Affect-aware tutors: recognising and responding to student affect

Beverly Woolf*

Department of Computer Science
University of Massachusetts
Amherst, USA
E-mail: bev@cs.umass.edu
*Corresponding author

Winslow Burleson

School of Computer Science and Informatics/Arts, Media
and Engineering
Arizona State University, USA
E-mail: winslow.burleson@asu.edu

Ivon Arroyo, Toby Dragon and David Cooper

Department of Computer Science
University of Massachusetts
Amherst, USA
E-mail: ivon@cs.umass.edu
E-mail: dragon@cs.umass.edu
E-mail: dcooper@cs.umass.edu

Rosalind Picard

Affective Computing Research Group
M.I.T. Media Laboratory, USA
E-mail: picard@media.mit.edu

Abstract: Theories and technologies are needed to understand and integrate the knowledge of student affect (*e.g.*, frustration, motivation and self-confidence) into learning models. Our goals are to redress the cognitive versus affective imbalance in teaching systems, develop tools that model student affect and build tutors that elicit, measure and respond to student affect. This article describes our broad approach towards this goal and our three main objectives: develop tools for affect recognition, interventions in response to student affect, and emotionally animated agents.

Keywords: affective recognition and feedback; multimodal sensing; motivation; learning companions; learner modelling; intelligent tutors; affect-aware.

Reference to this paper should be made as follows: Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D. and Picard, R. (2009) 'Affect-aware tutors: recognising and responding to student affect', *Int. J. Learning Technology*, Vol. 4, Nos. 3/4, pp.129–164.

Biographical notes: Beverly Park Woolf is a Research Professor and Director of the Center for Knowledge Communication at the Computer Science Department, University of Massachusetts Amherst, USA. Her team develops and deploys intelligent tutors that combine cognitive analysis of learning with artificial intelligence, network technology and multimedia. These tutors often incorporate learner models, intelligent interfaces and inference mechanisms, and have been deployed in education and industry in a variety of disciplines (chemistry, psychology, physics, geology, art history, mathematics and economics). One of these educational systems is used by more than 100 000 students per semester across 100 colleges. Dr. Woolf published the book *Building Intelligent Interactive Tutors* (Elsevier, 2009) along with over 200 articles on the subject of intelligent educational technology. She is a fellow of the American Association of Artificial Intelligence.

Winslow Burleson is an Assistant Professor of Human Computer Interaction and has joint appointments in the School of Computing, Informatics, and Decision Systems Engineering (CIDSE) and School of Arts, Media, and Engineering (AME) at Arizona State University, USA. He is the Founding Director of the Motivational Environments research group, and author of over 50 scientific publications and 10 patents. He received his PhD from the MIT Media Lab, working with the Affective Computing and Life Long Kindergarten research groups and the Entrepreneurial Management Unit at Harvard Business School on creativity research. He is a Kavli Fellow and serves on NAS and NSF committees. He has a BA in Biophysics from Rice University and an MSE in Mechanical Engineering from Stanford University, where he taught product design, creativity and visual thinking. NSF, NASA-JPL, Deutsche Telekom, iRobot, LEGO, Microsoft and Motorola support his research. He has worked at IBM Research, NASA-SETI Institute, the Space Telescope Science Institute and UNICEF.

Ivon Arroyo is a Research Scientist at the Computer Science Department, University of Massachusetts Amherst, USA. Over the past ten years, she has carried out research on how students learn mathematics with intelligent multimedia tutoring systems at the K-12 level in public school settings. She is a PI or co-PI of NSF and US Department of Education research grants that attempt to find principles for the design of digital learning environments that enhance affective and cognitive outcomes, particularly for girls and students with learning disabilities. She holds a doctorate in Math and Science Education and a Master's and BS in Computer Science, and is the author of more than 50 research articles at the forefront of education and computer science. She is a Fulbright Fellow and an elected member of the Executive Committee of the International Society of Artificial Intelligence in Education.

Toby Dragon is a PhD candidate in Computer Science at the University of Massachusetts Amherst, USA focusing on intelligent systems in education. He received his MS in Computer Science from the University of Massachusetts in 2006 and his BA in Computer Science from Ithaca College in 2003. His research interests include intelligent tutoring systems with a focus on tutoring for ill-defined domains and affective computer tutors. He explores the use of both simulation dynamics and expert knowledge systems to guide students through authentic problem spaces where no single correct answer is sought and no single correct solution path is implied. He has published work on affective computing, ill-defined domains and intelligent tutors.

David Cooper is a PhD candidate in Computer Science at the University of Massachusetts Amherst, USA focusing on computational models of emotion. He holds an MS in Computer Science from the University of Massachusetts Amherst, and a BS in Cognitive Science from Carnegie Mellon University. He has published work on affective computer tutors, cognitive agent architectures, emotional robotics, evolutionary agent learning and sensor data fusion. His research interests include biologically inspired computation, emotional and cognitive models of human interaction, and sensor integration for computer awareness.

Rosalind W. Picard is a Professor at the MIT Media Lab, Founder and Director of the Affective Computing Group, and leader of the Autism and Communication Technology Initiative at MIT in Cambridge, Massachusetts, USA. She is the author of around 200 scientific articles. She is best known for her pioneering work in image and video content-based retrieval (the original Photobook system), for developing texture models and machine learning for their combination (Society of Models) and for her book *Affective Computing* (MIT Press, 1997), which helped launch a field by that name. She was a member of the Technical Staff at AT&T Bell Labs in Holmdel, New Jersey (1984–1987) and has consulted at many companies including Apple, IRobot, BT and Motorola. Her current research interests focus on the development of technology to help people comfortably and respectfully measure and communicate affective information, as well as on the development of models of affect that improve decision making and learning.

1 Introduction

1.1 Vision and challenge

Intelligent tutors provide individualised teaching in multiple domains and demonstrate learning gains similar to or greater than those provided by human tutors (Woolf, 2009; Fletcher, 1996; Koedinger *et al.*, 1997; Shute and Psotka, 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, marginalising or ignoring affect. If computers are to interact naturally, with humans, they must recognise affect and express social competencies (Picard *et al.*, 2004). The role of affect in instruction is at best in its infancy. One obvious next frontier in computational instruction is to systematically examine the relationship(s) between student affective state and learning outcomes (Shute, 2008).

When humans use affect within one-to-one teaching relationships, the result is very powerful. In their research on 'thin slices', Ambady and Rosenthal demonstrated that based on a short segment of video, as little as six seconds of a teacher's first interactions with a student, participants could predict that teacher's effectiveness and student end-of-term grades based on the teacher's exhibited use of affect (Ambady and Rosenthal, 1992). Wentzel (1997) has shown that caring bonds between middle school children and their teachers are predictive of learners' performance.

This research looks at the role new technology plays in recognising and responding to affect. We describe research to measure and support the affective dimension of learning in classrooms in ways that were not previously possible. 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.

This section describes theories of affect, learning and human emotion. It looks at the constellation of student behaviours that are labelled as emotion, examines how they relate to learning and provides a brief overview of computer recognition and response to student affect. Section 2 describes three approaches to affect recognition (human observation, hardware sensors and machine learning techniques). Section 3 describes responses to a student's cognitive-affective state from within an intelligent tutoring system. Section 4 describes emotional embodied pedagogical agents and the final section provides a discussion and view of future work.

1.2 Theories of affect, learning and human emotion

Modelling student emotion has become increasingly important for computational teaching systems. Teachers have long recognised 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 in learning (*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 (Burleson 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). Numerous studies addressed emotions involved in learning, *e.g.*, emotions can paralyse a student's ability to retain information (Baddeley, 1986; Lepper and Chabay, 1988; Mandler, 1984; Kort *et al.*, 2001).

Human emotion is completely intertwined with cognition in guiding rational behaviour, including memory and decision making and emotion has been named as one of the 12 major challenges for the field of cognitive science (Norman, 1981). Emotion and cognitive functions are inextricably integrated into the human brain (Cytowic, 1989). Emotional skills have been shown to be more influential than cognitive abilities for personal, career and scholastic success (Goleman, 1995). For instance, in the comparison of impulsivity and verbal IQ as predictors of future delinquent behaviour, impulsivity was twice as powerful a predictor (Block, 1995). Recent findings suggest that when basic mechanisms of emotion are missing, intelligent functioning is hindered.

Acceptance of ideas about emotion in learning is based largely on intuition and generalised 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.

Nearly a hundred definitions of emotion have been categorised (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). Emotion is often defined as an intuitive feeling derived from one's circumstance, mood or relation with others. Most studies of emotion do not include the phenomena observed in natural learning situations, such as interest, boredom, or surprise. Rather, emotion definitions emphasise 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).

Motivation is one emotion strongly linked to learning and has been defined as an inner drive that causes a person to act with direction and persistence (Merriam-Webster, 2009; Webster, 1984). 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 towards 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 through frustration to a greater extent (Burleson and Picard, 2004; 2007).

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 and the importance of having an audience that cares (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 behaviour (*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 characterised by elements of negative affect and defined as a feeling of being out of control, a lack of concentration, inability to maintain focused attention, mental fatigue and distress (Burleson 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 and Brass, 2003).

The concept of affect is often distinguished from that of emotion. Affect usually refers to a broader category than emotion (*e.g.*, including states such as interest or boredom, and phenomena such as motivation. 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 recognise 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.

1.3 Computer categorisation and recognition of emotion

We identify a subset of emotions that we intend to recognise in student behaviour and for which intelligent tutors will provide interventions during learning. This selection of emotion is based on both cognitive and affective analyses. We begin with Paul Ekman's categorisation 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 realise that these emotions are appropriate for a 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 to Ekman's categorisation a cognitive component that is present in educational settings, thus initiating what we call 'cognitive-affective' terms, see Table 1. For each of Ekman's emotion we created a scale, resulting in four orthogonal bipolar axes of cognitive-affect. 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 those categories.

Table 1 Cognitive-affective terms based on human face studies

Ekman's categorisation	*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"

Sources: Ekman *et al.* (1972) and Ekman (1999)

Our methodology is to evaluate learning in classrooms while students work with intelligent tutors and develop models of student affect along with tools that recognise 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).

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 contribution is to dynamically collect cognitive and affective information within classrooms, detect a need for interventions and determine which interventions are most successful for individual students and contexts (*e.g.*, problem or affective state). The tutor then 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.

Table 2 Case studies of students' cognitive-affective mechanisms

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

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 behaviour?
- How accurate are different machine learning methods (*e.g.*, Bayesian Networks, hidden Markov models) at predicting affect from student behaviours?
- 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:

1 affect recognition
2 interventions provided for students in response to affect
3 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 methods and tools that are likely different from techniques that will be used once the field has matured and reached its steady state, *e.g.*, unobtrusive sensors and invisible machine learning techniques to measure student affect online. However, at this early research stage, we use a variety of obtrusive techniques until we can efficiently predict affect with the automatic techniques alone.

In one technique described below we invited trained human observers to label students' affect. Although this technique is labour and time-intensive, it provides several advantages, such as identifying high-level student learning behaviours 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 and Lester, 2006). Before and after completing the tutoring session (a matter of several days) students are presented with well-established emotion instruments to measure long-term changes in their motivation, self-confidence and boredom, see Table 3. These older instruments are used because they have been validated and used by hundreds of people.

Table 3 Independent behavioural variables and dependent variables to measure student affect

Dependent variables ⇒⇒⇒	*Frustration*	*Motivation/flow*	*Confidence*	*Boredom*	*Fatigue*
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?')
Behavioural variables					
Behaviours that help predict the dependent variables	*Sensing data (camera, pressure sensitive chair, skin conductance glove, sensitive mouse)*				
	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

Another obtrusive technique is student self-report, which typically requires interrupting the student during the learning experience or afterwards (via video) to ask about their feelings (Graesser *et al.*, 2007). Both methods can be unreliable as they are obtrusive, time consuming and result in variance in reliability (Picard *et al.*, 2004). 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 behaviour related to poor self-monitoring and/or poor use of meta-cognitive resources.

We triangulate among four techniques (human observations, sensors readings, machine learning and student self-reports) in an attempt to resolve towards agreement with the realisation that we may be far away from realising any consensus. We intend to empirically identify which methods are more successful in the recognition of student affect in specific contexts. We also analyse the dependency of specific behavioural 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 we have used to recognise student affect: human observations, a platform of hardware sensors and machine learning techniques.

2.1 Human observation to recognise affect

Our first experiment 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 behaviour once students are used to the observer's presence (Rodrigo *et al.*, 2009). We trained researchers to conduct unobtrusive quantitative field observations and to note students' behaviour while using intelligent tutors. Observers identified variables that represented emotions and desirable/undesirable states linked to student learning and physical behaviours linked to affect states. Human observation are a useful exploratory strategy since observers can intuitively discern high-level behaviours 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 behaviours that are worth observing and then sensors are used to gather this behavioural data in bulk, see Section 2.2. 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 (Dragon *et al.*, 2008). 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 behavioural and subjective variables. Prior to the experiment, observers studied videos of students using Wayang, an intelligent tutor, (Arroyo *et al.*, 2007), 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 observers moved on to the next student. Because students may experience several behaviours/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 behaviour were observed, as shown in Table 6.

Behavioural 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 behaviour (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, 2007). On-task students might be reading/thinking about the problem, talking to a friend about the problem, or writing a solution on paper.

Emotional indicators

Distinguishing one emotion from another is very difficult, especially using only facial expressions and body movement. For example, research shows that neither frustration nor boredom is clearly distinguished from a neutral emotion 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 analyse physical activity as an expression of arousal, calling values below a baseline low arousal and those above it high 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 high arousal is related to being excited and joyful
- positive valence and low arousal is related to being concentrated or satisfied
- negative valence and high arousal is related to being frustrated or angry
- negative valence and low 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 behaviour. 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 we created another variable, *Desirability Value*, which is both task- and emotion-dependent (on/off-task, valence and arousal), see Table 4. Labelling 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, low 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 many other things, such as time spent on the task; sometimes breaks are important to sustain learning.

Table 4 Valence and arousal indicators and their desirability for learning

Valence	*Arousal*	*On/Off task*	*Example student behaviour*	*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

2.1.2 Results of classroom observations

We computed correlations between emotion indicators and intermediate emotion/task-based state variables and analysed the correlation between these state-based variables and student behaviours (Dragon *et al.*, 2008). Students were detected to be on-task 76% of the time, slightly lower than previous findings regarding off/on-task behaviour with software learning environments (Baker, 2007).

Table 5 shows the frequencies of different emotional states among observed students. 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.

Table 5 Frequency of emotion indicators and desirable learning states

Emotion indicators: valence and arousal	*Frequency*	*Percent (%)*
+ valence and – arousal (concentrated, satisfied)	148	58
+ valence and + arousal (excited, joyful, actively engaged)	85	34
– valence and + arousal (frustrated, angry)	16	6
– valence and – arousal (bored, tired)	5	2
Total	254	100
Desirable state		
Highly desirable	181	73
Not desirable	61	24
Medium desirable	7	3

Correlation between emotion indicators and learning/attitudes

We analysed whether we can use emotional indicators and other state variables to predict learning and motivation, the variables we want to optimise.

- *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* (characterised 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.

- *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 behaviour. 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 behaviour (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 the Desirability Value (*i.e.*, it predicted posttest scores and posttest self-concept in mathematics to a similar extent as did on/off-task behaviour). 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.

Table 6 Correlations between student behaviour and emotion states

Behaviour	*Valence*	*Arousal*	*On task?*	*Desirability Value*	*Talk*
CHAIR MOVEMENT	–.467 (0.46*)	.420 (.000***)	–.140 (.027*)	–.154 (.015*)	–
	(N = 252)	(N = 252)	(N = 249)	(N = 247)	
CHAIR MIDDLE	.148 (.018*)	.107 (090)	–.002 (.974)	–.003 (.967)	–
	(N = 252)	(N = 252)	(N = 249)	(N = 247)	
HEAD MOVE	–.224 (.000***)	.345 (.000***)	–.417 (.000***)	–.435 (.000***)	
	(N = 249)	(N = 249)	(N = 246)	(N = 244)	
HEAD SIDE	–.195 (.002**)	.247 (.000***)	–.325 (.000***)	–.337 (.000***)	–
	(N = 254)	(N = 254)	(N = 251)	(N = 249)	
HEAD MOVE SIDE	–.270 (.000***)	.230 (.000***)	–.422 (.000***)	–.443 (.000***)	–
	(N = 249)	(N = 249)	(N = 246)	(N = 244)	
HEAD MIDDLE	.202 (.000***)	–.186 (.000***)	.427 (.000***)	.436 (.000***)	–
	(N = 254)	(N = 254)	(N = 251)	(N = 249)	
HEAD UP	–.097 (.123)	.062 (.326)	–.214 (.001**)	–.235 (.000***)	–
	(N = 254)	(N = 254)	(N = 251)	(N = 249)	
TALK	–.117 (.064)	.304 (.000***)	–.644 (.000***)	–.628 (.000***)	–
	(N = 251)	(N = 251)	(N = 251)	(N = 249)	
SOUND	–.075 (.248)	.370 (.000***)	–.388 (.000***)	–.379 (.000***)	–
	(N = 242)	(N = 242)	(N = 241)	(N = 239)	
SMILE	–.086 (.185)	.313 (.000***)	–.430 (.000***)	–.420 (.000***)	.485 (.000***)
	(N = 240)	(N = 240)	(N = 237)	(N = 235)	(N = 237)
NEUTRAL	.142 (028*)	–.238 (.000***)	.395 (.000***)	.409 (.000***)	–.285 (.000***)
	(N = 240)	(N = 240)	(N = 237)	(N = 235)	(N = 237)

Notes: *** 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).

Pearson correlations among student behaviour (chair and head positions), emotion indicators (valence and arousal), the Desirability Value and student talk.

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

Correlations between student behaviour and emotion states

Several correlations were discovered (and indicated in shades of grey) among student behaviour (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 do not 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. Lastly, 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 behaviour 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 do not 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 behaviour, and desirable state for learning.

Students indicate off-task behaviour 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 does not 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 behaviour ($R = -.430$ 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 behaviour. 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 (low arousal) and not talking.

2.2 Hardware sensors to recognise student affect

Our second method for recognising student affect is through a research platform of unobtrusive hardware sensors. The computer assesses a constellation of patterns from these sensors and relates them to students' affective state. Clearly sensors cannot 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 behaviour (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 recognise facial muscle movements associated with a smile, and the smile-detection software then helps reason about the probability the person is actually happy.

Recent research has focused on recognising 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 recognisers have been shown to perform comparably to humans trained in recognising facial actions (Cohn *et al.*, 1999). Combining visual information with other modalities can provide improved results (Chen *et al.*, 1998; Kapoor *et al.*, 2004).

Our research platform includes four sensors (Figures 1–4):

1 a facial expression system

2 posture analysis seat

3 pressure mouse

4 wireless skin conductance sensor.

This hardware platform (with the exception of the camera) was manufactured at Arizona State University, in collaboration with MIT, based on validated instruments developed by the Affective Computing group at the MIT Media Lab. Pre-production prototypes of each sensor were developed and 25 sets manufactured for simultaneous use in classrooms in Fall 2008 (Cooper *et al.*, 2009; Arroyo *et al.*, 2009a–b). They were then integrated into the Wayang Intelligent Tutor. Sensors collect constant streams of data in parallel, allowing for much more consistent observation than a human ever could accomplish.

Figure 1 Affective state detection from facial expression analysis (see online version for colours)

Notes: Detection, tracking and affective state recognition from facial expression video provided by Logitech QuickCam STX with MindReader software.

Source: El Kaliouby and Robinson (2005)

Figure 2 Posture analysis seat sensor (see online version for colours)

Source: Burleson (2006, see Section 3.1)

Figure 3 Pressure mouse sensor (see online version for colours)

Source: Burleson (2006, see Section 3.1)

Figure 4 Skin conductance sensor (see online version for colours)

Source: Reynolds and Picard (2004)

2.2.1 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 or 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 behaviour (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 colour 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.

2.2.2 Posture analysis seat sensor

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 analysing movement from this posture analysis chair. The Learning Companion (LC) system discussed in Section 4 used a 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 behaviours to *learn* which behaviours tended to accompany states such as interest and boredom.

2.2.3 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 20 ms intervals from each of the six sensors.

2.2.4 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 (Reynolds and Picard, 2004; 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 towards learning and tends to accompany significant, new, or attention-getting events (Boucsein, 1992).

Information from these four sensors is analysed along with cognitive activities from the tutor (time spent 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 client was located on a separate computer to processes and relay skin conductance data. The client program on each student's computer is initialised 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 synchronise 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).

The four wireless sensors described above were evaluated with nearly 100 students in Fall 2008. Summaries of student physiological activity, in particular data streams from facial detection software, helped to predict more than 60% of the variance of student emotional states (Arroyo *et al.*, 2009a–b; Cooper *et al.*, 2009). Stepwise regression was performed with each of the emotions as the dependent variable, using tutor and sensor features as the independent variables (Cooper *et al.*, 2009). Results from the regression show that the best models for *confident*, *frustrated*, and *excited* came from the examples where all of the sensor data was available and the best model for *interested* came from the subset of examples with mouse data available. To facilitate dynamic feedback to the tutor about student emotion, the available sensor and tutor features were placed into a classifier and reported when a student was likely to report a high value of a particular emotion. To test the efficiency of this idea, we created a classifier from each linear model and identified the results from the best classifier of each emotion in terms of accuracy of its prediction. The accuracy of the emotion *frustrated*, for example was 89% using all sensors and the accuracy of the prediction of *interested* using only the mouse was 72.67% (Cooper *et al.*, 2009).

2.3 Machine learning techniques to recognise affect

Our third and final method for recognising student affect is using machine learning techniques. These techniques are very versatile and have been used with intelligent tutors to answer a variety of questions (Woolf, 2009): Is the student engaged? Is the student motivated? What type of intervention should be tried next? How is learning progress measured? How can student success be recognised? 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 cognition and affect (Conati *et al.*, 2002; Murray and VanLehn, 2000; Baffes and Mooney, 1996; Mayo and 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 a student's inappropriate task strategies, procedural errors, or misconceptions.

We used Bayesian networks to infer affect based on a student's observed problem-solving behaviour and estimations from surveys filled out by prior students (Arroyo and Woolf, 2005). Networks were used to discover links between affect (revealed in a post-survey) and observable behaviour (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' behaviours, attitudes and perceptions exist and correlations between help requests and learning have been shown to be consistent with other authors' findings (Wood and Wood, 1999; Renkl, 2002).

In another study, machine learning techniques were used to show that disengagement negatively correlates with performance gain (Johns and 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). Interventions provided for students were generated based on a probabilistic model consisting of four variables: student proficiency, motivation, evidence of motivation and response to a problem (Johns and Woolf, 2006). Motivation was represented as a dynamic variable that changed during a session as students became more or less engaged with the material. Motivation was modelled as a dynamic, discrete variable and proficiency as a static, continuous variable. These assumptions are based on a student's tendency to exhibit different behavioural patterns over the course of a tutoring session. We investigated three types of motivation: motivated, unmotivated (abusing hints) and quickly guessing. This tutor predicted the probability of a correct student response with up to 75% accuracy (Johns and Woolf, 2006). It was tested dynamically with high school students using the Wayang tutor (Arroyo *et al.*, 2007), 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 modelled as a dynamic, discrete variable and proficiency as a static, continuous variable. These assumptions are based on a student's tendency to exhibit different behavioural patterns over the course of a tutoring session.

3 Interventions that respond to students' cognitive-affective state

The second area of this research is to evaluate interventions that are provided for students based on their cognitive-affective state. Interventions adapted to both a student's cognition and affect can be powerful. 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, behaviour 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 behavioural variables and which interactions between variables impact student behaviour 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 provide for the student.

3.1 Intelligent tutors

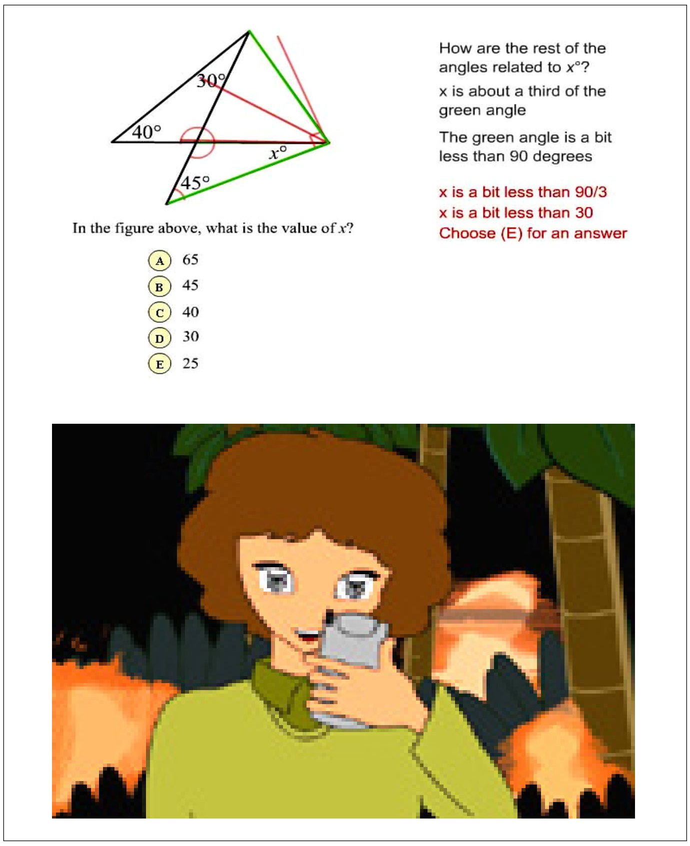
All four affect-recognition methods described above (human observation, sensor readings, machine learning and students self-report) have been evaluated in Wayang Outpost, an intelligent tutor that infers a student's cognitive skills and reasons about which type of hints are best to present in each context. This tutor teaches mathematics (geometry, statistics) and prepares students for standardised state exams[1] (Arroyo *et al.*, 2004; 2005; 2007). The theme and setting of the tutor is a research station on the island of Borneo and the tutor 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 two 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. 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 customises the choice of hint type for individual students based on their cognitive profile, gender, spatial ability, and mathematics fact retrieval speed.

The tutor facilitates the task of logging, pre/post-testing and data collection. Decision making is performed with a database-backed Java servlet with a Flash interface. Students are presented with customised 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 5 Wayang outpost (see online version for colours)



Notes: 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).

We implemented and evaluated two animated affective agents that work with students as LCs, Figures 6 and 7 (Arroyo *et al.*, 2009b) 'Jake' and 'Jane' are amusing and friendly study partners who offer advice and encouragement while reflecting on the range of the student's own emotions. Affective characters are useful as they act out the student's emotion and express full sentences of cognitive, meta-cognitive and emotional feedback, outlined in Table 7. These LCs are non-intrusive – they work on their own computers to solve the problem and react only after the student has answered the question. They mirror and animate student emotion (acting frustrated, bored, or confused). Both agents were extended to multiple ethnicities (Hispanic and Black) by modifying their face module, hair texture and skin colour.

Figure 6 Affect-aware agents are integrated in the Wayang Geometry Tutor (see online version for colours)

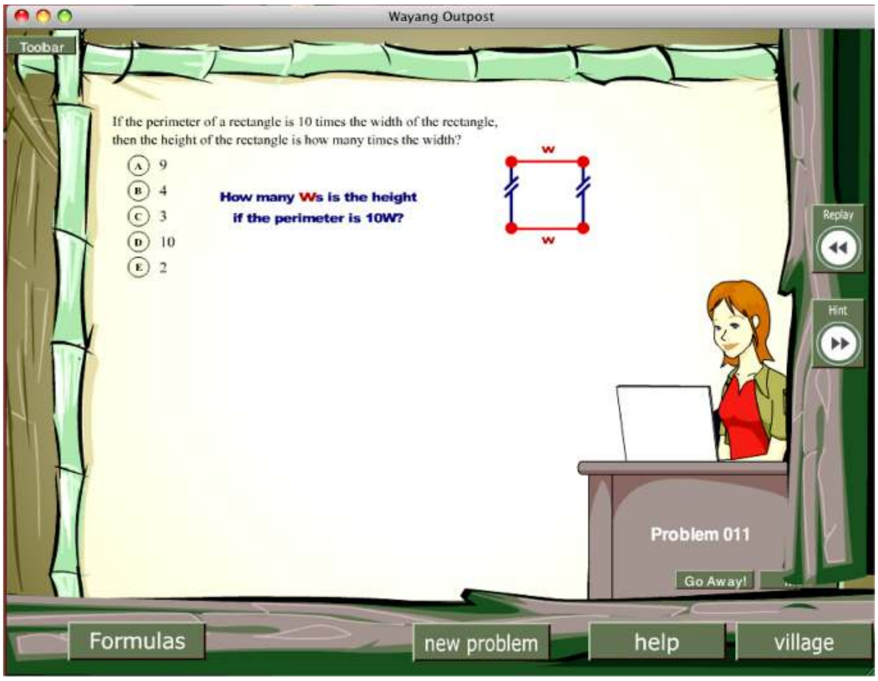
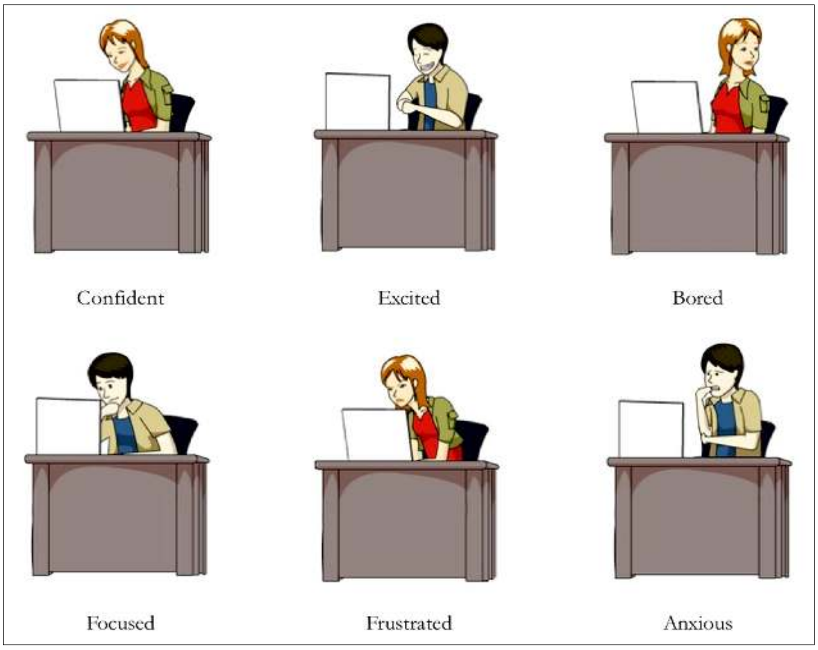


Figure 7 Emotional pedagogical agents show a range of emotion (see online version for colours)



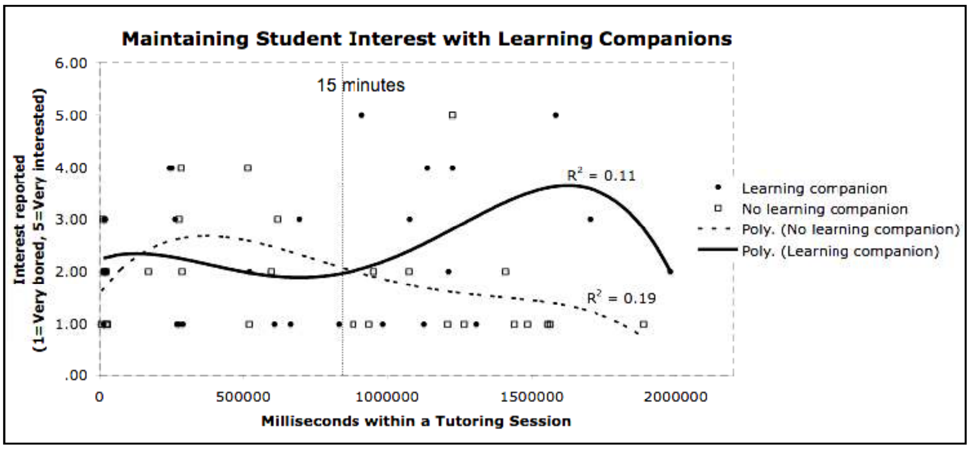
Note: Pedagogical agents act out their emotions and talk with the student expressing full sentences of cognitive, meta-cognitive and emotional feedback, see Table 7.

Table 7 Responses of emotional animated agents

Agent's goal	*Example agent intervention*
Mirror student emotional state visually, as a way to *empathise* 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.	*Agent says* "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.	*Agent says* "Students sometimes get bored with this problem. Are you bored?"
Acknowledge student emotion if it is negative. Provide a helpful hint.	*Agent says* "Some students are frustrated by this problem." "Let's look at some similar worked out problems."
Meta-cognitive response about students' progress and about good learning habits.	*Agent says* "Congratulations! You are getting more questions right than before. Do you see that from the chart?"

We measured the impact of these LCs on student motivation and achievement and integrated controlled exploration of their communicative factors (facial expression and mirroring postures) as the student/agent relationship developed (Arroyo *et al.*, 2009b). Students frequently bring baggage of negative attributions in their self-perception of their mathematics ability. If a tutor recognises that a student is frustrated and supports her or him, the student may persist longer and move beyond frustration. Characters were perceived as mentors, someone who is together with the student against the computer, but who is more knowledgeable than the student most of the time (not always) both cognitively and emotionally (Woolf *et al.*, 2009). Cognitive, meta-cognitive and emotional feedback was outlined in the form of messages, *e.g.*, that attribute failure to something different than lack of inherent ability and empathise with students to help them cope with frustration and anxiety. This allowed us to study the benefits of feedback at key moments of student disengagement and frustration. These agents were evaluated in the classroom along with the sensors described in Section 2.2.

Figure 8 Learning companions make mathematics seem more interesting



Notes: Students with LCs reported more interest in mathematics sessions. Lines represent best-fit curves, which in this case are fourth degree polynomials.

Empirical studies showed that students who used LCs increased their math value (*e.g.*, questions such as 'Mathematics is an important topic'), self-concept (*e.g.*, 'I am good in mathematics') and mastery orientation (Arroyo *et al.*, 2009b). Students frequently become more bored (less interested) towards the end of any instructional session. Yet the student using LCs maintained higher levels of interest and reduced boredom after 15 min of tutor use, see Figure 8. They reported a higher mean confidence, interest and excitement. Despite the fact these results were not significant, this relative advantage for LCs indicates that they might alleviate students' boredom as the session progresses.

3.2 Interventions based an a student's affective state

The Wayang intelligent tutor used a variety of heuristic strategies to respond to student affect (providing text messages, mirroring student actions). For example, policies such as agent responses listed in Table 8 were applied when a specific emotional state was detected. Machine learning optimisation 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.

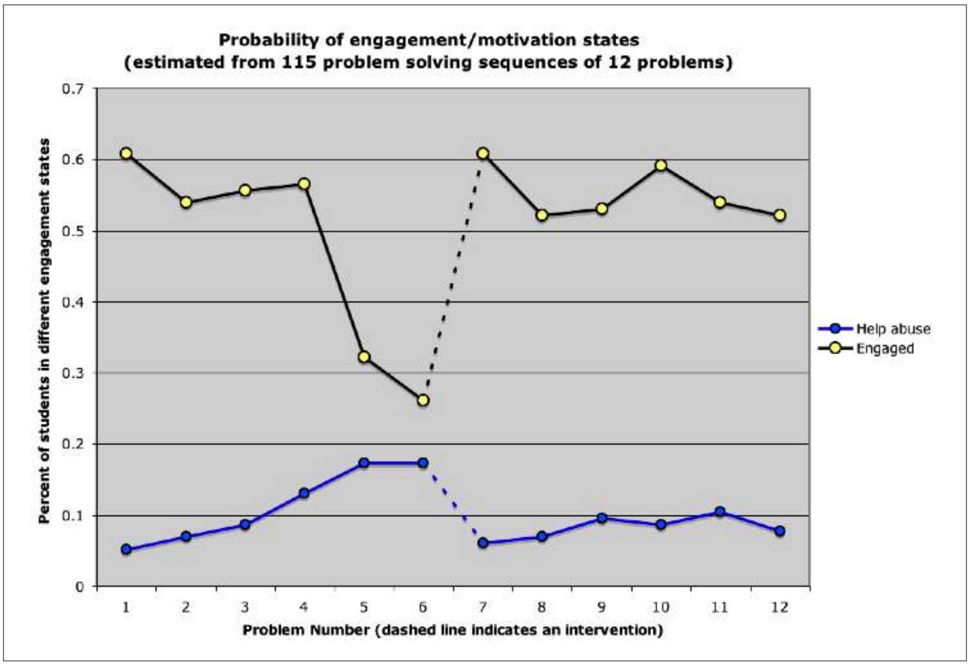
Table 8 Interventions generated by a pedagogical agent based on student affect

High student frustration	*Low student motivation*	*Low student confidence*
Agent looks concerned and provides an empathetic responses: "That was frustrating. Let's move to something easier." It gives student control: "Would you like to choose the next problem? What kind of problem would you like?"	Agent mirrors low motivation and changes its voice, motion and gestures; it may present graph, hints, adventures.	Agent provides encouragement; links performance to student effort and attributes failure to external issue (hard problem) and success to internal issues (you are doing great).
Boredom because student cannot do the work	*Boredom because work is too easy*	*Fatigue*
Agent moves to an easier topic and identifies material that the student can accomplish.	Agent mirrors boredom and increases the challenge level of the activity; it provides empathy messages: "Maybe this is boring? Would you like to move to something more challenging?"	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.

However, the interventions listed in Table 8 need to be evaluated with numerous students in a variety of contexts. For example, mirroring student emotion (see Section 4) 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 her frustration. Rather, a look of concern, or appearing subdued in response, and certainly not smiling, may be an appropriate response to frustration.

We used offline unsupervised learning to help the system learn optimal policies from student data. The following are a few examples of 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 behaviour was recognised after the 6th problem, see Figure 9, top graph. Interventions included either performance graphs with accompanying messages or tips that suggested more productive learning behaviour. 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 behaviour based on digital intervention. Once interventions were presented on-target engaged student behaviour returned (top line) and hint abuse (quickly asking for hints) subsided.

Figure 9 Students became motivated by an intervention (see online version for colours)



Notes: Student engagement declined for six problems (top). An intervention was presented after the sixth problem and then engagement improved. Help abuse (defined as quickly asking for hints to see the answer) was reduced after the intervention.

In the current research, we measure the impact 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 (Stuck). Once a particular emotion is teased out, this information is synchronised with the learning tools to train classifier algorithms. We are investigating interventions for students who exhibit self-confidence, frustration, boredom and self-concept.

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?').

We are using the hardware/software research platform described above to measure the impact of interventions on students who appear frustrated (feelings, thoughts, and behaviours 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 behaviour that could be frustration (arousal and hyperactive behaviour) or boredom (gaming but with low arousal). In conjunction with a student's activity behaviour (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.

We are investigating student self-concept (students' assessment of their own performance in a discipline). 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 implementation and evaluation of emotional embodied animated pedagogical agents, as discussed above in Section 3.1. 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 the curriculum for the individual. Many research issues remain to be addressed. Do human-like LCs (that actually help the student in the learning process) improve student's self-concept and attitudes towards the topic? Does the presence of LCs affect students' learning? Are LCs that resemble a student's gender/ethnicity more effective? This section discusses pedagogical agents, their potential impact on learning and highlights agents we have developed.

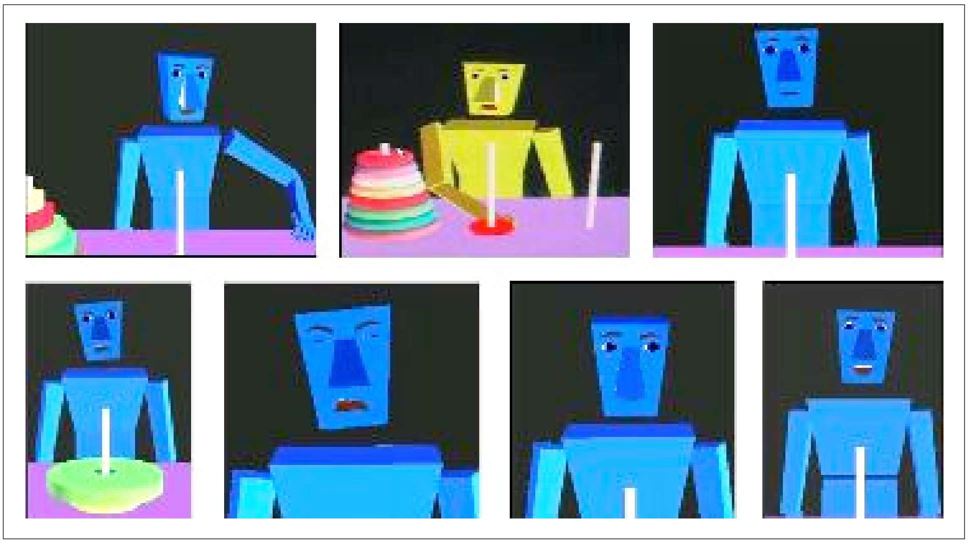
Learning is enhanced when human empathy or support is present (Graham and Weiner, 1996; Zimmerman, 2000). The presence of someone who cares, or at least appears to care, can be motivating (Wentzel, 1997). Various studies have linked interpersonal relationships between teachers and students to motivational outcomes over the long term (Picard *et al.*, 2004; Pianta, 1992; Wentzel and Asher, 1995). Can this noted 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 an apology (Prendinger *et al.*, 2003; Prendinger and Ishizuka, 2005) and relational skills improve long-term ratings of caring, trust, respect, desire to keep working (Bickmore and Picard, 2004).

Animated agents often engage students with dramatic graphics and dynamic colourful animations, see Figures 6, 7 and 10. The added advantage of integrating agents within intelligent tutors is to provide a controllable level of challenge for students facing problems (Burelson and Picard, 2004; 2007). Intelligent tutors adjust material to challenge the individual student. Csikszentmihalyi (1990) 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.

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 do not feel embarrassed if they ask the same question over and over again (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.

In a separate experiment, we developed a LC that automatically recognised and responded to student frustration (Burleson, 2006; Kapoor *et al.*, 2007). The agent was a collaborator on the side who solved the Towers of Hanoi puzzle with the child, Figure 10. This companion helped children to improve their strategies for solving the puzzle. 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 that might indicate a successful learning experience (Burleson, 2006; Burleson and Picard, 2007). One goal was to identify students who encountered frustration and to teach them how to persevere and increase their ability and desire to engage in learning.

Figure 10 Affective learning companion (see online version for colours)



Source: From Burleson (2006)

Two non-verbal interactions conditions were developed: hardware sensor driven 'mirroring' interventions and pre-recorded interventions. Unobtrusive multimodal real-time sensors, the predecessors of those described in Section 2.2, were used to sense a student's affective state and were coupled with the LC. The companion mirrored non-verbal social behaviours that influence persuasion, liking, and social rapport and responded to frustration with empathetic or task-support dialogue. This research sought to provide students with awareness of their affective state, help them understand failure and develop motivation to move onward.

The system used a Posture Analysis Seat, developed for medical and automotive applications (Burleson, 2006; Tekscan, 1997) and used pattern recognition techniques to *learn* which student behaviours 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. Classifier algorithms predicted student frustration with 79% accuracy (Kapoor *et al.*, 2007). 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 behaviours (Mota and Picard, 2003). Both sets of results are conservative, as the system was trained on a small set of data.

5 Discussion and future work

This article described a variety of methods used in intelligent tutors to recognise 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 behaviours linked to emotional states. We described correlations between low-level observations (*i.e.*, chair movement) and higher-level observations (on-off task behaviour) 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 identified variables that are useful predictors of learning and affective outcomes and wireless sensors that predict affective states related to learning.

The research described here provides a collection of models, tools and metaphors to understand student affect. Such methods need further improvement, *e.g.*, we need to integrate a wide variety of behavioural information from hardware sensors and to increase the accuracy of inferences about affect and 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 they 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. For example, 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 hope to recognise and help students cope with states of negative valence and to support their return to on-task behaviour.

Another goal is to support a student's meta-affective state, or reflection about their emotion. For example, we intend to build tutors that generate long-term pedagogical and emotional decisions and view a series of student behaviours, not simply a single-shot action. Affect recognition can significantly improve a tutor's long-term planning if teaching is sometimes directed at helping students move beyond Stuck into Flow.

Additionally we are evaluating the impact of the presence of gendered characters on student motivation and achievement within learning environments. 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 theories about agents 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 hope to evaluate the hypothesis 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).

Acknowledgements

This research was funded in part by awards from the National Science Foundation and the US Department of Education:

- NSF 0705554, IIS/HCC *Affective Learning Companions: Modelling and supporting emotion during teaching*, Woolf and Burleson (PIs) with Arroyo, Barto, and Fisher
- NSF 0411776, ROLE, *Learning to Teach: The Next Generation of Intelligent Tutor Systems*, Woolf (PI) with Arroyo, Barto, Mahadevan and Fisher
- US Dept. of Educ. #R305a0664 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.

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. We thank Jerry Chan and William Ryan for development of the animated agents and Kasia Muldner and Robert Christopherson for helpful comments on this paper.

References

Ambady, N. and Rosenthal, R. (1992) 'Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis', *Psychological Bulletin*, Vol. 111, pp.256–274.

Ames, C. (1992) 'Classrooms: goals, structures, and student motivation', *Journal of Educational Psychology*, Vol. 80, pp.260–267.

Arroyo, I., Beal, C.R., Murray, T., Walles, R. and Woolf, B.P. (2004) 'Web-based intelligent multimedia tutoring for high stakes achievement tests', in J.C. Lester, R.M. Vicari and F. Paraguaçu (Eds.) *Proceedings of the Intelligent Tutoring Systems, 7th International Conference, ITS 2004*, Springer, pp.468–477.

Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K. and Christopherson, R. (2009a) 'Emotion sensors go to school', *Proceedings of the International Conference on Artificial Intelligence in Education*, IOS Press, July, pp. 17–24.

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S. and Woolf, B. (2007) 'Repairing disengagement with non invasive interventions', *International Conference on Artificial Intelligence in Education*, IOS Press, pp.195–202.

Arroyo, I. and Woolf, B. (2005) 'Inferring learning and attitudes from a Bayesian network of log file data', *Proceedings of the Twelfth International Conference on Artificial Intelligence in Education*, IOS Press, pp.33–40.

Arroyo, I., Woolf, B. and Beal, C. (2005) 'Gender and cognitive differences in help effectiveness during problem solving', *Technology, Instruction, Cognition and Learning*, Vol. 3, No. 1.

Arroyo, I., Woolf, B.P., Royer, J.M. and Tai, M. (2009b) 'Affective gendered learning companion', *Proceedings of the International Conference on Artificial Intelligence and Education*, IOS Press, pp.41–48.

Ashford, S.J., Blatt, R. and VandeWalle, D. (2003) 'Reflections on the looking glass: a review of research on feedback-seeking behavior in organizations', *Journal of Management*, Vol. 29, pp.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*, Vol. 96, No. 3, pp.523–535.

Baddeley, A. (1986) *Working Memory*, Oxford Psychology Series No. 11, Oxford, UK: Clarendon Press.

Baffes, P. and Mooney, R. (1996) 'Refinement-based student modeling and automated bug library construction', *Journal of Artificial Intelligence in Education*, Vol. 7, No. 1, pp.75–116.

Baker, R., Corbett, A.T. and 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*, pp.1059–1068.

Bandura, A. (1977) 'Self-efficacy: toward a unifying theory of behavior change', *Psychological Review*, Vol. 84, pp.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 and K. Obermayer (Eds.) *Advances in Neural Information Processing Systems*, Cambridge, MA: MIT Press, Vol. 15.

Bickmore, T. and Picard, R.W. (2004) 'Establishing and maintaining long-term human-computer relationships', *Transactions on Computer-Human Interaction*, Vol. 12, No. 2, pp.293–327.

Block, J. (1995) 'On the relation between IQ, impulsivity and deliquency', *Journal of American Psychology*, Vol. 104, pp.395–398.

Boucsein, W. (1992) *Electrodermal Activity*, New York: Plenum Press.

Bransford, J.D., Brown, A.L. and Cocking, R.R. (Eds.) (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: National Academy Press.

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, http://affect.media.mit.edu/pdfs/06.burleson-phd.pdf.

Burleson, W. and 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*, 31 August, Maceio – Alagolas, Brasil.

Burleson, W. and Picard, R.W. (2007) 'Evidence for gender specific approaches to the development of emotionally intelligent learning companions', Special issue on Intelligent Educational Systems, *IEEE Intelligent Systems*, Vol. 22, No. 4, pp.62–69.

Chen, L.S., Huang, T.S., Miyasato, T. and 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. and Kanade, T. (1999) 'Automated face analysis by feature pont tracking has high concurrent validity with manual FACS coding', *Psychophysiology*, Vol. 36, pp.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*, Vol. 12, No. 4, pp.371–417.

Cooper, D.G., Arroyo, I., Woolf, B.P., Muldner, K., Burleson, W. and Christopherson, R. (2009) 'Sensors model student self concept in the classroom', Submitted to the *International Conference on User Modeling and Adaptive Presentation*, June.

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', Special issue on Intelligent Educational Systems, *IEEE Intelligent Systems*, Vol. 22, No. 4, pp.53–61.

Dragon, T., Arroyo, I., Woolf, B.P., Burleson, W., El Kaliouby, R. and Eydgahi, H. (2008) 'Viewing student affect and learning through classroom observation and physical sensors', *Intelligent Tutoring Systems*, pp.29–39.

Dweck, C.S. (1986) 'Motivational processes affecting learning', *American Psychologist*, Vol. 41, pp.1040–1048.

Dweck, C.S. (1999) *Self-Theories: Their Role in Motivation, Personality and Development*, Philadelphia: Psychology Press.

Dweck, C.S. and Leggett, E.L. (1988) 'A social-cognitive approach to motivation and personality', *Psychological Review*, Vol. 95, pp.256–273.

Ekman, P. (1999) *Facial Expressions*, New York: John Wiley & Sons Ltd.

Ekman, P., Friesen, W.V. and Ellsworth, P. (1972) *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*, New York: Pergamon Press.

El Kaliouby, R. (2005) 'Mind-reading machines: automated inference of complex mental states', Unpublished PhD thesis, University of Cambridge.

El Kaliouby, R. and 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, R.W. (2006) 'An exploratory social-emotional prosthetic for autism spectrum disorders', *International Workshop on Wearable and Implantable Body Sensor Networks*, Cambridge, MA: MIT Media Lab, p.3.

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. and Brass, D. (2003) 'Exploring the social ledger: negative relationships and negative asymmetry in social networks in organizations', Special issue: Building Effective Networks, *Academy of Management Review*, Chicago, Illinois.

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

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

Graham, S. and Weiner, B. (1996) 'Theories and principles of motivation', in D. Berliner and R. Calfee (Eds.) *Handbook of Educational Psychology*, New York: Macmillan, pp.63–84.

Isen, A.M. (2000) 'Positive affect and decision making', in M. Lewis and J. Haviland (Eds.) *Handbook of Emotions*, New York: Guilford.

Johns, J., Mahadevan, S. and Woolf, B. (2006) 'Estimating student proficiency using an item response theory model', in M. Ikeda, K.D. Ashley and T.W. Cahn (Eds.) *ITS 2006, Lecture Notes in Computer Science*, Berlin-Heidelberg: Springer-Verlag, pp.453–462.

Johns, J. and Woolf, B. (2006) 'A dynamic mixture model to detect student motivation and proficiency', *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)*, Menlo Park, CA: AAAI Press, pp.2–8.

Kapoor, A., Burleson, W. and Picard, R. (2007) 'Automatic prediction of frustration', *International Journal of Human Computer Studies 2007*, Vol. 65, No. 8, pp.724–736.

Kapoor, A. and 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, DC.

Kapoor, A., Picard, R.W. and Ivanov, Y. (2004) 'Probabilistic combination of multiple modalities to detect interest', Paper presented at the *Int. Conf on Pattern Recognition*, Cambridge, England.

Keller, J.M. (1987) 'Strategies for stimulating the motivation to learn', *Performance and Instruction*, Vol. 26, No. 8, pp.362–632.

Keller, J.M. (Ed.) (1983) *Motivational Design of Instruction*, Hillsdale, NJ: Erlbaum.

Klein, J., Moon, Y. and Picard, R.W. (2002) 'This computer responds to user frustration: theory, design, results, and implications', *Interacting with Computers*, Vol. 14, No. 2, pp.119–140.

Kleinginna, P.R., Jr. and Kleinginna, A.M. (1981) 'A categorized list of emotion definitions, with suggestions for a consensual definition', *Motivation and Emotion*, Vol. 5, No. 4, pp.345–379.

Koedinger, K.R., Anderson, J.R., Hadley, W.H. and Mark, M.A. (1997) 'Intelligent tutoring goes to school in the big city', *International Journal of Artificial Intelligence in Education*, Vol. 8, No. 1, pp.30–43.

Kort, B., Reilly, R. and 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. and Chabay, R.W. (1988) 'Socialising the intelligent tutor: bringing empathy to computer tutors', *Learning Issues for Intelligent Tutoring Systems*, pp.242–257.

Lepper, M.R. and Hodell, M. (1989) 'Intrinsic motivation in the classroom', in C. Ames and R.E. Ames (Eds.) *Research on Motivation in Education*, New York: Academic Press, Vol. 3, pp.73–105.

Lester, J., Converse, S., Kahler, S., Barlow, T., Stone, B. and Bhogal, R. (1997) 'The persona effect: affective impact of animated pedagogical agents', Paper presented at the *Conference on Computer-Human Interfaces*, New York, New York.

Mandler, G. (1984) *Mind and Body: Psychology of Emotion and Stress*, New York, NY: WW Norton & Co.

Marsella, S.C. and Johnson, W.L. (2003) 'Interactive pedagogical drama for health interventions', in U. Hope, F. Verdejo and J. Kay (Eds.) *Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies*, Amsterdam: IOS Press, pp.341–348.

Mayo, M. and Mitrovic, A. (2001) 'Optimizing ITS behavior with bayesian networks and decision theory', in W. van der Linden and R. Hambleton (Eds.) *International Journal of Artificial Intelligence in Education*, 1997 edition, New York, NY: Springer-Verlag, Vol. 12, pp.124–153.

McDaniel, B.T., D'Mello, S.K., King, B.G., Chipman, P., Tapp, K. and Graesser, A.C. (2007) 'Facial features for affective state detection in learning environments', in D.S. McNamara and J.G. Trafton (Eds.) *Proceedings of the 29th Annual Cognitive Science Society*, Austin, TX: Cognitive Science Society, pp.467–472.

McQuiggan, S. and Lester, J. (2006) 'Diagnosing self-efficacy in intelligent tutoring systems: an empirical study', in M. Ikeda, K. Ashley and T.W. Chan (Eds.) *Eighth International Conference on Intelligent Tutoring Systems*, Jhongli, Taiwan.

Merriam-Webster (2009) *Merriam-Webster Online*, http://www.merriam-webster.com/dictionary/motivation.

Mota, S. and 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. and Helsel, F.I. (2007) 'Improving basic mathematics instruction: promising technology resources for students with special needs', *Technology in Action*, Vol. 2, No. 5, pp.1–6.

Murray, R.C. and 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', *Perspectives on Cognitive Science*, Hillsdale, NJ: Erlbaum, pp.265–295.

Ortony, A., Clores, G.L. and 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*, Vol. 66, pp.543–578.

Piaget, J. and 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*, San Francisco: Jossey-Bass, Vol. 57.

Picard, R.W. (2006) 'Building an affective learning companion, invited keynote talk', Paper presented at the *International Conference of Intelligent Tutoring Systems*, Taiwan.

Picard, R.W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D. and Strohecker, C. (2004) 'Affective learning – a manifesto', *BT Technical Journal*, Vol. 2, No. 4, pp.253–269.

Prendinger, H. and Ishizuka, M. (2005) 'The empathic companion: a character-based interface that addresses users' affective states', *Applied Artificial Intelligence*, Vol. 19, Nos. 3–4, pp.267–285.

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.

Reeves, B. and 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*, Vol. 12, pp.529–556.

Reynolds, C. and Picard, R.W. (2004) 'Affective sensors, privacy, and ethical contracts', *Proceedings of CHI*, Vienna, Italy, April, pp.1103–1106.

Rodrigo, M.M.T., Baker, R.S., Jadud, M.C., Amara, A.C.M., Dy, T., Espejo-Lahoz, M.B.V., Lim, S.A.L., Pascua, S.A.M.S., Sugay, J.O. and Tabanao, E.S. (2009) Affective and behavioral predictors of novice programmer achievement', *Proceedings of the 14th ACM-SIGCSE Annual Conference on Innovation and Technology in Computer Science Education*, pp.156–160.

Schunk, D.H. (1989) 'Self-efficacy and achievement behaviors', *Educational Psychology Review*, Vol. 1, pp.173–208.

Shute, V. and Psotka, J. (1995) 'Intelligent tutoring systems: past, present, and future', *Handbook of Research on Educational Communications and Technology*, pp.570–600.

Shute, V.J. (2008) 'Focus on formative feedback', *Review of Educational Research*, Vol. 78, No. 1, pp.153–189.

Slater, D. (2000) 'Interactive animated pedagogical agents', *International Conference on Computers and Human Interaction*, Atlanta, Georgia.

Strauss, M., Reynolds, C., Hughes, S., Park, K., McDarby, G. and Picard, R.W. (2005) 'The handwave bluetooth skin conductance sensor', *Affective Computing and Intelligent Interaction Conference*, Beijing, China, pp.699–706.

Sweller, J., van Merrienboer, J. and Paas, F. (1998) 'Cognitive architecture and instructional design', *Educational Psychology Review*, Vol. 10, pp.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 Processs*, Cambridge, MA: Harvard University Press (Originally published 1930, Oxford University Press, New York).

Webster (1984) *Webster's New World Dictionary*, Simon and Schuster.

Wentzel, K. (1997) 'Student motivation in middle school: the role of perceived pedagogical caring', *Journal of Educational Psychology*, Vol. 89, No. 3, pp.411–419.

Wentzel, K. and Asher, S.R. (1995) 'Academic lives of neglected, rejected, popular, and controversial children', *Child Development*, Vol. 66, pp.754–763.

Wood, H. and Wood, D. (1999) 'Help seeking, learning, and contingent tutoring', *Computers and Education*, Vol. 33, pp.153–169.

Woolf, B. (2009) *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-learning*, San Francisco, CA: Elsevier Inc., Morgan Kauffman.

Wundt, W. (1902) *Outlines of Psychology*, London: Williams & Norgate.

Zacks, J.M., Tversky, B. and Iyer, G. (2001) 'Perceiving, remembering, and communicating structure in events', *Journal of Experimental Psychology, General*, Vol. 130, pp.29–58.

Zimmerman, B.J. (2000) 'Self-efficacy: an essential motive to learn', *Contemporary Educational Psychology*, Vol. 25, pp.82–91.

Note

1 See http://althea.cs.umass.edu/wayang/wayangindex.html.