

POSTURE AS A PREDICTOR OF LEARNERS' AFFECTIVE ENGAGEMENT: BOREDOM AND FLOW

Sidney K. D'MELLO, Roger S. TAYLOR, Kristy TAPP, Brandon KING, and Arthur C. GRAESSER

Institute for Intelligent Systems, University of Memphis

{sdmello, rstaylor, kmsnyder, bgking, a-graesser}@memphis.edu

Abstract

This research demonstrates the utility of automatically monitoring a student's posture to track the affective states of boredom and flow, which have been shown to influence learning. After a tutoring session with AutoTutor, the affective states of the student were rated by the learner, a peer, and two trained judges. Our results indicated that the affective state of flow was manifested though a heightened pressure exerted on the seat of a pressure sensitive chair. Boredom, in turn, was associated with an increased change in pressure on the seat, perhaps indicative of a state of restlessness.

Introduction

The task of maintaining a student's engagement in educational activities is extremely challenging. Establishing and maintaining the engagement of learners is especially critical in situations with high degrees of learner control, such as in distance education, computer-based tutoring, and informal learning environments. For instance, with web-based instruction, individuals are one-mouse-click-away from ending the session. There have been several different approaches taken to address this problem, such as collaborative learning (e.g., Palincsar & Brown, 1984), educational games and simulations (e.g., Ferrari, Taylor, & VanLehn, 1999), and inquiry learning (e.g., Chinn & Malhotra, 2002). One technique which can help to maintain a student's motivation and engagement is to structure instruction so that it matches a student's zone of proximal development (Brown, Ellery, & Campione, 1998; Vygotsky, 1986). However, determining the appropriate level of difficulty is a non-trivial task. In addition, each individual has their own specific zone of proximal development – which changes over time in response to instruction – so that the goal of providing training that hits this 'moving target' is exceptionally difficult.

An exciting new alternative involves the use of *emotionally-sensitive* intelligent tutoring systems (ITSs). ITSs have been designed to assist learners in the active construction of knowledge, particularly at deeper levels of comprehension (Johnson, 2001). ITSs typically provide one-on-one tutoring, which is known to be a powerful method of promoting knowledge construction (Cohen, Kulik, & Kulik, 1982). Emotionally-sensitive ITSs attempt to incorporate the learner's emotions (i.e., affect, or affective state) into their pedagogical strategies in order to enhance learning (D'Mello, Craig, Gholson, Franklin, Picard, & Graesser, 2005).

A fundamental question that emerges in the design of an emotionally-sensitive ITS involves the manner with which to identify the affective states that accompany learning processes. For instance, a recent study found that increased levels of boredom were negatively correlated with learning of computer literacy (Craig, Graesser, Sullins, & Gholson, 2004). Alternatively, increased levels of engagement (or the flow experience, Csikszentmihalyi, 1990) were positively correlated with learning.

In the present study, we sought to evaluate the utility of tracking the affective states of students during their interactions with an ITS called AutoTutor. AutoTutor is an intelligent tutoring system that helps learners construct explanations by interacting with them in natural language (Graesser, Chipman, Haynes, & Olney, 2005). A variety of non-intrusive sensing devices were used to monitor a learner's affective states during the tutoring session. While there has been a well established body of research in detecting affect by monitoring facial features (Ekman & Friesen, 1978) or tracking speech patterns (Forbes-Riley & Litman, 2004), the research on emotion detection from *posture* is much more sparse, particularly in the context of learning. The most relevant work that has been done in this area is that by Mota and Picard (2003) in which they used posture sensors to measure children's interest levels while playing a computer game.

This paper focuses on the relationship between a learner's *posture patterns* and the affective states of *boredom*

(low engagement) and *flow* (high engagement), viewing them as approximate endpoints on a continuum of engagement.

Method

Participants

The participants were 28 undergraduate students from a mid-south university who participated for extra course credit.

Materials

AutoTutor. AutoTutor was designed to simulate a human tutor while conversing with students in natural language. It begins by presenting a challenging question to the student. The typical response from the learner, however, is usually very brief. The system then uses a series of prompts and hints to elicit more elaborate responses from the learner.

Body Pressure Measurement System (BPMS). The BPMS system, developed by Tekscan™ (1997), consists of a thin-film pressure pad (or mat) that can be mounted on a variety of surfaces. The pad is paper thin with a rectangular grid of sensing elements. Each sensing element provides a pressure output in mmHg. The setup used in this study involved the use of one sensing pad placed on the seat of a Steelcase™ Leap Chair and another placed on the back of the chair (see Figure 1).

The output of the BPMS system consists of 38x41 matrix (rows x columns) with each cell in the matrix monitoring the amount of pressure as reported by the corresponding sensing element. Therefore, in accordance with our setup, at each sampling instance (1/4 second) matrices corresponding to the pressure in the back and the seat of the chair were recorded for future, offline analyses.

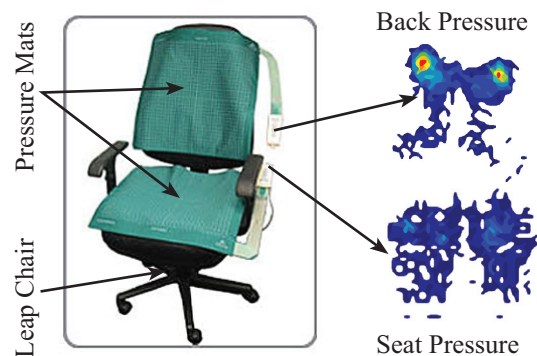


Figure 1. The Body Pressure Measurement System

Procedure

The study was divided into two phases. The first phase was a standard pretest – intervention – posttest design. During the pretest, the participants completed a 36-item assessment on the topic of computer literacy. This was followed by the instruction intervention in which the participants used the AutoTutor system for 32 minutes on one of three randomly assigned topics in computer literacy (Hardware, Internet, Operating Systems). During the tutoring session, a video of the participants' face, their posture pressure patterns, and a video of their computer screen were recorded. Lastly, after completing the tutoring session, the participants completed a 36-item posttest assessment.

The second phase involved *affect judgments* by the learner, a peer, and two trained judges. The affect judging session proceeded by displaying video streams of both the learner's screen and face, which were captured during the AutoTutor session. The raters were instructed to make judgments on what affective states were present in each 20-second interval at which the video automatically paused, as well as any other affective states they observed. Four sets of emotion judgments were made for the observed affective states of each participant's AutoTutor session. First, for the *self* judgments, the participant watched his or her own AutoTutor session immediately after having interacted

with it. Second, for the *peer* judgments, participants came back a week later to watch and rate another participant's session on the same topic in computer literacy. Finally, two *trained* judges – who had been trained on how to detect facial action units according to Ekman's Facial Action Coding System (Ekman & Friesen, 1978) – independently rated all of the sessions.

Data Treatment

Analysis of Agreement among Judges. Interrater reliabilities (Cohen's Kappa) computed across the six rater pairs were: self-peer (.08), self-trained judge1 (.14), self-trained judge2 (.16), peer-trained judge1 (.14), peer-trained judge2 (.18), and trained judge1-trained judge2 (.36). Statistical analyses reported in Graesser et al., 2006, revealed that the trained judges had the highest agreement while the novice judges (*self-peer*) had near zero agreement. The interrater reliability for the remaining four pairs was in between.

Extraction of Posture Features. At each sampling point (1/4 second) the BPMS system provides a spatial map of the pressure exerted on the seat and the back of the chair. By averaging across each of the 16,128 sensing elements on the back and seat pads one obtains the net pressure exerted on the back and seat of the chair. This forms the basis of the first two posture features: the *back pressure* and the *seat pressure* which represent the net pressure exerted by the learner on the back and seat of the chair. However, since we are primarily interested in posture patterns during episodes of boredom and flow (as indicated by the self, peer, or 2 trained judges) we computed the back and seat pressure features while learners experienced these emotions. In addition to these features that assessed the average pressure, we also introduced two features that attempted to measure the rate of change in pressure exerted by the learner on the back and seat of the chair during an emotional episode. That is, by computing the rate of change in pressure 2 seconds before and 2 seconds after the learner experienced an emotion, we were able to operationally define a measurement of the amount of body activity of the learner when he or she experienced an emotion. These features are termed the *back change* and *seat change* for the back and seat respectively.

Data Selection. Three data sets were constructed by temporally integrating the 4 posture features with the emotion judgments of the raters. Specifically, the four posture features (independent variables) were assessed in predicting the emotion of the learner (dependent variable). The first two models consisted of posture features aligned with judgments of the affective states of boredom and flow provided by the self ($N_{\text{BOREDOM}} = 483$; $N_{\text{FLOW}} = 593$) and the peer ($N_{\text{BOREDOM}} = 582$; $N_{\text{FLOW}} = 605$). The third model was constructed by considering affect ratings where both trained judges agreed on whether the learner was experiencing boredom or flow ($N_{\text{BOREDOM}} = 268$; $N_{\text{FLOW}} = 224$).

Results

The analyses proceeded by first exploring the relationships between the posture features and the affective states. We then investigated the reliability associated with a computer automatically discriminating between the states of boredom and flow.

The Relationship between Posture and Boredom and Flow

Descriptive statistics (means and 95% confidence interval) for the four posture features for boredom and flow for each of the three data sets (self, peer, trained-judges-agree) are presented as Figure 2. The results indicate that boredom is accompanied by an increase in the pressure exerted on the back of the chair (see Figure 2a). This pattern was statistically significant for the data sets in which the affect judgments were provided by the peer and the trained judges. In contrast, the affective state of flow appears to be manifested by a heightened pressure exerted on the seat of the chair (Figure 2b). This relationship was statistically significant across all three data sets.

It appears that the change in pressure exerted on the back of the chair is quantitatively similar for boredom and flow. Therefore, this feature does not appear to be very useful in discriminating these emotions (Figure 2c). However, Figure 2d indicates that boredom is typically accompanied by heightened level of activity on the seat of the chair (i.e., fidgeting). This trend was observed for the data sets where the affect judgments were provided by the peer and the trained judges but not the self judgments.

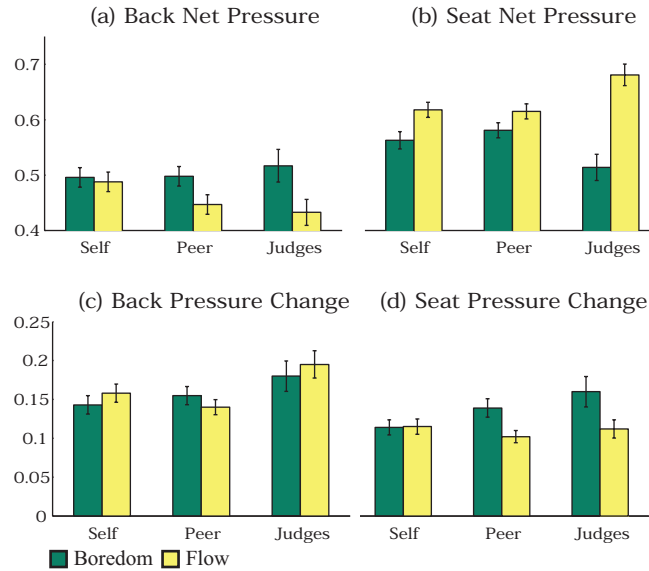


Figure 2: Descriptive statistics for posture features for the data sets

Automatic Detection of Boredom and Flow from Posture

The efficacy of an emotionally-sensitive Intelligent Tutoring System is determined in part by the accuracy with which the system can classify the learners' affective states. Therefore, we conducted a discriminant analysis on each of the three data sets in order to determine the reliability with which a computer could automatically discriminate between boredom and flow.

The classification accuracies (kappas) associated with boredom and flow as well as the overall accuracy is presented in Table 1. We note that if one focuses on the overall classification accuracy, the discriminant function constructed on the dataset where both trained judges agreed on an affect judgment ($\kappa=.357$) outperformed those constructed on the data sets of the novice judges (self and peer). This is particularly due to the fact that discriminant functions constructed on the self and peer datasets detected boredom at the chance rate. These results indicate that that posture is a viable channel to automatically segregate boredom from flow.

Table 1: Classification accuracies

Affect Judge	Classification Accuracy (Kappa)		
	Boredom	Flow	Overall
Self	0	.647	.175
Peer	.07	.370	.221
Judges	.377	.336	.357

Discussion

This exploratory research provides new findings on the relationship between a learner's posture and affective states. Our results indicated that boredom is manifested in two distinct forms. The first is consistent with the preconceived notion of boredom in which a learner stretches out, lays back, and simply disengages. However, a counter-intuitive finding is that *boredom* was associated with a form of restlessness manifested by rapid changes in pressure on the seat of the chair. It is not clear as to whether these two bodily expressions of boredom are isolated, combine, or interact during experience. Finer grained analyses would be required to tease apart these alternatives.

The affective state of *flow* was associated with a heightened pressure in the seat of the chair with minimal movement. This may imply that the learner is mentally engaged in absorbing the material and thereby devotes

a smaller amount of cognitive processing towards trivial bodily motion as explained by the proclivity-arousal framework.

One of the goals laid out in the *affective learning manifesto* (Picard, et al., 2004, p. 254) was that of “building new models and learning systems that incorporate affect, as a foundation for both new approaches to education and more effective machine learning”. We hope that the research presented herein makes a contribution toward this important goal.

Acknowledgements

We thank our research colleagues at the University of Memphis and MIT, as well as Steelcase Inc. for providing us with the Tekscan Body Pressure Measurement System at no cost. This research was supported by the National Science Foundation (REC 0106965 , ITR 0325428, REESE 0633918). Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

References

- Brown, A. L., Ellery, S., & Campione, J. C. (1998). Creating zones of proximal development electronically. In J. G. Greeno and S. V. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 341-367). Mahway, NJ: Lawrence Erlbaum.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175-218.
- Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A metaanalysis of findings. *American Educational Research Journal*, 19, 237-248.
- Craig, S., A. Graesser, J. Sullins, & B. Gholson. 2004. Affect and learning: An exploratory look into the role of affect in learning. *Journal of Educational Media* 29:241-250.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper-Row: NY.
- D’Mello, S. K., S. D. Craig, B. Gholson, S. Franklin, R. Picard, & A. C. Graesser. 2005. 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*, 7-13. New York: AMC Press.
- Ekman, P., & W. V. Friesen. 1978. *The facial action coding system: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press.
- Ferrari, M., Taylor, R., & VanLehn, K. (1999). Adapting Work Simulations for Schools. *Journal of Educational Computing Research*. 21 (1), 25-53.
- Forbes-Riley, K. and D. Litman. 2004. Predicting Emotion in Spoken Dialogue from Multiple Knowledge Sources. In Proceedings of the *Human Language Technology Conference* of the North American Chapter of the Association for Computational Linguistics, 201-208. Boston, MA.
- Graesser, A.C., P. Chipman, B. C. Haynes, and A. Olney. 2005. AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions in Education* 48: 612-618.
- Johnson, L. (2001). Pedagogical Agent Research at CARTE. *AI Magazine*, 22(4) 85-94.
- Mota, S. & R. W. Picard. 2003. Automated Posture Analysis for Detecting Learner’s Interest Level. *Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, CVPR HCI*, June, 2003.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension –monitoring activities. *Cognition and Instruction* 1(2), 117-175
- Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazel, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., & Strohecker, C. (2004). Affective learning – a manifesto. *BT Technology Journal*, 22(4), 253-269.
- Tekscan (1997). *Tekscan Body Pressure Measurement System User’s Manual*. Tekscan Inc., South Boston, MA, USA.
- Vygotsky, L. S. (1986). *Thought and Language*. Cambridge, MA: MIT Press. (Original work published 1934).