Modeling Trainee Affective and Cognitive State Using Low Cost Sensors

Meredith Carroll, Christina Kokini, Roberto Champney
Design Interactive, Inc., Oviedo, FL
meredith, christina, roberto@designinteractive.net

Sven Fuchs
Fraunhofer FKIE, Wachtberg, Germany
sven.fuchs@fkie.fraunhofer.de

Robert Sottilare, Benjamin Goldberg
U.S. Army Research Laboratory
Human Research and Engineering Directorate
Orlando, FL
robert.sottilare,
benjamin.s.goldberg@us.army.mil

ABSTRACT

Given the limited time available for training and increased emphasis on self-directed learning in the military, it is essential to develop methods to improve training effectiveness with minimal impact to instructor resources. Training practitioners have attempted to achieve this through the incorporation of automated systems such as Intelligent Tutoring Systems to augment instructor time by emulating human tutors. However, these systems have yet to reach training effectiveness levels that rival those of human tutors. A review of the literature indicates a significant share of the performance gap between computer-based tutoring and human tutors lies in the ability of the humans to be aware of and responsive to the learner's cognitive/affective states. Even so, human tutors have only limited perception of the trainee's cognitive/affect states. When instantiated in a training system, perceptive abilities may allow computer-based tutors to go beyond the abilities of human instructors. It is thus imperative that the trainee model within these systems incorporate both the trainee's performance as well as a diagnosis of their affective and cognitive state.

This paper presents a theoretical framework for the creation of a trainee model that incorporates affective and cognitive state of trainees based on inputs from low-cost, non-intrusive sensors. This framework has theoretical foundations in learning science and physiological measurement and could drastically increase the diagnostic capability of current intelligent training systems. Implementation of this framework could transform adaptive training based on cognitive/affective states from a cost prohibitive endeavor to a goal well within reach. It is hypothesized that a trainee model based on lower cost sensors will account for a significant portion of the variance measured by benchmark sensors/systems that prove expensive or invasive.

ABOUT THE AUTHORS

Meredith B. Carroll, PhD is a Senior Research Associate at Design Interactive, Inc. and has been involved in design, development and evaluation of performance assessment tools and virtual training tools for the Office of Naval Research, the Air Force Research Laboratory, and the Army's Research, Development and Engineering Command. Her work focuses primarily on individual and team performance assessment, including physiological and behavioral measurement, performance diagnosis and training remediation through feedback and training adaptation. She has also performed extensive work conducting task analyses, designing virtual training environments and performance assessment tools and conducting training effectiveness evaluations. Her research has focused on human/team performance and training in complex systems in aviation and military domains, with focuses on perceptual skills and decision making. She received her B.S. in Aerospace Engineering from the University of Virginia, her M.S. in Aviation Science from Florida Institute of Technology and her Ph.D. in Applied Experimental and Human Factors Psychology from the University of Central Florida.

Christina Kokini is a Research Associate at Design Interactive, and has been involved in design, development and evaluation of virtual training tools for the Office of Naval Research, and the Army's Research, Development and Engineering Command. Her work focuses on training system design, development, and usability, including designing functionality into products to support user requirements, as well as conducting usability and training effectiveness evaluations. She holds a Master's degree from Penn State University in Industrial Engineering with a Human Factors Option, where her research focused on the direct effect of contextual characteristics on the perceived usability of a

product. She also has a Bachelor's degree from Purdue University in Industrial Engineering. Prior to working at Design Interactive, she had a variety of experience in human factors, product design, manufacturing, and logistics, working for companies such as IBM, GE, Caterpillar, and SA Technologies.

Roberto Champney, PhD is a Senior Research Associate at Design Interactive, Inc. and is currently supporting training system design and evaluation for the office of Naval Research's Human Performance, Training and Education (HPT&E) Program. His work focuses on Task Analysis, design and evaluation of Training Environments and Interactive Systems, Human Performance Assessments and Training Management tools and conduction of training effectiveness evaluations. His research has focused on human performance and training in military domains, design and assessment of interactive systems, usability and user experience involving human emotions. Roberto holds an M.S. in Human Engineering and Ergonomics, and a Ph.D. in Industrial Engineering from the University of Central Florida.

Sven Fuchs is a researcher in the Human-Systems Engineering group at the Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE) in Wachtberg, Gerrmany, where he currently develops methods and tools for optimizing communication processes on naval platforms. Prior to joining FKIE, he was a Senior Research Associate at Design Interactive, Inc. where his research focused on multimodal design science and development of dynamic adaptation frameworks and adaptation strategies to optimize operator performance, training effectiveness, and efficiency. He was involved with several SBIR efforts related to Augmented Cognition and adaptive training and also served as a project lead of interface and usability design efforts for a variety of industry clients. Sven co-invented three patent-pending approaches related to human-systems interaction in military environments. Previously, Sven completed an interdisciplinary degree in Computer Science and Media from Flensburg University of Applied Science in Germany and was a Fulbright scholar at DePaul University, Chicago, where he earned an M.S. in Human-Computer. Sven also brings 5 years of experience in the creative industries where he worked in marketing, multimedia production and advertising.

Robert A. Sottilare, Ph.D. is the Associate Director for Science & Technology within the U.S. Army Research Laboratory - Human Research & Engineering Directorate (ARL-HRED). Dr. Sottilare has over 25 years of experience as both a U.S. Army and Navy training & simulation researcher, engineer and program manager. He leads STTC's international program and participates in training technology panels within both The Technical Cooperation Program (TTCP) and NATO. He has a patent for a high resolution, head mounted projection display (U.S. Patent 7,525,735) and his recent publications have appeared in the Journal for Defense Modeling and Simulation, the NATO Human Factors and Medicine Panel's workshop on Human Dimensions in Embedded Virtual Simulation and the Intelligent Tutoring Systems Conference. Dr. Sottilare is a graduate of the Advanced Program Managers Course at the Defense Systems Management College at Ft. Belvoir, Virginia and his doctorate in modeling & simulation is from the University of Central Florida. The focus of his current research program is in machine learning, trainee modeling and the application of artificial intelligence tools and methods to adaptive training environments.

Benjamin Goldberg is a member of the Learning in Intelligent Tutoring Environments (LITE) Lab at the U.S. Army Research Laboratory's (ARL) Simulation and Training Technology Center (STTC) in Orlando, FL. He has been conducting research in the Modeling and Simulation community for the past 3 years with a focus on adaptive learning and how to leverage Artificial Intelligence tools and methods for adaptive computer-based instruction. Mr. Goldberg is a Ph.D. student at the University of Central Florida and holds an M.S. in Modeling & Simulation. Prior to employment with ARL, he held a Graduate Research Assistant position for two years in the Applied Cognition and Training in Immersive Virtual Environments (ACTIVE) Lab at the Institute for Simulation and Training. Recent publications include a proceedings paper in the 2010 International Conference on Applied Human Factors and Ergonomics and a poster paper presented at the 2010 Spring Simulation Multiconference.

Modeling Trainee Affective and Cognitive State Using Low-Cost Sensors

Meredith Carroll, Christina Kokini, Roberto Champney
Design Interactive, Inc., Oviedo, FL
Meredith, Christina, Roberto@designinteractive.net

Sven Fuchs
Fraunhofer FKIE, Wachtberg, Germany
sven.fuchs@fkie.fraunhofer.de

Robert Sottilare, Benjamin Goldberg
U.S. Army Research Laboratory
Human Research and Engineering Directorate
Orlando, FL
robert.sottilare,
benjamin.s.goldberg@us.army.mil

INTRODUCTION

Military organizations within the North Atlantic Treaty Organization (NATO) are emphasizing the need for alternative training solutions to current practices that allow for more self-directed learning and reduce the strain on instructor and training support resources (NATO RTO, 2009). Practitioners have attempted to achieve this through the incorporation of automated computer-based systems such as Intelligent Tutoring Systems (ITS) that aim to personalize instruction based on performance, competency, and the individual needs of the learner (Self, 1999). In essence, they have attempted to emulate the strategies utilized by expert human tutors in the hopes of attaining similar effectiveness as one-to-one tutoring (e.g., Bloom's (1984) "2 Sigma Problem"). ITSs are computer-based instructional environments that apply artificial intelligence (AI) technologies (tools and methods) to deliver personalized training experiences geared to maximize instructional effectiveness. The aim of an ITS is to deliver individualized or tailored feedback and/or content manipulations (e.g., changes in flow or challenge level) based on current and predicted trainee performance, cognition (e.g, readiness to learn) and affect (e.g, confusion, boredom). Through the application of automated diagnosis ITSs allow the instructor to be augmented/replaced with adaptive interventions based on models of user, expert and instructor performance. This facilitates the optimal selection of instructional strategies, and the delivery of timely, focused feedback and future content manipulations with the aim of improving knowledge and skill acquisition.

ITSs have proven to be an effective and reliable tool for improving learning (e.g., VanLehn et al., 2005; Stottler, Harmon, and Michalak, 2001; Sherry, Feary, Polson, and Palmer, 2000) in well-defined domains (e.g., algebra and physics). However, they have yet to reliably reach training effectiveness levels that rival those of human instructors. This, in part, is due to ITSs' inability to monitor and assess a trainee's cognitive and affective state in real-time. A majority of current systems apply performance-based models to assess trainee progression and adapt content based on deviations between trainee interactions and an expert model of task performance. It is believed that a significant portion of the performance gap between current ITSs and human tutors lies in the capability of humans to be aware and responsive to learners' affective states (c.f., Sarrafzadeh, Alexander, Dadgostar, Fan and Bigdel, 2006). By perceiving the influencing factors of cognition and affect on learning, an intelligent training system can truly adapt to the changing needs of the learner.

The importance of real-time state assessment is clear, but instantiating such practices into an ITS is challenging. At its core, human affective state detection is a classification problem and requires collection of data unique to the individual learner (Li and Ji, 2005). There are a number of approaches for collecting affective state data, primarily through physiological and behavioral markers extracted from a source signal. A constraint, however, is such sensing techniques are invasive, they produce noisy unlabeled data, and they are often very costly to implement on a large scale. For ITS technology to be exercised as a practical training alternative it is necessary to first explore the capabilities of low-cost non-intrusive sensors that can inform the affective state of the trainee. The goal of this effort is to develop a framework that integrates low-cost unobtrusive sensors that can feed the trainee model with accurately classified affective and cognitive states that will guide feedback and content manipulations.

STATE OF THE SCIENCE

Training performance is not solely the result of skills and abilities, but also the result of a trainee's cognitive and affective learning state. In fact the best tutors empower the students to work on the learning challenges while the tutor interjects to minimize frustration and confusion (Merrill et. al., 1992). Void of an understanding of these factors, an intelligent training system cannot truly adapt to the changing needs of the learner, ensuring they are not only physically and cognitively ready to learn, but also emotionally ready and are maintained in those states throughout learning.

Affective States and Learning

Literature has established that affective states may influence learning by both enhancing or hindering learning and retention (Small, Dodge, and Jiang, 1996; Burleson and Picard, 2004; Woolf et al., 2009). Affective states or phenomena are composed of emotions, attitudes, moods, and affective traits (Davidson, Scherer, and Goldsmith, 2003). This effort focused on emotions because they are induced by a particular event, and tend to be the most short-lived and the most easily influenced of the affective phenomena, and hence provide the greatest opportunity for diagnosis and mitigation. Emotions can be defined as brief episodes of coordinated changes (brain, autonomic, and behavioral) to facilitate a reaction to a significant event. The complex relationship between emotions and learning can be characterized by the effect emotions have on the learning process (e.g., attention, encoding, recall). For example, boredom leads to lower retention and less ability to apply information (Small, Dodge and Jiang, 1996) and is negatively correlated with learning gains (D'Mello, Graesser, and Taylor, 2007). On the other hand, joy leads to significant increases in intellectual gains and performance (Fredrickson, 1998). Impacts on learning were found for a variety of emotions including anger and anxiety (Woolf et al., 2009; Burleson and Picard, 2004), frustration (McQuiggan et al., 2007), shame (Ingleton, 2000) and surprise (Holland and Gallagher, 2006). Additionally, when the trainee lacks motivation, it hinders creativity and flexibility in problem solving, as well as leading to withdrawal from learning (Woolf et al., 2009). A list of affective states that have been found to impact learning are summarized in Table 1

Table 1. Affective States that Impact Learning

• Anger/ Frustration • Boredom • Confidence • Confusion • Fear/ Anxiety	• Joy • Motivation • Sadness • Shame • Surprise • Wonderment/ Awe

In order to provide a more standardized set of affective states, non-orthogonal states that are actually different intensity levels within the same emotional category were consolidated into one emotion. These consolidations included both anger and frustration as well as fear and anxiety (Scherer, 2005).

Cognitive States and Learning

It has also been proposed that cognitive state monitoring during training may be useful to identify a trainee's readiness to learn (Stevens, Galloway and Berka, 2007). Impact of cognitive state on the learning process has been demonstrated for a variety of states, including engagement, cognitive workload, and drowsiness. For example, high engagement reflects attentional focus (Dorneich et al., 2004). Alternatively, there are a number of prominent states that generally negatively impact training performance since they reduce attentional resources that facilitate learning and retention (e.g., drowsiness leads to a drop in attention (Small, Dodge and Jiang, 1996; Neri, Dinges, and Rosekind, 1997)). Furthermore, mental workload can cause delays in information processing or cause users to ignore or misinterpret incoming information (Ryu and Myung, 2005). Divided attention is associated with reductions in memory performance (Craik et al., 1996). Distraction can result in the acquisition of knowledge that can be applied less flexibly in new situations (Foerde et al., 2006). The cognitive states that have been found to impact learning are summarized in Table 2.

Table 2. Cognitive States that Impact Learning

• Attention
• Distraction
• Drowsiness
• Engagement
• Flow
• Workload

Measurement of States that Impact Learning

It is evident that monitoring and adapting training based on the learner's performance is simply not enough. Intelligent training must take into account learner's affective and cognitive state to ensure training content is adapted accordingly. The first step in achieving this is the measurement of critical affective and cognitive states. However, the challenge with this is that these states are not easily measured objectively.

Traditionally, affective and cognitive states have been measured subjectively through self-report questionnaires such as the Emotional State Questionnaire (ESQ) and the NASA-TLX (to assess workload). However, such methods interrupt the flow of task performance, which can impact emotions themselves (e.g., frustrate a learner who was immersed in the training scenario). Furthermore, as cognitive states and emotions are short-lived and easily influenced, these assessment methods are time sensitive, meaning that if too much time passes between the induction of a particular state and the assessment, a new state may already be experienced. For this reason, and because people do not always accurately understand or remember the emotions they have felt, these methods are not completely reliable. To overcome these challenges, there is a need to incorporate non-invasive measures that can rapidly access these affective and cognitive states.

Measurement of Affective States

Multiple physiological approaches based on numerous sensors have been identified over recent years that can be used to rapidly assess components of affect. For instance, the current state of electroencephalography (EEG) suggests it is effective in measuring the general level of arousal of the brain (Gratton et al., 2008). Eye-tracking can also be used to assess emotional arousal via pupilometry (Partala and Surakka, 2003). Feature extractions from speech and facial expression classifiers have been utilized to assess valence (positive or negative nature of the emotion) and arousal components (Woo Kim, Jin, Fuchs, and Fouad, 2010). Using a web camera to distinguish facial expressions has also been used to determine specific emotions such as anger, disgust, fear, joy, and surprise (Woolf et al., 2009). Electrodermal sensors can be used to measure galvanic skin response (GSR), which has been linked to variations in emotion (Critchley, 2002) and emotional response (Bradley, Moulder, and Lang, 2005), indicative of emotions such as anxiety, frustration (Scheirer et al., 2002), and boredom (Merrifield, 2010). Cardiovascular measures such as heart rate can be used to determine levels of arousal (Jang et al., 2002; Hoover and Muth, 2004), and changes in heart rate have been found to occur during periods of anger, fear, (Lisetti and Nasoz, 2004), and boredom (Merrifield, 2010). A combination of heart rate and variability have further proven useful in discriminating emotional states (Jang et al., 2002) as well as the level of stress induced by specific aspects of a test environment, such as aircraft takeoff and landing (Cacioppo, Berntson, Sheridan, and McClintock, 2000; Kramer, 1991; Wilson, 1992). Blood pressure is another cardiovascular measure that has been used to evaluate emotional response (Roberts and Weerts, 1982). Similarly, respiration rate and volume have been used in the evaluation of stress (Wientjes, 1992). Additionally, changes in posture indicative of boredom (e.g., people tend to lean back when they are bored) can be measured using a pressure sensors embedded on the seat and the back of a chair (Woolf et al., 2009; D'Mello, Chipman and Graesser, 2007).

Measurement of Cognitive States

Physiological sensors have also been used to measure cognitive state. Changes in posture, measured using chair pressure sensors are indicative of levels of engagement in that people tend to lean forward in their seats when they are engaged (Mota and Picard, 2003) or when they are experiencing the state of flow, a state of highly focused engagement where skill level matches challenge level of a task (D'Mello, Chipman, and Graesser, 2007). EEG can be used to assess a person's level of attention and engagement, including cognitive processes such as information-gathering, visual scanning, and sustained attention and EEG engagement indices are associated with increasing demands for visual processing and allocation of attention to both auditory and haptic stimuli (Berka et al., 2007). Furthermore, EEG Gamma-Band Response (GBR) is modulated by attention (Tiitinen et al., 1993). EEG workload indices have shown increases with working memory load and with increasing difficulty level of mental arithmetic and other problem-solving tasks (Berka et al., 2007). Eye-tracking can also be used to assess cognitive workload via blink rates (e.g., Scerbo et al., 2001), pupil amplitude variation (Ahlstrom and Friedman-Bern, 2006), pupil dilation (Pomplun and Sunkara, 2003), and saccade peak velocity (Di Stasi et al., 2010), attention via number of fixations on each area of interest (Hyona, Radach, and Deubel, 2003), and drowsiness via blink rate and blink duration (Ryu and Myung, 2005) and saccade peak velocity (Di Stasi et al., 2010). Electrodermal sensors measure GSR, which has been linked to variations in engagement (Mandryk, 2005). Similarly, cardiovascular measures such as heart rate can be used to determine levels of engagement (Jang et al., 2002; Hoover and Muth, 2004).

These findings provide a foundation for identifying non-intrusive measures that could be used to rapidly assess trainee affective and cognitive state to inform a trainee model.

THEORETICAL FRAMEWORK

This effort aimed to identify low-cost, non-intrusive physiological sensors to measure affective and cognitive state in order to inform the trainee model. Although an ideal trainee model would account for all positive and negative affective and cognitive states, there were several constraints which bound this efforts problem space. First, in order to achieve the greatest benefit, those states that have been shown to have the most significant impact on learning were targeted. Second, in order for the state to be mitigated resulting in enhanced learning, it must be measurable, and hence only states that had validated measures were included. Third, in order to meet our objective of low-cost, non-intrusive measurement, the state must be measureable via sensors that meet these criteria. As such, in defining the theoretical framework, first the affective and cognitive states that provided the greatest opportunity for enhancing learning were identified, followed by determinations of the types of physiological data indicative of these states and availability of low-cost, non-intrusive sensors to collect this data.

Target Affective States

Five affective states were identified from the eleven states in Table 1 as providing significant opportunity for enhancing learning, including 1) anger/frustration, 2) fear/anxiety, 3) boredom, 4) motivation, and 5) confusion (e.g., Woolf et al., 2009; Burleson and Picard, 2004; Craig et al., 2004). There was significant support for inclusion of these five states and additional supporting data is summarized below.

McQuiggan, Lee, and Lester (2007) found that anxiety and frustration divert attention from the task at hand, impeding learning. Furthermore, students who are anxious or angry do not learn as well because they do not take in information efficiently (Woolf et al., 2009; Burleson and Picard, 2004). Alternatively, motivation is positively correlated with learning. If student motivation is sustained throughout periods of disengagement, students can persevere through frustration to a greater extent (Woolf et al., 2009; Burleson and Picard, 2004). Similarly, confusion was found to be significantly positively correlated with learning gains (r=0.33, $p<0.05$; Craig et al., 2004). In fact, Craig et al. (2004) found an effect size on learning of 0.64, observed when confusion was present versus absent, suggesting that some level of confusion is critical for optimal learning. D'Mello, Taylor, and Graesser (2007) found that when learners are confused, they are less likely to become disengaged and transition into boredom, which is significantly negatively correlated with learning (r=-0.39, $p<0.05$; Craig et al., 2004). Boredom leads to lower retention and less ability to apply information (Small, Dodge, and Jiang, 1996).

However, in determining the types of physiological data indicative of these states, only three of these states have shown correlation with data from physiological sensors. For example, Lisetti and Nasoz (2004) found that heart rate values for a fearful participant increased, whereas heart rates decreased when the participant was angry. Heart rate has also been shown to be correlated with boredom (Merrifield, 2010). According to Woolf et al. (2009), facial expressions can be used to detect fear and anger. Additionally, posture has been used to detect frustration (Kapoor et al., 2007) and boredom (D'Mello, Chipman and Graesser, 2007). Given this, three target affective states were identified from the eleven states in Table 1 for inclusion in the trainee model: Anger/Frustration, Fear/Anxiety and Boredom. These states and their impact on learning are summarized in Table 3. It should be noted that some states were excluded due to lack of validated methods to induce the state, and hence inability to validate measurement effectiveness.

Table 3. High Priority Affective States

Affective State	Description
Anger/ Frustration	Negative, high arousal emotion that occurs when an event or the actions of self or other prevent one from achieving a goal (Ortony, Clore, and Collins, 1988)
Fear/ Anxiety	Negative, high arousal emotion that occurs at the prospect of a negative event with consequences to oneself (Ortony, Clore, and Collins, 1988)
Boredom	Negative, low-arousal emotion that occurs when a situation is construed to be monotonous or dull (Merrifield, 2010)

Based on correlations with physiological data found in the literature, it is hypothesized that anger/frustration can be assessed by five sensors, including a heart rate monitor and skin conductance sensor (Lisetti and Nasoz, 2004), pressure mouse and chair pressure sensors (Kapoor et al., 2007), and web camera (Woolf et al., 2009. Fear/anxiety is correlated with data from three sensors: a heart rate monitor (Lisetti and Nasoz, 2004), skin conductance sensor (Schierer et al., 2002), and web camera (Woolf et al., 2009). It is hypothesized that boredom can be assessed by four sensors, including a heart rate monitor and skin conductance sensor (Merrifield, 2010), and chair and mouse pressure sensors (Woolf et al., 2009). Low-cost (i.e., less than $500), non-intrusive (i.e., do not touch the body or sit comfortably on the body, and do not impede task performance) versions of sensors which could assess these states were identified. Because posture is correlated with both anger/frustration (Kapoor et al., 2007) and boredom (D'Mello, Chipman and Graesser, 2007), it was hypothesized that a motion detector that sits in front of the computer screen and detects when an object moves closer or further away may also be correlated with these affective states. Preliminary testing confirmed this hypothesis. Those sensors that provided the greatest opportunities for capturing all target states were selected, including the motion detector. The selected sensors, the states measurable by these sensors and approximate cost are summarized in Table 4.

Table 4. Sensors Used to Measure Affective States

States	Sensor	Cost
Anger/ Frustration, Boredom	Motion Detector	~$100
Anger/ Frustration, Fear/ Anxiety, Boredom	Heart Rate Monitor	~$100
Anger/ Frustration, Boredom	Chair Pressure Sensors	~$200

Target Cognitive States

Three cognitive states were identified from the six states in Table 2 as providing significant opportunity for enhancing learning, including 1) attention, 2) engagement and 3) workload. There was significant support for inclusion of these three states and additional supporting data is summarized below. Divided attention is associated with large reductions in memory performance and small increases in reaction time during encoding and larger increases in reaction time during recall (Craik et al., 1996). Furthermore, lower retention and less learning result when attention is diverted from the task at hand (McQuiggan, Lee, and Lester, 2007; Small, Dodge, and Jiang, 1996). Disengagement is negatively correlated with learning (Woolf, Burelson, and Arroyo, 2007) and performance gains (Johns and Woolf, 2006). High workload has been found to be detrimental to performance improvements (Gonzalez, 2005). Additionally, Mykityshyn, Fisk, and Rogers (2002) found that when older people were taught how to use a blood glucose meter using either a manual or a video, and then given a knowledge test immediately after and a retention test 2 weeks later, the manual group performed worse on the both tests, and reported higher subjective workload ratings than the video group.

Further, all three of these states have shown correlation with data from physiological sensors. Changes in human gamma-frequency oscillations are associated with changes in attention (Jensen, Kaiser, and Lachaux, 2007; Tiitinen et al., 1993). Attention can also be directly measured through the number of fixations on areas of interest (Hyona, Radach, and Deubel, 2003). Berka et al. (2007) created an EEG engagement index that reflects processes that involve information-gathering, visual scanning, and sustained attention and is associated with increasing demands for visual processing and allocation of attention to both auditory and haptic stimuli. There is also evidence that supports correlation between postures and level of engagement (Mota and Picard, 2003). Jensen et al. (2002) found that oscillations in the alpha band increase with increased mental workload. There is also a strong correlation between pupil amplitude variation and the amount of cognitive resources used to perform a task (Ahlstrom and Friedman-Bern, 2006). Furthermore, in visual performance tasks, saccade peak velocity varies with the subject's state of mental workload (Di Stasi et al., 2010). Given this, three target cognitive states were identified from the six states in Table 2 for inclusion in the trainee model: Attention, Engagement and Workload. These states are summarized in Table 5.

Table 5. High Priority Cognitive States

Cognitive State	Description
Attention	Sustained attention, vigilance (Sarter et al., 2001
Engagement	Level of cognitive processes related to information gathering and sensory processing (Berka et al., 2007)
Workload	mental workload, level of cognitive processes related to central executive function (Berka et al., 2007)

Attention can be assessed by two sensors, EEG (Tiitinen et al., 1993) and an eye-tracker (Hyona, Radach, and Deubel, 2003). Engagement can be assessed by EEG (Berka et al., 2007), and potentially chair pressure sensors, based on the correlation between engagement and posture (Mota and Picard, 2003). Workload can also be assessed by EEG (Berka et al., 2007) and an eye-tracker (Di Stasi et al., 2010) Low-cost (i.e., less than $500), non-intrusive (i.e., do not touch the body or sit comfortably on the body, and do not impede task performance) versions of sensors that could assess these states were identified. Those sensors that provided the greatest opportunities for capturing all target states were selected. The selected sensors, the states measurable by these sensors and approximate cost are summarized in Table 6. The low-cost eye-tracker is not an off-the-shelf model; the cost is estimated based on the cost of components such as a camera and IR lights.

Table 6. Sensors Used to Measure Cognitive States

States	Sensor	Cost
Engagement	Chair Pressure Sensors	~$200
Attention, Engagement, Workload	EEG	~$200
Attention, Workload	Eye-tracker	~$500

Informing the Trainee Model

Incorporation of these target states into the trainee model increases the granularity of the model diagnoses, thereby expanding the opportunities for individualizing training. Typically, intelligent training systems are limited to adaptation based on performance. Such measures are limited in their ability to discriminate within the “good” or “bad” performance categories. For example two trainees may both reach a good decision; however, the amount of effort it took to reach this decision or the amount of anxiety in doing so might differ significantly (Klein, 2008). A trainee model that incorporates assessment of affective and cognitive states would allow training to be tailored to these trainees differently to optimize opportunities for learning. The goal would be to incorporate adaptation strategies that aim to keep learners in both cognitive and affective states that are optimal for learning. Figure 1 illustrates the target range for affective states to optimize learning opportunities. It is hypothesized that learning opportunities are optimized when learners have positive levels of valence and high to moderate levels of arousal, such as when excited, delighted, happy, or calm. Figure 2 illustrates the target range for cognitive states to optimize learning opportunities. It is hypothesized that learning opportunities are optimized when learners are moderately to highly engaged and under moderate workload.

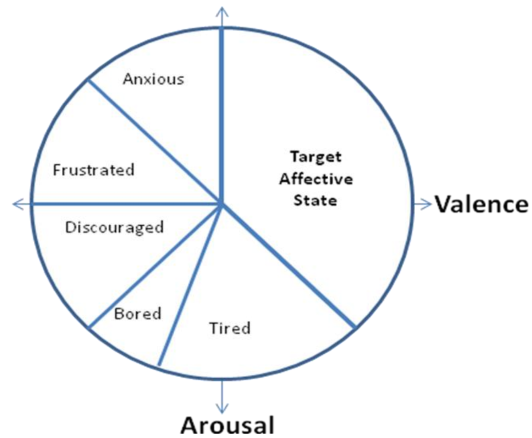

Figure 1. Target Affective States (modified based on Barrett and Russell (1999))

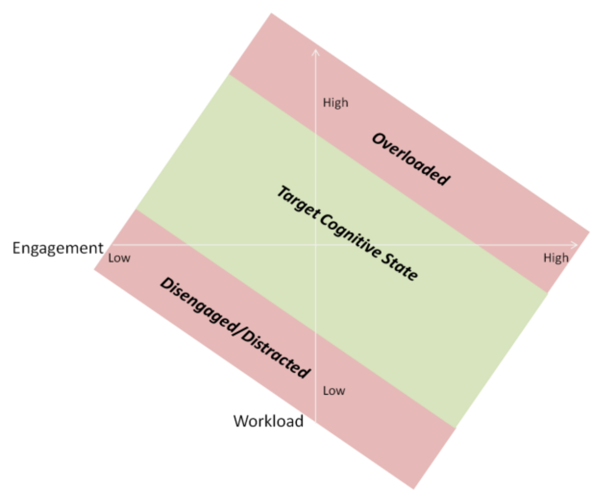

Figure 2. Target Cognitive States

When a learner is operating outside of these target affective and cognitive ranges, mitigation techniques should be employed in an attempt to optimize learning state and ensure learning opportunities are not lost. Table 7 provides a set of example non-optimal combinations of performance, and affective and cognitive states that may require mitigation. Also included are hypothesized diagnoses for what these combinations of factors mean for learning state.

Table 7. Notional Diagnosis Examples

Performance	Anger	Fear	Boredom	Workload	Engagement	Attentive	Diagnosis
Good	Low	Low	High	Low	Low	Low	Learner is not being challenged
Good	High	Low	High	Low	Low	Low	Learner is frustrated with ease of training
Poor	High	Low	Low	Low	Low	Low	Learner is annoyed with the task and disengaged or distracted
Poor	High	Low	Low	Low	High	High	Learner is frustrated with the task
Poor	High	High	Low	High	High	High	Learner is overloaded and feels overwhelmed by the task
Poor	Low	High	Low	High	High	High	Learner is worried about task performance or disturbed by task content
Poor	Low	High	Low	Low	Low	Low	Learner is distracted from task
Poor	Low	Low	High	Low	Low	Low	Learner is disengaged from task
Poor	Low	Low	Low	High	High	High	Learner is overloaded by task

BENEFITS

The goal of this research was to enhance capabilities for self-directed learning. The primary benefits of improving the perception of computer-based tutors are: training effectiveness, efficiency, flexibility and accessibility, and decreased training support costs. Just as with human tutors, providing better, more focused information to make instructional decisions will result in more effective decisions. Improvements to sensing techniques and machine-based classifiers to assess trainee state and select instructional strategies will alleviate unproductive time during training.

Better sensing and classification techniques will also allow more complex, ill-defined tasks to be trained using computer-based one-to-one instruction. This will allow military training organizations to focus their instructional resources on collective training where computer-based tutoring techniques remain immature (Sottilare, 2010).

FUTURE WORK

Future work on this effort will include the development and validation of machine learning classifiers of cognitive state to inform the trainee state model. Data to facilitate development and validation of the classifier will be collected both with laboratory participants (e.g., undergraduates) and active duty military participants (i.e., West Point Cadets).

Other potential future work will include development of strategies for mitigating trainee negative learning states. Techniques for successfully mitigating such learning states have been developed for ITS, but they must be evaluated with numerous students in a variety of contexts (Woolf et al., 2009) to ensure validity. Further, the accuracy of the trainee learning state model could be improved by including the assessment of additional cognitive and affective states.

The results of this work will be fed into the Army Research Laboratory's Generalized Intelligent Framework for Tutoring (GIFT), which is being developed to assess computer-based tutoring technology and improve the authoring of ITS.

CONCLUSION

The next generation of training systems must support self-directed learning that requires little or no instructor support. However, for these training systems to remain effective, it is critical that they are able to recognize and adapt to trainee cognitive and affective state, similarly to the way a human instructor would. By utilizing low-cost, non-intrusive physiological sensors to assess affective and cognitive state and analyzing the data in real-time with machine learning algorithms, this work has the opportunity to enhance ITS diagnosis and remediation, leading to more effective and efficient training systems.

ACKNOWLEDGEMENTS

This research was sponsored by the U.S. Army Research Laboratory - Human Research and Engineering Directorate under contract W91CRB-10-C-0327.

REFERENCES

Ahlstrom, U., and Friedman-Bern, F.J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36(7), 623-636.

Barrett, L.F., and Russell, J.A. (1999). Structure of current affect. *Current Directions in Psychological Science, 8,* 10–14.

Berka C, Levendowski DJ, Lumicao MN, Yau A, Davis G, Zivkovic VT, Olmstead RE, Tremoulet PD, Craven, PL. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation Space and Environmental Medicine*; 78(5, Suppl.):B231-B244.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4-16.

Bradley, M. M., Moulder, B., and Lang, P. J. (2005).When good things go bad: The reflex physiology of defense. *Psychological Science*, 16, 468–473.

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

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., and McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin*, 126, 829–843.

Craig, S. D., Graesser, A. C., Sullins, J., and Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media*, 29(3), 241-250.

Craik, F. I., Govoni, R., Naveh-Benjamin, M., and Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology General*, 152(2), 159-180.

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *Neuroscientist*, 8(2), 132–142.

Davidson, R. J., Scherer, K. R., and Goldsmith, H. H. (2003). *Handbook of Affective Sciences*. New York: Oxford University Press.

Di Stasi, L.L., Renner, R., Staehr, P., Helmert, J.R., Velichkovsky, B.M., Cañas, J.J., Catena, A. et al. (2010). Saccadic peak velocity sensitivity to variations in mental workload. *Aviation, Space, and Environmental Medicine*, 81(4), 413-417.

D'Mello, S., Chipman, P., Graesser, A. (2007). Posture as a predictor of learner's affective engagement. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, 905–910, Cognitive Science Society.

D'Mello, S. K., Taylor, R., and Graesser, A. C. (2007). Monitoring Affective Trajectories during Complex Learning. In D. S. McNamara and J. G. Trafton (Eds.), *In: Proceedings of the 29th Annual Cognitive Science Society*, pp. 203--208. Cognitive Science Society, Austin.

Dorneich, M. et al. (2004). DARPA Improving Warfighter Information Intake under Stress - Augmented Cognition Concept Validation Experiment (CVE) Analysis Report for the Honeywell Team. Prepared under contract: DAAD16-03-C-0054.

Foerde, K., Knowlton, B. J., and Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the United States of America*, 103(31), 11778-11783.

Fredrickson, B. (1998). What good are positive emotions? *Review of General Psychology*, 2(3), 300-319.

Gonzalez, C. (2005). The relationship between task workload and cognitive abilities in dynamic decision making. *Human Factors*, 47(1), 92–101.

Gratton, G., Kramer, A.F., and Fabiani, M. (2008). Brain Sensors and Measures. In D.D. Schmorrow and K.M. (Eds). Augmented Cognition: A Practitioner's Guide. Santa Monica, CA: Human Factors and Ergonomics Society.

Holland, P. C. and Gallagher, M. (2006). Different Roles for Amygdala Central Nucleus and Substantia Innominata in the Surprise-Induced

Enhancement of Learning. *The Journal of Neuroscience*, 26(14), 3791-3797.

Hoover, A., and Muth, E. (2004). A real-time index of vagal activity. *Journal of Human Computer Interaction, 17 (2),* 197-209.

Hyona, J., Radach, R., and Deubel, H. (2003). *The mind's eye: cognitive and applied aspects of eye movement research*, Oxford, England, Elsevier.

Ingleton, C. (2000). Emotion in learning - a neglected dynamic, In R. James, J. Milton and R. Gabb (Eds.), *Research and Development in Higher Education*, *Cornerstones of Higher Education.* Melbourne, 22, 86-99.

Jang, D.P., Kim, I.Y., Nam, S.W., Wiederhold, B.K., Wiederhold, M.D. and Kim, S.I. (2002). Analysis of Physiological Response to Two Virtual Environments: Driving and Flying Simulation. *Cyberpsychology and Behaviour, 5(1)*, 11-18.

Jensen, O., Gelfand, J., Kounios, J., Lisman, J.E., (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12,* 877–882.

Jensen, O., Kaiser, J., Lachaux, J.-P., 2007. Human gamma-frequency oscillations associated with attention and memory. *Trends in Neuroscience 30*(7), 317–324.

Johns, J. and Woolf, B. (2006). A Dynamic Mixture Model to Detect Student Motivation and Proficiency, *In Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)* 2-8. Menlo Park, Calif.: AAAI Press

Kapoor, A., Burleson, W., and Picard, R. W. (2007). Automatic prediction of frustration. *International Journal of Human-Computer Studies, 65*, p. 724-736.

Klein, G. (2008). Naturalistic Decision Making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 456-460.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple task performance* (pp. 279–328).Washington, DC: Taylor and Francis.

Li, X. and Ji, Q. (2005). Active affective state detection and user assistance with dynamic bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 35*(1), 93-105.

Lisetti, C.L., Nasoz, F. (2004). Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *Journal on Applied Signal Processing*, 11, 1672-1687.

Mandryk, R. L. (2005). *Modeling user emotion in interactive play environments: A fuzzy physiological approach*. Unpublished Ph.D. thesis, Simon Fraser University, Burnaby, Canada.

McQuiggan, S., Lee, S., Lester, J. (2007). Early prediction of student frustration. Affective Computing and Intelligent Interaction. 698-709.

Merrifield, C. (2010). Characterizing the Psychophysiological Signature of Boredom. Unpublished masters thesis, University of Waterloo, Waterloo, ON.

Merrill, D. C., Reiser, B. J., Ranney, M., and Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences.* Vol. 2. No. 3. pp. 277-305.

Mota, S., and Picard, R.W., (2003). Automated posture analysis for detecting learner's interest level. In: Workshop on computer Vision and Pattern Recognition for Human-Computer Interaction, June 2003.

Mykityshyn, A. L., Fisk, A., D., and Rogers, W. A. (2002). Learning to Use a Home Medical Device: Mediating Age-Related Differences with Training. *Human Factors*, *44*(3), 354-364.

North Atlantic Treaty Organisation (NATO) Research and Technology Organisation (RTO) (2009). Human Dimensions in Embedded Virtual Simulation. *In Proceedings of the Human Factors and Medicine Panel Research Workshop* (RTO-MP-HFM-169), October 2009, Orlando, Florida.

Neri, D.F., Dinges, D.F., and Rosekind, M.R. (1997). Sustained carrier operations: Sleep loss, performance, and fatigue countermeasures. Moffett Field, California: NASA Ames Research Center. Accessed 24 July 2009 from: http://humanfactors.arc.nasa.gov/zteam/PDF_pubs/Nimitz1997.pdf

Ortony, A., Clore, G. L., Collins, A. (1988). *The Cognitive Structure of Emotions.* Cambridge: Cambridge University Press.

Partala, T. and Sukarra V. (2003) Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59 (1-2), 185-198.

Pomplun, M., and Sunkara, S. (2003). Pupil dilation as an indicator of cognitive workload in Human-Computer Interaction. *In Proceedings of HCI International 2003*: 3, 542-546. Mahwah, NJ: Lawrence Erlbaum Associates.

Roberts, R. J., and Weerts, T. C. (1982), Cardiovascular responding during anger and fear imagery. *Psychological Reports*. 50, 219–230.

Ryu, K. and Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, 35(11), 991-1009.

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., Bigdeli, A. (2006). See Me, Teach Me: facial expression and gesture recognition for intelligent tutoring systems. *Innovations in Information Technology*. pp. 1-5

Sarter, M., Givens, B., and Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain Research Reviews.* 35, 146-160.

Scerbo, M.W, Freeman, F.G., Mikulka, P.J., Parasuraman, R., Nocero, F.D., Prinzel, L.J., III (2001) *The Efficacy of Psychophysiological Measures for Implementing Adaptive Technology*. NASA/TP-2001-211018

Scheirer, J., Klein, J., Fernandez, R., and Picard, R.W. (2002). Frustrating the User on Purpose: A Step Toward Building an Affective Computer. *Interaction with Computers*. *14*(2), 93-118.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695-729.

Self, J. (1999). The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. *International Journal of Artificial Intelligence in Education, 10*(3-4), 350-364.

Sherry, L., Feary, M., Polson, P., and Palmer, E. (2000). Autopilot Tutor: Building and Maintaining Autopilot Skills. *Proceedings Int. Conf. on Human Computer Interaction–AERO, Toulouse, France*.

Small, R. V., Dodge, B. J., and Jiang, X. (1996). Dimensions of interest and boredom in instructional situations. *In Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology* (18th, Indianapolis, IN).

Sottilare, R. (2010). Toward the Development of an Intelligent Tutoring System for Distributed Team Training through Passive Sensing. *In Proceedings of the Intelligent Tutoring Systems (ITS) 2010 Conference*, Pittsburgh, Pennsylvania, June 2010.

Stevens, R., Galloway, T., and Berka, C. (2007). EEG-Related Changes in Cognitive Workload, Engagement and Distraction as Students Acquire Problem Solving Skills. User Modeling Conference. Athens, Greece.

Stottler, D., Harmon, N., and Michalak, P. (2001). Transitioning An ITS Developed For Schoolhouse Use To The Fleet-TAO ITS, A Case Study. *Paper presented at the The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC*), Orlando, Fl.

Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., Na¨a¨ta¨nen, R., 1993. Selective attention enhances the auditory 40-Hz transient response in humans. Nature 364, 59–60

VanLehn, K., Lynch, C., Schultz, K., Shapiro, J. A., Shelby, R. H., Taylor, L., et al. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence and Education, 15*(3).

Wientjes, C. J. E. (1992). Respiration in psychophysiology: Methods and applications. *Biological Psychology*, 34, 179–204.

Wilson, C. R. (1992). Applied use of cardiac and respiration measures: Practical considerations and precautions. *Biological Psychology, 34*, 163–178.

Woo Kim, J., Jin, G., Fuchs, S., Fouad, H. (2010). RADIS: real time affective state detection and induction system. Presented at IHCI 2010, Freiburg, Germany.

Woolf, B., Burelson, W., and Arroyo, I. (2007). Emotional intelligence for computer tutors. *Supplementary Proceedings of the 13th International Conference of Artificial Intelligence in Education*, 6-15.

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D. and Picard, R. (2009). Affect-aware tutors: recognising and responding to student affect, *International Journal of Learning Technology*, *4*(3/4), pp.129–164.