Objective Measurement of Student Affect to Optimize Automated Instruction

Dr. Elizabeth Sheldon Biddle
Sonalysts, Inc.
12501 Research Parkway, Suite 260
Orlando, Florida 32826
esheldon@sonalysts.com

Dr. Linda Malone
Industrial Engineering and Management Systems
University of Central Florida
4000 Central Florida Boulevard
Orlando, Florida 32816-2450
lmalone@mail.ucf.edu

Dr. Dennis McBride
Institute for Simulation & Training
University of Central Florida
3280 Progress Drive
Orlando, Florida
dmcbride@ist.ucf.edu

Abstract

Student responses to various instructional methods differ due to individual differences. Additionally, these responses are observed in the student's affect. In human instructor learning environments, the human instructor is able to adapt his teaching method based on observable signals of the student's affect (e.g., wandering gaze, slumped shoulders, facial expressions, etc.). However, in an intelligent tutoring environment, the system is not able to infer the student's affect and consequently, the instruction is tailored solely on the student's performance. There are methods for automatically obtaining objective measures of affect. As such, an affective component has been designed to enhance the student model in an ITS in providing a more comprehensive diagnosis of the student's performance (e.g., discriminating between lack of knowledge and boredom or frustration). This paper describes an experiment that was conducted in support of the development of the affective component.

1. Introduction

Affect (Spielberger, 1966), personality (Matthews, 1999), and motivation (Ackerman, Kanfer, & Goff, 1995) influence a student's experience in a learning environment. The effectiveness of different methods of instruction differs between individual students (Sternberg, 1997) as does the student's preference for a given instructional method.

Instructional technologies, such as intelligent tutoring systems (ITSs), have been developed to respond to the individual student's needs. ITSs primarily tailor instruction based on the student's rote performance of the task, but rarely consider the student's attitudes and emotions. However, there is evidence (Chi, Siler, Jeong, Yamauchi, & Hausman, 2001) that one-on-one human tutoring yields improved instructional outcomes over traditional and technological instructional methods because human instructors are able to vary the manner in which they provide instruction based upon their observations of the student's affect (*e.g.*, a lack of attentiveness or an expression of confusion).

Therefore, the recommendation (Rickel & Johnson, 1999) that research regarding affective modeling of the student is needed for the realization of truly interactive learning environments is of little surprise. The concept of an affective component to evaluate student affect (Sheldon, 2001) was designed as a tool that enables an ITS to tailor its instruction to optimize the student's experience. An experiment was conducted to evaluate the feasibility of developing an affective component for an ITS. This experiment is summarized in the remainder of this article.

2. Background

As described previously, a student's unique set of skills, personality traits, and prior experience impacts his interaction in a learning environment (Matthews, 1999). A student's response (*e.g.,* like/dislike, enjoyment, disdain, *etc.*) to an instructional event reflects his unique characteristics and are exhibited via his affect. A human instructor senses an individual's affect by observing the student's body language, facial expressions, gaze, and vocal intonation.

A typical ITS does not possess this capability, although there are a few ITS architectures that have been developed to include student affect or attitude in the student model. For instance, researchers (*e.g.*, Conati, 2002; Conati & Zhou, 2002; Kapoor, Mota, & Picard, 2001; Murray & VanLehn, 2000) have designed ITS applications that sense student affect to determine an optimal instructional intervention strategy. Additionally, there has been investigation regarding the inclusion of motivation in the student model (Del Soldato & duBoulay, 1995; Mitrovic, Djordjevic-Kajan, & Stoimenov, 1996; Tong & Agelides, 1999) by querying the student periodically.

There is sufficient evidence (*e.g.*, Cacioppo & Petty, 1986; Fowles, 1986; Grossman, 1992; Levine, 1986; Scherer, 1993) that physiological measures can be used to identify affect. There is also evidence that a moderate level of arousal, which can be obtained via physiological measurement and is indicative of a positive affective state, is associated with optimal performance (Edwards, 1999). Further, increased levels of anxiety or distress interfere with learning (*e.g.*, Bandura, 1997; Spielberger, 1966).

Therefore, an affective component (Figure 2.1) was proposed (Sheldon, 2001) to enhance performance evaluation in an ITS. The affective component is designed to use physiological or behavioral measures of affect to evaluate a student's responses to instructional events (*e.g.*, feedback, presentation of material, *etc.*). The evaluation completed by the affective component is compared to both the student's performance of the instructional task and the student's interactions with the ITS. The result of this comparison is used to recommend a modification to the instructional intervention provided by the ITS so as to optimize the student's performance.

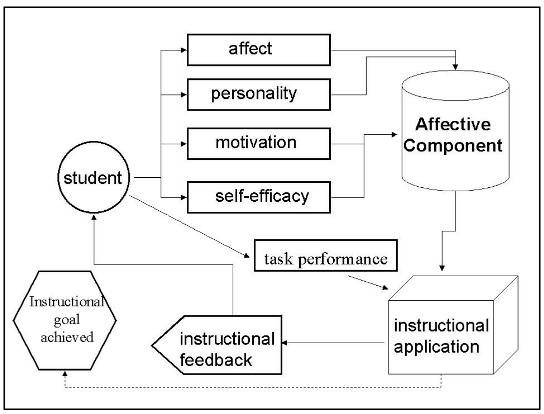

Figure 2.1. Affective Component (Sheldon, 2001).

3. Experiment

A feasibility study was conducted to evaluate the use of physiological measures to evaluate a student's interactions in a computerized instructional environment, such as an ITS. The contributions of instructional feedback styles, personality, and affect on the learning outcomes for a management training application were investigated. The training application was a discrete-event simulation regarding crisis management procedures performed by the principal in a public school system.

The crisis management training scenario presented a situation (*e.g.*, a student reports that another student brought a gun on campus) and then asked the participant, who assumed the role of the principal, to select one of four actions. The subsequent situation reflected the participant's action choice, and the scenario continued. Instructional feedback was automatically provided after each selection to inform the participant whether or not his response was correct and an explanation of why it was or was not the correct action to take. The positive and negative feedback types were audio, verbal feedback. The feedback was differentiated as positive or negative based on the instructor's vocal inflection and the phrasing of the sentences to provide an affective connotation, such as "good job" for the positive feedback and "no, that is wrong" for the negative feedback. For the neutral feedback condition, the feedback was provided as text and was devoid of any emotive phrases.

Physiological measures (skin conductance level, muscle tension, and heart rate) were recorded to provide an objective evaluation of changes in the participants' affect in response to the feedback from their physiology at a state of rest. A physiological, two-dimensional (arousal/engagement vs. distress) model of affect (Frankenhauser, 1986) was used to evaluate changes in the participant's affect. Skin conductance measured arousal, muscle tension evaluated distress, and heart-rate was used to evaluate arousal and distress, depending on the context of the situation. A measure of the participant's physiology at rest was obtained with an average of each 15 sec interval of the two-minute baseline recording, and their physiological responses to the feedback were measured by the average physiological activity (*e.g.*, heart rate) for the 15 sec period following each presentation of feedback. Change scores were computed for each of the physiological measures by taking the difference between the scenario and baseline averages.

The Eysenck Personality Questionnaire-Revised (EPQ-R: Eysenck & Eysenck, 1994) was administered to obtain data regarding the participants' personality characteristics because affective responses are influenced by personality and interaction with the environment (Matthews, 1999), such as the feedback. Finally, knowledge retention was assessed with a multiple-choice post-test on the material presented in the crisis management training scenario.

Finally, the Feedback Test was administered to the participants in the positive and negative feedback groups to determine if they perceived the feedback they received to be positive or negative, respectively. The test consisted of twenty pre-recorded sentences of which ten were spoken with a negative vocal inflection and ten were spoken with a positive vocal inflection. Additionally, the sentences included the emotive phrases from the positive and negative feedback, respectively. Additionally, there was a question that asked the participants to rate the feedback that they received in the crisis management training scenario as positive or negative.

4. Results

The purpose of the affective component is to include affect in the student model so that instruction can be tailored to optimize student performance. Therefore, the data collected in this experiment were used to build a model of student performance as a function of the feedback type (positive, negative, neutral), physiological measures of affect, and personality.

Regression analysis was conducted to build a model of training performance, as measured by the knowledge retention test. The independent variables included in the analysis were feedback type (positive, negative, or neutral), EPQ-R scores (Psychoticism, Extraversion, Neuroticism, and Lie Scale), and each of the eight physiological measures (change in mean and variance SCL, SCR, muscle tension, and HR for each participant), and gender. As there have been observed gender differences in physiological responses and personality traits (Eysenck, 1994), which was observed in the neutral feedback group, gender was included in the model. Additionally, the score for the EPQ-R Lie Scale was included in the model as it provided a measure of motivation (Sheldon, 2001). The interaction terms were included based upon the significant relationships identified in separate analyses (Sheldon, 2001). The final model was significant (R2 = .265, S = .094, p = .027*), thus supporting the feasibility of modeling student performance based on his affect, personality, and instructional interactions (*e.g.,* feedback).

Analysis of the Feedback Test indicated that the many of the participants receiving negative feedback did not report it, although analysis of their physiological measures indicated that they were responding to the feedback in a different manner than the participants who received positive feedback. Only 63% of the participants in the negative feedback group correctly identified the feedback they received as negative. A test of proportions (Ho: p = .5; Ha: p > .5) was performed using the normal approximation to the binomial to examine the relationship between feedback type and the participant's report of the type feedback they received, and the result obtained for the negative feedback group was not significant. This was an important finding that substantiated the need for objective measures of student interaction and corroborated previous findings that have provided an indication of the poor validity of subjective reports of affect.

5. Implications

The results of this experiment provide evidence that student performance is influenced by a function of his instructional interactions, affect, and personality. Specifically, the analysis demonstrated that feedback type, personality traits, and affect could be used to predict the participants' performance on the Knowledge Retention Test. It was also demonstrated that the physiological measures used in this experiment (skin conductance level, skin conductance response, muscle tension, and heart rate) were useful for evaluating the participants' reactions to the feedback provided during the crisis management training scenario.

The next step in this research is to develop a prototype affective component and integrated with an ITS. Sonalysts, Inc.'s simulation-based intelligent tutoring (Expert Train) application varies its feedback on three levels (indication of an error, hint, or answer and explanation). The Expert Train architecture permits the integration of a prototype affect component. This will enable us to conduct experimentation to compare the learning outcomes of instruction provided by the current version of Expert Train to a version of Expert Train with the affective component enhancement. Such experimentation will enable evaluation of the benefits of adapting instruction based on the student's affective responses to the environment.

Reference

Ackerman, P.L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied*, 1(4), 270-304.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W.J. Freeman & Co.

Cacioppo, J.T. & Petty. R.E (1986). Social processes. In, Coles, M.G., Donchin, E., & Porges, S.W. (Eds.), *Psychophysiology: Systems, Processes, and Application* (pp. 646-679). New York: Guilford Press.

Conati C. (2002). Probabilistic Assessment of User's Emotions in Educational Games . In: Journal of Applied Artificial Intelligence, special issue on " Merging Cognition and Affect in HCI", vol. 16.

Conati C. and Zhou X. (2002). Modeling Students' Emotions from Cognitive Appraisal in Educational Games . In Proceedings of ITS 2002, 6th International Conference on Intelligent Tutoring Systems, Biarritz, France.

Chi, M.T.T., Siler, S.A., Jeong, H., Yamauchi, T., & Hausman, R.G. (2001). Learning from human tutoring. *Cognitive Science*, 25(4), pp. 471-533.

del Soldato, T. & du Boulay, B. (1995). Implementations of motivational tactics in tutoring systems. *Journal of Artificial Intelligence in Education*, 6 (4), 337-378.

Edwards, D. (1999). *Emotion and Motivation: Evolutionary, Physiological, Cognitive, and Social Influences*. Thousands Oaks, CA: Sage Publications.

Eysenck, H. & Eysenck, S. (1994). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: EdITS/Educational and Testing Service.

Fowles, D.C. (1986). The eccrine system and electrodermal activity. In, Coles, M.G., Donchin, E., & Porges, S.W. (Eds.), *Psychophysiology: Systems, Processes, and Applications* (pp. 51-96). NewYork: Guilford Press.

Frankenhaeuser, M. (1986). A psychobiological framework for research on human stress and coping. In, Appley, M. & Trumbull, R. (Eds.), Dynamics of Stress: Physiological, Psychological, and Social Perspectives (pp. 101-116). New York: Plenum Press.

Grossman, P. (1992). Respiratory and cardiac rhythms as windows to central and autonomic biobehavioral regulation: Selection of window frames, keeping the panes clean and viewing the neural topography. *Biological Psychology*, 34, 131-161.

Kanfer, R. & Heggestad, E. (1999). Individual differences in motivation: Traits and self-regulatory skills. In Ackerman, P., Kyllonen, P., & Roberts, R. (Eds.), *Learning and Individual Differences: Process, Trait, and Content Determinants* (pp. 293-314). Washington, DC: American Psychological Association.

Kapoor, A., Mota, S., & Picard, R.W. (2001). Towards a learning companion that recognizes affect. In, Proceedings from Emotional and Intelligent II: The Tangled Knot of Social Cognition, AAAI Fall Symposium (November).

Kyllonen, P.C. & Shute, V. J. (1989). A taxonomy of learning skills. In, Ackerman, P., Sternberg, R., & Glaser, R. (Eds.), *Learning and Individual Differences: Advances in Theory and Research* (pp. 117-163). New York: W.H. Freeman & Company.

Levine, P. (1986). Stress. In, Coles, M.G., Donchin, E., & Porges, S.W. (Eds.), *Psychophysiology: Systems, Processes, and Application* (pp. 331-353). New York: Guilford Press.

Matthews, G. (1999). Personality and skill: A cognitive-adaptive framework. In Ackerman, P., Kyllonen, P., & Roberts, R. (Eds.), *Learning and Individual Differences: Process, Trait, and Content Determinants* (pp. 251-270). Washington, DC: American Psychological Association.

Mitrovic, A., Djordjevic-Kajan, S., & Stoimenov, L. (1996). INSTRUCT: Modeling students by asking questions. *User Modeling and User-Adapted Interaction*, 6, 273-302.

Murray, C.R. & VanLehn, K. (2000). DT Tutor: A decision-theoretic, dynamic approach for optimal selection of tutorial actions. In, Gauthier, G., Frasson, C., & VanLehn, K. (Eds.), Intelligent Tutoring Systems, 5th International Conference, ITS 2000 (pp. 153-162). New York: Springer.

Rickel, J. & Johnson, W. L. (1999). Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence*, 13, 343-382.

Scherer, K. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion*, 7 (1), 1-41.

Sheldon, E. (2001). *Virtual Agent Interactions* (Doctoral dissertation). Orlando: University of Central Florida.

Spielberger, C.D. (1966). *Anxiety and Behavior*. New York: Academic Press.

Sternberg, R. (1997). *Thinking Styles*. New York: Cambridge University Press.

Tong, A. & Agelides, M. (1999). Formalising tutoring strategy selection in multimedia tutoring systems. *Information and Software Technology*, 41, 67-90.

Acknowledgement

This research was supported by a grant from the Cognitive, Neural and Biomolecular Science and Technology Division, Office of Naval Research under Contract No. N00014-00-1-0077, Contract Authority No. 10 USC 2358 and USC 6304.