Mining Bodily Patterns of Affective Experience during Learning

Sidney D'Mello and Art Graesser
[sdmello, a-graesser]@memphis.edu
Institute for Intelligent Systems, University of Memphis

Abstract. We investigated 28 learners' postural patterns associated with naturally occurring episodes of boredom, flow/engagement, confusion, frustration, and delight during a tutoring session with AutoTutor, a dialogue-based intelligent tutoring system. Training and validation data were collected in a learning session with AutoTutor, after which the learners' affective states (i.e., emotions) were rated by the learner, a peer, and two trained judges. An automated body pressure measurement system was used to capture the pressure exerted by the learner on the seat and back of a chair during the tutoring session. We extracted 16 posture-related features that focused on the pressure exerted along with the magnitude and direction of changes in pressure during emotional experiences. Binary logistic regression models yielded medium sized effects in discriminating the affective states from neutral. An analysis of the parameters of the models indicated that the affective states were manifested by three unique postural configurations and a general increase in movement (when compared to neutral).

1 Introduction

Intelligent Tutoring Systems (ITSs) have emerged as valuable tools to promote active learning by capitalizing on the benefits of one-on-one tutoring in an automated fashion. Although ITSs have typically focused on the learner's cognitive states, they can be far more than mere cognitive machines. ITSs can be endowed with the ability to recognize and respond to learners' affective states. Consequently, the last decade has witnessed a burst of research activities aimed at developing ITSs that are responsive to learners' affective states in addition to their cognitive states [1-6]. Much of the research has focused on developing systems to detect learner affect automatically because an ITS cannot respond to learners' affective states if it cannot detect their affective states.

State-of-the-art affect detection systems have overlooked posture as a serious contender when compared to facial expressions and acoustic-prosodic features (see reviews by [7, 8]), so an analysis of posture merits a closer examination. There apparently are some benefits to using posture as a means to diagnose the affective states of a learner. One compelling reason is that human bodies are large and have multiple degrees of freedom, thereby providing them with the capability of assuming a myriad of unique configurations [9]. These static positions can be concurrently combined and temporarily aligned with a multitude of movements, all of which makes posture a potentially ideal affective communicative channel [10, 11]. Perhaps the greatest advantage to posture-based affect detection is that gross body motions are ordinarily unconscious, unintentional, and thereby not susceptible to social editing, at least compared with facial expressions, speech intonation, and some gestures [12].

There has been some recent research aimed at developing posture-based affect detection systems [13-15]. The results of these studies have indicated that posture is an important channel to detect the affective states that occur during learning sessions. Although these studies have provided an indication of *how accurate* posture-based affect detectors are, left unanswered is the important question of *how does* the body convey affect through articulations of posture and modulations of movement. For example, we have previously reported that supervised classifiers can discriminate confusion, frustration, boredom, engagement/flow, and delight from neutral with an average accuracy of 74% (baseline = 50%) [15], but we are unaware of the manner in which these states are expressed through the body (e.g. forward leans, arms akimbo, general fidgeting).

Hence, taking a step back from affect detection, the present paper focuses on applying data mining techniques to uncover relationships between body position, arousal, and affect. Learner affect and posture data were obtained from an earlier study [16] where 28 students were tutored on computer literacy topics with AutoTutor, an ITS with conversational dialogues [17]. Learner affect was measured via a retrospective affect judgment protocol in which the learners' affective states were rated by the learner, a peer, and two trained judges (described in section 3). An automated body pressure measurement system was used to capture the pressure exerted by the learner on the seat and back of a chair during the tutoring session. Bodily patterns of affective experience were mined with binary logistic regressions. We begin with a description of the theoretical framework that links body posture and affect.

2 Theoretical Framework Linking Posture and Affect

Several researchers have proposed that affective experience can be productively analyzed by initially considering the underlying dimensions of valence (pleasant/unpleasant) and arousal (active/sleepy) [18, 19]. After this initial classification of emotions is achieved, the emotion is analyzed further with respect to the task, environment, and social world (i.e., an *appraisal* is performed). We adopt the first-stage valence-arousal framework in uncovering the expressive nature of body language in communicating affect.

One challenge to adopting the valence-arousal framework for learning environments is that it does not neatly fit as well as the analyses of the basic emotions (anger, fear, sadness, enjoyment, disgust, and surprise) which naturally align on a valence scale [20]. A simple valence-based categorization is more challenging for the affective states that are prominent during learning. For example, while prolonged, hopeless confusion can be considered to be a negative emotion, it is not necessarily the case that brief experiences of confusion share the same negative connotation. In fact, confusion is a state that has been positively correlated with deep thinking and learning [21, 22], and some learners, such as academic risk takers, like to be challenged with tasks that create short-term confusion, followed by a resolution [23]. Although it is difficult to fathom that learners derive pleasure from being confused, there are theories of learning that suggest complex relations between affect and cognition that extend beyond a simple valence dimension.

Therefore, we relied on an *attentive-arousal* framework [24] to interpret relationships between posture and the affective states of the learner. One can think of heightened pressure in the seat as resonating with a tendency to position one's body towards the source of stimulation (i.e., high attentiveness since the learner is positioning his or her body towards the learning interface, or a short distance between the nose and the screen). On the other hand, an increase in pressure on the back of the chair suggests that the learner is leaning back and detaching himself or herself from the stimulus (low attentiveness). Arousal was operationally defined by the rate of change of pressure exerted on the back and the seat of the pressure sensitive chair (described below) and is similar to the degree of movement [18].

3 Data Collection

The participants were 28 undergraduate students from a mid-south university who participated for extra course credit. The study was divided into two phases. The first phase was a standard pretest–intervention–posttest design. The participants completed a pretest with multiple-choice questions, then interacted with the AutoTutor system for 32 minutes on one of three randomly assigned topics in computer literacy (Hardware, Internet, Operating Systems), and then completed a posttest. During the tutoring session, the system recorded a video of the participant's face, their posture patterns with the Body Pressure Measurement System (BPMS), and a video of their computer screen.

The BPMS system, developed by Tekscan™ consisted of a thin-film pressure pad (or mat) that can be mounted on a variety of surfaces. The pad was paper thin with a rectangular grid of sensing elements. Each sensing element provided a pressure output in mmHg. Our setup had one sensing pad placed on the seat of a Steelcase™ Leap Chair and another placed on the back of the chair (see Figure 1A).

The output of the BPMS system consisted of 38 × 41 matrix (rows × columns) with each cell in the matrix monitoring the amount of pressure as reported by the corresponding sensing element (see Figure 1B). 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 offline analyses.

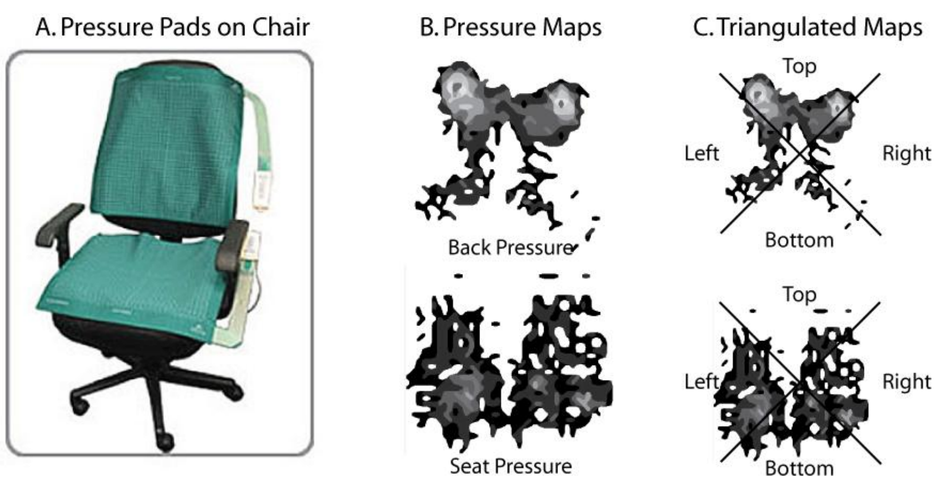

Figure 1: Body Pressure Measurement System

The second phase of the study 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. Judges were instructed to make judgments on what affective states were present at any moment during the tutoring session by manually pausing the videos (spontaneous judgments). They were also instructed to make judgments at each 20-second interval where the video automatically stopped (fixed judgments); these fixed points are particularly useful to compute interrater reliability among judges (see [16]). A list of the affective states and definitions was provided for all judges. The states were boredom, confusion, flow, frustration, delight, neutral and surprise. The selection of emotions was based on previous studies of AutoTutor and other learning environments [25].

Judgments were provided by the learners themselves (self judgments), by untrained peers, and by two trained judges. These trained judges had been trained on how to detect facial action units according to Ekman's Facial Action Coding System [20] and on characteristics of AutoTutor's dialogue (i.e., contextual cues).

4 Results and Discussion

4.1 Computing Posture Features

Several features were computed by analyzing the pressure maps of the 28 participants recorded in the study. We computed six pressure-related features and two features related to the pressure coverage for both the back and the seat, yielding 16 features in all. Each of the features was computed by examining the pressure map (called the *current frame*) during an emotional episode (i.e., when an emotion judgment was made). The pressure related features include the *average pressure,* which measures the average pressure exerted on the chair and the *top* pressure which was the pressure on the topmost segment of the pad (see Figure 1C). The *prior change* and *post change* measure the difference between the average pressure in the current frame and the frame three seconds earlier and later respectively. The *reference change* measures the difference between the average pressure in the current frame and the frame for the last known affective rating. Finally, the *average pressure change* measures the mean change in the average pressure across a predefined window, typically 4 seconds, that spans two seconds before and two seconds after an emotion judgment. The two coverage features examined the proportion of non-negative sensing units (*average coverage*) on each pad along with the mean change of this feature across a 4-second window (*average coverage change*). Please see [15] for detailed description of the features.

Each feature vector was associated with an emotion category on the basis of each of the four human judges' affect ratings. More specifically, each emotion judgment was temporally bound to each posture based feature vector. This data collection procedure yielded four ground truth models of the learner's affect (self, peer, two trained judges), so we were able to construct four labeled data sets. When aggregated across each 32-minute session for each of the 28 participants, we obtained 2967, 3012, 3816, and 3723 labeled data points for the self, peer, trained judge 1, and trained judge 2, respectively.

4.2 Affect-Neutral Discrimination from Posture Features

We conducted a series of binary logistic regression analyses in order to systematically explore relationships between the posture features and the various affective states. The analyses served two purposes. First, we were interested in determining the extent to which the five affective states of interest could be predicted from the various posture features. Second, the statistically significant predictors (positive or negative) of the logistic regression models were used to isolate commonalities in posture configurations that accompany affective experience through the body.

We conducted five binary logistic regression analyses for each of the four data sets (self, peer, 2 trained judges). Each logistic regression analysis was conducted to segregate each of the five affective states from neutral. Therefore, the criterion (dependent) variable for each logistic regression analysis was the affective state (1 or 0 if present or absent, respectively) whereas the predictor variables were the set of posture features.

Statistically significant relationships ($p < .05$) were discovered for all of the models. When R^2 values were averaged across judges and affective states, the posture features explained 11% of the variance in discriminating each affective state from neutral. For the affective states of boredom, delight, flow, and frustration, the posture features explained 13% 12%, 14%, and 11% of the variance respectively. The weakest model was obtained for confusion with the posture features explaining only 6% of the variance.

However, when one considers the best model across judges (e.g., self for boredom and confusion, peer for flow and frustration, and judge 2 for delight), 19%, 10%, 16%, 19%, and 14% of the variance was explained for boredom, confusion, delight, flow, and frustration, respectively. More succinctly, 16% of the variance was explained on average. This result is close to a medium sized effect [26], and indicates that posture is indeed a viable channel for affect detection.

Table 1. Summary of binary logistic regression analyses with posture features as predictors of each affective state from neutral for data sets based on affective judgments of the four judges

	Affect Judge							
Affective State	*Self*		*Peer*		*Judge 1*		*Judge 2*	
	χ^2	$R^{2\,b}$	χ^2	$R^{2\,b}$	χ^2	$R^{2\,b}$	χ^2	$R^{2\,b}$
Boredom	196.87	.19	139.43	.12	91.44	.09	148.61	.10
Confusion	105.71	.10	87.05	.08	53.35	.03	44.01	.03
Delight	42.24	.09	29.82	.09	94.13	.12	10.22	.16
Flow	142.96	.13	235.40	.19	118.92	.09	142.90	.13
Frustration	101.94	.12	102.61	.14	73.33	.08	71.63	.11

[b] Computed in accordance with the Nagelkerke R^2 which is a pseudo R^2 [27].

Next, we consulted the numerical directions (i.e. signs, + and -) of the statistically significant coefficients of the logistic regression models in order to explore relationships between body position, movement, and affect. Although all 16 features were used as predictors for the logistic regression analyses, we focus on the average pressure and the change in pressure; these are the most interpretable features and can be theoretically aligned within the attentive-arousal framework.

The logistic regression analyses were used to discriminate between each affective state versus neutral, hence, a statistically significant predictor implies that the feature is heightened (significant positive predictor) or suppressed (significant negative predictor) during the emotional experience when compared to neutral. For example the *back average pressure change* feature was a significant positive predictor for the boredom-neutral logistic regression, so the episodes of boredom were accompanied by an increase in movement on the back when compared to neutral.

Table 2. Significant predictors for the multiple regression models for emotions in each data set

		Boredom				**Confusion**				**Delight**				**Flow**				**Frustration**			
Sensor	**Feature**	*SF*	*PR*	*J1*	*J2*	*SF*	*PR*	*J1*	*J2*	*SF*	*PR*	*J1*	*J2*	*SF*	*PR*	*J1*	*J2*	*SF*	*PR*	*J1*	*J2*
Back	Pressure			+	+		-	-			-	-		-	-	-	-		-		
	Change		-									+	+			+	+		-		
Seat	Pressure		-	-	-						+	+	+	+		+	+	-	+		+
	Change		+	+	+		+	+				+				-			+	+	

Notes. SF: Self Judgments, PR: Peer Judgments, J1: Trained Judge1, J2: Trained Judge2
+ or - indicates that the feature is a positive or negative predictor in the logistic regression model at $p < .05$ significance level. Empty cells indicate that a feature was not a statistically significant predictor for the respective emotion.

A number of relationships surface when one considers the significant predictors of the affective states in which at least two judges agreed. The requirement that two judges agree on the significance and direction of each predictor is motivated by a desire to establish a degree of convergent validity in exploring the posture-affect relationships. By requiring that the features need to be significant predictors of affect for at least half of the judges models ensures that, to some extent, they generalize across judges.

Boredom. Our results suggested that during episodes of boredom, the learners leaned back and presumably disengaged from the learning environment (low attentiveness indicated by an increase in pressure on back and a significant decrease in pressure on the seat). Experiences of boredom were also accompanied by an increase in the rate of change of pressure exerted on the seat. Therefore, heightened arousal was associated with the boredom experience, presumably as learners mentally disengage from the tutor and divert their cognitive resources to fidget around and alleviate their ennui.

Some may view the heightened arousal accompanying boredom to conflict with the preconceived notion of boredom in which a learner stretches out, lays back, and simply disengages. Our results suggest that the learner lays back, disengages, but is aroused. Furthermore, this pattern of increased arousal accompanying disengagement (or boredom) replicates a previous study by Mota and Picard [14]. They monitored activity related posture features and discovered that children fidget when they were bored while performing a learning task on a computer.

Delight and Flow. In contrast to boredom, learners experiencing the positive emotions of delight and flow demonstrate increased attentiveness towards the learning environment by leaning forward. Learners experiencing these emotions also demonstrate heightened arousal on the back of the chair – at least when compared to the neutral state.

Confusion and Frustration. Similar to delight and flow, learners experiencing confusion and frustration tend to lean forward. However, it appears that during episodes of delight and flow learners lean forward at a steeper inclination than with confusion and frustration. We arrived at this conclusion because the increase in the pressure exerted on the seat of the chair was accompanied by a commensurate decrease in pressure exerted on the back of the chair for delight and flow. On the other hand, confusion was accompanied by a decrease in pressure exerted on the back of the chair without any accompanying increase on the seat. Similarly for frustration the increase in pressure on the seat was devoid of a notable (statistically significant) decrease on the back. We suspect that the pattern of body position with confusion and frustration indicates that learners are in an upright position when they experience these states, as opposed to the forward lean that seems to accompany experiences of delight and flow.

Arousal. In addition to boredom, it also appears that experiences of delight, flow, confusion, and frustration are accompanied by significant arousal, either on the back or the seat of the pressure sensitive chair. The arousal that is affiliated with confusion and frustration occurs on the back while there is an increase in movement on the seat when the emotions of delight and flow are experienced. While the significance of the location of the arousal (back or seat) during experiences of these is unclear, what is important is that all affective experiences (including boredom) were accompanied by heightened arousal when compared to neutral. In summary, the experience of each affective state is accompanied by a significant increase in arousal, at least when compared to the neutral baseline.

General Discussion

We discovered relationships between body position, degree of movement, and learners' affective states. Our findings are in line with an attentive-arousal or engagement-arousal framework (see section 2). With respect to the attentiveness dimension, it appears that there are three bodily configurations that are associated with the affective states. These include heightened *attentiveness,* which is manifested by a forward lean when the positive emotions of delight and flow are experienced. On the other hand, bored learners tend to lean back, presumably in a state of *disengagement* (boredom).

States such as confusion and frustration occur when learners confront contradictions, anomalous events, obstacles to goals, salient contrasts, perturbations, surprises, equivalent alternatives, and other stimuli or experiences that fail to match expectations [28]. Learners are in a state of cognitive disequilibrium, with more heightened physiological arousal, and more intense thought. Our results suggest that the bodily corollary to the mental state of cognitive disequilibrium is an *alert* position, where the learner sits upright and pays attention.

The single major finding with respect to the arousal dimension is that each of the affective experiences is accompanied by a significantly higher arousal when compared to a neutral baseline (i.e. no emotion). This finding has important theoretical implications because some of our colleagues view some of these emotions (i.e., flow, confusion, etc.) as cognitive states, whereas other researchers would classify them as either emotions or affect states. We have traditionally agreed with the latter group because we hypothesize that the single major discriminator of an affective state over a cognitive state is that the affective state is accompanied by enhanced physiological arousal (compared with neutral). Our results indicate that in most cases there is a significant increase in bodily movements (bodily arousal) during the experience of emotional episodes indicating that both cognitive and affective processes are at play. Therefore, it might be the case that the term cognitive-affective state is the most defensible position for mental states such as confusion, flow, and frustration.

We acknowledge that the aforementioned relationships between body posture and affect ignore individual differences in affect expression. In ideal circumstances, from a statistical point of view, the landscape of postural configurations would be evenly distributed among the 28 different students. However, this claim is implausible and it is therefore important to contrast contributions of individual learners versus generalizable posture features in predicting affect. Our results focused on broad patterns observed across all learners and should be interpreted with a modicum of caution. We are currently addressing these concerns by building models that attempt to separate variance explained by individual student characteristics versus variance explained by posture features above and beyond individual differences.

References

[1] A. Kapoor and R. Picard, "Multimodal affect recognition in learning environments," in *Proceedings of the 13th annual ACM international conference on Multimedia*, Hilton, Singapore, 2005.

[2] C. Conati and H. Maclaren, "Empirically building and evaluating a probabilistic model of user affect," *User Modeling and User-Adapted Interaction,* vol. 19, pp. 267-303, 2009.

[3] K. Forbes-Riley, M. Rotaru, and D. Litman, "The relative impact of student affect on performance models in a spoken dialogue tutoring system," *User Modeling and User-Adapted Interaction,* vol. 18, pp. 11-43, Feb 2008.

[4] S. McQuiggan, B. Mott, and J. Lester, "Modeling self-efficacy in intelligent tutoring systems: An inductive approach," *User Modeling and User-Adapted Interaction,* vol. 18, pp. 81-123, Feb 2008.

[5] S. D'Mello, R. Picard, and A. Graesser, "Towards an affect-sensitive AutoTutor," *Intelligent Systems, IEEE,* vol. 22, pp. 53-61, 2007.

[6] I. Arroyo, B. Woolf, D. Cooper, W. Burleson, K. Muldner, and R. Christopherson, "Emotion Sensors Go To School," in *Proceedings of 14th International Conference on Artificial Intelligence In Education*, V. Dimitrova, R. Mizoguchi, B. Du Boulay, and A. Graesser, Eds. Amsterdam: IOS Press, 2009, pp. 17-24.

[7] Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 31, pp. 39-58, Jan 2009.

[8] M. Pantic and L. Rothkrantz, "Toward an affect-sensitive multimodal human-computer interaction," *Proceedings of the IEEE,* vol. 91, pp. 1370-1390, Sep 2003.

[9] N. Bernstein, *The co-ordination and regulation of movement*. London: Pergamon Press, 1967.

[10] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *Journal of Nonverbal Behavior,* vol. 28, pp. 117-139, Sum 2004.

[11] J. Montepare, E. Koff, D. Zaitchik, and M. Albert, "The use of body movements and gestures as cues to emotions in younger and older adults," *Journal of Nonverbal Behavior,* vol. 23, pp. 133-152, Sum 1999.

[12] P. Ekman and W. Friesen, "Nonverbal Leakage and Clues to Deception," *Psychiatry,* vol. 32, pp. 88-&, 1969.

[13] G. Castellano, M. Mortillaro, A. Camurri, G. Volpe, and K. Scherer, "Automated Analysis of Body Movement in Emotionally Expressive Piano Performances," *Music Perception,* vol. 26, pp. 103-119, Dec 2008.

[14] S. Mota and R. Picard, "Automated Posture Analysis for Detecting Learner's Interest Level," in *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03. Conference on*. vol. 5, 2003, pp. 49-49.

[15] S. D'Mello and A. Graesser, "Automatic detection of learners' affect from gross body language," vol. 23, pp. 123 - 150, 2009.

[16] A. Graesser, B. McDaniel, P. Chipman, A. Witherspoon, S. D'Mello, and B. Gholson, "Detection of emotions during learning with AutoTutor," in *28th Annual*

Conference of the Cognitive Science Society, Vancouver, Canada, 2006, pp. 285-290.

[17] A. Graesser, S. L. Lu, G. Jackson, H. Mitchell, M. Ventura, A. Olney, and M. Louwerse, "AutoTutor: A tutor with dialogue in natural language," *Behavioral Research Methods, Instruments, and Computers,* vol. 36, pp. 180-193, 2004.

[18] J. Russell, "Core affect and the psychological construction of emotion," *Psychological Review,* vol. 110, pp. 145-172, 2003.

[19] L. Barrett, B. Mesquita, K. Ochsner, and J. Gross, "The experience of emotion," *Annual Review of Psychology,* vol. 58, pp. 373-403, 2007.

[20] P. Ekman and W. Friesen, *The Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press, 1978.

[21] S. Craig, A. Graesser, J. Sullins, and J. Gholson, "Affect and learning: An exploratory look into the role of affect in learning," *Journal of Educational Media,* vol. 29, pp. 241-250, 2004.

[22] A. Graesser, P. Chipman, B. King, B. McDaniel, and S. D'Mello, "Emotions and learning with AutoTutor," in *13th International Conference on Artificial Intelligence in Education* R. Luckin, K. Koedinger, and J. Greer, Eds. Amsterdam: IOS Press, 2007, pp. 569-571.

[23] D. Meyer and J. Turner, "Re-conceptualizing emotion and motivation to learn in classroom contexts," *Educational Psychology Review,* vol. 18, pp. 377-390, Dec 2006.

[24] P. Bull, *Posture and Gesture*. Oxford Pergamon Press, 1987.

[25] R. Baker, S. D'Mello, M. Rodrigo, and A. Graesser, "Better to be frustrated than bored: The incidence and persistence of affect during interactions with three different computer-based learning environments," *International Journal of Human-Computer Studies,* vol. 68 (4), pp. 223-241, 2010.

[26] J. Cohen, "A power primer," *Psychological Bulletin,* vol. 112, pp. 155-159, Jul 1992.

[27] N. J. D. Nagelkerke, "A Note On A General Definition Of The Coefficient Of Determination," *Biometrika,* vol. 78, pp. 691-692, 1991.

[28] A. Graesser, S. Lu, B. Olde, E. Cooper-Pye, and S. Whitten, "Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down," *Memory and Cognition,* vol. 33, pp. 1235-1247, 2005.