated agent often engages students with dramatic graphics and dynamic colorful animations, see Figures 6, 7 and 9. The added advantage of integrating such agents within intelligent tutors is to provide a controllable level of challenge for students facing problems (Burlson and Picard, 2004). Intelligent tutors can adjust material to be challenging for each student. Csikszentmihályi (1990) has found that people become most deeply engaged in activities that are challenging, but not overwhelming. When the work does become frustrating, learning is improved by agents showing empathy in the context of frustration.

#### **4.1 Mirroring student affect**

In a separate experiment, a learning companion was developed that automatically recognized and responded to student frustration (Burlson, 2006; Kapoor et al., 2007). The agent was a player or collaborator on the side of the student who was engaged in the Towers of Hanoi activity, Figure 9. This companion helped children learn and to learn how to learn better. It supported them to explore options, by occasionally prompting with questions or feedback, and by watching and responding to aspects of the affective state of the child. It watched especially for signs of frustration and boredom that may precede quitting, for signs of curiosity or interest that tend to indicate active exploration, and for signs of enjoyment and mastery, which might indicate a successful learning experience. One goal was to identify students who encountered frustration and teach them how to persevere and increase their ability and desire to engage in learning.

Two non-verbal interactions conditions were developed: sensor driven 'mirroring' interactions and pre-recorded interactions. Non-invasive multimodal real-time sensors were used to sense a student's affective state and were coupled with the learning companion capable of supporting learners by engaging in real-time responsive expressivity. The companion was coordinated with the software to mirror non-verbal social behaviors that influence persuasion, liking, and social rapport and responded to frustration with empathetic or task-support dialogue. Classifier algorithms predicted student frustration with 79% accuracy (Burlson, 2006). This research developed a theory for using affective sensing and appropriate relational agent interactions to support perseverance through failure. It focused on meta-cognitive awareness and personal strategies. For example, it sought to provide students with awareness of their affective state, help them understand failure and develop motivation to move onward.

The system captured many student movements relevant to education, specifically postural analysis to recognize frustration. It used an extremely expensive Posture Analysis Seat, developed for medical and automotive applications (Burlison, 2006; Tekscan, 1997) and used pattern recognition techniques to *learn* which student behaviors tended to accompany states such as interest and boredom. Students were reminded to push a button when they became frustrated. Sensor readings were used to predict when students might push the frustration button. The system achieved an accuracy of 76% on affect category recognition from chair pressure patterns, and 88% on nine ‘basic’ postures that were identified as making up the affective behaviors (Mota and Picard, 2003). Both sets of results are conservative, as the system was trained on a small set of data. These results were highly significant, confirming that there is strong evidence of affective information in the postural moves of a child. These results show that elements of affect can be measured with results significantly higher than random.

## 5 Discussion

This article described a variety of methods used in intelligent tutors to recognize and respond to student affect. We identified emotion indicators (valence and arousal) that combined with on- and off-task variables to represent desirable/undesirable states linked with student learning, as well as physical behaviors linked to emotional states. We described correlations between low-level observations (i.e. chair movement) and higher-level observations (on-off task behavior) and then between these higher-level observations and student learning and attitudes. We discussed pedagogical agents that respond to students learning and attitudes towards learning

Hardware sensors provide information about how students perform and when students are in non-productive states so computational tutors can provide appropriate interventions. Sensors also inform us whether given interventions are working or not. With this in mind, low cost portable sensors are being used in natural classroom settings. We have identified variables that are useful predictors of learning and affective outcomes and sensors predict students’ affective states related to learning.

We have unveiled several interesting findings: fluctuating states of emotion and on/off task behavior help predict posttest performance and attitudes/motivation; student states are expressed with specific behaviors that can be automatically detected with sensors; and a detector for strong/weak learning behavior was identified. As a result of these findings we identified how sensors can predict and reflect student learning, see Table 9.

**Table 9. Guide to interpreting sensor data and predicting learning**

<b>Desirable Learning States</b>	<b>Emotion/task indicators</b>	<b>Biologic indicators</b>	<b>Sensors to use</b>
<b>Most desirable</b> (Joy, Aha moment, Concentrated Actively engaged)	+ Valence AND On-task	Lean on hand; Little chair/head movement; Sit in middle of chair; Head in middle; Neutral face;	Chair sensors Camera
<b>Medium desirable</b> (Frustrated, angry)	-- Valence + Arousal	Head movement; Chair movement; Squeezing of mouse	Camera, Pressure mouse; Chair Sensors
<b>Least desirable</b> (Bored, tired)	Off-task OR -- Valence -- Arousal	Talking; Large chair movement; Head movement; Head to side or head up; Smile	Skin conductance; Camera; Chair sensors; Microphone

## 6 Future Work

The research described here provides a collection of models, tools and metaphors to understand student affect. Such methods need further improvement, especially to integrate a wide variety of

behavioral information from hardware sensors, increase the accuracy of inferences about affect and to refine the interventions. The goal is to fully elicit, sense, communicate, measure and respond to students' affect. Future work consists of predicting desirable/undesirable learning states and student attitudes towards learning. Moreover, because certain states (e.g., negative valence and high levels of arousal) are unproductive for post-tutor assessments of learning/attitudes, such states will prompt interventions. At that point the tutor must decide which interventions are most successful for individual students and context (e.g. topic, emotional state). Finally, we intend to better understand the nature of data from different sensors. The camera provides very high-level judgments and uses its own inference engine to decide emotional states, whereas all other sensors provide relatively raw data. We are developing machine-learning algorithms that relate these data sets to learners' diverse emotional states. Using all of these techniques, we plan to recognize and help students cope with states of negative valence and support their return to on-task behavior.

Another goal of this research is to support student reflection about their emotion and to increase teaching efficiency. We intend to build tutors that can generate long-term pedagogical decisions and view a series of student actions, not simply a single-shot action. Affect recognition can significantly improve a tutor's long-term planning if teaching is sometimes directed toward eliciting long-term experiences, which might entail sacrifices in immediate student performance.

Additionally we are evaluating the impact of the presence of gendered characters on student motivation and achievement within a learning environment. The intent is to integrate controlled exploration of communicative factors (facial expression, empathy, mirroring postures) as they impact learning, human interaction and relationship development. An integrated tutor-agent can bring many of these movements under precise control. This is not to say that the inferences, movements and interventions of a tutor-agent can exactly replace those of people, nor that such theories can exactly map to the human-human environment; however, this level of control does allow for careful testing of hypotheses (Picard, 2006). In the long term, we will evaluate several predictions, specifically that affective tutors perceived as 'caring' will help students persevere longer through frustrating learning episodes, better increase motivation and contagiously excite learners with passion for a topic, leading to greater effort to master the topic (Picard, 2006).

## **7 Acknowledgement**

We acknowledge contributions to the system development from Rana el Kaliouby, Ashish Kapoor, Ken Perlin, Selene Mota and Carson Reynolds. We also thank Joshua Richman, Roopesh Konda, and Assegid Kidane at ASU for their work on sensor development and manufacturing and we thank Kasia Muldner and Robert Christopherson for helpful comments on this paper.

This research was funded in part by awards from the National Science Foundation and the U.S. Department of Education: i) NSF 0705554, IIS/HCC *Affective Learning Companions: Modeling and supporting emotion during teaching*, Woolf and Burleson (PIs) with Arroyo, Barto, and Fisher; ii) NSF 0411776, ROLE, *Learning to Teach: The Next Generation of Intelligent Tutor Systems*, Woolf (PI) with Arroyo, Barto, Mahadevan and Fisher; and iii) Dept. of Educ. to Woolf, B. P. (PI) with Arroyo, Maloy and the Center for Applied Special Technology (CAST), *Teaching Every Student: Using Intelligent Tutoring and Universal Design to customize the mathematics curriculum*.

Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

## 8 References

- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*(111), 256-274.
- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology*, 80, 260-267.
- Arroyo, I., Beal, C.R., Murray, T., Walles, R., Woolf, B.P. (2004). Web-Based Intelligent Multimedia Tutoring for High Stakes Achievement Tests. In J. C. Lester, R. M. Vicari & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems, 7th International Conference, ITS 2004*, (pp. 468-477). Maceiò, Alagoas, Brazil, Proceedings. Lecture Notes in Computer Science 3220: Springer 2004.
- Arroyo, I., Woolf, B. (2005). Inferring Learning and Attitudes from a Bayesian Network of Log File Data. In *Proceedings of the Twelfth International Conference on Artificial Intelligence in Education*.
- Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S., Woolf, B. (2007). *Repairing Disengagement With Non Invasive Interventions*. International Conference on Artificial Intelligence in Education, Marina del Rey, CA.
- Arroyo, I., Woolf, B., & Beal, C. (2005). Gender and Cognitive Differences in Help Effectiveness During Problem Solving. *Technology, Instruction, Cognition and Learning*, 3 (1).
- Ashford, S. J., Blatt, R., & VandeWalle, D. (2003). Reflections on the Looking Glass: A Review of Research on Feedback-Seeking Behavior in Organizations. *Journal of Management*, 29, 773-799.
- Azevedo, R. and Cromley, J. (2004). Does Training on Self-Regulated Learning Facilitate Students, Learning With Hypermedia? *Journal of Educational Psychology Review, American Psychological Association*, 96 (3), 523-535.
- Baffes, P., Mooney, R. (1996). Refinement-based student modeling and automated bug library construction. *Journal of Artificial Intelligence in Education*, 7 (1), 75-116.
- Baker, R., Corbett, A. T., & Koedinger, K. (2004). *Detecting Student Misuse of Intelligent Tutoring Systems*. Proceedings of the Seventh International Conference on Intelligent Tutoring Systems, pp 531-540
- Baker, R. S. J. d. (2007). Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems. *Proceedings of ACM CHI 2007: Computer-Human Interaction*, 1059-1068.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavior change. *Psychological Review*(84), 191-215.
- Bartlett, M. S., Littlewort, G., Braathen, B., Sejnowski, T.J., and Movellan, J.R. (2003). A Prototype for Automatic Recognition of Spontaneous Facial Actions. In S. Becker, Obermayer, K. (Ed.), *Advances in Neural Information Processing Systems* (Vol. 15). Cambridge, MA: MIT Press.
- Bickmore, T., Picard, R. W. (2004). Establishing and Maintaining Long-Term Human-Computer Relationships. *Transactions on Computer-Human Interaction*, 12(2), 293-327.
- Block, J. (1995). On the Relation Between IQ, Impulsivity and Delinquency. *Journal of American Psychology*, 104, 395-398.
- Boucsein, W., (1992) *Electrodermal activity*. New York: Plenum Press.
- Bransford, J. D., Brown, A.L. and R. R. Cocking (Ed.) 2000, *How People Learn: Brain, Mind, Experience and School, Committee on Developments in the Science of Learning, NRC Commission on Behavioral and Social Sciences and Education*. Washington DC:

Naitonal Academy Press.

- Burleson, W., Picard, R.W. (2004). *Affective Agents: Sustaining Motivation to Learn Through Failure and a State of Stuck, Social and Emotional Intelligence in Learning Environments Workshop*. 7th International Conference on Intelligent Tutoring Systems, August 31, 2004, Maceio - Alagolas, Brasil.
- Burleson, W. (2006). *Affective Learning Companions: Strategies for Empathetic Agents with Real-Time Multimodal Affective Sensing to Foster Meta-Cognitive Approaches to Learning, Motivation, and Perseverance*. MIT PhD thesis, Available at <http://affect.media.mit.edu/pdfs/06.burleson-phd.pdf>.
- Chen, L. S., Huang, T.S., Miyasato, T., Nakatsu, R. (1998). *Multimodal Human Emotion/Expression Recognition*. Proceedings of International Conference on Automatic Face and Gesture Recognition, Nara, Japan.
- Cohn, J. F., Zlochower, A.J., Lien, J., Kanade, T. (1999). Automated Face Analysis by Feature Pont Tracking Has High Concurrent Validity with Manual FACS Coding. *Psychophysiology* (36), 35-43.
- Conati, C., Gertner, A. and VanLehn, K. (2002). Using Bayesian Networks to Manage Uncertainty in Student Modeling. *User Modeling and User-Adapted Interaction*, 12(4), 371-417.
- Csikszentmihalyi, M. (1990). *Flow, the Psychology of Optimal Experience*. New York, NY: Harper Perennial.
- Cytowic, R. E. (1989). *Synesthesia: A Union of the Senses*. New York: Springer Verlag.
- D'Mello, S. K., Picard, R. W., and Graesser, A. C. (2007). Towards an Affect-Sensitive AutoTutor. *pecial issue on Intelligent Educational Systems IEEE Intelligent Systems*, 22 (4), 53-61.
- Dweck, C. S. (1986). Motivational Processes Affecting Learning. *American Psychologist*, 41, 1040-1048.
- Dweck, C. S. (1999). *Self-Theories: Their role in motivation, personality and development*. Philadelphia: Psychology Press.
- Dweck, C. S., & Leggett, E. L. (1988). A Social-Cognitive Approach to Motivation and Personality. *Psychological Review*, 95, 256-273.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings*. New York: Pergamon Press.
- Ekman, P. (1999). *Facial Expressions*. New York: John Wiley & Sons Ltd.
- El Kaliouby, R. (2005). *Mind-Reading Machines: Automated Inference of Complex Mental States*. Unpublished Ph. D. thesis, University of Cambridge.
- El Kaliouby, R., Robinson, P. (2005). *Generalization of a vision-based computational model of mind-reading*. First Int. Conf. Affective Computing and Intelligent Interaction, Beijing, China.
- El Kaliouby, R., Teeters. A., and Picard. (2006). R.W. An Exploratory Social-Emotional Prosthetic for Autism Spectrum Disorders. In *International Workshop on Wearable and Implantable Body Sensor Networks* (pp. 3). Cambridge, MA: MIT Media Lab.
- Fletcher, J. D. (1996). *Does this stuff work? Some findings from applications of technology to education and training*. Conference on Teacher Education and the Use of Technology Based Learning Systems. Society for Applied Learning Technology
- Giuseppe, L., Brass, D. (2003). Exploring the Social Ledger: Negative Relationships and Negative Asymmetry in Social Networks in Organizations. *Academy of Management Review, Special issue: Building Effective Networks*.
- Goleman, D. (1995). *Emotional Intelligence*. NY: Bantam Books.
- Graesser, A. C. , Chipman, P., King, B., McDaniel, B., and D'Mello, S (2007). Emotions and Learning with AutoTutor. *13th International Conference on Artificial Intelligence in Education (AIED 2007)*. R. Luckin, K. Koedinger, and J. Greer (Eds), (pp 569-571). IOS Press.

- Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In Berliner, D. & Calfee, R (Eds), *Handbook of Educational Psychology*. 63-84. New York: Macmillan.
- Isen, A. M. (2000). Positive Affect and Decision Making. In M. Lewis, Haviland, J. (Ed.), *Handbook of Emotions*. New York: Guilford.
- Johns, J., Woolf, B. (2006). A Dynamic Mixture Model to Detect Student Motivation and Proficiency. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)* (pp. 2-8). Menlo Park, CA: AAAI Press.
- Johns, J., Mahadevan, S., & Woolf, B. (2006). Estimating Student Proficiency Using an Item Response Theory Model. In M. Ikeda, K. D. Ashley & T. W. Cahn (Eds.), *ITS 2006, Lecture Notes in Computer Science* (pp. 453-462.). Berlin-Heidelberg: Springer-Verlag.
- Kapoor, A., Picard, R. W., Ivanov, Y. (2004). *Probabilistic combination of multiple modalities to detect interest*. Paper presented at the Int. Conf on Pattern Recognition, Cambridge, England.
- Kapoor, A., Burleson, W., Picard, R. (2007). Automatic Prediction of Frustration. *International Journal of Human Computer Studies* 2007, 65(8), 724-736.
- Kapoor, A., & Picard, R. W. (2002) *Real-Time, Fully Automatic Upper Facial Feature Tracking*. Paper presented at the Proceedings of the 5th International Conference on Automatic Face and Gesture Recognition, Washington, D.C.
- Keller, J. M. (Ed.). (1983). *Motivational design of instruction*. Hillsdale, NJ: Erlbaum.
- Keller, J. M. (1987). Strategies for stimulating the motivation to learn. *Performance and Instruction*, 26 (8), 362-632.
- Klein, J., Y. Moon, and R.W. Picard ( 2002). This Computer Responds to User Frustration: Theory, Design, Results, and Implications. *Interacting with Computers* 14, No 2, pp. 119-140.
- Kleinginna Jr, P. R., Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 5(4), 345-379.
- Koedinger, K. R., Anderson, J. R., Hadley, W. H., Mark, M.A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education* 8(1), 30-43.
- Kort, B., Reilly, R., Picard, R. W. (2001). External representation of learning process and domain knowledge: affective state as a determinate of its structure and function. *Workshop on Artificial Intelligence in Education (AI-ED 2001), San Antonio*.
- Lepper, M. R., Chabay, R. W. (1988). Socialising the intelligent tutor: bringing empathy to computer tutors. *Learning issues for intelligent tutoring systems*, 242-257.
- Lepper, M. R., Hodell, M. (1989). Intrinsic Motivation in the Classroom. In C. Ames & R. E. Ames (Eds.), *Research on Motivation in Education* (Vol. 3, pp. 73-105). New York: Academic Press.
- Lester, J., Converse, S., Kahler, S., Barlow, T., Stone, B., & Bhogal, R. (1997) *The persona effect: Affective impact of animated pedagogical agents*. Paper presented at the Conference on Computer-Human Interfaces, New York, NY.
- Mandler, G. (1984). *Mind and Body: Psychology of Emotion and Stress*. New York, NY: W W Norton & Co.
- Marsella, S. C., Johnson, W.L. (2003). Interactive pedagogical drama for health interventions. In U. Hope, F. Verdejo, and J. Kay, J. (Eds.), *Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies* (pp. 341-348). Amsterdam: IOS Press.
- Mayo, M., & Mitrovic, A. (2001). Optimizing ITS Behavior with Bayesian Networks and Decision Theory. In W. van der Linden, Hambleton, R. eds. 1997 (Ed.), *International Journal of Artificial Intelligence in Education* (Vol. 12, pp. 124-153). New York, NY: Springer-Verlag.
- McDaniel, B. T., D'Mello, S. K., King, B. G., Chipman, P., Tapp, K., Graesser, A. C. (2007). Facial Features for Affective State Detection in Learning Environments. In D. S.

- McNamara, Trafton, J. G. (Ed.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 467-472). Austin, TX: Cognitive Science Society.
- McQuiggan, S. & Lester, J. (2006). *Diagnosing Self-Efficacy in Intelligent Tutoring Systems: An Empirical Study*. Eighth International Conference on Intelligent Tutoring Systems, M. Ikeda, K. Ashley, T.W. Chan (Eds), Jhongli, Taiwan
- Mota, S., Picard, R. W. (2003) Automated posture analysis for detecting learner's interest level. *1st IEEE Workshop on Computer Vision and Pattern Recognition, CVPR HCI 2003*, (TR 574).
- Murray, B., Silver-Pacuila, H. & Helsel, F.I. (2007). Improving Basic Mathematics Instruction: Promising technology resources for students with special needs. *Technology in Action*, 2 (5), 1-6; 8.
- Murray, R. C., & VanLehn, K. (2000). *DT Tutor: A decision-theoretic, dynamic approach for optimal selection of tutorial actions*. Paper presented at the Fifth International Conference on Intelligent Tutoring Systems, Montreal, Canada.
- Norman, D. A. (1981). Twelve Issues for Cognitive Science. In *Perspectives on Cognitive Science* (pp. 265-295). Hillsdale, NJ: Erlbaum.
- Ortony, A., Clores, G. L., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge, MA: Cambridge University Press.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66, 543-578.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Pianta R C (1992) Beyond the parent: the role of other adults in children's lives, *New Directions in Child Development*, 57, Jossey-Bass, San Francisco (1992).
- Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., Strohecker, C. (2004). Affective Learning--A Manifesto. *BT Technical Journal*, 2(4), 253-269.
- Picard, R. W. (2006). *Building an Affective Learning Companion, Invited Keynote Talk*. Paper presented at the International Conference of Intelligent Tutoring Systems, Taiwan.
- Prendinger, H., Mayer, S., Mori, J., and Ishizuka, M., 2003, Persona Effect Revisited, *Intelligent Agents*, 4th International Workshop, Germany, Lecture Notes in Computer Science 2792 Springer, pp 283-291.
- Prendinger, H., Ishizuka, M. (2005) The Empathic Companion: A Character-Based Interface that Addresses Users' Affective States. *Applied Artificial Intelligence* 19 (3-4): 267-285.
- Reeves, B., & Nass, C., 1998, *The media equation: How people treat computers, television and new media like real people and places*, New York, CSLI
- Renkl, A. (2002). Learning from Worked-Out Examples: Instructional Explanations Supplement Self Explanations. *Learning and Instruction*, 12, 529-556.
- Reynolds, C. and Picard, R.W., 2004, Affective Sensors, Privacy, and Ethical Contracts," *Proceedings of CHI*, April 2004, Vienna, Italy, pp. 1103-1106.
- Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review*, 1, 173-208.
- Shute, V., Psofka, J. (1995). Intelligent tutoring systems: Past, Present, and Future. *Handbook of Research on Educational Communications and Technology*. 570-600.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.
- Slater, D., 2000, Interactive animated pedagogical agents, International Conference on Computers and Human Interaction, Atlanta, GA
- Strauss, M., Reynolds, C., Hughes, S., Park, K., McDarby, G., & Picard, R. W. (2005). The Handwave Bluetooth Skin Conductance Sensor. In *Affective Computing and Intelligent Interaction Conference* (pp. 699-706). Beijing, China.
- Sweller, J., van Merriënboer, J., Paas, F. (1998). Cognitive architecture and instructional design.

*Educational Psychology Review*, 10, 251-296.

- Tekscan. (1997) *Tekscan Body Pressure Measurement System User's Manual*. South Boston, MA: Tekscan Inc.
- Vail, P. (1994). *Emotions: the on/off switch for learning*: Modern Learning Press.
- Vroom, V. H. (1964). *Work and Motivation*. New York: Wiley.
- Vygotsky, L. (1962). *Thought and Language*. Cambridge, MA: The MIT Press.
- Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press, (Originally published 1930, Oxford University Press, New York).
- Wentzel, K. (1997). Student Motivation in Middle School: The Role of Perceived Pedagogical Caring. *Journal of Educational Psychology*, 89 (3), pp 411-419.
- Wentzel K and Asher S R, 1995, Academic lives of neglected, rejected, popular, and controversial children', *Child Development*, 66, pp 754—763 (1995).
- Wood, H., Wood, D. (1999). Help Seeking, Learning, and Contingent Tutoring. *Computers and Education*, 33, 153-169.
- Woolf, B. (2008). *Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning*. San Francisco, CA: Elsevier Inc., Morgan Kaufman.
- Wundt, W. (1902). *Outlines of Psychology*. London: Williams & Norgate.
- Zacks, J. M., Tversky, B., Iyer, G. (2001). Perceiving, Remembering, and Communicating Structure in Events. *Journal of Experimental Psychology, General* (130), 29-58.
- Zimmerman, B. J. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, 25, 82-91.