Using Student Mood And Task Performance To Train Classifier Algorithms To Select Effective Coaching Strategies Within Intelligent Tutoring Systems (its)

2009

Robert Sottilare
University of Central Florida

USING STUDENT MOOD AND TASK PERFORMANCE TO TRAIN CLASSIFIER ALGORITHMS TO SELECT EFFECTIVE COACHING STRATEGIES WITHIN INTELLIGENT TUTORING SYSTEMS (ITS)

by

ROBERT A. SOTTILARE
M.S. Modeling & Simulation, University of Central Florida, 2006
B.S.E.E. University of Central Florida, 1984

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Modeling and Simulation
in the Department of Modeling and Simulation
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2009

Major Professor: Michael Proctor

ABSTRACT

The ultimate goal of this research was to improve student performance by adjusting an Intelligent Tutoring System's (ITS) coaching strategy based on the student's mood. As a step toward this goal, this study evaluated the relationships between each student's mood variables (pleasure, arousal, dominance and mood intensity), the coaching strategy selected by the ITS and the student's performance. Outcomes included methods to increase the perception of the intelligent tutor to allow it to adapt coaching strategies (methods of instruction) to the student's affective needs to mitigate barriers to performance (e.g. negative affect) during the one-to-one tutoring process.

The study evaluated whether the affective state (specifically mood) of the student moderated the student's interaction with the tutor and influenced performance. This research examined the relationships, interactions and influences of student mood in the selection of ITS coaching strategies to determine which strategies were more effective in terms of student performance given the student's mood, state (recent sleep time, previous knowledge and training, and interest level) and actions (e.g. mouse movement rate).

Two coaching strategies were used in this study: Student-Requested Feedback (SRF) and Tutor-Initiated Feedback (TIF). The SRF coaching strategy provided feedback in the form of hints, questions, direction and support only when the student requested help. The TIF coaching strategy provided feedback (hints, questions, direction or support) at key junctures in the learning process when the student either made progress or failed to make progress in a timely fashion.

The relationships between the coaching strategies, mood, performance and other variables of interest were considered in light of five hypotheses. At alpha = .05 and beta at least as great as .80, significant effects were limited in predicting performance. Highlighted findings

include no significant differences in the mean performance due to coaching strategies, and only small effect sizes in predicting performance making the regression models developed not of practical significance. However, several variables including performance, energy level and mouse movement rates were significant, unobtrusive predictors of mood.

Regression algorithms were developed using Arbuckle's (2008) Analysis of MOment Structures (AMOS) tool to compare the predicted performance for each strategy and then to choose the optimal strategy. A set of production rules were also developed to train a machine learning classifier using Witten & Frank's (2005) Waikato Environment for Knowledge Analysis (WEKA) toolset. The classifier was tested to determine its ability to recognize critical relationships and adjust coaching strategies to improve performance. This study found that the ability of the intelligent tutor to recognize key affective relationships contributes to improved performance. Study assumptions include a normal distribution of student mood variables, student state variables and student action variables and the equal mean performance of the two coaching strategy groups (student-requested feedback [SRF] and tutor-initiated feedback [TIF]). These assumptions were substantiated in the study.

Potential applications of this research are broad since its approach is application independent and could be used within ill-defined or very complex domains where judgment might be influenced by affect (e.g. study of the law, decisions involving risk of injury or death, negotiations or investment decisions). Recommendations for future research include evaluation of the temporal, as well as numerical, relationships of student mood, performance, actions and state variables.

This dissertation is dedicated to Shannon, my wife of 27 years. She is my inspiration. Her love and patience allowed me to dedicate the time and effort needed to complete this dissertation.

ACKNOWLEDGMENTS

The success of a project of this size and complexity involves contributions of many people. My wife, Shannon, has been a rock. My love and thanks go to her and my children, Joseph and Lynn for their many sacrifices during my mid-life college career.

I am indebted to Dr. Michael Proctor, my committee chair and advisor. I previously had the pleasure of watching him guide several Army officers to their advanced degrees and I knew his guidance and motivational style would keep me on track. When I decided to return to school, Dr. Proctor was the first person I called. Thank you, Dr. Proctor for your support and patience during this process.

Thanks to Dr. Stephen Goldberg. Dr. Goldberg provided tireless support not only as a committee member, but as a mentor and a friend. He definitely influenced the psychological content of this dissertation. See old engineers can learn new tricks!

I also want to recognize the contributions of my other committee members, Dr. Thomas Atkinson, Dr. Peter Kincaid and Dr. Michael Matthews for their guidance and support. Dr. Atkinson's course on instructional system design really solidified my goals to improve performance through this research. Dr. Kincaid has been supportive throughout my graduate career. Dr. Matthews and LTC James Merlo at the United States Military Academy helped me through the IRB process and secured participants for my experiment when I thought it might be a long struggle to secure the numbers I needed. Thank you both for you friendship and support.

I have been very fortunate to have many role models provide motivation throughout this process. Three immediately come to mind. One model is Dr. Brian Goldiez, who also had the courage to return to school and complete his PhD after the tender age of forty.

Another wonderful role model is my Mom, Terri, who returned to school to earn a degree and a second career in nursing years ago. Her tenacity and zest for learning set an example I carry with me every day.

A third role model is my cousin, Dr. Manuel Francisco. His steadiness and camaraderie helped me remain “even” through this process. Thanks, Manny for your hospitality and allowing me to blow off steam when I really needed it.

Last, but not least, I would like to acknowledge the contributions of my colleagues at the Simulation and Training Technology Center. Thanks to the “lunch bunch” for their counsel and sometimes for just listening. Thanks to COL Langhauser, Beth, Angel, Neal, Chris and John for their patience with all my distractions. Thanks to Neal, Maria and Tere for being such great role models.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Chapter One Summary

In this chapter, research goals along with the scope and motivation for this research including the concept of an adaptable intelligent tutoring system (ITS) are considered.

Research Goals and Objectives

An ongoing goal in the research and development of ITS is to increase their adaptability to better serve student needs (Hernandez, 2006; Sottilare, 2006; Loftin, et al, 2004; and Heylen, 2003). "The basic tenet of intelligent tutors is that information about the user (e.g. knowledge, skill level, personality traits, mood level or motivational level) can be used to modify the presentation of information so that learning proceeds more efficiently." (Johnson and Taatgen, 2005).

The purpose of this research was to identify methods to increase the intelligent tutor's perception of the student's affective state during the one-to-one tutoring process. Personality preferences, mood and emotions are known collectively as "affect", which is the "general term for feelings, emotions or moods which includes the conscious subjective aspect of feeling" (Gebhard, 2005).

The primary objective of this research was to evaluate the relationships, interactions and influences of student mood in the selection of effective coaching strategies with regards to student performance. A secondary objective was to develop methods for the unobtrusive inference of the student's mood based on student actions during the performance assessment as a comparison to the self-reported method which was also used in this study.

Critical to conducting student performance research involving student mood and ITS coaching strategies are techniques to unobtrusively infer mood. If mood can be inferred, then one can adapt coaching strategy to the mood to hopefully optimize learning and performance. This study focused on performance. The notion of a "learning companion" explored by Kim, et al (2007) was used in this research to form the notion of an adaptive tutor and the notion of Zimmerman, et al (2003) for unobtrusive inference of mood.

Kim explored pedagogical agents as learning companions (PALs). A PAL attempts to "facilitate social relations with a learner" by enhancing the trust level between the learner and the PAL through knowledgeable, helpful and motivating interactions. Kim, et al (2007), Dempsey & van Eck, (2003), van Eck & Dempsey, (2002) and Aleven & Koedinger, (2000) indicated that giving the student more control through coaching strategies like student-requested feedback (SRF) might not be the "best solution for weak learners". Kim warns that many learners' rely heavily on help functions built into tutoring systems and therefore focus too much on completing the lesson and too little effort on learning the material.

Zimmerman, et al (2003) did not focus on adapting coaching strategies within tutoring systems to the student's affective state to but rather sought to infer student affect through unobtrusive methods like computer keystrokes and mouse movements.

In this research Zimmerman's notion of unobtrusive evaluation of affect was extended to measure student mood as was Kim's learning companion notion to make it adaptable to the student mood previously obtained. . The resulting student performance was then tested to determine how student mood influences performance. The research considered students with low, moderate and high levels of competency in both "high control" (e.g. SRF) and "low control" (e.g. tutor-initiated feedback, TIF) coaching strategies while also factoring the student's mood into the decision to provide feedback in the form of information, direction or support.

Scope of Research

This study evaluated the use of techniques to assess student mood and the influence of mood and coaching strategies on task performance in adult learners. The evaluation of performance and mood of children and adolescents (age < 18 years old) is not part of this study.

Mood was selected for this study over emotions and personality preferences because mood is less transient than emotions and less static than personality preferences (Kshirsagar and Magnenat-Thalmann, 2002) and mood has "a great influence on human's cognitive functions" (Morris, 1989) and thereby learning.

Emotions are a psychological state or process that functions in the management of goals and needs of an individual (Broekens & Degroot, 2004). Personality is long-term affect and reflects individual differences in cognitive and social processes. The Five Factor Model (FFM) of personality (McCrae and John, 1992) defines behavior traits in terms of openness, conscientiousness, extraversion, agreeableness and neuroticism; all of which tend to be relatively stable in adult populations. Mood, "an affective state distinguished from emotions and personality in terms of [duration], influence and cause" (Gebhard, 2005).

Moods generally have a moderate duration as opposed to emotions (short duration) and personality (long duration). Moods have a subtle influence on cognition which may go unnoticed even by human tutors. Davidson (1994) argues that "emotions bias action, whereas moods bias cognition" which indicates that emotions may inhibit appropriate actions or cause unintended actions during the learning process, but mood may affect perception and reasoning.

Moods are generally of unknown cause and may be realized as the cumulative effect of a series of emotional events or general feeling. This is as opposed to emotions which are "usually bound to a specific event, action or object, which is the cause of this emotion" (Gebhard, 2005; Becker, 2001).

Mood has been used within ITS (Core, et al, 2006; Heylen, et al, 2003; and Graesser, et al, 2001). However, it has generally been applied to virtual characters to represent the tutor's affect vice representing affect in student models where it could be used to influence tutor decisions, strategies and interactions with the student. This study provided an opportunity to incorporate mood in the ITS student model and use it to optimize coaching strategies. Personality preferences and emotions were not considered as part of this research.

For this study, task performance was considered the measure of skills in terms of accuracy during the application of knowledge. This research was limited to measuring the mood and performance of individuals vice teams. This more complex interaction design problem space was left for future research. This study was also primarily concerned with perceiving mood state vice understanding the root cause of the mood state (e.g. fatigue, sleepiness or other physical constraints). It is accepted that fatigue and other physical factors can influence mood, but the goal of this study was to identify the mood state and then select the best coaching strategy given that mood to get the best performance results.

Motivation for Research

Linnenbrink and Pintrich (2002) found that many students experience some confusion when confronted with information that does not fit their current knowledge base, but those in a generally positive affective state will adapt their known concepts to assimilate the new information. Students in a generally negative affective state will usually reject this new information. This infers the need for tutors (human or otherwise) to be able to perceive and address the affective state of the student and adapt instruction to optimize the assimilation of new information and enhance performance.

Adaptable ITS

According to (Rodrigues, Novais & Santos, 2005), an intelligent tutor "must be capable of dynamically adapting and monitoring each student" by delivering content in a way that adapts to their particular personality and learning style preferences. The intelligent tutor advises the student about methods to learn the content and helps the student adopt an appropriate study schedule. The ITS monitors the student's progress and provides real-time diagnostic help. Like human tutors, it is desirable that ITS have the capability to coach or guide the student through the learning experience by providing a customized program of support and direction. An "intelligent" tutor should choose the content and the method of instruction based on both the student's pedagogical state (i.e. knowledge and skill level) and his affective state (i.e. emotions, mood and personality) in the same way "an experienced human tutor manages the emotional state of the student to motivate him and to improve the learning process." (Hernandez, et al, 2006).

Expert human tutors adapt to students by observing affective cues and then using this information to guide changes in tutoring style, tempo or coaching strategies (e.g. directive or supportive). Like a human tutor, ITS need to gather information about the student and use this information to build a student model that guides the application of coaching strategies. The amount of information needed to accurately assess the affective state of the student can result in poor prediction accuracy if too little information is gathered or a distraction if the student is frequently queried for information. Intelligent tutors that have the ability to perceive student affect are frequently cumbersome (e.g. large number of sensors), expensive (e.g. costly sensor technology), ineffective (e.g. inaccurately predicts the student's affective state) or lack strategies to address the student's affective needs. Based on the current capabilities of ITS, it is desirable to identify unobtrusive method to assess the student's affective state so as not to disrupt the learning process.

One-to-One Tutoring

ITS are typically used to provide a one-to-one tutoring experience. The value of one-on-one tutoring vice group tutoring (i.e. traditional classroom teaching) has been documented among students who work one-to-one with expert human tutors and often score 2.0 standard deviations higher than students in a conventional classroom (Bloom, 1984). ITS will eventually have the potential to post similar results when they have the same capabilities as human tutors to perceive the affective state of their students to determine whether they are engaged, bored, frustrated or even hostile.

Another advantage of tutoring over classroom settings is that in one-to-one tutoring settings, students ask approximately 26.0 questions per hour versus less than 0.20 questions per hour in the classroom (Dillon, 1988; Graesser & Person, 1994). This higher rate of interaction provides additional learning opportunities for weaker students. Good students ask fewer questions, but these questions tend to be "deep-reasoning questions" (Person & Graesser, 2003).

Loftin, et al (2004) assert that "while one-to-one human tutoring is still superior to ITS in general, this approach is idiosyncratic and not feasible to deliver to [any large population] in any cost-effective manner." It is a goal of this research to improve the adaptability of ITS so the intelligent tutor can interact one-to-one with each student and provide a cost-effective, yet tailored learning experience not possible with one-to-one human tutoring. Given the need to provide efficient and cost-effective training to large, distributed populations (e.g. the United States military), it is important to explore methods of increasing ITS perception and thereby their ability to provide adaptable one-to-one training.

Efficiently Measuring Mood

Many ITS tend to focus on the performance profile of the student and primarily track the progress of quantifiable knowledge and skill acquisition relative to an ideal student model for a specific domain (e.g. mathematics or physics). The acquisition and use of information about the student's affective state is fuzzier and much more difficult to acquire without significant effort. Today, the most widely used measurement methods include physiological measures (e.g. eye dilation, skin conductance or heart rate) and self-reporting instruments (e.g. surveys and questionnaires) which are used to infer the student's affective state. Physiological measures tend to be time consuming to set up and expensive. The drawback for self-reporting instruments is that they are intrusive since they require the student to stop to take the survey and may detract from the learning process.

Based on the high cost of data collection, it is also important to explore methods of streamlining the data collection process required to develop student models within ITS. This research attempts to do this by evaluating the relationship between affective attributes influencing performance (e.g. mood) and the type and timing of feedback provided by the tutor (e.g. during or after tutoring). Understanding how these attributes influence desirable learning outcomes (e.g. increased task performance) will enable the tutor to select the coaching feedback strategy needed to generate the most effective outcome (i.e. the highest positive change in task performance – time and accuracy). An efficient method of selecting intelligent tutor coaching strategies based on a minimal set of student data (e.g. mood and task performance results) is a goal of this research.

Emotionally Intelligent Tutors

ITS that lack adaptability may also adversely impact the student's affective state resulting in negative emotions like frustration or boredom. The frequency, timing and content of interventions provided by the ITS may result in negative emotions and adversely affect learning. Frequent interventions can cause frustration. Mistimed interventions can provide too much or too little information resulting in dependency on the tutor for knowledge or withdrawal due to boredom. Frustration can result from tutoring significantly above or below the student's knowledge level. Sessink, et al (2007) assert that "most traditional learning material [for ITS] targets the 'average student', and is suboptimal for students who lack certain prior knowledge, or students who have already attained some of the course objectives". Further, ITS should match "adequate challenge with skill in service of Flow, or optimal experience" (Csikszentmihalyi, 1990) within the learning system.

Alexander, et al, (2003) maintain that "an important factor in the success of human one-to-one tutoring is the tutor's ability to identify and respond to affective cues given by the student". Picard (2006) asserts that "No matter how intelligent a tutor or companion is, it will eventually become annoying if it does not have emotional intelligence." Emotional intelligence may be defined as "a set of skills hypothesized to contribute to the accurate appraisal and expression of emotion in oneself and in others, the effective regulation of emotion in self and others, and the use of feelings to motivate, plan, and achieve in one's life" (Salovey & Mayer, 1990). In other words, "intelligent" tutors that have emotional intelligence are capable of:

- recognizing the student's emotional state (e.g. motivated, engaged, frustrated);

- making the student aware of his affective state (e.g. emotional state, mood) so he can participate in managing his affective state;
- providing options (e.g. strategies) for the student to manage/control his affective state;
- using emotion to motivate the student to achieve established tutoring objectives.

In summary, the modeling of affective attributes within tutoring systems is recognized as a key element in the learning process (Bickmore & Picard, 2004; Burleson & Picard, 2004; Johnson, et al, 2000 and Picard, et al, 2004), but the emotional intelligence of ITS remains limited due to the dynamic range of human emotions that must be recognized by the tutor and then associated with an appropriate strategy or set of strategies to be implemented. Picard (2006) suggests that "simplest set" of emotions for an ITS to recognize are the emotions each of us is born with: pleasure, boredom and frustration.

In general, research evaluating "emotionally intelligent" tutors and their link to effective coaching strategies has not always shown significant differences in outcomes with tutors that consider emotional intelligence. For example, Baylor, et al (2004) noted that "while the learning scores for students working with the agents with motivational support were higher than the students who worked with agents without motivation support, the difference was not statistically significant". Thus additional research is needed. This appears to include research into other factors (i.e. frequency and type of feedback or competence level) which are necessary to demonstrate significant improvement in task performance.

CHAPTER TWO: LITERATURE REVIEW

Chapter Two Summary

This chapter reviews literature on ITS concepts (components, practices and applications), learner needs, the relationship of mood and learning, and research that specifically utilizes affect within ITS. Specific examples of mood and ITS coaching strategies are examined. Potential approaches, techniques and results to build upon are discussed. Research gaps are identified and research questions are posed based on those gaps.

A Review of Intelligent Tutoring System (ITS) Concepts

The basic approaches and components of an ITS are well established in the literature (Woolf, 1992; Beck, 1996; Corbett, 1997; Clancey, 1981; Shute, 1990; Anderson, 1993; and Lehman, 2006) and detailed below along with common practices and applications.

ITS Concepts

ITS components vary in name and function, but in general, ITS contain four major components as identified by Woolf (1992): the student model, the pedagogical module, the domain knowledge module, and the communication module. Beck (1996) identified a fifth component, the expert model, which Woolf included as part of the domain knowledge module. These components, their functions and interactions are described below.

Student Model or Performance History Model

The student model has generally been a record of the student's knowledge state (Corbett, 1997). It stores information specific to each individual learner including a history of performance and other pertinent data. This could include personality preference information or other state information. The student model also records observable actions and may (through some fuzzy logic) infer non-observable states (i.e confusion, boredom or other emotions). "Since the purpose of the student model is to provide data for the pedagogical module of the system, all of the information gathered should be able to be used by the tutor [pedagogical module]." (Beck, 1996)

Pedagogical Module or Instructional Planner

This component provides a model of the instruction process and contains logic for making decisions about when to review information, when to present new topics or concepts. The sequencing of topics is controlled by the pedagogical module. Once the topic has been selected, a problem must be generated for the student to solve and then feedback is provided on the student's performance. As noted above, the student model is used as input to this component, so the pedagogical decisions reflect the differing needs of each student (Beck, 1996).

Domain Knowledge

This component contains information the tutor uses to instruct the student. It is critical that the domain be accessible by other parts of the ITS. "One related research issue is how to represent knowledge so that it easily scales up to larger domains. Another open question is how to represent domain knowledge other than facts and procedures, such as concepts and mental models." (Beck, 1996) This component contains items like generic instructional strategies, databases of scenarios and diagnostics.

Communications or Interface Module

This component controls interactions with the learner, including the dialogue and how the material should be presented to the student in the most effective way. This selection of presentation format is driven by the selection of instructional strategies in the pedagogical module. The communications module may also include some type of natural language understanding function to support verbal interaction with the student.

Expert Model

This component is also known as the Cognitive Model of Ideal Student Behaviors. The expert model is similar to the domain knowledge in that it is a model of how someone skilled in a particular domain represents the knowledge. Generally, it takes the form of a runtime expert model (i.e. one that is capable of solving problems in the domain). (Clancey, 1981) "By using an expert model, the tutor can compare the learner's solution to the expert's solution, pinpointing the places where the learner had difficulties." (Beck 1996)

ITS Practices and Applications

Below are several approaches to the development of intelligent tutoring systems. Each of these approaches generally supports a single learning style. In the literature, it was rare to find a tutor that encompassed more than two of these approaches or adapted to individual preferences. Given the variance in human personality, an adaptable tutor that encompassed all of these approaches and others might be desirable.

Human emulation of a tutor

This approach uses natural language processing to interact with the student and may use some type of virtual human (i.e. embodied conversational agent). This approach is similar to dealing with a human in one-to-one tutoring, but is very difficult to model given the requirement to provide real-time reactions (verbal and non-verbal) to student inquiries. Success with this type of tutor has been limited to small, well-defined knowledge domains. The cost for this type of approach remains higher than other approaches mainly because of the lack of tools to develop content quickly and the inability to easily apply this technology across multiple knowledge domains (Core, et al, 2006) . The human emulation is flexible enough to be applied with student of differing learning styles. As artificial intelligence technology matures and software tools make it quicker to develop and easier to apply in a variety of domains, human emulation will become the tutoring method of choice in representing, perceiving and reacting to the student's affective needs.

Bug Detection

"There are classically two components in a student model: an overlay of the domain expert knowledge and a bug catalog, which is a set of misconceptions or incorrect rules." (Corbett, 1997) In a bug detection scheme, the tutor corrects errors by explaining what the error is (e.g. the student is using the rules properly, but the problem is that it is the wrong rule is being applied). A drawback to this approach is that too frequent intervention by the tutor can detract from the learning experience and could frustrate the student. This intelligent tutoring method is an approach where students learn only what is provided by the tutor and are asked to apply specific rules in broader contexts.

Exploratory systems (discovery worlds, micro worlds)

Exploratory systems are environments that "place less emphasis on supporting learning through explicit instruction and more on providing the learner with the opportunity to explore the instructional domain freely, acquiring knowledge of relevant concepts and skills in the process" (Shute, 1990). A drawback to this approach is that learning may be time intensive and very inefficient. Given sufficient time, this approach may be very appealing for some learners.

Smithtown, which provides a guided discovery of economics, is an example of an exploratory system that employs artificial intelligence methods to assist students in beginning economics courses to improve their problem-solving skills. (Raghavan and Katz, 1989).

Model Tracing

A cognitive model of the task to be taught is developed through a task analysis. Student progress is assessed by ``tracing'' the student's task actions (i.e., matching user and application events against the task model). The student is permitted to consult task model as needed. This approach seems to be the most prevalent and tied closely to cognitive models like ACT-R (Anderson, 1993) and SOAR (Lehman, 2006).

Constructivism

Constructivism is a philosophy of learning where "individuals actively construct and reconstruct their own reality in an effort to make sense of their experience" (Prince and Felder, 2006). The student reflects on new information and filters information based on student's knowledge, prior experiences, values, beliefs, preconceptions and misconceptions, prejudices, and fears. Information that aligns with those filters is accepted and integrated (learned) and stored for later use. Information that does not fit is generally rejected. In a "constructivist" approach, the ITS provides opportunities for the student to participate and manage the instructional process. There are no standardized curricula, tests or grades. Instead, constructivism promotes the use of customized curricula based on the student's prior knowledge and emphasizes hands-on problem solving and reflection. Constructivism is an inductive learning process in which the focus is on the student and the "big picture". The learning process allows the student to frame the body of knowledge being explored in their own way and prepares the learner to reflect and take broad concepts and apply them in experiential learning contexts to demonstrate understanding. Given the current state of tutor perception models, a student could spend a lot of time in a constructivist environment frustrated or bored without any timely intervention. This environment may be appealing to a more reflective learner.

The limitations noted for each intelligent tutoring type above serves as rich environment for future ITS research in affective perception and coaching strategies. The sections that follow further define learner needs, current research capability gaps, research needs and research questions.

A Review of Learner Needs

According to a National Research Council (2005) report, four learning environment perspectives are required to support the learner's needs:

- learner-centered: encourages attention to preconceptions, and begins instruction with what students think and know
- knowledge-centered: focuses on what is to be taught, why it is taught, and what mastery looks like
- assessment-centered: emphasizes the need to provide frequent opportunities to make students' thinking and learning visible as a guide for both the teacher and the student in learning and instruction
- community-centered: encourages a culture of questioning, respect, and risk taking

According to the report "How People Learn" (National Research Council, 2000) being able to understand relationships within concepts (e.g. the relationship between the structure and function) will increase the likelihood that learners will be able to use what they have learned to solve new problems and thereby demonstrate evidence of transfer of learning. The report outlined four factors that influence 'transfer of learning' the ability to extend what has been learned in one context to new contexts (Byrnes, 1996). These factors are important to performance and include:

- providing multiple contexts for the original learning
- representing problems at higher levels of abstraction
- overlapping the original domain of learning and the new one to a high degree
- implementing dynamic processes that require learners to actively choose and evaluate strategies, consider resources, and receive feedback

The learning factors cited above define the attributes of theoretically ideal learning systems and are independent of how the instruction is delivered or moderated (either via human tutor or intelligent tutor). Intelligent tutor design should consider these learning system needs and develop associated characteristics to optimize learning (and transfer of learning) to provide an environment that addresses both the pedagogical and affective needs of the student.

The Relationship between Mood, Learning and Performance

The following definitions are provided to facilitate discussion of the relationships between mood, learning and performance. Mood is a person's state of mind or general expression of attitude (mood, Encarta, 2008). According to Mehrabian (1996) mood has three vectors: pleasure which ranges from pleasure to displeasure and distinguishes the positive-negative affective quality of mood (Klesen, 2002); arousal which ranges from high physical activity and mental alertness to calm which refers to low physical activity and mental sluggishness (Klesen, 2002); and dominance which is scaled from dominance to submissiveness and is defined in terms of control versus lack of control (Klesen, 2002).

Learning (learning, Encarta, 2007) is the relatively permanent change in, or acquisition of, knowledge, skills, understanding, or behavior. For purposes of this research, task performance will equate to a demonstration of learning or competence. With definitions and context set for mood, learning, competence and task performance, specific citations are provided below to demonstrate the strong link between learning and mood.

Vermunt (1996) describes three types of learning activities: cognitive, affective and metacognitive or regulative. These activities directly relate affect and learning. Cognitive learning activities include those activities that people use to process learning content and include relating, structuring, analyzing, concretizing, applying, memorizing, critical processing and selecting. Affective learning activities include those activities that people use to cope with emotions that arise during the learning process and include attributing, motivating, concentrating, judging oneself, appraising, exerting effort, generating emotions and expecting. Metacognitive learning activities include those activities that people use to regulate cognitive and affective learning activities. Examples of metacognitive activities include orienting, planning, monitoring, testing, diagnosing, adjusting, evaluating and reflecting.

The influence of affect on human cognition and behavior is well established in research domains that include creative problem solving (Isen, 2000), interaction design (Norman, 2002), motivation (Erez and Isen, 2002), attention (Forgas and Bower, 1987), memory (Lee and Sternthal, 1999), and social interaction (Berkowitz, 1993).

The Network Theory of Affect (Bower, 1981; Bower and Forgas, 2000) specified that mood could have an observable effect on memory (and thereby learning) in that:

- memory is facilitated when mood state at learning matches mood state at recall

- material with affective tone that is congruent with current mood is most easily retrieved from memory
- material with affective tone that is congruent with current mood is most easily learned
- affectively intense material is learned best.

Gold & van Buskirk, (1975) demonstrated that high levels of arousal, a vector of mood, enhanced performance on memory tasks. Isen, et al (1987) conducted four experiments that indicated that positive affect improved performance on two tasks requiring creative problem solving skills. Linnenbrink and Pintrich (2002) found that while most students experience some confusion when confronted with information that does not fit their current knowledge, those in a generally positive affective state will adapt their known concepts to assimilate it, whereas students in a generally negative affective state will reject the new knowledge.

Increasing motivation is dependent on increasing arousal (Robbins, 1997). An optimal level of arousal, an element of mood (Mehrabian, 1996) and key element of motivation is necessary for effective learning to take place (Holzinger, 2000). Cordova and Lepper (1996) found that students that were exposed to "motivationally embellished" educational software had higher levels of intrinsic motivation and were more deeply engaged in the learning process and learned more during the same period of time.

The relationship between mood and performance is well established in the literature. Isen (2003) states that people in a "good" mood generally do not "distort or ignore useful negative, threatening, or disconfirming information in an effort to maintain their good mood". In other words, people in a good mood are open to taking in information even when that information does not align with their natural filters (e.g. values). This bodes well for the learning process in that critical concepts are not generally filtered out by the student based on the method of instruction.

Isen (2003) also asserts that the notion that positive affect disrupts thinking is incorrect and that the literature supports just the opposite view. Positive mood facilitates "cognitive flexibility" and is an important element in the ability to be innovative, creative and a good problem solver (Carnevale and Isen, 1986; Estrada, et al 1994; Estrada, et al, 1997; George and Brief, 1996; Greene and Noice, 1988; Hirt, etal, 1996; Isen, 1999; Isen, Daubman, and Nowicki, 1987; Isen, Johnson, Mertz and Robinson, 1985; Isen, Rosenzweig, & Young, 1991; Kahn and Isen, 1993; Lee & Sternthal, 1999; Staw and Barsade, 1993).

A Review of ITS Research

This section reviews recent research vectors in ITS with primary focus on artificial intelligence research and development applied to the education and training domain. There are several issues that have been drivers for recent research in ITS. These include: high development costs, lack of interoperability, restrictive delivery platform requirements, difficulty of sharing materials and benchmarking and high maintenance costs (Rodrigues, 2005). Below are several areas of recent research thrusts in ITS:

Ontology

Ontology is defined as "a controlled vocabulary that describes objects and the relations between them in a formal way, and has a grammar for using the vocabulary terms to express something meaningful within a specified domain of interest. The vocabulary is used to make queries and assertions. Ontological commitments are agreements to use the vocabulary in a consistent way for knowledge sharing." (Browne, 2001). They include "structured ontologies or upper models that define and organize pedagogically relevant attributes of knowledge for classes of domains, enabling the writing and sharing of instructional strategies in terms of these attributes." (Rodrigues, 2005) "The systematic development of a formal ontology must be pursued, and the results of this effort widely disseminated. This type of effort will serve to focus attention on this critical "missing piece" and generate the necessary discussions within the Intelligent Tutoring System research community to achieve a reasonable degree of consensus" (Loftin, et al, 2004).

Tutoring in ill-defined domains

Research in this area attempts to extend ITS capabilities beyond the structured domains (e.g. math, physics, economics) concentrated on in the past. The ability of tutors to establish metrics for evaluating learner performance against expert behavior in ill-defined domains such as cultural awareness, art interpretation and legal reasoning remains an area open for additional research (Lane, et al, 2006).

Architectures

Loftin, et al (2004) argued that "a study is required to map current [ITS] capabilities to a selected training/education domain. This mapping will then identify the small number of architectures that must be supported during application development". Rodrigues, et al (2005) identified the need for "architectures and protocols involving collaborating processes or shared knowledge bases which address issues of modularity and reusability."

ITS Adaptability

The need for cognitive and affective adaptability in ITS is well established (Funk & Conlan, 2002; Johnson and Taatgen, 2005; Rodrigues, 2005; Loftin, et al, 2004; Alexander, 2003; Hernandez, 2006; Roll, 2005). Recently, significant research has been focused on the affective aspects of intelligent tutoring and specifically on learner modeling and its relationship to motivation. The idea of adapting to the learner to provide a personalized experience includes techniques like user profiling, content management and web mining in which statistical methods and data mining processes are applied to web log files to identify unique learner behaviors and patterns of behavior (Rodrigues, 2005). Research to enhance these techniques and develop novel ones should remain a focus for the foreseeable future.

Virtual Humans Research

"Research on the value of virtual humans as an adjunct to or element of an Intelligent Tutoring Systems should be conducted (Loftin, et al, 2004). Specifically, understanding which elements of human tutoring need to be duplicated in virtual humans as part of ITS require additional research. Current work on virtual humans includes the modeling of multi-dimensional personality traits that include emotions. Additional research using virtual humans to support training in areas such as leadership, cultural awareness, and negotiation tactics is underway, but limited. Open research issues include how to build agents that can reason about their emotions and models of the world. (Core, et al, 2006).

Research on the Use of Affective Attributes within ITS

The following provide examples of research in which affect is incorporated into the intelligent tutoring process:

Neji and Ben Ammar (2007) investigated an intelligent tutoring system that included an embodied conversational agent. In addition to the two-way conversational input and output, the agent behavior was informed on the emotional state of the student through a machine vision system. The machine vision system perceived changes in facial expressions of the student and based on distances between facial landmarks classified the expression as one of six universal emotional states (joy, sadness, anger, fear, disgust and surprise) or a neutral expression. Emotional state was then used in the ITS to determine which tutoring strategy (e.g. sympathizing or non-sympathizing feedback, motivation, explanation, steering). The internal state of the agent is based on the PECS (Physical conditions, Emotional state, Cognitive capabilities and Social status) architecture proposed by Schmidt (2000).

A significant drawback to Neji and Ben Ammar's (2007) "Affective e-Learning Framework" is the cost of the vision system which limits its deployability in teaching large, distributed populations (e.g. military organizations). While their approach provides key components (emotional sensing and perception, selection of instructional strategies and interactions based on learner emotional state and the PECS architecture) for an adaptable tutoring system, it does not assess:

- whether the ITS' perception of the affective state of the learner aids the intelligent tutor in selecting appropriate instructional strategies that result in enhanced learning outcomes or performance;
- the influence of other affective variables (e.g. mood components like pleasure and arousal) have on learner outcomes or how these affective variables might influence each other.

Another key study involving affect, was the notion for unobtrusive inference of mood put forth by Zimmerman, et al (2003). However, Zimmerman, et al did not focus on adapting coaching strategies within ITS to the student's affective state but rather sought to infer student affect through unobtrusive methods like computer keystrokes and mouse movements.

Kim, et al (2007) explored pedagogical agents as learning companions (PALs). A PAL attempts to "facilitate social relations with a learner" by enhancing the trust level between the learner and the PAL through knowledgeable, helpful and motivating interactions. Kim, et al (2007), Dempsey & van Eck, (2003), van Eck & Dempsey, (2002) and Aleven & Koedinger, (2000) indicated that giving the student more control through coaching strategies like student-requested feedback (SRF) might not be the "best solution for weak learners". Kim warned that many learners' rely heavily on help functions built into tutoring systems and therefore focus too much on completing the lesson and too little effort on learning the material.

The research in this dissertation drew upon the notion of a "learning companion" explored by Kim, et al (2007) for the notion of adapting coaching strategy to either a student-requested feedback (SRF) approach or a tutor-initiated feedback approach (TIF). This research also drew on Zimmerman, et al (2003) for the notion of unobtrusive inference of mood builds upon Neji and Ben Ammar's (2007) in that it also uses affect to moderate tutor interactions (e.g. coaching strategies). The primary differences in my approach and Neji's is: 1) mood will be used vice the six emotional states and neutral; 2) the tutor interactions will be based on two coaching strategies: student-requested feedback (SRF) which provides information and feedback only when the student asks for it and tutor-initiated feedback (TIF) in which the tutor decides what the student needs to know and what type of encouragement to provide (if any). These are important differences from previous research in that:

- mood offers a more stable view of affect than emotions and is in fact a moderator of emotions, and this is important in focusing instruction on major issues that might interfere with learning vice oscillating from moment to moment based on emotion;
- mood offers a more fluent view of affect than the very static view that personality preferences offer and this is important in looking at a interactive learning process between the ITS and the student.

Further, this research considered the key work by D'Mello, etal (2006) in predicting affective states when students emote aloud. Frequent conversation patterns significantly predicted students' affective states (i.e. confusion, eureka, frustration) and provided feedback, pumps, hints and assertions to influence the student's progress. The primary drawbacks to this approach compared to the goals of this dissertation were the requirement for students to "emote aloud" which have some of same drawbacks as other self-report methods and may be incompatible with students with lower openness scores in personality assessments like the Big Five Personality Test; the small number of participants that could participate because of the labor intensive nature of the data collection and analysis. Based on these limitations, D'Mello's approach was abandoned in favor of unobtrusive methods that allowed for high numbers of participants and broadly supported personality traits in the general population.

Research on Pedagogical Agents and Motivation

This vector consists of research that uses artificially intelligent agents to support educational and training constructs. Loftin, et al (2004) noted that research "should be initiated to (1) investigate means to measure learner motivation within an Intelligent Tutoring System and (2) develop mechanisms to enhance learner motivation through scenario creation and feedback from the Intelligent Tutoring System".

Pedagogical agents and motivation research includes the use of agents to: influence students' interest and motivation (Rosenberg-Kima, et al, 2007); support students as affective learning companions (Burleson & Picard 2004); simulate instructional roles (Baylor & Kim, 2005); create socially intelligent tutors (Heylen, et al, 2003); and to support learners with math anxieties (Baylor, et al, 2004).

Okonkwo and Vassileva (2001) designed an emotional pedagogical agent, a tutoring persona with emotional attributes that acts in an interactive learning environment using Ortony's (1986) "Cognitive Structure of Emotion model" and McCrae and John's (1992) five-factor model of personality. Okonkwo and Vassileva (2001) investigated the persuasive impact of the emotional pedagogical agent on a group of learners. The results of incorporating an emotional pedagogical agent were mixed in that they did not show any significant performance gain in learning, but did change the way students perceived the learning process and increased their level of engagement/motivation. This finding is also supported by Van Mulken, Andre and Muller (1998) whose research also concluded that the "presence of a persona has no significant impact on the users' understanding". This research contrasts with the research proposed by this dissertation in that Okonkwo and Vassileva (2001) focused on evaluating the response of the learner to an emotional agent, but did not address the perception of learner's affect by the intelligent tutor or use the learner's affective state to influence instructional strategy.

Student Modeling

The vector consists of research in which the student's attributes are represented in the student model for the purpose of making the ITS more responsive and adaptable to the student's needs, preferences and capabilities. This research includes Bayesian modeling of student interactions to predict future behavior (VanLehn and Niu, 2001); and open student modeling (Zapata-Rivera, et al, 2007), where the learner model is "a visible and interactive part of the learning environment" (Lane, 2006).

Zapata-Rivera's open student model work utilizes Evidence-Centered Design (ECD) (Mislevy, Steinberg & Almond, 2000; Mislevy, Steinberg & Almond, 2003) as a methodology that emphasizes an "evidence-based chain of reasoning, with the goal of ensuring the validity of [student] assessment results" and their approach is evidence-based interaction with open student models (EI-OSM). This approach is composed of three models: a proficiency model (also known as a student model); a task model, which describes the assessment tasks; and an evidence model that defines which observations are necessary to support claims of proficiency. Based on the evaluation of EI-OSM, Zapata-Rivera, et al (2007) assert that EI-OSM "offers an appropriate environment for developing tools aimed at enhancing student reflection and critical thinking skills."

The linking of student modeling, tasks and proficiency definitions offer a disciplined systems engineering approach to tool development, but offer little in the way of affective representation in the student model. Subsequent sections below discuss research specifically focused on affect representation in student models.

Research in Modeling Student Affect in ITS

This vector consists of research that uses affective attributes (i.e. emotion, mood or personality) in the modeling of the affective state of the student. This research includes: the perception of dialog and posture to determine affect (D'Mello and Graesser, 2007); the integration of affect sensors in ITS (D'Mello, et al, 2005); the development of probabilistic models of affect (Conati & McLaren, 2004); and the inference of user goals based on inferred affective state of the student (Zhou & Conati, 2003).

Van Labeke, Brna and Morales, R. (2007) propose a five layer open learner model that includes metacognition, motivation/affect, competency, conceptual and procedural errors (CAPES) and the subject domain. The research in this dissertation is intended to contribute to the overall body of research in motivation and affect, and extends and supports Van Lebeke's xLM research.

<u>Research in Modeling Student Mood in ITS</u>

Mood is now used extensively within ITS (Core, et al, 2006; Heylen, et al, 2003; and Graesser, et al, 2001), however, it is generally applied to virtual characters representing the tutor's mood state vice representing the student's mood in the student model where it could be used to influence tutor decisions, strategies and interactions with the student.

Woolf, Burelson and Arroyo (2007) pursued a hardware-based solution to recognize student emotion. Their research apparatus included four sensors (camera, posture sensing devices, skin conductance wristband, and a pressure sensitive mouse). The information from the sensors and the participants' interactions during a performance test were used to infer valance and arousal and were applied to train classifier algorithms. Feedback was provided during the performance test to determine their effectiveness in supporting the participant. The research conducted under this dissertation seeks to build upon Woolf, Burleson and Arroyo's work, but intends to focus on less intrusive means of capturing mood states in lieu of their "heavy" hardware approach. A "lighter" approach will enable the transition of mood inference outside the laboratory to an application environment (e.g. a classroom or semantic web application).

Metacognition and Self-Regulated Learning Research

This vector consists of research that explores the process of perceiving your own thought processes and having active control over those cognitive processes during the problem solving process (Lane 2007).

Roll, et al (2005, 2007) conducted research on self-regulated learning in which an analysis of students' actions across two different tutors found that a help-seeking model was domain independent and students' behaviors were "consistent across classrooms, age groups, domains, and task elements".

Wagster, et al (2007) have taken a novel approach to metacognition by providing opportunities for reflection as the student tutors a virtual human (Betty) within the ITS (Betty's Brain). This approach of 'learning by teaching' examines student behavior changes as they "observe, practice, and then internalize self-regulation skills".

Interaction Design Research

This broad area of research considers ITS interaction design based on the student's requirements (i.e. effectiveness, efficiency, safety, utility, learn-ability or memor-ability) to positively affect the student's experience goals which may include affective factors like enjoyment (Sharp, 2007) as they interact with systems. A 'system' in the human system interaction and ITS context is a multi-function tutoring machine with an interface through which the user perceives (e.g. sees, hears, tastes, feels or smells), manipulates information, judges this information and reacts to it with the goal being increased knowledge or skills.

This vector includes research on the integration of perceptual and cognitive modeling for adaptive human-computer interaction (Duric, 2002); task navigation (Pearce & Luckin, 2007); and the use of pedagogical agents to persuade/influence the student (Okonkwo and Vassileva, 2001). Research in interaction design was considered in the development of the training environment that supported the performance test under this study.

Research Gaps and Questions

Several authors (Gratch, 2000; Gratch & Marsella, 2001; Heylen, et al, 2003; Kim, 2003; and dozens of others) have incorporated affect into the intelligent tutor model so the tutor can display emotion. A smaller group of researchers (Picard & Klein, 2002; Alexander, 2003; Bickmore & Picard, 2004; Gratch & Marsella, 2004; Picard, 2004; Kim 2005; Baylor, et al, 2004; Conati & McLaren, 2004) have incorporated affect into student models so that tutors can interact with emotional intelligence and adapt to the changing needs of students. The number of researchers who have used knowledge of student affect to implement instructional strategies with ITS is even smaller (Burleson & Picard, 2004; Anolli, et al, 2005; Gebhard, 2005; D'Mello, et al, 2005; Gratch, et al, 2006; Kim & Baylor, 2006). Finally, the number of researchers who have used knowledge of student affect to implement effective instructional strategies within ITS is less than a handful (Hernandez, et al, 2006; D'Mello, et al, 2006; D'Mello, et al, 2007).

Understanding student affect has been shown as a key to implementing appropriate coaching strategies. DeVincente (2003) argued that "the available theories of motivation in education are not specific enough and are of limited usefulness" in detecting motivation, an affective element in the learning process (Vermunt, 1996). He stresses the need for empirical studies that can help extract a more formal analysis of motivation and its relationship to other variables. Mood, a component of affect, is defined in terms of its relationship to motivation and the pleasure-arousal-dominance vectors described by Mehrabian (1996).

As noted by Gijbels, et al (2005), "more general factors such as prior academic achievement or Grade Point Average (GPA) (Snelgrove & Slater, 2003; Young, 1993; Zeegers, 2001), self-confidence (Watkins & Biggs, 1996) and academic self-efficacy (Pintrich & de Groot, 1990) are potential moderators in the relation between students' approaches to learning and students' quantitative learning outcomes [e.g. task performance] which should be a subject of future research".

Research on the application of coaching strategies for organizational development and one-to-one mentoring is well documented (Hofstede, 1980; House, et al, 1999; Hersey and Blanchard, 1969; and Randolph and Posner, 1979). Persistence, the desire of the student to continue engaging with the learning system (i.e. human or intelligent tutor) is a significant issue in any tutoring context. The use of coaching strategies may be an effective means to manage mood and improve persistence.

Soller (2001) asserted that intelligent tutors (also known as personal learning assistants) should be designed to build the confidence level of students and encourage them through feedback, questioning, hinting and other tutor interaction mechanisms to participate and persist in the learning process. Student confidence is "actually about feeling comfortable with uncertainty" in the learning process (self-confidence, 2009). Bandura (1982) argued that confidence influences performance positively and negatively through emotions where successful performances result in positive emotions (e.g. joy or optimism) and improved confidence while repeated failures generally result in negative emotions (e.g. frustration, fear or anger) and lower confidence. Bandura (1982) also determined that "the higher the level of induced self-efficacy, the higher the performance accomplishments and the lower the emotional arousal". Finally, Bandura (1997) also noted that the knowledge of the task to be performed and shorter times between self-efficacy ratings and task performance influenced effect size.

In order to meet Soller's objective of a tutor that builds the confidence level of students, intelligent tutors need the ability to recognize various levels of confidence and the associated affective states of the student. Unfortunately, intelligent tutors are not yet as adept as human tutors in recognizing affect. Picard (2006) identified the design limitations of intelligent tutors including their inability to: 1) accurately recognize the student's affective state; 2) respond appropriately to the student's affective state; and 3) understand when and how to appropriately express emotion to build trust and motivate the student.

Based on the needs identified by Soller (2001), Hernandez (2006), Sottilare (2006), Loftin, et al (2004), Heylen (2003) and Johnson & Taatgen (2005); and the ITS design limitations noted by Picard (2006), the following problem statement and associated research questions were posed to direct design, experimentation and analysis:

"*ITS have limited ability to unobtrusively perceive and then use information about the student's affective state to adapt ITS coaching strategies to optimize the student performance*".

- What methods are available to unobtrusively determine student mood?
- What relationship between mood, coaching strategies and task performance is expected?
- Are student actions (e.g. mouse movement rates) significant predictors of student mood?
- Are student mood, student state variables (e.g. energy level) and a selected ITS coaching strategies significant predictors of task performance?
- Is there a significant difference in task performance measures based on different coaching strategies?

CHAPTER THREE: METHODS

Chapter Three Summary

This chapter profiles the experiment participants, challenges and objectives, hypotheses, the apparatus used to stimulate the participants and collect data, the experimental procedure and the data analysis methods employed during this study.

Participants

Cadets from the United States Military Academy (USMA) were the target population for this study. Cadets at USMA are typically 18-22 years old. The population is approximately 85% male and 15% female. The experimental design focused on the target population with low-moderate competence in tactical combat casualty care (TC3). To this end, freshman cadets (Plebes) from the PL100 General Psychology course at USMA were offered extra course credit to participate in this study.

Freshmen enrollment at USMA was estimated to be 1400 for the 2008-2009 academic year and approximately 700 were available to participate in this study during the Fall 2008 semester. One hundred-thirty one (131) PL100 cadets participated in the study which produced 124 instances of usable data. Of the 124 participants, 108 were males (age M = 18.79, SD = 0.98) and 16 were females (age M = 18.38, SD = 0.62). The 124 participants in this study are approximately 17.7% of the available participants.

Experiment Challenges

Data collection and algorithm selection and performance were noteworthy challenges during this study:

Data collection

A prime consideration in developing a model to evaluate the relationships between the student mood, coaching strategies and performance was data acquisition. As noted by Horvitz (2007): "Designs for introducing intelligent reasoning into the world often depend critically on acquiring a case library of rich data that can be used to build predictive systems". Data collection methods were considered carefully to insure the type of data, the amount and the quality of the data was sufficient to provide statistically significant results. The automation of data collection was a critical priority to reduce time in data analysis.

Algorithm selection and performance

A second consideration in developing an affective learner model was the selection of a method to evaluate variable relationships. Regression analysis was one method used. Another method involved the evaluation of fifty (50) machine learning classifiers to determine which would provide the best predictive performance for the selected problem domain and representations (attributes and outcomes). Classifiers are algorithms used to determine relationships between variables and generally fall into one of two categories based (nominal data or numerical data) although some hybrid classifiers allow the use of both nominal and numerical data. The testing of the very large number of classifiers available was accomplished using WEKA (Witten and Frank, 2005), a machine learning toolset.

Experimental Objectives

Four experimental objectives were put forth in this study. The first objective was determine the relationship between student mood variables, the ITS coaching strategy used and student performance. The second objective was to determine the influence student state variables (e.g. energy level and apriori knowledge) and student actions (e.g. mouse movement rate, control selection rate or help request frequency) have on student performance. The third objective was to determine the influence student state variables and student actions have in unobtrusively predicting mood. The fourth objective was to determine if the coaching strategy used and its associated method of instruction affected student performance. Five hypotheses were tested to help meet these objectives.

Hypotheses Under Test

The variables associated with each hypothesis are reviewed and the rationale for their use in this study is also provided.

Hypothesis "A": Assessing Predictors of Performance

Hypothesis "A" was posed to determine if the coaching strategy used in the performance assessment and its associated method of instruction affected student performance. Kim, et al (2007), Dempsey & van Eck, (2003), van Eck & Dempsey, (2002) and Aleven & Koedinger, (2000) indicated that giving the student more control through coaching strategies like student-requested feedback (SRF) might not be the "best solution for weak learners" Although the participants were considered to have low apriori knowledge of the subject matter, they are classified as "strong learners". Based on the literature reviewed above, it is hypothesized that "*the mean student performance in the SRF coaching strategy group will be significantly higher than the mean student performance in the TIF coaching strategy group*".

Hypotheses "B", "C" and "D": Assessing Predictors of Performance

In order to assess the minimum number of variables needed to accurately predict student performance, three hypotheses (B, C and D) were posed and tested in this research. Hypothesis "B" evaluated whether student mood variables (e.g. pleasure or dominance) alone were significant predictors of student performance for a given ITS coaching strategy. Based on Bandura's (1977) social learning theory and the literature reviewed above, it is hypothesized that "*student mood variables are predictors of student performance for a given ITS coaching strategy*".

Hypothesis "C" evaluated whether student state and student action variables were significant predictors of student performance for a given ITS coaching strategy. Based on the literature reviewed above, it is hypothesized that "*student state variables and student action variables are predictors of student performance for a given ITS coaching strategy*".

Hypothesis "D" evaluated whether student action variables alone were significant predictors of student performance for a given ITS coaching strategy. Based on the literature reviewed above, it is hypothesized that "*student action variables are predictors of student performance for a given ITS coaching strategy*".

The dependent variable for Hypotheses "B", "C" and "D" was <u>student performance</u>. Performance, the act of accomplishing something such as a task or action, was chosen to demonstrate the practical application of coaching strategies and mood within ITS. The independent variable was the ITS coaching strategy which was manipulated to be either the SRF or TIF coaching strategy. Predictor variables for Hypothesis "B" included student mood variables. Predictor variables for Hypothesis "C" included student state variables and student action variables. Predictor variables for Hypothesis "D" included student action variables only. The variables for Hypotheses "B", "C" and "D" are shown in Table 1.

Table 1: Hypotheses "B", "C" and "D" Experiment Variables

Variable	Variable Type	Measurement Method	Measurement Scale
Performance Score	Dependent	Tested	Interval (0-100)
Coaching Strategy	Independent	Randomly Selected	Nominal (SRF or TIF)
Student Mood Variables			
Initial Pleasure	Predictor	Self Reported	Interval (1-9)
Initial Arousal	Predictor	Self Reported	Interval (1-9)
Initial Dominance	Predictor	Self Reported	Interval (1-9)
Initial Mood Intensity	Predictor	Derived from intial pleasure, arousal and dominance scores	Interval
Final Pleasure	Predictor	Self Reported	Interval (1-9)
Final Arousal	Predictor	Self Reported	Interval (1-9)
Final Dominance	Predictor	Self Reported	Interval (1-9)
Final Mood Intensity	Predictor	Derived from final pleasure, arousal and dominance scores	Interval
Student State Variables			
Sleep	Predictor	Self Reported	Interval (3 - 9)
Energy Level	Predictor	Self Reported	Interval (1-5)
Computer Confidence Level	Predictor	Self Reported	Interval (1-5)
Self-Assessed First Aid Knowledge	Predictor	Self Reported	Interval (1-5)
Topic Training Experience	Predictor	Self Reported	Interval (1-5)
Topic Interest Level	Predictor	Self Reported	Interval (1-5)
Initial Knowledge	Predictor	Tested	Interval (0-100)
Final Knowledge	Predictor	Tested	Interval (0-100)
Knowledge Improvement	Predictor	Derived from initial and final knowledge scores	Interval (0-100)
Student Action Variables			
Help Request Frequency (SRF Only)	Predictor	Captured	Interval
Mouse Movement Rate	Predictor	Captured	Interval
Action Rate	Predictor	Captured	Interval

Mood variables were selected as predictor variables over other affective variables (personality or emotion) based on their duration, influence and cause. Mood generally has a moderate duration whereas emotions have a short duration and personality attributes have a long duration (Gebhard, 2005). Mood was selected over personality and emotion because mood provided the opportunity to observe changes in the student's affective state during a single training session, but did not change easily based on a single positive or negative incident. Mood was also selected because of its subtle influence on cognition which tends to go unnoticed even by human tutors and should be accounted for in the learning process. Davidson (1994) argued that "emotions bias action, whereas moods bias cognition" which indicates that mood may affect perception and reasoning during learning. Except for mood intensity which was derived, mood variables were measured on a 9-point scale using Lang's (1980) Self Assessment Manikin (SAM) shown in Appendix F.

"Mood intensity" was derived as the vector sum in the three dimensional space formed by pleasure, arousal, and dominance mood dimensions. The magnitude of mood intensity for a given subject is assumed to be the vector sum calculated by equal weighting of the mood dimension levels of the subject. Mood intensity (initial and final) were calculated as shown in Equation 1.

$$\textbf{Mood intensity} = \textbf{SQRT}\ (\textbf{pleasure}^2 + \textbf{arousal}^2 + \textbf{dominance}^2) \qquad \textbf{(1)}$$

Student state variables (e.g. the amount of sleep experienced the previous night, energy level, computer confidence level, a priori knowledge level of the topic and interest in the topic.) and student action variables (e.g. mouse movement rates and action rates) were assumed to vary normally. Sleep was self-reported as the number of hours of sleep the previous night rounded to the nearest whole hour. Other predictor variables (energy level, computer confidence level and interest in the topic) were reported by the participant and measured on a 5-point scale. Mouse movement rates were captured and assigned as the average pixel movements per second during the performance assessment. Action rates, the selection of controls over time, were captured and assigned as the average number of controls selected per minute during the performance assessment. Knowledge (initial and final) survey scores were translated to 100 point interval scales to match the performance scale. Knowledge improvement was defined as simply the difference between final and initial knowledge.

Hypothesis "E": Assessing Predictors of Mood

Hypothesis E was posed to determine the significant unobtrusive predictors of mood. Hypothesis "E" evaluated whether student state variables, student action variables and student performance were unobtrusive predictors of student mood variables. Based on Zimmerman, et al (2003) and the literature reviewed above, it is hypothesized that "*student state variables, student action variables and student performance are predictors of student mood independent of coaching strategy*".

In Hypothesis "E", mood variables (pleasure, arousal, dominance and mood intensity) were evaluated as dependent variables to determine if mood could be predicted by a combination of student state variables (e.g. the amount of sleep experienced the previous night, energy level, computer confidence level, a priori knowledge level of the topic and interest in the topic) and student action variables (e.g. mouse movement and action rates). No independent variables were manipulated as part of this hypothesis. Variables associated with Hypothesis "E" are shown in Table 2.

Table 2: Hypothesis “E” Experiment Variables

Variable	Variable Type	Measurement Method	Measurement Scale
Student Mood Variables			
Initial Pleasure	Dependent	Self Reported	Interval (1-9)
Initial Arousal	Dependent	Self Reported	Interval (1-9)
Initial Dominance	Dependent	Self Reported	Interval (1-9)
Initial Mood Intensity	Dependent	Derived from intial pleasure, arousal and dominance scores	Interval
Final Pleasure	Dependent	Self Reported	Interval (1-9)
Final Arousal	Dependent	Self Reported	Interval (1-9)
Final Dominance	Dependent	Self Reported	Interval (1-9)
Final Mood Intensity	Dependent	Derived from final pleasure, arousal and dominance scores	Interval
Performance Score	Predictor	Tested	Interval (0-100)
Student State Variables			
Sleep	Predictor	Self Reported	Interval (3 - 9)
Energy Level	Predictor	Self Reported	Interval (1-5)
Computer Confidence Level	Predictor	Self Reported	Interval (1-5)
Self-Assessed First Aid Knowledge	Predictor	Self Reported	Interval (1-5)
Topic Training Experience	Predictor	Self Reported	Interval (1-5)
Topic Interest Level	Predictor	Self Reported	Interval (1-5)
Initial Knowledge	Predictor	Tested	Interval (0-100)
Final Knowledge	Predictor	Tested	Interval (0-100)
Knowledge Improvement	Predictor	Derived from initial and final knowledge scores	Interval (0-100)
Student Action Variables			
Help Request Frequency (SRF Only)	Predictor	Captured	Interval
Mouse Movement Rate	Predictor	Captured	Interval
Action Rate	Predictor	Captured	Interval

Apparatus

The content of this study was delivered to the participants via a multimedia (e.g. static pictures, video, concept animation and audio) presentation on laptop computers. One laptop was used for each participant and Microsoft PowerPoint was used to present the content. Visual Basic for Applications (VBA) was used to develop the interactive controls, to develop the coaching strategy logic, to timestamp events and to record each participant’s interaction and survey data. The data collection and stimuli included questionnaires, surveys, a TC3 training course, and a TC3 performance assessment. The experiment procedure is shown in Figure 1 and further discussed below.

Experimental Procedure

Prior to starting the experiment, each participant was provided a USMA-approved participant consent form (see Appendix B) to read and sign. One was kept for USMA's files and one was provided to the participant. The UCF-approved participant consent form (see Appendix A) was provided to the participant electronically just prior to the start of the experiment. A hardcopy of the form was also available.

The experiment employed the 'between-subject' design which randomly distributed participants between two coaching strategy groups (SRF and TIF). Coaching strategies are part of the ITS pedagogical model that interacts with the ITS student model to determine instruction content and presentation. As mentioned earlier, coaching strategy selection is an added layer on top of the TC3 interactive multimedia MS PowerPoint training package. For the purposes of this experiment, we limited the scope of the training to a subset of the U.S. Army's TC3 course and focused solely on training the hemorrhage control task. The information presented in the course was interwoven with coaching points and practice questions.

The SRF coaching strategy used a tutor in the form of a picture of a medic as the mechanism for the participant to request help. When the SRF tutor control was selected the tutor provided direction, a hint or encouragement that the participant was progressing in the appropriate direction. The tutor only provided help when asked.

The TIF coaching strategy only provided information that the experimenter deemed necessary to meet the TC3 performance objectives as noted in the U.S. Army TC3 course.

Unlike the SRF coaching strategy, there was no option in the TIF strategy for the participant to request help from the coach. The coach provided feedback to the student when thresholds were met that determined that the student was out of line with the correct solutions. If the students' performance was below satisfactory levels, encouragement, hints or direction were provided. If the student was at or above satisfactory levels, feedback was provided to let the student they were "on track".

Participants took part in the experiment in groups of eight or less. Upon entering the training room, each participant was instructed to select one of eight available seats by a table with a laptop computer on it (Step 1 in Figure 1). Four of the laptops provide the SRF coaching strategy and the other four provide the TIF coaching strategy. The strategies on each laptop were changed between training sessions so as to avoid any possible bias.

Demographic information (age, gender, academic year and academic department) and student state information (energy level, amount of sleep, level of confidence in using a computer, general first aid experience and interest level in TC3) was collected through a biographical survey (Step 2 in Figure 1) composed of eleven (11) questions shown in Appendix E. All surveys, training and assessments were administered via the multimedia environment. Measurement scales for the demographics and all other variables in this research are described in Table 1.

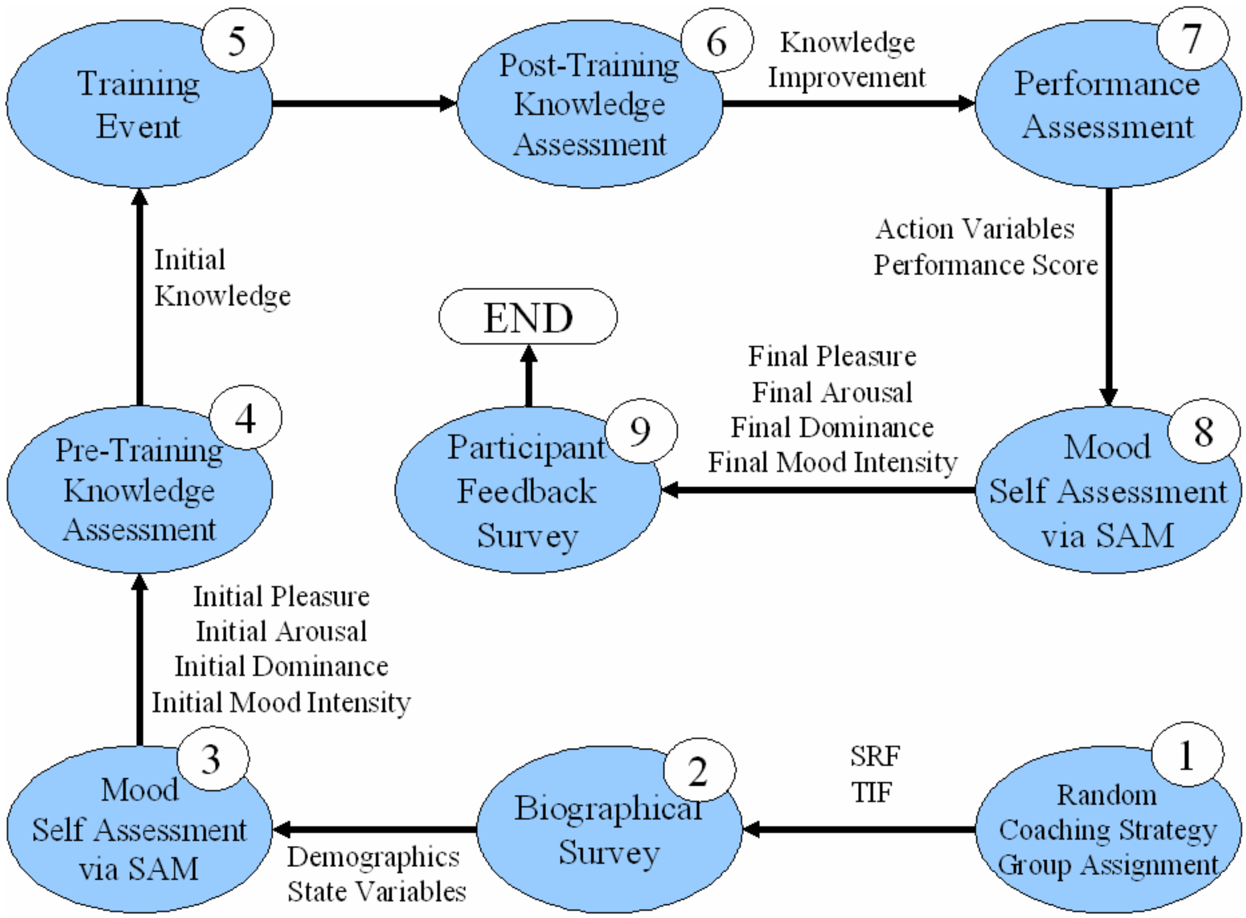


Figure 1: Experimental Procedure

Next, information on the initial mood variables was collected through the Self-Assessment Manikin (SAM) survey (Lang, 1980) (Step 3 in Figure 1). The SAM survey shown in Figure 2 and in Appendix F was administered twice to each participant: once prior to the training course and once at the completion of the TC3 performance assessment (Step 8 in Figure 1). SAM is a graphical survey with three sets of pictures representing pleasure, arousal and dominance in Mehrabian's (1996) mood model. Each of the three picture sets vary from happy to sad, excited to calm, and in control to powerless corresponding to pleasure, arousal and dominance respectively. Mood intensity was measured as a scalar of the three mood components (see Equation 1).

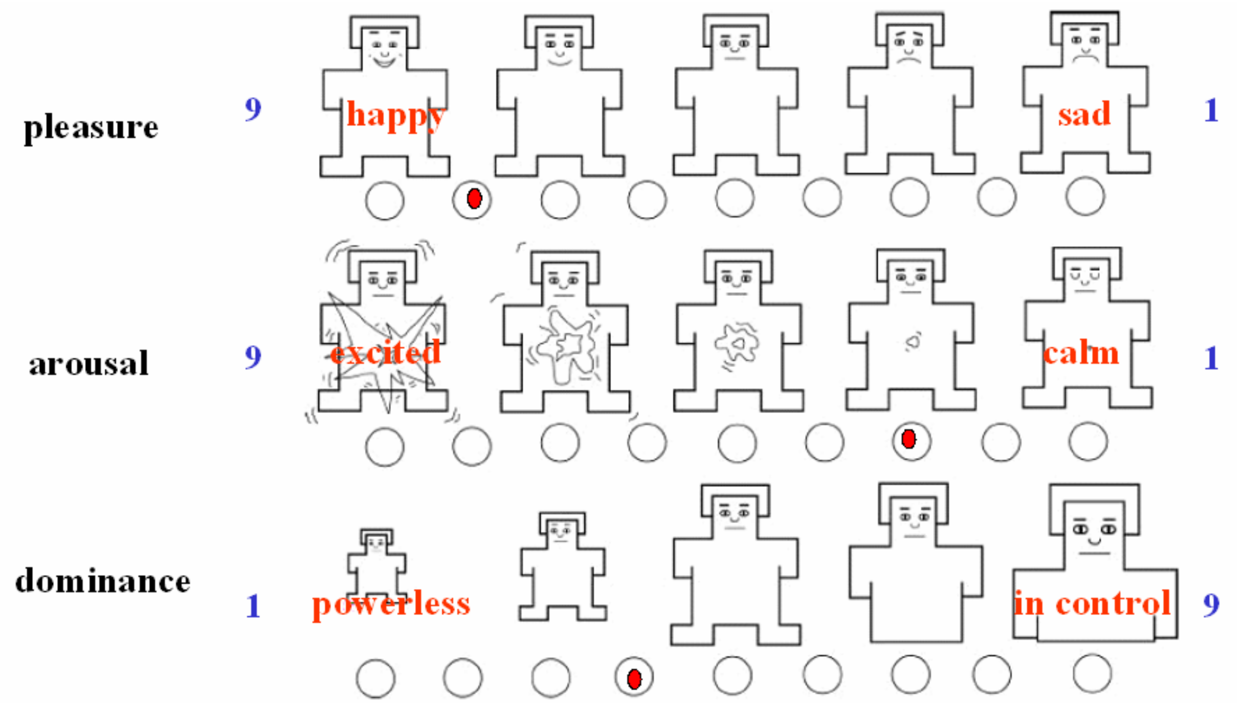


Figure 2: Self-assessment Manikin (SAM) (Lang, 1985)

After the initial mood assessment, a knowledge pre-training assessment (Step 4 in Figure 1), shown in Appendix G composed of fifteen (15) questions on a 100-point interval scale, was administered to each participant to ascertain each participant's knowledge of TC3 prior to the delivery of instruction. A knowledge post-training assessment (Step 6 in Figure 1) was administered after the delivery of instruction.

Next, the TC3 course was delivered (Step 5 in Figure 1). This training course was composed of interactive multimedia content based on a subset of the U.S. Army's TC3 course that focused on hemorrhage control. The information presented in the course was interwoven with coaching points and practice questions. The goal of the course was to teach the key objectives for TC3. All participants received identical training course content and the method of instruction and feedback was identical for all participants during the TC3 training course.

After the TC3 course and final knowledge assessment, the performance assessment (Step 7 in Figure 1) was delivered to each participant. The performance assessment differed from the TC3 course in that the course focused on acquiring knowledge (facts and principles) while the performance test focused on applying knowledge and demonstrating skill through decisions and actions. A screen capture from the TC3 performance assessment, an interactive simulation that required perception, judgment and action by the participant, is shown in Figure 3.

Figure 3: Screen capture from TC3 performance assessment for the SRF coaching strategy

Student performance was measured using a 100-point scale during the conduct of a TC3 training event in which the participant treated an injured virtual soldier. The ITS tracked student actions like mouse movement rates through polygonal traps (shown in Figure 4) that captured the number of pixels the cursor moved during the performance assessment.

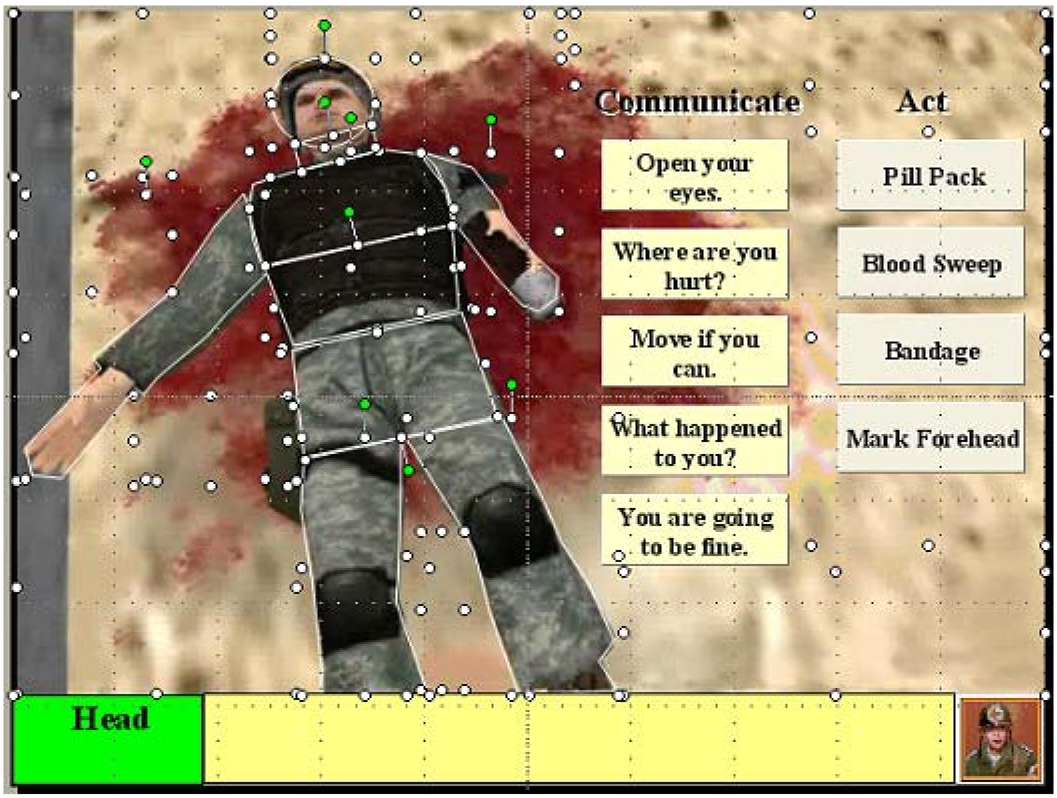


Figure 4: Capturing action in TC3 performance assessment

The multi-media training package captured key parameters related to the health of the virtual Soldier and used these to provide a performance score. Some of these key parameters included the amount of blood lost by the virtual Soldier, the number of body parts the participant performed a blood sweep on, the application of pressure bandage and tourniquet, marking the time on the forehead of the virtual Soldier once a tourniquet was applied, number of times the participant communicated with the virtual Soldier and the appropriate application of medication to control pain.

The performance assessment was used to evaluate the participant's skill (the application of knowledge) acquired in the TC3 training course. A hemorrhage scenario tested the participant's knowledge and the use of casualty care tools (e.g. a tourniquet, medication or a pressure bandage). The scenario modeled the casualty's blood loss over time. The performance assessment score was based on the participant's accuracy (percentage of correct actions vs. required actions), speed (time to complete in seconds) and the scenario outcome (e.g. "stable soldier" or "deceased soldier").

Finally, the Participant Feedback Survey (Step 9 in Figure 1), composed of fourteen (14) questions was used to assess the experimental process and the student's perception of the intelligent tutor, and to provide the participant an opportunity to give feedback to the principal investigator. The participant feedback survey questions shown in Appendix H were used to assess the participant's perception of the intelligent tutor's trustworthiness, competence and supportiveness. A comparison of the participant's performance and perceptions of the tutor was conducted using a Student's t-Test. The t-Test compared participants at the same level of performance (e.g. high, moderate or low) across the two coaching strategies. The results of this analysis are discussed in the results section below.

Data Analysis Methods

A goal of this research was to create an intelligent coaching feedback selection algorithm or classifier algorithm based on student mood, student performance and other student state variables and student actions. Data analysis was accomplished primarily using four tools: Arbuckle's (2008) Analysis of MOment Structures (AMOS) tool was used to compare the predicted performance for each strategy and then choose an optimal strategy. A set of production rules were developed to train a machine learning classifier using Witten & Frank's (2005) Waikato Environment for Knowledge Analysis (WEKA) toolset. Finally, post-hoc power calculations and effect size calculations were completed using Soper's (2009) online statistical calculators.

CHAPTER FOUR: DATA AND ANALYSIS

Chapter Four Summary

This chapter reviews the results of this study including descriptive statistics, hypotheses test results from the regression analysis using AMOS and the "regression by discretization using the J48 decision tree" model developed using WEKA.

Descriptive Statistics

Descriptive statistics (mean, standard deviation and range) are provided for the total sample, the SRF and the TIF coaching strategy groups in Tables 3, 4 and 5 respectively. The correlation of each variable with the primary dependent variable, performance, is provided in the last column of each table.

Table 3: Descriptive & Performance Correlation Statistics for the Total Study Sample

Variables	*Mean*	*Standard Deviation*	*Range*	*Correlation with Performance*
Sleep	5.838	1.206	6	0.102
Energy Level	2.927	0.055	4	0.037
Computer Confidence	4.137	0.747	3	0.123
TC3 Training Experience	3.630	1.885	5	-0.004
Self-Assessed First Aid Knowledge	2.944	0.747	4	0.019
TC3 Interest Level	3.266	0.980	4	0.158
Initial Pleasure	6.032	1.223	6	0.088
Initial Arousal	4.048	1.743	8	0.022
Initial Dominance	6.073	1.380	6	0.289
Initial Mood Intensity	9.651	1.712	9.289	0.186
Final Pleasure	5.363	1.630	7	0.224
Final Arousal	5.282	1.641	7	0.117
Final Dominance	5.782	1.641	7	0.393
Final Mood Intensity	9.706	1.976	10.343	0.353
Help Request Frequency	[redacted]			
Mouse Movement Rate	197.540	55.447	307	0.251
Action Rate	12.433	3.278	20.048	0.144
Initial Knowledge	49.435	11.533	60	0.160
Final Knowledge	73.911	12.123	80	0.107
Knowledge Improvement	24.476	12.523	60	0.263
Performance	80.992	12.210	66	1.000

Table 4: Descriptive & Performance Correlation Statistics for SRF Coaching Strategy Group

Variables	*Mean*	*Standard Deviation*	*Range*	*Correlation with Performance*
Sleep	6.048	1.165	6	0.096
Energy Level	2.935	0.698	4	-0.048
Computer Confidence	4.161	0.729	3	0.085
TC3 Training Experience	3.532	1.922	5	0.049
Self-Assessed First Aid Knowledge	2.968	0.768	3	0.159
TC3 Interest Level	3.194	1.038	4	0.367
Initial Pleasure	5.839	1.190	6	0.156
Initial Arousal	4.000	1.699	7	-0.011
Initial Dominance	5.935	1.341	6	0.338
Initial Mood Intensity	9.422	1.608	8.184	0.238
Final Pleasure	5.081	1.529	7	0.357
Final Arousal	5.194	1.491	7	0.033
Final Dominance	5.661	1.619	7	0.580
Final Mood Intensity	9.399	1.903	10.343	0.462
Help Request Frequency	4.048	3.919	14	0.273
Mouse Movement Rate	209.161	61.286	287	0.325
Action Rate	12.633	2.744	12.226	0.182
Initial Knowledge	48.629	11.386	55	0.226
Final Knowledge	73.548	13.501	80	0.260
Knowledge Improvement	24.919	11.682	50	0.081
Performance	80.774	12.704	62	1.000

Table 5: Descriptive & Performance Correlation Statistics for TIF Coaching Strategy Group

Variables	*Mean*	*Standard Deviation*	*Range*	*Correlation with Performance*
Sleep	5.629	1.218	6	0.117
Energy Level	2.919	0.522	3	0.160
Computer Confidence	4.113	0.770	2	0.163
TC3 Training Experience	3.726	1.857	5	-0.066
Self-Assessed First Aid Knowledge	2.919	0.731	3	-0.139
TC3 Interest Level	3.339	0.922	4	-0.097
Initial Pleasure	6.226	1.234	6	0.014
Initial Arousal	4.097	1.799	8	0.055
Initial Dominance	6.210	1.416	6	0.240
Initial Mood Intensity	9.879	1.793	9.033	0.136
Final Pleasure	5.645	1.690	7	0.098
Final Arousal	5.371	1.785	7	0.193
Final Dominance	5.903	1.667	6	0.199
Final Mood Intensity	10.012	2.015	8.476	0.246
Mouse Movement Rate	185.919	46.569	208	0.172
Action Rate	12.232	3.749	20.048	0.120
Initial Knowledge	50.242	11.715	45	0.089
Final Knowledge	74.274	10.668	50	0.267
Knowledge Improvement	24.032	13.393	60	0.135
Performance	81.210	11.795	62	1.000

When considering the information in the tables, the lack of a difference in performance between the two coaching strategy groups may be striking. It infers that coaching strategy does not play a role in achieving the level of performance observed. But closer examination of the underlying variables reveals that the coaching strategy acts independently on each subject to affect performance.

Performance scores were categorized as very high (90-100), high (80-89), moderate (70-79), low (60-69) and very low (below 60). Percentages of participants in each category for each coaching strategy group are shown in Table 6.

Table 6: Distribution of Performance Scores across Coaching Strategy Groups

	SRF	*TIF*
Very High	23%	15%
High	44%	56%
Moderate	16%	19%
Low	11%	2%
Very Low	6%	8%

In order to better understand what underlying variables were active in each coaching strategy environment, we performed a multiple regression analysis of appropriate variables to predict performance. A further regression analysis was also conducted to determine predictors of student mood independent of the coaching strategy used.

Hypotheses Test Results

Hypothesis "A" Test Results: Assessing Performance Differences in Coaching Strategy Groups

A Student's t-Test was conducted to test the hypothesis that "*the mean student performance in the SRF coaching strategy group was significantly higher than the mean student performance in the TIF coaching strategy group*". There was no significant effect for coaching strategies, $t(61) = -0.198$, $p = 0.844$. No support was found for this hypothesis. There is no significant difference in the mean performances of the SRF group and TIF group at $\alpha = 0.05$ which indicates that the coaching strategies and their associated methods of instruction did not affect performance, and that individual differences in performance were due to other factors like mood or other student state variables (e.g. sleep and energy levels) or student actions (e.g. mouse movement). This result also indicates that coaching strategies were correctly applied and misapplied in both groups.

Hypothesis "B" Test Results: Predicting Performance with Student Mood Data

To determine the relationship between mood, coaching strategies and performance a multiple regression analysis was conducted to test the hypothesis that "*student mood variables are predictors of student performance for a given ITS coaching strategy*". A structural equation modeling tool called the Analysis of Moment Structures (AMOS) was used to perform the regression analysis.

For the F test of the multiple R^2, a large effect size (ES) was expected and $f^2 = 0.35$ (from Table 1 of Cohen, 1992). Accordingly, a set of eight (8) potential predictors (initial pleasure, final pleasure, initial arousal, final arousal, initial dominance, final dominance, initial mood intensity and final mood intensity) were evaluated for each coaching strategy group and this indicated that the required minimum sample size for each group was 50 for a Power = 0.80.

As shown in Equation 2, the mood variable, final dominance, was the only significant predictor ($p < 0.01$) of performance in the SRF coaching strategy group. Final dominance was assessed immediately after the performance test was completed. This infers that participants in the SRF coaching strategy group were in control and that their confidence was strengthened by their ability to access the tutor to obtain help to answer questions. We believe that the high influence of final dominance is a good indication that participants were confident enough to recognize when they needed help and secure in asking for it.

$$\textbf{Performance}_{\textbf{SRF}} \textbf{ = 55.023 + (4.549 * final dominance)} \qquad \textbf{(2)}$$

Final dominance significantly predicted performance, $\beta = 0.000143$, $t(60) = 5.51$, $p < .01$. Final dominance also explained a significant proportion of variance in performance scores, $R^2 = .34$, $F(1, 60) = 30.36$, $p < .01$. Calculated effect size = 0.51 (large) which confirms the assumption of a large effect size per Cohen (1992). There is sufficient power and effect to indicate the practical, as well as the statistical significance, of final dominance as a predictor of performance in the SRF coaching strategy group. The mean absolute error for the SRF regression equation (see Equation 2) was 7.93 (11.13%).

For the TIF group, none of the eight (8) mood variables significantly predicted performance. Failure to support this hypothesis for the TIF group leaves us with no options for comparing predicted performances based on each coaching strategy.

Hypothesis "C" Test Results: Predicting Performance with Student State and Action Data

To determine the relationship between student state variables, student action variables, coaching strategies and performance a multiple regression analysis was conducted to test the hypothesis that "*student state variables and student action variables are predictors of student performance for a given ITS coaching strategy*". AMOS was again used to perform the regression analysis.

For the F test of the multiple R^2, a large effect size (ES) was expected and $f^2 = 0.35$ (from Table 1 of Cohen, 1992). Accordingly, a set of twelve predictors were used for each coaching strategy group and this indicated that the required minimum sample size for each group was 61 for a Power = 0.80.

As shown in Equation 3, TC3 interest level (the student's interest in the topic), and help request frequency (the number of times student requested feedback from the tutor) were significant predictors of performance in the SRF coaching strategy group.

$$\mathbf{Performance_{SRF} = 63.702 + (4.307 * TC3\ interest\ level) + (0.819 * help\ request\ frequency)} \qquad \mathbf{(3)}$$

TC3 interest level significantly predicted performance, β = 0.0580, t(59) = 3.01, $p < .01$. Help request frequency significantly predicted performance, β = 0.0580, t(59) = 2.16, p = .035.

TC3 interest level and help request frequency explained a significant proportion of variance in performance scores, $R^2 = .20$, $F(2, 59) = 7.28$, $p < .01$. Calculated effect size = 0.25 which indicates a medium to large effect size. There is sufficient power and effect to indicate the practical, as well as the statistical significance, of TC3 interest level and help request frequency as significant predictors of performance in the SRF coaching strategy group. The mean absolute errors for the SRF regression equation (see Equation 3) was 8.95 (12.76%).

As shown in Equation 4, only final knowledge (the student's knowledge assessment score after the TC3 training course and prior to the performance test) was a significant predictor of performance in the TIF coaching strategy group.

$$\textbf{Performance}_{\textbf{TIF}} = \textbf{59.283} + (\textbf{0.295} * \textbf{final knowledge}) \qquad \textbf{(4)}$$

Final knowledge significantly predicted performance, β = 0.4147, t(60) = 2.15, p = .036. Final knowledge explained a small proportion of variance in performance scores, $R^2 = .07$, $F(1, 60) = 4.61$, $p = .03$. Calculated effect size = 0.08 and post-hoc power is low (0.5853) which indicates little practical significance. The minimum group size required to meet the minimum acceptable power (0.80) for this relationship is calculated to be 104 participants and the group size was 62. No comparison of the predicted performances based on each coaching strategy was conducted based on the low practical significance of this hypothesis for the TIF group.

Hypothesis "D" Test Results: Predicting Performance with Student Action Data

To determine the relationship between coaching strategies, performance and student action variables (e.g. mouse movement rate, help request frequency and action rate) a multiple regression analysis was conducted to test the null hypothesis that "*student action variables are predictors of student performance for a given ITS coaching strategy*". Once again, AMOS was used to perform the regression analysis.

For the F test of the multiple R^2, a large effect size (ES) was expected and $f^2 = 0.35$ (from Table 1 of Cohen, 1992). Accordingly, a set of three predictors were used for the SRF coaching strategy group and a set of two predictors was used for the TIF coaching strategy group. This indicated that the required minimum sample size for each group was 34 in the SRF group and 30 in the TIF group for a Power = 0.80. Regression equations were developed to predict SRF and TIF performance respectively using AMOS.

As shown in Equation 5, only mouse movement rate was significant predictor of performance in the SRF coaching strategy group.

$$\textbf{Performance}_{\textbf{SRF}} = \textbf{66.671} + (\textbf{0.067} * \textbf{mouse movement rate}) \qquad \textbf{(5)}$$

Mouse movement rate significantly predicted performance, $\beta = 0.2129$, $t(60) = 2.64$, $p < .01$. Mouse movement rate explained a small, but significant proportion of variance in performance scores, $R^2 = .11$, $F(1, 60) = 7.10$, $p < .01$. Calculated effect size = 0.12 which medium effect size. While the relationship between mouse movement rate and performance is significant, the post-hoc power is slightly low (0.79). The minimum group size required to meet the minimum acceptable power (0.80) is 64 participants. Based on low power and low R^2, we determined that mouse movement rate is statistically significant, but may not be of practical significance.

For the TIF coaching strategy group, help request frequency was not an option and neither mouse movement rate nor action rate was a significant predictor of performance in the TIF coaching strategy group. .

The mean absolute errors for the SRF regression equation (see Equation 5) was 8.95 (13.03%). However the failure to find support for this hypothesis in the TIF group left no option for comparison of performance based on the two strategies.

Hypothesis "E" Test Results: Predicting Mood with Student State, Student Action and Student Performance Data

Mood was only self-assessed twice in this experiment so as not to influence or interfere with the learning process. The drawback to this was that mood state was not available throughout the learning process. This is why we also investigated student state, student action and performance variables to assess their ability to predict mood. The relationships evaluated were done without respect to their temporal relationships. This study found that pleasure, dominance and mood intensity could be predicted through combinations of student state, student actions and performance data. This hypothesis was evaluated across all participants and without respect to coaching strategy group.

To determine the relationship between mood variables and other independent variables (e.g. sleep or energy) a multiple regression analysis was conducted to test the null hypothesis that there is not a predictive relationship at $\alpha = 0.05$ between student state variables, student action variables and student mood variables. Once again, AMOS was used to perform the regression analysis.

For the F test of the multiple R^2, a large effect size (ES) was expected and $f^2 = 0.35$ (from Table 1 of Cohen, 1992). Accordingly, a maximum set of twelve predictors were used for each mood variable and this indicated that the required minimum sample size for each group was 61 for a Power = 0.80. Eight hypotheses were tested; one for each mood variable. Regression equations were developed if significant relationships existed between mood variables and/or student state, student action or performance variables.

For initial pleasure and final arousal, no statistically significant relationships with student state, student action or performance variables were identified. For initial arousal, significant main effects exist, but with sufficient power and effect to be of practical significance. For initial dominance, initial mood intensity, final pleasure, final dominance and final mood intensity, support for these hypotheses was found. The regression statistics for mood attributes with a significant main effect are shown in Table 7.

Table 7: Mood attributes with a Significant Main Effect Without Respect to Coaching Strategy

					Score Prediction		Variance Prediction			Effect
Model	**No. of Significant Predictors**	**Predictors**	**N**	**Post-hoc Power (1-β)**	**t**	**p**	**R2**	**F**	**Sig F**	**f^2**
Initial Pleasure	1	Energy Level	124	.81	t(122) = 2.75	< .01	.06	F(1,122) = 7.55	< .01	.06
Initial Arousal	1	Energy Level	124	.63	t(122) = 2.26	.025	.04	F(1,122) = 5.13	.025	.04
Initial Dominance	3	Self-Assessed First Aid Knowledge	124	> 0.99	t(120) = 2.95	< .01	.19	F(3,120) = 9.43	< .01	.24
		Mouse Movement Rate			t(120) = -2.67	< .01				
		Performance			t(120) = 4.02	< .01				
Initial Mood Intensity	4	Action Rate	124	0.99	t(119) = -2.37	.02	.17	F(4,119) = 6.14	< .01	.20
		Energy Level			t(119) = 2.40	.02				
		Self-Assessed First Aid Knowledge			t(119) = 3.11	< .01				
		Performance			t(119) = 2.40	.02				
Final Pleasure	3	Action Rate	124	.98	t(120) = -2.7	< .01	.15	F(3,120) = 6.14	< .01	.18
		Self-Assessed First Aid Knowledge			t(120) = 2.22	< .03				
		Performance			t(120) = 2.96	< .01				
Final Arousal	1	Energy Level	124	.63	t(122) = 2.14	.03	.04	F(1,122) = 4.57	.03	.04
Final Dominance	2	TC3 Interest Level	124	.99	t(121) = 2.11	.04	.18	F(2,121) = 13.71	< .01	.22
		Performance			t(121) = 4.40	< .03				
Final Mood Intensity	2	TC3 Interest Level	124	0.99	t(121) = 2.32	.02	.16	F(2,121) = 11.66	< .01	.19
		Performance			t(121) =3.82	< .01				

Mood regression equations are shown in Equations 6-10.

$$\textbf{Initial Dominance} = 2.744 + (0.448 * \text{self-assessed first aid knowledge}) + (0.038 * \text{performance}) – (0.006 * \text{mouse movement rate}) \quad (6)$$

$$\textbf{Initial Mood Intensity} = 5.254 + (0.560 * \text{energy level}) + (0.598 * \text{self-assessed first aid knowledge}) + (0.028 * \text{performance}) – (0.105 * \text{action rate}) \quad (7)$$

$$\textbf{Final Pleasure} = 2.835 + (0.413 * \text{self-assessed first aid knowledge}) + (0.034 * \text{performance}) – (0.116 * \text{action rate}) \quad (8)$$

$$\textbf{Final Dominance} = 0.844 + (0.294 * \textbf{TC3 interest level}) + (0.049 * \textbf{performance}) \qquad (9)$$

$$\textbf{Final Mood Intensity} = 4.201 + (0.394 * \textbf{TC3 interest level}) + (0.052 * \textbf{performance}) \qquad (10)$$

<u>Composite Model Using AMOS and Hypothesis Results</u>

Based on the results of the hypothesis testing, three composite models were constructed that included performance, mood, student state and student action relationships. The three composite models for SRF, TIF and Both Coaching Strategies are shown in Figure 5,

As shown in Figure 5, the composite model for the SRF coaching strategy group explained a significant proportion of variance in performance scores, R^2 = .65; initial dominance, R^2 = .25; initial mood intensity, R^2 = .26; final pleasure, R^2 = .24; final dominance, R^2 = .11; and final mood intensity, R^2 = .25. Regression weights and standardized regression weights for the SRF Composite Model are shown in Table 8 and Table 9 respectively.

Table 8: Regression Weights for SRF Composite Model

Table 9: Standardized Regression Weights for SRF Composite Model

Figure 6 and

Figure 7 respectively. The purpose of the composite models was to analyze the variable relationships and their effects on each other as part of a system. The composite models were constructed using AMOS and only statistically significant variables generated from the Hypotheses B-E were used to predict performance and mood variables.

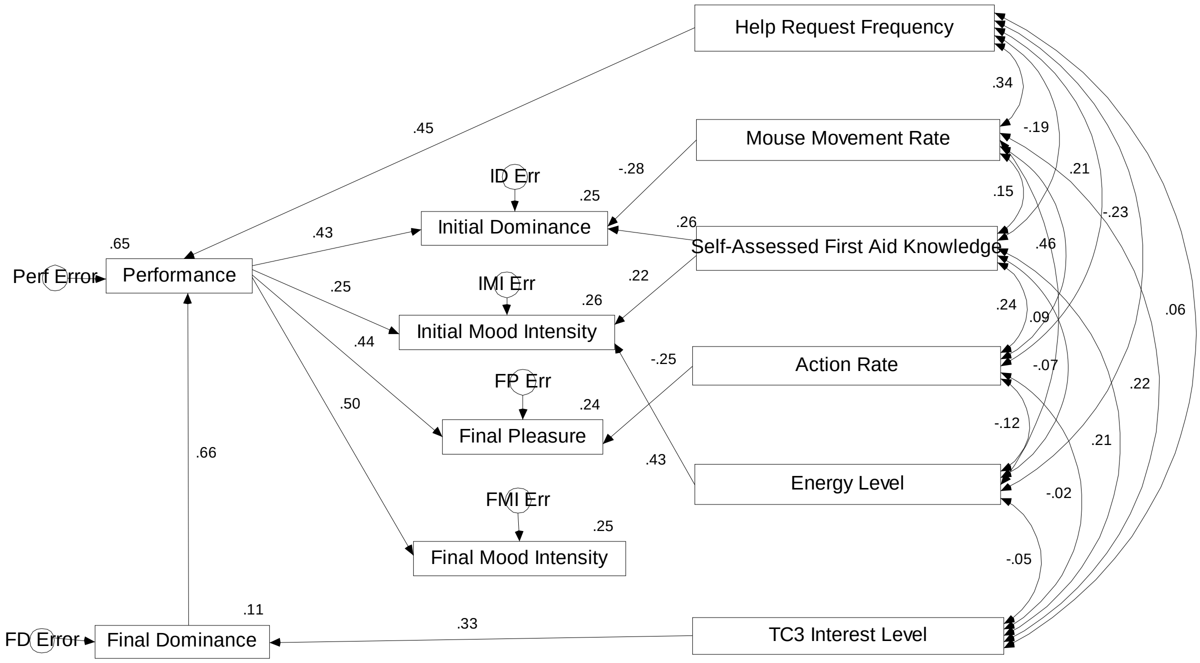


Figure 5: SRF Composite Model

As shown in Figure 5, the composite model for the SRF coaching strategy group explained a significant proportion of variance in performance scores, R^2 = .65; initial dominance, R^2 = .25; initial mood intensity, R^2 = .26; final pleasure, R^2 = .24; final dominance, R^2 = .11; and final mood intensity, R^2 = .25. Regression weights and standardized regression weights for the SRF Composite Model are shown in Table 8 and Table 9 respectively.

Table 8: Regression Weights for SRF Composite Model

			Estimate	S.E.	C.R.	P
Final Dominance	<---	TC3 Interest Level	.519	.188	2.753	.006
Performance	<---	Final Dominance	5.800	.661	8.770	***
Performance	<---	Help Request Frequency	1.641	.273	6.005	***
Initial Mood Intensity	<---	Performance	.028	.012	2.279	.023
Final Pleasure	<---	Performance	.049	.012	3.921	***
Final Mood Intensity	<---	Performance	.069	.015	4.548	***
Initial Dominance	<---	Mouse Movement Rate	-.006	.003	-2.441	.015
Initial Dominance	<---	Self-Assessed First Aid Knowledge	.466	.198	2.353	.019
Initial Dominance	<---	Performance	.041	.011	3.786	***
Initial Mood Intensity	<---	Energy Level	.986	.254	3.888	***
Final Pleasure	<---	Action Rate	-.146	.064	-2.267	.023
Initial Mood Intensity	<---	Self-Assessed First Aid Knowledge	.448	.229	1.952	.051

Table 9: Standardized Regression Weights for SRF Composite Model

			Estimate
Final Dominance	<---	TC3 Interest Level	.332
Performance	<---	Final Dominance	.661
Performance	<---	Help Request Frequency	.453
Initial Mood Intensity	<---	Performance	.254
Final Pleasure	<---	Performance	.440
Final Mood Intensity	<---	Performance	.503
Initial Dominance	<---	Mouse Movement Rate	-.280
Initial Dominance	<---	Self-Assessed First Aid Knowledge	.264
Initial Dominance	<---	Performance	.429
Initial Mood Intensity	<---	Energy Level	.433
Final Pleasure	<---	Action Rate	-.254
Initial Mood Intensity	<---	Self-Assessed First Aid Knowledge	.216

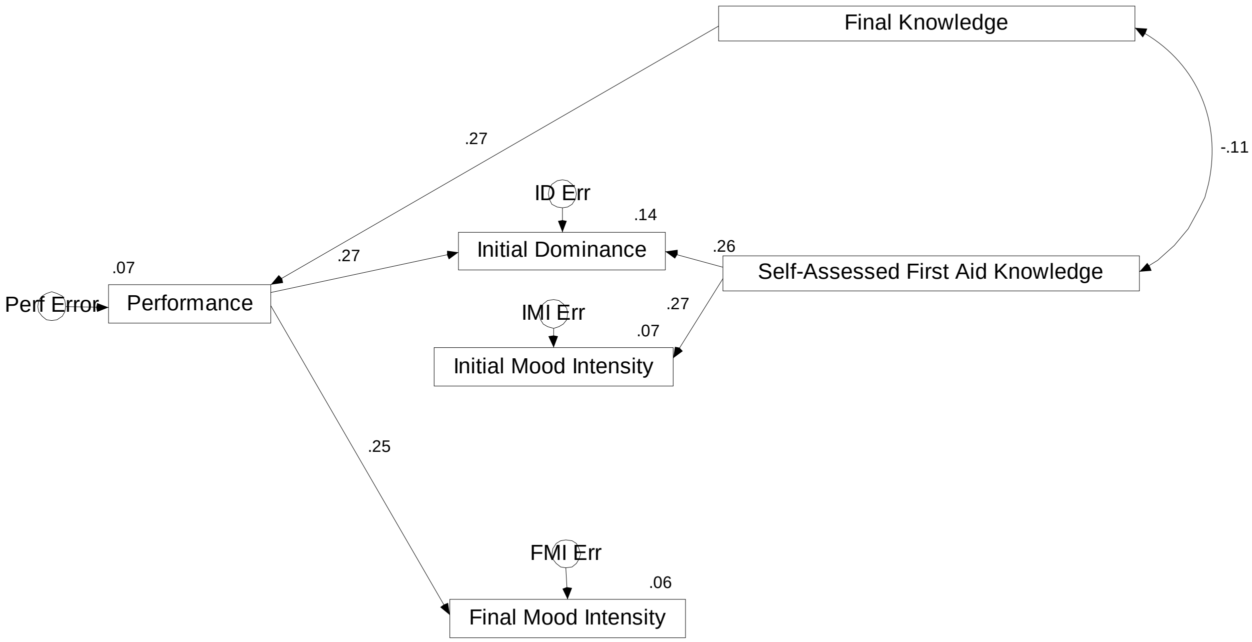


Figure 6: TIF Composite Model

As shown in Figure 6, the composite model for the TIF coaching strategy group explained a small, but significant proportion of variance in performance scores, $R^2 = .07$; initial dominance, $R^2 = .14$; initial mood intensity, $R^2 = .07$; and final mood intensity, $R^2 = .25$. Regression weights and standardized regression weights for the TIF Composite Model are shown in Table 10 and

Table 11 respectively.

Table 10: Regression Weights for TIF Composite Model

			Estimate	S.E.	C.R.	P
Performance	<---	Final Knowledge	.295	.136	2.164	.030
Final Mood Intensity	<---	Performance	.042	.021	1.985	.047
Initial Dominance	<---	Self-Assessed First Aid Knowledge	.505	.232	2.176	.030
Initial Dominance	<---	Performance	.033	.014	2.302	.021
Initial Mood Intensity	<---	Self-Assessed First Aid Knowledge	.657	.303	2.170	.030

Table 11: Standardized Regression Weights for TIF Composite Model

			Estimate
Performance	<---	Final Knowledge	.267
Final Mood Intensity	<---	Performance	.246
Initial Dominance	<---	Self-Assessed First Aid Knowledge	.259
Initial Dominance	<---	Performance	.274
Initial Mood Intensity	<---	Self-Assessed First Aid Knowledge	.268

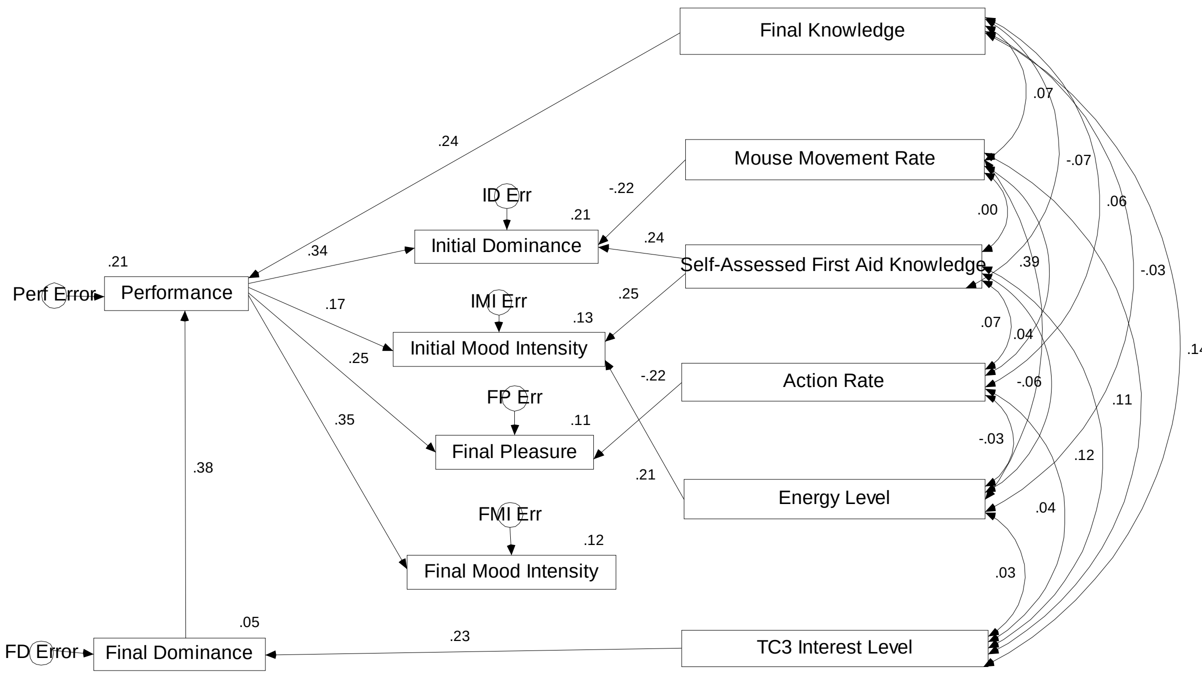


Figure 7: Composite Model for Both Coaching Strategies

As shown in Figure 7, the composite model for the both coaching strategy groups explained a significant proportion of variance in performance scores, $R^2 = .21$; initial dominance, $R^2 = .21$; initial mood intensity, $R^2 = .13$; final pleasure, $R^2 = .11$; final dominance, $R^2 = .05$; and final mood intensity, $R^2 = .12$. Regression weights and standardized regression weights for the Composite Model both strategies are shown in Table 12 and Table 13Table 9 respectively.

Table 12: Regression Weights for Composite Model for Both Coaching Strategies

			Estimate	S.E.	C.R.	P
Performance	<---	Final Knowledge	.295	.136	2.164	.030
Final Mood Intensity	<---	Performance	.042	.021	1.985	.047
Initial Dominance	<---	Self-Assessed First Aid Knowledge	.505	.232	2.176	.030
Initial Dominance	<---	Performance	.033	.014	2.302	.021
Initial Mood Intensity	<---	Self-Assessed First Aid Knowledge	.657	.303	2.170	.030

Table 13: Standardized Regression Weights for Composite Model for Both Coaching Strategies

			Estimate
Performance	<---	Final Knowledge	.267
Final Mood Intensity	<---	Performance	.246
Initial Dominance	<---	Self-Assessed First Aid Knowledge	.259
Initial Dominance	<---	Performance	.274
Initial Mood Intensity	<---	Self-Assessed First Aid Knowledge	.268

Construction of a classifier model using WEKA

In order to explore the potential for improving the accuracy of the equations developed under the multiple regression analysis, approximately fifty (50) machine learning classifiers in the WEKA Explorer toolset (Witten, 2005) were evaluated using the experimental dataset. Based on the accuracy results (mean absolute error = 3.787 [5.58%] compared to 7.93 (11.13%) for the SRF equation generated in Hypothesis "B") and a high coefficient of determination (R^2 = 0.742 compared to the best R^2 = 0.65 generated in the SRF composite model), the "regression by discretization" method using the J48 tree classifier was chosen to build the prototype classifier. The "regression by discretization" method can use any classifier on a copy of the numerical data that has the class attribute (in this case, performance) divided into equally sized bins. The dataset in this study was divided into 25 discrete bins. The J48 tree classifier is a class for generating a pruned or unpruned C4.5 decision tree (Quinlan, 1993). In this case, a pruned tree (size = 89) with 45 leaves was created and is shown in part in

Figure 8. Nominal data for coaching strategy and gender was converted to numerical states and injected into this model. This provided a distinct advantage over the regression analysis using AMOS which was unable to handle hybrid data or digital data without errors.

```
Initial Mood Intensity <= 7.81
|  Final Dominance <= 3
|  |  Initial Pleasure <= 4: Very Low
|  |  Initial Pleasure > 4: Very Low
|  Final Dominance > 3
|  |  Final Pleasure <= 4
|  |  |  Coaching Strategy (SRF = 1) <= 0: Moderate
|  |  |  Coaching Strategy (SRF = 1) > 0
|  |  |  |  Initial Mood Intensity <= 7.07: Moderate
|  |  |  |  Initial Mood Intensity > 7.07: Low
|  |  Final Pleasure > 4
|  |  |  Final Knowledge <= 85: Moderate to High
|  |  |  Final Knowledge > 85: High
Initial Mood Intensity > 7.81
|  Gender (Male = 1) <= 0
|  |  TC3 Training <= 3
|  |  |  Coaching Strategy (SRF = 1) <= 0: Very Low
|  |  |  Coaching Strategy (SRF = 1) > 0: Very Low
|  |  TC3 Training > 3
|  |  |  Computer Confidence <= 4
|  |  |  |  Final Arousal <= 5: Low
|  |  |  |  Final Arousal > 5: High
|  |  |  Computer Confidence > 4: Very High
```

Figure 8: Classifier developed through "regression by discretization"

The classifier output was numerical data that was translated to the performance ranges for very high (90-100), high (80-89), moderate (70-79), low (60-69) and very low (below 60). This output provides a mechanism to project performance based on student state, student mood, coaching strategy and gender. Action variables were not significant in the classifier model and not included in the classifier model. If incorporated in an ITS, the performance prediction output would trigger the ITS to change strategies assuming the change in strategies would result in a higher predicted performance.

The Student's Perception of the Intelligent Tutor

Participants were asked to rate the trustworthiness, competence and supportiveness of the intelligent tutor. The participants were asked rate statements regarding the tutor's trustworthiness, competence and supportiveness from 1 (strongly disagree) to 5 (strongly agree). A series of t-Tests were conducted to evaluate differences in the student's perception of the tutor.

Trustworthiness

For participants with high and very high performance scores, their agreement that the intelligent tutor was trustworthy was significantly lower in the SRF group ($M = 3.80$, $SD = 0.81$) than in the TIF group ($M = 4.20$, $SD = 0.55$), $t = -2.63$, $p < .01$, $d = .58$. The effect size (Cohen's d) indicates a medium effect and indicates the SRF participants perceived the tutor to be less trustworthy than the TIF participants.

Competence

For participants at all levels of performance scores, their agreement that the intelligent tutor was credible and competent was not significantly different in the SRF group than in the TIF group. For participants at all levels of performance scores, their agreement that the intelligent tutor provided useful feedback was not significantly different in the SRF group than in the TIF group.

Supportiveness

For participants at all levels of performance scores, their agreement that the intelligent tutor was encouraging and motivating was not significantly different in the SRF group than in the TIF group. For participants at all levels of performance scores, their agreement that the intelligent tutor provided timely guidance and feedback was not significantly different in the SRF group than in the TIF group.

For participants with moderate performance scores, their agreement that the intelligent tutor was annoying was significantly lower in the SRF group ($M = 1.70$, $SD = 0.82$) than in the TIF group ($M = 2.50$, $SD = 0.90$), $t = -2.17$, $p = .02$, $d = .93$. The effect size (Cohen's d) indicates a large effect and indicates that moderate performing SRF participants perceived the tutor to be less annoying than moderate performing TIF participants.

For participants with low and very low performance scores, their agreement that the intelligent tutor was annoying was significantly higher in the SRF group ($M = 3.00$, $SD = 0.89$) than in the TIF group ($M = 2.17$, $SD = 0.41$), $t = 2.63$, $p < .01$, $d = .43$. The effect size (Cohen's d) indicates a medium effect and indicates that low and very low performing SRF participants perceived the tutor to be more annoying than low and very low performing TIF participants.

CHAPTER FIVE: DISCUSSION AND RECOMMENDATIONS

Chapter Five Summary

This chapter reviews study conclusions and discusses lessons learned and recommendations for future research.

Conclusions

The test results of Hypothesis "A", "*the mean student performance in the SRF coaching strategy group will be significantly higher than the mean student performance in the TIF coaching strategy group*", indicated no difference in the mean performances of the SRF group and TIF group at $\alpha = .05$. From this we can rule out any affect from the coaching strategy, content or method of instruction. This means that any individual differences in performance were due to other factors like mood-coaching strategy compatibility (or incompatibility), low energy, lack of sleep, lack of focus or other reasons.

Individual differences were noted across the coaching strategy groups. The root cause of individual differences were not specifically identified, but a comparative analysis of the student state, mood, performance and action variables did reveal some potential influences. A t-Test was used to compare these variables and significant differences ($\alpha = .05$) were noted for the amount of sleep, final pleasure values and mouse movement rates. Sleep and final pleasure were significantly lower in the SRF group. Mouse movement rate was higher in the SRF group. Final pleasure was highly correlated (0.357) with performance in the SRF group.

Razzaq, et al (2007), and Murray and VanLehn (2006) reported differences in the performance scores in proactive coaching strategies (e.g. TIF) that were not realized in this study. The knowledge improvement and performance scores of initial low competency (bottom one-third), initial moderate competency (middle third) and initial high competency (top third) indicated no significant difference in performance.

For Hypothesis "B", "*student mood variables are predictors of student performance for a given ITS coaching strategy*", final dominance significantly predicted performance in the SRF group. Initial dominance came closest (p = .06) to predicting performance in the TIF group, but there was insufficient support for this hypothesis. Therefore, we conclude that mood variables alone aree not a significant predictors of performance within both strategies for the subject population and a comparison of predicted performance is not possible using regression analysis alone.

For Hypothesis "C", "*student state variables and student action variables are predictors of student performance for a given ITS coaching strategy*", TC3 Interest Level, a student state variable, significantly predicted performance in the SRF group. Help Request Frequency, student action variable, significantly predicted performance in the SRF group. Final Knowledge significantly predicted performance in the TIF group, but were not of practical significance due to small effect size. Therefore, we conclude that student state and student action variables alone are not significant predictors of performance within both strategies for the subject population and a comparison of predicted performance is not possible using regression analysis alone.

For Hypothesis "D", "*student action variables are predictors of student performance for a given ITS coaching strategy*", mouse movement rate significantly predicted performance in the SRF group, but were not of practical significance due small effect size. No student action variables were significant predictors of student performance in the TIF group. Student action variables by themselves are not significant predictors of performance within both strategies for the subject population and a comparison of predicted performance is not possible using regression analysis alone.

Given the results from Zimmerman, et al (2003), it was expected that there would be a relationship between mouse movement rates and arousal. However, the test results of Hypothesis "E", "*student state variables, student action variables and student performance are predictors of student mood independent of coaching strategy*" found no significant predictors of arousal (either initial arousal or final arousal). Performance and its effect on mood variables were evident in Hypothesis "E" and there were several significant predictors of mood including:

- initial dominance was significantly predicted by self assessed first aid knowledge, mouse movement rate and performance
- initial mood intensity was predicted by action rate, energy level, self assessed first aid knowledge and performance
- final pleasure was significantly predicted by action rate, self assessed first aid knowledge, performance
- final dominance and final mood intensity were significantly predicted by TC3 interest and performance

In regard to the students' perception of the tutor, the participant feedback survey questions that addressed the participant's perception of the tutor's trustworthiness, competence and supportiveness, indicated no difference in the perceptions of the SRF and TIF coaching strategy groups as a whole. However, there were differences in the perceptions of the tutor's trustworthiness and supportiveness based on a comparison of different levels of performance across the two coaching strategies.

The student's trust in the tutor was not significantly different between the coaching strategy groups except for high and very high performers in the SRF group who perceived the tutor to be significantly less trustworthy than the TIF group. This may have been influenced the amount of information forthcoming from the TIF tutor without student prompting.

There were no significant differences in the students' perception of the tutor's competence between the coaching strategy groups. The student's perception of the tutor's supportiveness was not significantly different between the coaching strategy groups. However, moderately performing students in the SRF group perceived the tutor to be less annoying that the TIF group. This could potentially be due to lack of control in initiating the amount and timing of feedback. Students from both groups and at all levels of performance agreed that the intelligent tutor was encouraging and motivating.

Mood variables in conjunction with coaching strategies, student state and student action variables were significant predictors of performance. This is evidenced by the AMOS SRF Composite models. AMOS provided a graphical method to identify significant predictors of performance and mood variables. The AMOS SRF Composite model explains 65% of the variance in performance within the SRF group while the AMOS TIF Composite model explains only about 7%. While the model is statistically significant, it is of no practical significance.

We also found evidence that a handful of variables (dominance, help request frequency, the amount of sleep the previous night, subject knowledge, mood intensity, interaction rates and student interest level) significantly predicted student performance. It was a surprise that other variables like pleasure and arousal were insignificant in predicting performance given the documented relationships between mood and performance in areas like sport psychology

(LeUnes & Burger, 1998; LeUnes, 2000), academic examinations (Thelwell, Lane and Weston, 2007) and management and organization (Jordan, Lawrence and Troth, 2006).

WEKA was used to create a tree classifier that can be applied within an ITS and used in future experiments. This classifier has potential to be used to predict performance in near-real-time throughout an experiment or training session versus a single discrete performance assessment at the end of the experiment.

The interaction and influence of student state variables including mood has been documented in this experiment, but the weight of these influences might also be impacted by variables not investigated in this study. A rigorous set of experiments should be undertaken to identify other influences, to more fully understand the relationships of other variables and to refine the models developed here.

Lessons-Learned

The development of a classifier model using WEKA provided a more flexible method to evaluate the relative importance of variables in predicting performance and by incorporating hybrid (numeric and nominal) data. This provided a more accurate picture of the influences of student variables including gender in the classifier model. The use of the graphical AMOS toolset provided a method to rapidly manipulate the interaction of model variables and assess significance, but had limitations in using nominal data which was critical to defining relationships of variables in this study.

The results of this study demonstrated the significant prediction of mood variables through student performance, student state information and student actions. Infrequent snapshots of student variables (e.g. initial and final knowledge, initial and final mood intensity) limit the ability of the models developed in this study to adapt to the student's needs in real-time or in near real-time. In order to support real-time assessment of the student's mood, state and performance, variables should be captured in real-time at key intervals. The models developed in this experiment could be adapted by assigning initial and final variables as proxies. In other words, final knowledge would become knowledge at time = t and initial mood intensity would become mood intensity at time = t_0 for the real-time model. Future experiments are required to determine the significance (statistical and practical) of these adapted models.

Certain student state variables (e.g. amount of sleep, TC3 interest level and energy level) were collected at the beginning of the experiment and only needed to be captured once. Other variables (e.g. mood variables) should have been captured throughout the experiment, but the methods used (e.g. SAM) made the frequent collection option unattractive and intrusive to the learning process. While student action variables were collected unobtrusively in real time, the temporal aspects of student mood variables in this experiment were limited to collecting mood twice and then using those mood states to test for relationships with action variables, other student state variables and performance.

	Pre-Performance Assessment	During Performance Assessment	Post-Performance Assessment
Dissertation Experiment (relational)	**Student State (excluding mood)** **Initial Student Mood (via SAM)**	**Student Actions** **Single Coaching Strategy**	**Final Student Mood (via SAM)** **Student Performance**
Future Experiments (relational & temporal)	**Student State (excluding mood)** **Student Mood Baseline at time = StartEx**	**Student Actions** **Student Mood at time = T** **Projected Student Performance at time = T** **Real-time Adjustment of Coaching Strategies**	**Student Mood at time = EndEx** **Student Performance at time = EndEx**

Figure 9: Temporal Approach to Data Collection

A further recommendation is to collect performance data throughout the training to assess the degree to which the coaching strategies are on track and to aid in predicting mood. In order to provide an accurate and timely assessment of the student's state, knowledge and performance, future experiments should consider assessing these variables frequently, but unobtrusively, during training as shown in Figure 9. Frequent collection of student state and mood data throughout the performance assessment will minimize the time between dominance ratings (an indicator of self-efficacy) and actual task performance theoretically resulting in higher effect size per Bandura (1997).

Recommendations for Future Research

The results of this study demonstrate the potential application of mood perception within ITS. Four specific recommendations are put forth for extending the research in this study: 1) conduct a within-subject study to evaluate temporal interaction between student and tutor by examining the potential of real-time perception of mood variables and their application to real-time coaching strategy decisions to aid in benchmarking both tutor perception limitations and student mood variability; 2) investigate the "appropriate timing of instructional support for weak students" per Kim, et al (2007); 3) increase the complexity of this study by increasing the complexity of the student-tutor interaction through other interaction strategies that include scaffolding; 4) expand this study and the within-subject study to broader populations to validate mood inference techniques across more diverse populations in terms of age, gender and ethnicity.

For possible consider of future researchers, we originally intended that mood attributes be independent variables which would be induced using the International Affective Picture System (IAPS). The IAPS provides a set of normative emotional stimuli for experimental investigations of affect and attention (Lang, Bradley & Cuthbert, 2005). During a pilot study with twenty (20) participants, sixty pictures were selected from the total IAPS set of 955 pictures to induce two distinct moods (high pleasure with high arousal and low dominance; and low pleasure with high arousal and low dominance) with ten (10) participants in each group. Each participant was exposed to each of the IAPS pictures for seven (7) seconds. Mood was assessed before and after the exposure to the IAPS subset using the Self-Assessment Manikin (SAM) (Lang, 1980). Only 2 of the 10 participants in the high pleasure with high arousal and low dominance group assessed themselves in the target mood. Only 1 of the 10 participants in the low pleasure with high arousal and low dominance group assessed themselves in the target mood. The pilot study results influenced us to eliminate mood induction and to make the assumption that mood variables were normally distributed. This assumption was validated in the main study.

APPENDIX A: UCF PARTICIPANT CONSENT FORM

This form is presented via laptop computer to each participant prior to the start of the experiment.

Informed Consent for an Adult in a Non-medical Research Study

Researchers at the University of Central Florida (UCF) study many topics. To do this we need the help of people who agree to take part in a research study. You are being invited to take part in a research study which will include about 80-100 people. You can ask questions about the research. You can read this form and agree to take part right now, or take the form home with you to study before you decide. You will be told if any new information is learned which may affect your willingness to continue taking part in this study. You have been asked to take part in this research study because you are a United States Military Academy (USMA) Cadet or a University of Central ROTC student. You must be 18 years of age or older to be included in the research study.

The person doing this research is Robert Sottilare, a doctoral candidate in the UCF Modeling and Simulation Program. Because the researcher is a doctoral dissertation student, he is being guided by Dr. Michael Proctor, a UCF faculty supervisor in Department of Industrial Engineering and Management Systems and Dr. Michael Matthews, a USMA professor in the Department of Leadership and Behavioral Sciences.

Study title: Dissertation research - Using student mood and task performance to train classifier algorithms to select effective coaching strategies within Intelligent Tutoring Systems (ITS)

Purpose of the research study: The purpose of this study is to evaluate the relationships between mood, coaching strategies and task performance to optimize the adaptability of intelligent tutoring systems.

What you will be asked to do in the study: You will be asked to take a biographical survey that requests demographic data (e.g. age, gender, education and college major), some other data (e.g. level of energy) and some background data to assess your level of medical combat casualty care expertise prior to the experiment.

You will also be asked to view a series of pictures from the International Affective Picture System (IAPS). Some of these images may be revolting, pleasant, puzzling or stimulating, but pose a very low risk to your mental health. After reviewing the IAPS pictures you will be asked to take another survey which will assess your mood. Then you will be exposed to a training module about combat casualty care in which you will receive some coaching on combat casualty care. At the end of the module your competency in combat casualty care will be assessed via a quiz. Finally, you will take another mood assessment survey and finally, a feedback survey on the study process.

This study will all be conducted by the principal investigator, Robert Sottilare. The data collection and training will take place on a laptop computer in a closed study room in Thayer Hall at USMA and will take less than an hour to complete. The data from these surveys will be used to generate a model of coaching strategies to be incorporated in intelligent tutoring systems. No data collected from the surveys will have your name associated with it. No one will be identified in the data or the published results. The raw data will not appear in the dissertation or any published results.

Voluntary participation: You should take part in this study only because you want to. There is no penalty for not taking part, and you will not lose any benefits. You have the right to stop at any time. Just tell the researcher or a member of the research team that you want to stop. You will be told if any new information is learned which may affect your willingness to continue taking part in this study.

Location: This study will be conducted at the United States Military Academy (USMA), West Point, NY, at Thayer Hall in the Department of Leadership and Behavior Sciences.

Time required: Less than one hour

Audio or video taping: This study does not include any audio or video taping.

Risks: The probability of any adverse reactions due to the tactical combat casualty care training or simulation content is low based on the previous exposure of twenty participants during the pilot study in August 2008.

Participants with combat experience or exposure to prior trauma are excluded from this study in order to minimize any adverse reaction to the Tactical Combat Casualty Care training and simulation used in this study.

In the event you feel that the study content is affecting you adversely during the study, immediately alert the principal investigator or another member of the research team and withdraw from the study immediately. There is no penalty for withdrawing from this study. If you have a problem after the study, please report this problem to the Research Coordinator at 845-938-5902.

Benefits: The direct benefit to you is the training you receive on how to provide combat casualty care. In addition you will learn about how research is conducted.

Compensation or payment: There is no compensation or other payment to you for taking part in this study. There is no compensation, payment or extra credit for taking part in this study.

Anonymous research: This study is anonymous. That means that no one, not even members of the research team, will know that the information you gave came from you.

Study contact for questions about the study or to report a problem: The principal investigator for this study is Robert Sottilare and he can be reached at (407) 384-3654 or by email at robert.sottilare@us.army.mil. Robert's dissertation committee chair and faculty supervisor is Dr. Michael Proctor, Associate Professor in the Department of Industrial Engineering and Management Systems and can be reached (407) 823-5296 or by email at mproctor@mail.ucf.edu. Dr. Michael Matthews, Professor of Engineering Psychology in the Department of Behavioral Sciences and Leadership, is the United States Military Academy sponsor for this study and can be reached at (845) 938-3696 or by email at mike.matthews@usma.edu.

IRB contact about your rights in the study or to report a complaint: Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (UCF IRB). For information about the rights of people who take part in research, please contact: Institutional Review Board, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901.

How to return this consent form to the researcher: A hardcopy of this form is available for you to take with you. Clicking on the button below constitutes your consent to participate in this study.

I consent to participate in this study.

APPENDIX B: USMA PARTICIPANT CONSENT FORM

Each participant and the principal investigator signed these consent forms and they are maintained on file at USMA in the Department of Behavioral Sciences and Leadership.

VOLUNTEER AGREEMENT AFFIDAVIT
For use of this form, see AR 70-25 or AR 40-38; the proponent agency is OTSG
PRIVACY ACT OF 1974

Authority: 10 USC 3013, 44 USC 3101 and 10 USC 1071-1087

Principle Purpose: To document voluntary participation in the Clinical Investigation and Research Program. SSN and home address will be used for identification and locating purposes.

Routine Uses: The SSN and home address will be used for identification and locating purposes. Information derived from the study will be used to document the study; implementation of medical programs, teaching, adjudication of claims, and for the mandatory reporting of medical condition as required by law. Information may be furnished to Federal, State and local agencies.

Disclosure: The furnishing of your SSN and home address is mandatory and necessary to provide identification and to contact you if future information indicates that your health may be adversely affected. Failure to provide the information may preclude your voluntary participation in this investigational study.

PART A - VOLUNTEER AFFIDAVIT
Volunteer Subjects in Approved Department of the Army Research Studies

Volunteers under the provisions of AR 40-38 and AR 70-25 are authorized all necessary medical care for injury or disease which is the proximate result of their participation in such studies.

I, ______________________________________ SSN _________________ having full capacity to consent and having attained my ________ birthday, do hereby volunteer to participate in the research protocol **Using student mood, knowledge and task performance to train classifier algorithms to select effective coaching strategies within Intelligent Tutoring Systems (ITS)** under the direction of **Dr. Michael Matthews and Mr. Robert Sottilare** conducted at **the United States Military Academy, West Point, NY**.
THE IMPLICATIONS OF MY VOLUNTARY PARTICIPATION; THE NATURE, DURATION AND PURPOSE OF THE RESEARCH STUDY; THE METHODS AND MEANS BY WHICH IT IS TO BE CONDUCTED; AND THE INCONVENIENCES AND HAZARDS THAT MAY REASONABLY BE EXPECTED HAVE BEEN EXPLAINED TO ME BY **MR. SOTTILARE**.

I have been given an opportunity to ask questions concerning this investigational study. Any such questions were answered to my full and complete satisfaction. **Should any further questions arise concerning my rights on study-related injury I may contact the Patient Representative at Keller Army Community Hospital, (845) 938-5874.**

I understand that I may at any time during the course of this study revoke my consent and withdraw from the study without further penalty or loss of benefits; however, I may be required (military volunteer) or requested (civilian volunteer) to undergo certain examinations if, in the opinion of the attending physician, such examinations are necessary for my health and well-being. My refusal to participate will involve no penalty or loss of benefits to which I am otherwise entitled.

PART B - EXPLANATION OF WHAT IS TO BE DONE

State the reason you are asking the subject to enrolled in this study; i.e., diagnosed with???
Do not modify the last two sentences in this paragraph.
INTRODUCTION: You have been invited to participate in a clinical research study conducted at USMA because you are a military cadet at least 18 years of age. Participation is entirely voluntary. You may refuse to participate or withdraw from the study at any time without penalty or loss of benefits to which you are otherwise entitled.

Explain in layman's language why you are conducting this study.
PURPOSE: This study will evaluate the relationships between coaching strategies provided by an intelligent tutoring system (ITS), a computer-based coach and students' mood, knowledge and task performance. The purpose of this study is to improve the adaptability of ITS by understanding these relationships and selecting appropriate coaching strategies to optimize knowledge and skill development.

Explain in layman's language everything you will be asking subjects do, i.e., describe any procedures, extra visits, tests, medications, etc.
PROCEDURES:
This research is being conducted to evaluate the use of techniques to assess student mood and apply different coaching strategies to determine which strategies result in the best student task performance given their mood. If you agree to participate, you will be asked to take a biographical survey that requests demographic data (e.g. age, gender, education and college major), some other data (e.g. level of energy) and some background data to assess your level of medical experience. You will next be asked to answer some questions to assess your knowledge of tactical combat casualty care prior to the experiment.
Following the knowledge assessment you will be asked to take a survey which will assess your mood. Then you will be exposed to a training module about casualty movement and hemorrhage control during care under fire. During this training you will receive some coaching on combat casualty care.

At the end of the training module your knowledge about casualty movement and hemorrhage control during care under fire will be assessed. After the knowledge assessment, you will participate in a simulation in which you will apply the skills you learned during the training. Finally, you will take another mood assessment survey and then a feedback survey on the study process.

Explain possible benefit(s) to your subject from study participation, and/or state “You may receive no direct personal benefit from participating in your protocol.”
POTENTIAL BENEFITS: The benefits to you may include the training you receive on combat casualty care, the opportunity to practice what you have learned and the knowledge you acquire about research methods used during this study.

Explain ALL risks of procedures, medications, etc.; inconveniences, such as extra clinic visits; discomforts, such as embarrassment, anxiety, minimally invasive procedures, etc.
RISKS, INCONVENIENCES, AND DISCOMFORTS:

Delete this paragraph if subjects will not have blood drawn for your protocol.
The probability of any adverse reactions due to the tactical combat casualty care training or simulation content is low based on the previous exposure of twenty participants during the pilot study in August 2008.
Participants with combat experience or exposure to prior trauma are excluded from this study in order to minimize any adverse reaction to the Tactical Combat Casualty Care training and simulation used in this study.
In the event you feel that the study content is affecting you adversely during the study, immediately alert the principal investigator or another member of the research team and withdraw from the study immediately. There is no penalty for withdrawing from this study. If you have a problem after the study, please report this problem to the Research Coordinator at 845-938-5902.

Modify this paragraph to explain a subject’s alternative treatment options. If there are no alternatives you must state this.
ALTERNATIVES TO PARTICIPATION: Cadets not wishing to participate in this study may read and critique a scientific article for equivalent credit.

Inclusion of this statement depends on whether the MAMC IRB has designated the risk status of your protocol as Minimal Risk or Other Than Minimal Risk. Modify this paragraph if you will be offering compensation for study subject participation.
COMPENSATION: You will not be paid for your participation in this study. PL100 cadet volunteers will receive extra credit IAW PL100 the 2008-2009 PL100 course guide. They will receive up to 10 bonus points toward their final PL100 grade (total testing time is less than one hour). Cadets will complete the research participation worksheet (attached) and return it to their instructor to receive credit for their participation.

Identify associate institutions or protocol sponsors having access to study data at the end of the first sentence in this paragraph. Do not modify the other sentences within this paragraph.
CONFIDENTIALITY OF RECORDS: The data in this study will be anonymous. This means that no one, not even members of the research team, will know that the information you gave came from you. Data will be collected via surveys and interactions on a laptop computer, and

names and other identifiers will not be placed on surveys or other research data. Your individual data will be maintained on a password-protected laptop computer.

The case records from this study will be available for review by members of the Institutional Review Board (IRB) in the Department of Behavior Science and Leadership in Thayer Hall at USMA . All records will be kept in a confidential form. Only the research team conducting this study will have access to the records from this study. Information gained from this study may be used as part of a scientific publication, but you will in no way be personally identified.

Disposition of any blood, urine or tissue samples must be explained as either destroyed (when) or stored (where, how long, what for). Delete this paragraph if it is not applicable to your protocol.

Do not modify the following paragraph.
NEW FINDINGS: Significant findings that occur during this study that might affect your decision to participate in the study will be discussed with you. Any significant findings developed from this study will be available to you and may be obtained from the principal investigator.

Do not modify the following paragraph.
REMOVAL STATEMENT: Your participation in this study may be terminated without your consent if conditions occur which might make your continued participation dangerous or detrimental to your health. Any participant with combat experience or exposure to prior trauma will be excluded in order to minimize any adverse reaction to the Tactical Combat Casualty Care training and simulation used in this study.

Do not modify the following paragraph.
OTHER INFORMATION: You are encouraged to ask any questions, at any time, that will help you to understand how this study will be performed and/or how it will affect you. Contact information for the research team follows:

Dr. Michael Matthews, Professor of Engineering Psychology in the Department of Behavioral Sciences and Leadership, is the United States Military Academy principal investigator for this study and can be reached at (845) 938-3696 or by email at mike.matthews@usma.edu.

The co-investigator for this study is Robert Sottilare and he can be reached at (407) 384-3654 or by email at robert.sottilare@us.army.mil.

You may contact the Research Coordinator at 845-938-5902 if you have questions or comments regarding your rights as a participant in the research. This research has been reviewed according to USMA's procedures governing your participation in this research.

IF THERE IS ANY PORTION OF THIS EXPLANATION THAT YOU DO NOT UNDERSTAND, ASK THE INVESTIGATOR BEFORE AGREEING TO PARTICIPATE IN THIS STUDY.

You will be given a signed and dated copy of this consent document for your records.

I do ☐ do not ☐ (check one & initial) consent to the inclusion of this form in my outpatient medical treatment record.

SIGNATURE OF VOLUNTEER	DATE	PRINTED NAME OF VOLUNTEER
PERMANENT ADDRESS OF VOLUNTEER		

Name of person administering consent: _Robert Sottilare

Signature of person administering consent: ________________________ Date: __________

APPENDIX C: UCF IRB APPROVAL LETTER

University of Central Florida

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Notice of Expedited Initial Review and Approval

From : **UCF Institutional Review Board**
FWA00000351, Exp. 6/24/11, IRB00001138

To : **Robert A Sottilare**

Date : **August 07, 2008**

IRB Number: **SBE-08-05755**

Study Title: **Using student mood and task performance to train classifier algorithms to select effective coaching strategies within Intelligent Tutoring Systems (ITS)**

Dear Researcher:

Your research protocol noted above was approved by **expedited** review by the UCF IRB Chair on 8/6/2008. **The expiration date is 8/5/2009.** Your study was determined to be minimal risk for human subjects and expeditable per federal regulations, 45 CFR 46.110. The category for which this study qualifies as expeditable research is as follows:

7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Please obtain the United States Military Academy IRB approval / documents before beginning this research. Provide copies to the UCF IRB by submitting a Miscellaneous Attachment Form within this study account.

A **waiver of documentation of consent** has been approved for all subjects. Participants do not have to sign a consent form, but the IRB requires that you give participants a copy of the IRB-approved consent form, letter, information sheet, or statement of voluntary consent at the top of the survey.

All data, which may include signed consent form documents, must be retained in a locked file cabinet for a minimum of three years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained on a password-protected computer if electronic information is used. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

To continue this research beyond the expiration date, a Continuing Review Form must be submitted 2 – 4 weeks prior to the expiration date. Advise the IRB if you receive a subpoena for the release of this information, or if a breach of confidentiality occurs. Also report any unanticipated problems or serious adverse events (within 5 working days). Do not make changes to the protocol methodology or consent form before obtaining IRB approval. Changes can be submitted for IRB review using the Addendum/Modification Request Form. An Addendum/Modification Request Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at http://iris.research.ucf.edu .

Failure to provide a continuing review report could lead to study suspension, a loss of funding and/or publication possibilities, or reporting of noncompliance to sponsors or funding agencies. The IRB maintains the authority under 45 CFR 46.110(e) to observe or have a third party observe the consent process and the research.

On behalf of Tracy Dietz, Ph.D., UCF IRB Chair, this letter is signed by:

Signature applied by Janice Turchin on 08/07/2008 09:51:22 AM EDT

Janice m Turchin

IRB Coordinator

APPENDIX D: USMA IRB APPROVAL LETTER

DEPARTMENT OF THE ARMY
U.S. ARMY MEDICAL DEPARTMENT ACTIVITY
WEST POINT, NEW YORK 10996-1197

REPLY TO
ATTENTION OF:

MCUD-CI

27 October 2008

MEMORANDUM FOR Michael Matthews, PhD, Department of Behavioral Sciences and Leadership, USMA, WP, NY 10996

SUBJECT: Notice of Institutional Review Board (IRB) Expedited Review Approval of Protocol 09/003 Using student mood, knowledge and task performance to train classifier algorithms to select effective coaching strategies within Intelligent Tutoring Systems (ITS), PI: Michael Matthews, PhD

1. The IRB has approved your protocol, "Using student mood, knowledge and task performance to train classifier algorithms to select effective coaching strategies within Intelligent Tutoring Systems" effective 24 October 2008 by expedited review. Attached is the IRB-approved stamped informed consent.

2. The work unit number for your protocol is 09/003 and must be on all correspondence pertaining to the study (upper right hand corner). Your study has been judged to have the following risk category: Minimal Risk.

3. Original copies of signed consent forms are the responsibility of the principal investigator and are to be maintained by the principal investigator and must be signed by the research subject and the investigator or his designee.

4. You should inform the Clinical Research Protocol Coordinator to bring to the committee any changes in the protocol or volunteer agreement, any changes in investigators or any adverse events or reactions which may relate in any way to the study. Further, a research progress report must be submitted to the Clinical Research Protocol Coordinator upon completion of the study or by 23 October 2009, whichever comes first. Your records are subject to review at any time.

5. Should you have any questions, contact Jane Reddington (938-4821). Sign below and return one copy. Keep the second for your records.

Encl

MICHAEL E. DOYLE
LTC, MC
Chair, KACH IRB

By my signature below I acknowledge receipt and accept full responsibility for this research proposal.

____ Dated______________________

VOLUNTEER AGREEMENT AFFIDAVIT

For use of this form, see AR 70-25 or AR 40-38; the proponent agency is OTSG

PRIVACY ACT OF 1974

Authority:	10 USC 3013, 44 USC 3101 and 10 USC 1071-1087
Principle Purpose:	To document voluntary participation in the Clinical Investigation and Research Program. SSN and home address will be used for identification and locating purposes.
Routine Uses:	The SSN and home address will be used for identification and locating purposes. Information derived from the study will be used to document the study; implementation of medical programs, teaching, adjudication of claims, and for the mandatory reporting of medical condition as required by law. Information may be furnished to Federal, State and local agencies.
Disclosure:	The furnishing of your SSN and home address is mandatory and necessary to provide identification and to contact you if future information indicates that your health may be adversely affected. Failure to provide the information may preclude your voluntary participation in this investigational study.

PART A - VOLUNTEER AFFIDAVIT

Volunteer Subjects in Approved Department of the Army Research Studies

Volunteers under the provisions of AR 40-38 and AR 70-25 are authorized all necessary medical care for injury or disease which is the proximate result of their participation in such studies.

I, ________________________________ SSN _______________ having full capacity to consent and having attained my _______ birthday, do hereby volunteer to participate in the research protocol **Using student mood, knowledge and task performance to train classifier algorithms to select effective coaching strategies within Intelligent Tutoring Systems (ITS)** under the direction of **Dr. Michael Matthews and Mr. Robert Sottilare** conducted at **the United States Military Academy, West Point, NY**.

The implications of my voluntary participation; the nature, duration and purpose of the research study; the methods and means by which it is to be conducted; and the inconveniences and hazards that may reasonably be expected have been explained to me by **Mr. Sottilare**.

I have been given an opportunity to ask questions concerning this investigational study. Any such questions were answered to my full and complete satisfaction. **Should any further questions arise concerning my rights on study-related injury I may contact the Patient Representative at Keller Army Community Hospital, (845) 938-5874.**

I understand that I may at any time during the course of this study revoke my consent and withdraw from the study without further penalty or loss of benefits; however, I may only receive the extra credit based on my time spent to that point. My refusal to participate will involve no penalty or loss of benefits to which I am otherwise entitled.

APPENDIX E: BIOGRAPHICAL SURVEY

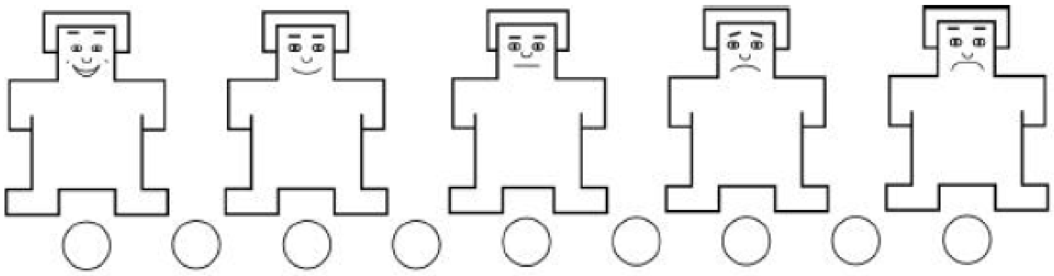
1. Your age in years is: 18___ 19___ 20___ 21___ 22___ 23___ 24___
2. Your gender is: M___ F___
3. Your class year is: freshman___ sophomore___ junior___ senior___
4. The Academic Department of your major is (e.g. Behavior Science and Leadership):

   ____
5. The number of hours of sleep you had last night was: 3___ 4___ 5___ 6___ 7___ 8___ 9___
6. Your present level of energy is (1-5 with 1 = low and 5 = high): ___
7. Your level of confidence in using a computer is (1-5 with 1 = low and 5 = high): __
8. Have you ever taken a course in First Aid? Yes___ No___
9. Have you ever taken a course in CPR? Yes___ No___
10. Your knowledge of first aid is (1-5 with 1 = low and 5 = high): __
11. Your present level of interest in CPR is (1-5 with 1 = low and 5 = high): ___

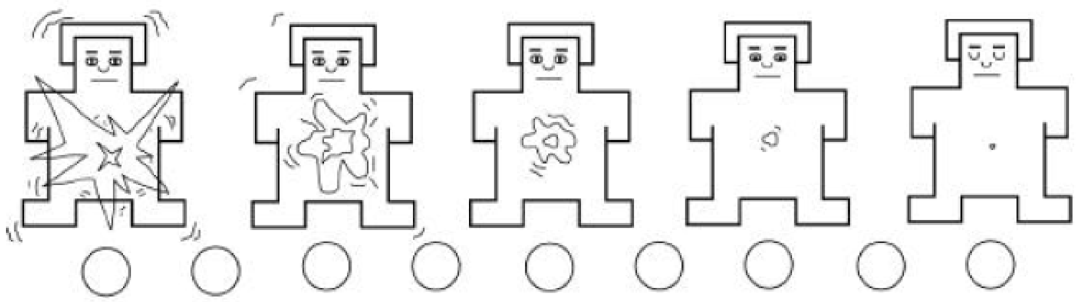
APPENDIX F: SELF ASSESSMENT MANIKIN (SAM) (Lang, 1985)

Please fill in the circle underneath the figure that most closely corresponds to your current mood:

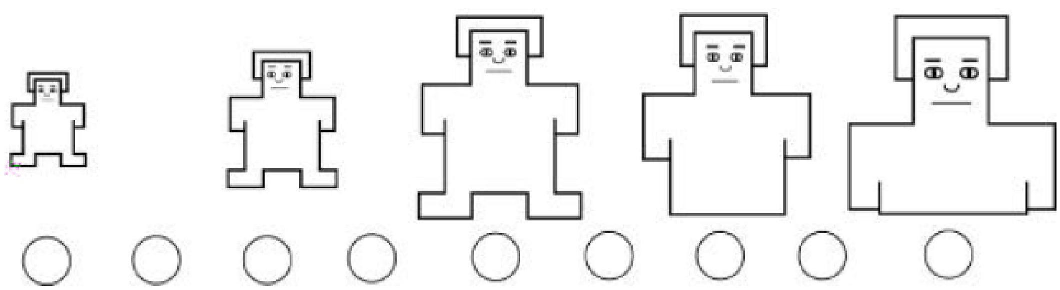
Please assess your mood in terms of **pleasure**



Please assess your mood in terms of **arousal**



Please assess your mood in terms of **dominance or feeling of control**



APPENDIX G: TACTICAL COMBAT CASUALTY CARE (TC3) KNOWLEDGE PRE-TEST AND POST-TEST

[The test questions below were provided electronically to each participant and the answers were automatically written to an anonymous participant file. The questions were used to evaluate the participant's Tactical Combat Casualty Care (TC3) competence level in regard to "Basic Care Under Fire knowledge' and "Hemorrhage Control" knowledge. The pre-test and post-test scores were compared to ascertain each participant's change in knowledge.]

Scoring notes: Responses were graded according to a rubric that provided 5 points for correct selections, 0 points for incorrect selections, and -5 points for incorrect selections in questions where the participant were asked to "select all that apply".

1. During casualty care under fire, you should:
    a. always drag casualties out of the line of fire
    b. never attempt to move a casualty because it is too dangerous
    c. use any means available to move the casualties and move as quickly as possible
    d. always administer care before attempting to move a casualty
2. Which of the following statements is true about "Care Under Fire"? Select all that apply.
    a. Medics should expect to return fire in a combat situation
    b. Casualties should return fire if able
    c. Airway management should be administered
    d. Medics should direct the casualty to move to cover and apply self aid if able
3. Typical signs and symptoms of hemorrhage may not be detectable. True or False.
4. Why is the use of a temporary tourniquet essential in a femoral artery disruption? (select all that apply)
    a. A temporary tourniquet can be rapidly applied
    b. A temporary tourniquet can quickly stop the bleeding
    c. A temporary tourniquet is easy to improvise
    d. A temporary tourniquet is most cost effective
5. Blood sweeps are used to detect hemorrhaging all over the body. True or False.

6. Blood sweeps are performed prior to measuring blood pressure or taking the casualty's pulse. True or False.

7. How long can you leave a tourniquet on without having to worry about loss of a limb?

   10 minutes___ 30 minutes___ 1 hour___ 2 hours___ 5 hours____

8. In care under fire, it is critical to definitively identify hemorrhaging as either arterial or venous. True or False.

9. At what blood pressure do casualties lose consciousness? Approximately…

   a. 25 mmHg
   b. 50 mmHg
   c. 75 mmHg
   d. 100 mmHg

10. Hemorrhage shock is secondary to hypovolemic shock. True or False.

11. Which of the following statements is true? Select all that apply.

    a. Do not attempt to salvage a casualty's rucksack, unless it contains items critical to the mission
    b. Always attempt to salvage a casualty's rucksack
    c. Don't waste time taking the casualty's weapon and ammunition
    d. Take the casualty's weapon and ammunition if possible

12. Pulse can be used to indicate the extent of blood loss. True or False.

13. The three phases of hemostasis are (pick one):

    a. vascular expansion, platelet plug formation and blood clotting (coagulation)
    b. vascular spasm, platelet plug formation and blood clotting (coagulation)
    c. vascular spasm, plasma flow and blood clotting (coagulation)

d. arterial aeration, platelet plug formation and blood clotting (coagulation)

14. How long does it take to bleed to death from a femoral artery disruption?

    a. 1-2 minutes

    b. 2-4 minutes

    c. 5-7 minutes

    d. 8-10 minutes

15. Hemorrhage control is the most important aspect of saving lives during Care Under Fire phase because: (select all that apply)

    a. A Soldier can go into hypovolemic shock and bleed to death in a very short period of time after injuring a large blood vessel

    b. Hemorrhage is the easiest thing to treat on the battlefield

    c. Hemorrhage is the leading cause of preventable death in combat

    d. Hemorrhage rarely leads to infection

APPENDIX H: PARTICIPANT FEEDBACK SURVEY

Directions: *The following statements concern your perception about your experience during the study that you just participated in. Please indicate the strength of your agreement with each statement below , utilizing a scale in which 1 denotes strong disagreement, 5 denotes strong agreement, and 2, 3, and 4 represent intermediate judgments. On the line after each statement, type a number from 1 to 5 from the following scale:*

1. Strongly disagree
2. Disagree
3. Neutral: neither agree nor disagree
4. Agree
5. Strongly agree

There are no "right" or "wrong" answers, so select the number that most closely reflects your experience. Take your time and consider each statement carefully. Once you have completed all questions click "Finished" at the bottom.

1. My overall experience during this study was pleasant: ______
2. My overall experience during this study was educational and I learned a lot: _____
3. The tutor provided useful feedback during the training: _____
4. The tutor provided timely feedback during the training: _____
5. The tutor was annoying: _____
6. The environment where the study was conducted was pleasant and comfortable: _____
7. The pictures I was shown affected my mood positively: _____
8. The procedure for this study was confusing: _____
9. I will be able to apply the skills I learned during the training part of this study: _____
10. I would participate again in a similar study: _____

LIST OF REFERENCES

Aleven, V., & Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help? In G. Gauthier, C. Frasson & K. VanLehn (Eds.) Intelligent Tutoring Systems: 5th International Conference, ITS 2000 (pp. 292-303). Berlin: Springer.

Alexander, S., Sarrafzadeh, A. & Fan, C. (2003). Pay Attention! The Computer is Watching: Affective Tutoring Systems. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003 (pp. 1463-1466). Chesapeake, VA: AACE.

Anderson, J. (1993). Rules of the Mind. Lawrence Erlbaum Associates, Mahwah, NJ, USA.

Anolli, L., Mantovani, F., Mortillaro, M., Vescovo, A., Agliati, A., Confalonieri, L., Realdon, O., Zurloni, V. and Sacchi, A. (2005). A Multimodal Database as a Background for Emotional Synthesis, Recognition and Training in E-Learning Systems. In Proceedings of the Affective Computing and Intelligent Interaction (ACII) Conference - 2005, Beijing, China.

Arbuckle, J. L. (2008). Analysis of Moment Structures (AMOS) 17.0.0 (Build 1404) Copyright 1983-2008. AMOS Development Corporation, Crawfordville, Florida.

Baylor, A. L., Shen, E. and Warren, D. (2004) "Supporting learners with math anxiety: The impact of pedagogical agent emotional and motivational support," presented at Workshop on "Social and Emotional Intelligence in Learning Environments," held at the International Conference on Intelligent Tutoring Systems., Maceió, Brazil, 2004.

Baylor, A. L. & Kim, Y. (2005). Simulating Instructional Roles through Pedagogical Agents. International Journal of Artificial Intelligence in Education.

Beck, J., Stern, M., and Haugsjaa, E. (1996) Applications of AI in Education, ACM Crossroads.

Becker P. (2001). Structural and Relational Analyses of Emotion and Personality Traits. In: Zeitschrift für Differentielle und Diagnostische Psychologie, 22, 3, 2001, 155-172.

Berkowitz, L. (1993). Aggression: Its causes, consequences, and control. McGraw-Hill, New York, USA.

Bickmore, T. and Picard, R.W. (2004). Towards Caring Machines, Proceedings of Conference on Human Factors in Computing Systems (CHI), 2004.

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

Bower, G.H. (1981). Mood and memory. American Psychologist. 1981 Feb Vol 36(2) 129-148.

Bower, G. H., & Forgas, J. P. (2000). Affect, memory, and social cognition. In E. Eich, J. F. Kihlstrom, G. H. Bower, J. P. Forgas, & P. M. Niedenthal (Eds.), Cognition and emotion (pp. 87–168). New York: Oxford University Press.

Broekens, J. and DeGroot, D. (2004). Scaleable and Flexible Appraisal Models for Virtual Agents. In Proceedings of CGAIDE, 2004.

Browne, G. and Jermey, J. (2001). Website Indexing 1st Edition. Auslib Press, Blackwood, South Australia, Australia.

Burleson, W. and R. Picard (2004). Affective agents: Sustaining motivation to learn through failure and a state of stuck. In Social and Emotional Intelligence in Learning Environments Workshop In conjunction with the 7th International Conference on Intelligent Tutoring Systems.

Byrnes, J.P. (1996). Cognitive Development and Learning in Instructional Contexts. Boston: Allyn and Bacon.

Carnevale, P.J.D., & Isen, A.M. (1986). The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation. Organizational Behavior and Human Decision Processes, 37, 1-13.

Clancey, W. and Letsinger, R. (1981). Tutoring Rules for Guiding a Case Method Dialog. In Proceedings of the Sixth IJCAI, Vancouver, B.C., Morgan-Kaufmann, San Mateo, California, pp. 829-835.

Cohen, J. (1992). Quantitative Methods In Psychology: A Power Primer. Psychological Bulletin, 1992, Vol 112, No 1, 155-159. American Psychological Association, Inc.

Conati, C. and Maclaren, H. (2004). Evaluating a probabilistic model of student affect. Proceedings of ITS'04, 55-66. Berlin: Springer-Verlag.

Corbett, A., Koedinger, K., and Anderson, J. (1997). Intelligent Tutoring Systems, Chapter 37 of Handbook of Human-Computer Interaction, Second, Completely Revised Edition by M. Helander, T. K. Landauer, P. Prabhu (Editors). Elsevier Science B. V

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning:
Beneficial effects of contextualization, personalization, and choice. Journal of Educational Psychology, 88(4), 715-730.

Core, M. G., Traum, D., Lane, H. C., Swartout, W., Marsella, S., Gratch, J., and van Lent, M. (2006) "Teaching negotiation skills through practice and reflection with virtual humans," Simulation, 82, 2006, pp. 685-701

Csikszentmihalyi, M. (1990). Flow: the psychology of optimal experience. Harper Perennial, New York, NY.

Davidson, R.J. (1994). On emotion, mood, and related affective constructs. In P. Ekman & R.J. Davidson (Eds.) The Nature of Emotion: Fundamental Questions. New York: Oxford University Press. 1994, 51-55.

De Vicente, A. (2003). Towards Tutoring Systems that detect students' motivation: an investigation. Ph.D. thesis, The University of Edinburgh, ICCS, School of Informatics.

D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R. W., Graesser, A. C.: Integrating affect sensors in an intelligent tutoring system. In: Affective Interactions: The Computer in the Affective Loop, Workshop at 2005 International conference on Intelligent User Interfaces. New York: AMC Press (2005) 7-13

D'Mello, S.K., Craig, S.D., Sullins, J. and Graesser, A.C. (2006) Predicting Affective States expressed through an Emote-Aloud Procedure from AutoTutor's Mixed-Initiative Dialogue. International Journal of Artificial Intelligence in Education.

D'Mello, S. and Graesser, A. (2007). Mind and Body: Dialogue and Posture for Affect Detection in Learning Environments. In Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED), Marina del Rey, CA.

Dempsey, J. V., & van Eck, R. (2003). Modality and placement of a pedagogical adviser in individual interactive learning. British Journal of Educational Technology, 34(5), 585-600.

Dillon, T.J. (1988). *Questioning and Teaching: A Manual of Practice*. New York: Teachers College Press.

Duric, Z., Gray, W., Heishman, R., Li, F., Rosenfeld, A., Schoelles, M. Schunn, C. and Wechsler, H. (2002). Integrating Perceptual and Cognitive Modeling for Adaptive and Intelligent Human–Computer Interaction. In Proceedings of the IEEE, VOL. 90, NO. 7, July 2002, pp. 1272-1289

Erez, A., and Isen, A.M. (2002). The Influence of Positive Affect on Components of Expectancy Motivation. Journal of Applied Psychology 87, 6, 1055-1067.

Estrada, C.A., Isen, A.M., & Young, M.J. (1994). Positive affect influences creative problem solving and reported source of practice satisfaction in physicians. Motivation and Emotion, 18, 285-299.

Estrada, C. A., Isen, A. M., & Young, M. J. (1997). Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians. Organizational and Human Decision Processes, 72, 117-135.

Forgas, J.P., and Bower, G.H. (1987). Mood effects on Person-Perception Judgments. Journal of Personality and Social Psychology 53, 1, 53-60.

Funk, P. and Conlan, O. (2002). Case-Based Reasoning and Knowledge Management to Improve Adaptability of Intelligent Tutoring Systems. In Proceedings: Workshop on Case-Based Reasoning for Education and Training at 6th European Conference on Case Based Reasoning, ECCBR2002, ed. Pedro A. González-Calero, pp 15-23, 4th September 2002, Aberdeen, Scotland.

Gebhard, P., (2005). ALMA - A Layered Model of Affect. In Proceedings of the Fourth International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS), Utrecht, Netherlands, July 25 to 29, 2005.

George, J.M., & Brief, A.P. (1996). Motivational agendas in the workplace: The effects of feelings on focus of attention and work motivation. In L.L. Cummings and B.M. Staw (Eds.), Research in organizational behavior: Vol. 18 (pp. 75-109). Greenwich, CT: J Press.

Gijbels, D., Van de Watering, G., Dochy, F. and Van den Bossche, P. (2005). The relationship between students' approaches to learning and the assessment of learning outcomes. European Journal of Psychology of Education 2005. Vol. XX. no. 4. 327-341

Gold, P. and van Buskirk R. (1975). Facilitation of time-dependent memory processes with posttrial epinephrine injections. Behavioral Biology, 13, 145-153.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal, 31,* 104-137.

Graesser, A.C., Person, N., Harter, D. and the Tutoring Research Group (2001). Teaching tactics and dialog in AutoTutor, 12, 257-279.

Gratch, J. (2000). Émile: Marshalling Passions in Training and Education. In Proceedings of the 4th International Conference on Autonomous Agents, June 2000

Gratch, J. & Marsella, S. (2001). Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In J.P. Miller, E. Andr, & S. Sen (Eds.), Proceedings of the Fifth International Conference on Autonomous Agents, 278-285. New York: ACM Press.

Gratch, J. and Marsella, S. (2004). Technical Details of a Domain-independent Framework for Modeling Emotion. Technical Report No: ICT TR 04.2004, University of Southern California, Los Angeles, CA, 2004

Gratch, J., Marsella, S., Mao, W. (2006). Toward a Validated Model of Emotional Intelligence. Twenty-First National Conference on Artificial Intelligence (AAAI06) (Boston, MA, July 16 - 20, 2006).

Greene, T.R., & Noice, H. (1988). Influence of positive affect upon creative thinking and problem solving in children. Psychological Reports, 63, 895-898.

Hernandez, Y., Noguez, J., Sucar, E. and Arroyo-Figueroa G. (2006). Incorporating an Affective Model to an Intelligent Tutor for Mobile Robotics. In Proceedings of the 36th Annual Frontiers in Education Conference, San Diego, California, USA.

Hersey, P. and Blanchard, K.H. (1969). Life Cycle Theory Of Leadership. Training Development, 23, 5, 26-34, 69 May

Heylen, D., Nijholt, A., op den Akker R. and Vissers, M. (2003), Socially Intelligent Tutor Agents. University of Twente, Computer Science, Enschede, Netherlands.

Hirt, E.R., Melton, R.J., McDonald, H.E. and Harackiewicz, J.M. (1996) Processing goals, task interest, and the mood-performance relationship: A mediational analysis. Journal of Personality and Social Psychology, 71, 245-261.

Hofstede, G. (1980). Motivation, leadership, and organizations: Do American theories apply abroad? Organizational Dynamics, 9, 42-63.

Holzinger, A. (2000). Basiswissen Multimedia, Band 2 Lernen: Kognitive Grundlagen multimedialer Informationssysteme. Würzburg: Vogel.

Horvitz, E. (2007) Machine Learning, Reasoning, and Intelligence in Daily Life: Directions and Challenges (Invited talk), Proceedings of Artificial Intelligence Techniques for Ambient Intelligence, Hyderabad, India.

House, R. J., Hanges, P., Ruiz-Quintanilla, S. A., Dorfman, P.W., Javidan, M., Dickson, M., Gupta, V., & 170 co-authors. (1999). Cultural influences on leadership and organizations: Project GLOBE. In W.F. Mobley, M. J. Gessner & V. Arnold (Eds), Advances in global leadership, Vol. 1 (pp.171-233). Greenwich, CT: JAI Press.

Isen, A.M., Johnson, M.M.S., Mertz, E., & Robinson, F.G. (1985). The influence of positive affect on the unusualness of word association. Journal of Personality and Social Psychology, 48, 1413-1426.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. Journal of Personality and Social Psychology, 52, 1122-1131.

Isen, A.M, Rosenzweig, A.S., & Young, M.J., (1991). The influence of positive affect on clinical problem solving. Medical Decision Making, 11, 221-227.

Isen, A.M. (1999). On the relationship between affect and creative problem solving. In S. Russ (Ed.), Affect, creative experience, and psychological adjustment (pp. 3-17). Philadelphia: Taylor & Francis.

Isen, A.M. (2000). Positive Affect and Decision Making, in M.Lewis & J.M. Haviland-Jones (eds.), Handbook of Emotions, second edition, The Guildford Press, 417- 435, 2000.

Isen, A. M. (2003). Positive Affect as a Source of Human Strength. In L. Aspinwall & U. Staudinger (Eds.), A Psychology of Human Strengths (pp. 179-195). The American Psychological Association, 2003, Washington, D.C.

Johnson, A. and Taatgen, N. (2005). Handbook of Human Factors in Web Design, Chapter 25 - User Modeling. Editors: Proctor, R., & Vu, K. New Jersey: Lawrence Erlbaum Associates, Inc.

Johnson, W.L., Rickel, J.W., Lester, J.C., "Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environment", International Journal of Artificial Intelligence in Education, 11, 2000, pp. 47-78.

Jordan, P.J., Lawrence, S.A. and Troth, A.C. (2006). The impact of negative mood on team performance. Journal of Management & Organization, Volume: 12, Issue: 2, Managing Emotions and Conflict in the Workplace, September 2006, 131-145.

Kim, Y. (2003). Pedagogical Agent as Learning Companion: Its Constituents and Educational Implications. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003 (pp. 2229-2236). Chesapeake, VA: AACE.

Kim, Y., Wei, Q., Xu, B., Ko, Y. and Ilieva, V. (2007). MathGirls: Toward Developing Girls' Positive Attitude and Self-Efficacy through Pedagogical Agents, International Journal of Artificial Intelligence in Education. R. Luckin et al. (Eds.) IOS Press, 2007, 51-58.

Kahn, B., & Isen, A.M. (1993). The influence of positive affect on variety seeking among safe, enjoyable products. Journal of Consumer Research, 20, 257-270.

Kim, Y. (2005). Empathetic Virtual Peers Enhanced Learner Interest and Self-Efficacy. Workshop on Motivation and Affect in Educational Software at the 12th InternationalConference on Artificial Intelligentce in Education, Amsterdam, The Netherlands.

Kim, Y., & Baylor, A. L. (2006). A social-cognitive framework for designing pedagogical agents as learning companions. Educational Technology Research & Development (ETR&D) Vol. 54, No. 6, pp. 569–590, 2006.

Kim, Y., Wei, Q., Xu, B., Ko, Y. and Ilieva, V. (2007). MathGirls: Toward Developing Girls' Positive Attitude and Self-Efficacy through Pedagogical Agents, International Journal of Artificial Intelligence in Education. R. Luckin et al. (Eds.) IOS Press, 2007, 51-58.

Klesen, M. (2002). Report on affective reasoning and cultural diversity. NECA: A Net Environment for Embodied Emotional Conversational Agents.

Krause, R. (2000). Affekt, Emotion, Gefühl, In: Merten W., Wandvogel B. Handbuch psychoanalytischer Grundbegriffe, Kohlhammer, 2000, 73-80

Kshirsagar, S., Magnenat-Thalmann, N. (2002). A multilayer personality model. In: Proceedingsof 2nd International Symposium on Smart Graphics, ACM Press (2002), 107–115

Lane, H.C. (2006). Intelligent Tutoring Systems: Prospects for Guided Practice and Efficient Learning. Whitepaper for the Army's Science of Learning Workshop, Hampton, VA. Aug 1-3, 2006.

Lane, H.C. (2007). Metacognition and the Development of Intercultural Competence. In Proceedings of the Workshop on Metacognition and Self-Regulated Learning in Intelligent Tutoring Systems at the 13th International Conference on Artificial Intelligence in Education (AIED), p. 23-32. Marina del Rey, CA.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), Technology in mental health care delivery systems (pp. 119–137). Norwood, NJ: Ablex.

Lang, P.J. (1985). The Cognitive Psychophysiology of Emotion: Anxiety and the Anxiety Disorders. Hillsdale, NJ; Lawrence Erlbaum.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6. University of Florida, Gainesville, FL.

learning. (2007). Retrieved September 6, 2007 from Encarta Online Dictionary: http://encarta.msn.com/dictionary_/learning.html

Lee, A., & Sternthal, B. (1999). The effects of positive mood on memory. Journal of Consumer Research, 26, 115.

Lehman, J., Laird, J. and Rosenbloom, P. (2006) A Gentle Introduction to SOAR, an Architecture for Human Cognition: 2006 Update. University of Michigan, Department of Electrical Engineering and Computer Science.

LeUnes, A. (2000). Updated bibliography on the Profile of Mood States in sport and exercise psychology research. Journal of Applied Sport Psychology, 12, 110-113.

LeUnes, A., & Burger, J. (1998). Bibliography on the Profile of Mood States in sport and exercise, 1971-1995. Journal of Sport Behavior, 21, 53-70.

Linnenbrink, E. A. and Pintrich, P. R. (2002). Motivation as an enabler for academic success. School Psychology Review, 31, 313-327.

Loftin, B., Mastaglio, T., and Kenney, P. (2004), Outstanding Research Issues in Intelligent Tutoring Systems, study commissioned by the Research Development and Engineering Command (RDECOM), Orlando, Florida, USA under contract N61339-03-C-0156, http://www.mymicsurveys.com/site/files/pub_4.pdf accessed October 23, 2006

McCrae R.R., and John O.P. (1992). An introduction to the five-factor model and its applications. Special Issue: The five-factor model: Issues and applications. Journal of Personality, 60: 175-215.

Mehrabian A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Current Psychology, 14 1996, 261-292

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2000). Evidence-centered assessment design: A Submission for the NCME Award for Technical or Scientific Contributions to the Field of Educational Measurement. Retrieved July 5, 2008 from http://www.ncme.org/about/awards/mislevy.html

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessment. Measurement: Interdisciplinary Research and Perspective, 1(1) 3-62.

mood. (2008). Retrieved January 31, 2008 from Encarta Online Dictionary: http://encarta.msn.com/dictionary_/mood.html

Morris W. N. (1989). Mood: The frame of mind. New York: Springer-Verlag.

Murray, C. and VanLehn, K (2006) A Comparison of Decision-Theoretic, Fixed-Policy and Random Tutorial Action Selection. In Ikeda, Ashely & Chan (Eds.). Proceedings of the 8th International Conference on Intelligent Tutoring Systems. Springer-Verlag: Berlin pp. 116-123. 2006.

National Research Council. (2000). How people learn: Brain, mind, experience, and school, Expanded edition. Committee on Developments in the Science of Learning and Committee on Learning Research and Educational Practice. J.D. Bransford, A. Brown, and R.R. Cocking (Eds.). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (2005). How Students Learn: History, Mathematics, and Science in the Classroom. Committee on How People Learn, A Targeted Report for Teachers, M.S. Donovan and J.D. Bransford, Editors. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Neji, M. and Ben Ammar, M. (2007). Agent-based Collaborative Affective e-Learning Framework. Electronic Journal of e-Learning Volume 5 Issue 2, pp. 123 – 134.

Norman, D. (2002). Emotion and Design: Attractive Things Work Better. Interactions, 9, 4, 36-42.

Okonkwo, C. and Vassileva, J. (2001) “Affective Pedagogical Agents and User Persuasion”. In C. Stephanidis (ed.) Proc. "Universal Access in Human - Computer Interaction (UAHCI)", 9th International Conference on Human- Computer Interaction, New Orleans, USA, Laurence Erlbaum, 397-401.

Ortony, A., Clore, G., Collins, A. (1986) "The Cognitive Structure of Emotions." Cambridge, MA: Cambridge University Press.

Pearce, D. and Luckin, R. (2007). The Principle of State Expansion in Task State-Space Navigation. In Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED), Marina del Rey, CA.

Person, N. K., & Graesser, A. C., & The Tutoring Research Group (2003). Fourteen facts about human tutoring: Food for thought for ITS developers. AI-ED 2003 Workshop Proceedings on Tutorial Dialogue Systems: With a View Toward the Classroom (pp. 335-344). Sydney, Australia.

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 Technology Journal, Vol 22 No 4, October 2004.

Picard, R. (2006). Building an Affective Learning Companion. Keynote address at the 8th International Conference on Intelligent Tutoring Systems, Jhongli, Taiwan. Retrieved from http://www.its2006.org/ITS_keynote/ITS2006_01.pdf

Pintrich, P.R., & De Groot, E.V. (1990). Motivation and self regulated learning components of classroom academic performance. Journal of Educational Psychology, 82(I), 33-40.

Prince, M.J. and Felder, R.M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. Journal of Engineering Education 95 (2): 123–38.

Quinlan, Ross (1993). "C4.5: Programs for Machine Learning", Morgan Kaufmann Publishers, San Mateo, CA.

Raghavan, K. and Katz, A. (1989). Smithtown: An Intelligent Tutoring System. Technological Horizons in Education, v17, p50-53 August 1989.

Randolph, W.A. and Posner, B.Z. (1979). Designing Meaningful Learning Situations in Management: A Contingency, Decision-Tree Approach. The Academy of Management Review, Vol. 4, No. 3 (Jul., 1979), pp. 459-467

Razzaq, L., Heffernan, N. and Lindeman, R. (2007). What Level of Tutor Interaction is Best? International Journal of Artificial Intelligence in Education. R. Luckin et al. (Eds.) IOS Press, 2007, 222-229.

Robbins, Trevor W. (1997). Arousal systems and attentional processes. Biological Psychology, 45, 57–71.

Rodrigues, M., Novais, P., and Santos, M.F. (2005). Future Challenges in Intelligent Tutoring Systems – A Framework. In Proceedings of the Third International Conference on Multimedia and Information & Communication Technologies in Education, Cáceres, Spain.

Roll, I., Baker, R., Aleven, V., McLaren, B. and Koedinger, K. (2005). Modeling Students' Metacognitive Errors in Two Intelligent Tutoring Systems. In Proceedings of the Tenth International Conference on User Modeling, Edinburgh, Scotland.

Roll, I., Aleven, V., McLaren, B. and Koedinger, K. (2007). Can Help Seeking Be Tutored? Searching for the Secret Sauce of Metacognitive Tutoring. In Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED), p. 203-210. Marina del Rey, CA.

Rosenberg-Kima, R.B., Plant, E.A., Baylor, A.L. and Doerr, C.E. (2007). Changing Attitudes and Performance with Computer-Generated Social Models. In Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED), p. 51-58. Marina del Rey, CA.

Salovey, P. & Mayer, J. D. (1990). Emotional intelligence. Imagination, Cognition, & Personality, 9, 185-211.

Schmidt, B. (2000). The Modelling of Human Behaviour. SCS-Europe BVBA, Ghent.

Sessink, O., Beeftink, H., Tramper, J. & Hartog, R. (2007). Proteus: A Lecturer-Friendly Adaptive Tutoring System. Journal of Interactive Learning Research. 18 (4), pp. 533-554. Chesapeake, VA: AACE.

Sharp, H., Rogers, Y., and Preece, J. (2007). Interaction design: beyond human-computer interaction, Second Edition. John Wiley and Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England.

Shute, V.J. and R. Glaser, A large-scale evaluation of an intelligent discovery world. Interactive Learning Environments, 1990.

Snelgrove, S., & Slater, J. (2003). Approaches to learning: Psychometric testing of a study process questionnaire. Journal of Advanced Nursing, 43(5), 496-505.

Soller, A.L. (2001). Supporting Social Interaction in an Intelligent Collaborative Learning System. International Journal of Artificial Intelligence in Education (2001), Volume 12, p. 40-62.

Soper, D.S. (2009) "The Free Statistics Calculators Website", Online Software, http://www.danielsoper.com/statcalc/ accessed January 3, 2009.

Sottilare, R. (2006). Modeling the Influences of Personality Preferences on the Selection of Instructional Strategies in Intelligent Tutoring Systems. Masters thesis. University of Central Florida, Orlando.

Staw, B. M., & Barsade, S. G. (1993). Affect and managerial performance: a test of the sadder-but-wiser vs. happier-and smarter hypotheses. Administrative Science Quarterly, 38, 304-331.

Thelwell, R., Lane, A.M. and Weston, N. (2007). Mood states, self-set goals, self-efficacy and performance in academic examinations. Personality and Individual Differences. Volume 42, Issue 3, February 2007, Pages 573-583.

van Eck, R., & Dempsey, J. (2002). The effect of competition and contextualized advisement on the transfer of mathematics skills in a computer-based instructional simulation game. Educational Technology Research and Development, 50(3), 23-41.

Van Labeke, N., Brna, P., & Morales, R. (2007). Opening up the Interpretation Process in an Open Learner Model. International Journal of Artificial Intelligence in Education. 17, 305-338.

VanLehn, K. and Niu, Z. (2001). Bayesian student modeling, user interfaces and feedback: A sensitivity analysis. International Journal of Artificial Intelligence in Education (2001), Volume 12, p. 154-184.

Van Mulken S., Andre, E. and Muller, J. (1998). "The Persona Effect: How Substantial Is It?" In H. Johnson, L. Nigay, and C. Roast, (Eds), People and Computers XIII: Proceedings HCI '98, Springer-Verlag, London, England.

Vermunt, J.D., (1996) Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis. Higher Education 31: 25-50, 1996. Kluwer Academic Publishers, Netherlands.

Wagster, J., Tan, J., Biswas, G. and Schwartz, D. (2007). How Metacognitive Feedback Affects Behavior in Learning and Transfer. Workshop on Metacognition and Self Regulated Learning. In Proceedings of the Artificial Intelligence in Education (AIED) Conference 2007.

Watkins, D., & Biggs, J. (Eds.) (1996). The Chinese learner: Cultural, psychological and contextual influences. Hong Kong: University of Hong Kong, Comparative Education Research Centre.

Witten, I.H. and Frank, E. (2005). "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, USA.

Woolf, B. 1992. AI in Education. Encyclopedia of Artificial Intelligence, Shapiro, S., ed., John Wiley & Sons, Inc., New York, pp. 434-444.

Woolf, B., Burelson, W., and Arroyo, I. (2007). Emotional intelligence for computer tutors. Affect and Learning Workshop. In Proceedings of the Artificial Intelligence in Education (AIED) Conference 2007.

Young, J.W. (1993). Grade adjustment methods. Review of Educational Research. (55(2), 151-165.

Zapata-Rivera, D., Hansen, E., Shute, V.J., Underwood, J.S. and Bauer, M. (2007). Evidence-based Approach to Interacting with Open Student Models. International Journal of Artificial Intelligence in Education, Special Issue (Part 2) "Open Learner Models: Future Research Directions" (editors Vania Dimitrova, Gord McCalla and Susan Bull), Volume 17, p. 273-303.

Zeegers, P. (2001). Student learning in science: A longitudinal study. British Journal of Educational Psychology 71 115-132.

Zhou X., Conati C. (2003). Inferring User Goals from Personality and Behavior in a Causal Model of User Affect. In Proceedings of IUI 2003.

Zimmermann, P. , Guttormsen, S., Danuser, B. and Gomez. P. (2003). Affective Computing - A Rationale for Measuring Mood with Mouse and Keyboard. International Journal of Occupational Safety and Ergonomics.