Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Facial Features for Affective State Detection in Learning Environments

Permalink

https://escholarship.org/uc/item/9w00945d

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN

1069-7977

Authors

McDaniel, Bethany
D'Mello, Sidney
King, Brandon
et al.

Publication Date

2007

Peer reviewed

Facial Features for Affective State Detection in Learning Environments

Bethany McDaniel (btmcdanl@memphis.edu)
Department of Psychology, University of Memphis

Sidney D’Mello (sdmello@memphis.edu)
Department of Computer Science, University of Memphis

Brandon King (bgking@memphis.edu)
Department of Psychology, University of Memphis

Patrick Chipman (pchipman@memphis.edu)
Department of Psychology, University of Memphis

Kristy Tapp (kmsnyder@memphis.edu)
Department of Psychology, University of Memphis

Art Graesser (a-graesser@memphis.edu)
Department of Psychology, University of Memphis

Abstract

This study investigated facial features to detect the affective states (or emotions) that accompany deep-level learning of conceptual material. Videos of the participants’ faces were captured while they interacted with AutoTutor on computer literacy topics. After the tutoring session, the affective states (boredom, confusion, delight, flow, frustration, and surprise) of the student were identified by the learner, a peer, and two trained judges. Participants’ facial expressions were coded by two independent judges using Ekman’s Facial Action Coding System. Correlational analyses indicated that specific facial features could segregate confusion, delight, and frustration form the baseline state of neutral, but boredom was indistinguishable from neutral. We discuss the prospects of automatically detecting these emotions on the basis of facial features that are highly diagnostic.

Keywords: Facial features; action units, affective states; emotions; learning; AutoTutor; classifying affect

Introduction

It is widely acknowledged that cognition, motivation, and emotions are three fundamental components of learning (Snow, Corno, & Jackson, 1996). Emotion has been viewed as source of motivational energy (Harter, 1981; Miserandino, 1996; Stipek, 1998), but it can also be viewed as a more complex independent factor that plays an explanatory role in both learning and motivation (Ford, 1992; Meyer & Turner, 2002). The link between emotions and learning has received more attention during the last decade in the fields of psychology, education, and computer science (Craig, Graesser, Sullins, & Gholson, 2004; Graesser, Jackson, & McDaniel, 2007; Kort, Reilly, & Picard, 2001; Picard 1997; Meyer & Turner, 2002).

Ekman and Friesen (1978) highlighted the expressive aspects of emotions with their Facial Action Coding System. This system specified how specific facial behaviors, based on the muscles that produce them, could identify “basic emotions”. Each movement in the face is referred to as an action unit (or AU). There are approximately 58 action units. These facial patterns were used to identify the emotions of happiness, sadness, surprise, disgust, anger, and fear (Ekman & Friesen, 1978; Elfenbein & Ambady, 2002). Doubts have been raised, however, that these six emotions are frequent and functionally significant in the learning process (D’Mello et al., 2006; Kapoor, Mota, & Picard, 2001). Some have challenged the adequacy of basing a theory of emotions on these “basic” emotions (Rozin & Cohen, 2003). Moreover, Ekman’s coding system was tested primarily on static pictures rather than on changing expressions over time..

There is some evidence for a different set of emotions that influence learning and cognition, specifically boredom (Csikszentmihalyi, 1990; Miserandino, 1996), confusion (Graesser & Olde, 2003; Kort, Reilly, & Picard, 2001), flow (i.e. engagement, Csikszentmihalyi, 1990), and frustration (Kort, Reilly, & Picard, 2001; Patrick et al., 1993). Curiosity and eureka (i.e. the “a-ha” experience) are also believed to accompany learning. A study was recently conducted to investigate the occurrence of these emotions, as well as Ekman’s basic emotions. The study used an emote-aloud procedure (D’Mello et al., 2006), a variant of the think-aloud procedure (Ericsson & Simon, 1993), as an online measure of the learners’ affective states during learning. College students were asked to express the affective states they were feeling while working on a task, in this case being tutored in computer literacy with AutoTutor. Using the emote-aloud method allowed for the on-line identification of emotions while working on the learning task. A sample of 215 emote-aloud observations were produced by 7 participants. The emotions of interest were listed and defined for the student before they started the learning task. The emotions came from both groups: (1) learning-specific emotions, i.e., boredom, confusion, curious, eureka, and frustration and (2) basic emotions, e.g., anger and disgust. The results indicated that the 5 learning specific emotions accounted for 91% of the total verbalized reports while the remaining 9% of the emote-aloud utterances were for the basic emotions. In addition, curiosity rarely occurred (3%) and eureka was often confused with delight when giving a correct answer or surprise when getting an unexpected feedback from the tutor. These affective states were therefore replaced with delight and surprise.

The present study focused on boredom, confusion, delight, flow, frustration, and surprise as the affective states that are relevant to learning. We acknowledge that this set of affective states is not exhaustive for all learning environments, but they were the most prominent while college students learn with AutoTutor (Graesser et al., 2006). Moreover, some of these affective states have been correlated with learning gains. Boredom is negatively correlated with learning, whereas confusion and flow are positively correlated with learning (Craig et al., 2004).

This study identified the action units that accompany the experience of these selected emotions. Once the representative AUs for the learning-specific affective states are determined, computer analyses can be used to automatically identify those AUs and make inferences on the emotions of the learner. If the affective states of a learner can be detected with a reasonable degree of accuracy, then intelligent tutoring systems, peer learning companions, and educational software in general can revise their pedagogical strategies by incorporating both the affective states of the learner and their cognitive states.

Methods

Participants

The participants were 28 undergraduate students who participated for extra course credit in a psychology course.

Materials

AutoTutor. AutoTutor is a fully automated computer tutor that simulates human tutors and holds conversations with students in natural language (Graesser, Chipman, Haynes, & Olney, 2005; Graesser et al., 1999). AutoTutor helps students learn Newtonian physics, computer literacy, and critical thinking by presenting challenging problems (or questions) from a curriculum script and engaging in a mixed-initiative dialog while the learner constructs an answer. In this study, students interacted with AutoTutor on topics concerning computer literacy.

Procedure

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 pressure patterns, and a video of their computer screen.

The second phase involved affect judgments by the learner, a peer, and two trained judges. A list of the affective states and definitions was provided for all four judges. The states were boredom, confusion, flow, frustration, delight, neutral and surprise. The selection of emotions was based on previous studies of AutoTutor (Craig et al., 2004; D'Mello et al., 2006; Graesser et al., 2006) that collected observational data and emote-aloud protocols while college students learned with AutoTutor.

In the affect judging session, participants viewed video streams of both the computer screen and face of the learner, during the AutoTutor session. The judges were instructed to make judgments on what affective states were present in each 20-second interval (called mandatory judgments); at these points the video automatically paused for their affect judgments. They were also instructed to indicate any affective states that were present in between the 20-second stops (voluntary judgments). Four sets of emotion judgments were made for each participant's AutoTutor session. First, in the self judgments, participants viewed their own AutoTutor session immediately after their learning session. Second, in the peer judgments, participants came back one week later to view and judge another participant's session on the same topic in computer literacy. Finally, two trained judges independently judged all of the sessions. The judges were trained on Ekman's Facial Action Coding System (Ekman & Friesen, 1978) and on characteristics of dialogue.

Data Treatment

Analysis of Agreement among Judges. Interjudge reliability was computed using Cohen's kappa for all possible pairs of judges: self, peer, trained judge 1, and trained judge 2. There were 6 possible pairs altogether. The kappas for the mandatory judgments were: self-peer (.06), self-judge 1 (.11), self-judge 2 (.13), peer-judge 1 (.11), peer-judge 2 (.15), and judge 1-judge 2 (.31). These kappa scores revealed that the trained judges had the highest agreement, the self and peer pair had lowest agreement, and the other pairs of judges were in between. It should be noted, however, that the kappa scores increased substantially [self-peer (.12), self-judge 1 (.31), self-judge 2 (.24), peer-judge 1 (.36), peer-judge 2 (.37), and judge 1-judge 2 (.71)] when we focused on voluntary judgments. Additional details on this analysis are reported in Graesser at al. (2006).

Data Selection. The above kappa scores indicate that an emotion labeling task is more difficult if judges are asked to make emotion judgments at regularly polled timestamps, rather than being able to stop a video display to make spontaneous judgments. The states at regular timestamps are much less salient so there is minimal information to base their judgments, compared with those points when affective states are detected by the judge. Therefore, we decided to focus on data points where the affect judgments were voluntarily provided by the judges. This approach is beneficial for two reasons. First, the increased kappa scores made us more confident in the validity of the emotion measurement task. Secondly, it is reasonable to assume that the facial expressions of learners are more animated at these voluntary points compared to regularly polled timestamps when their face is frequently neutral. This hypothesis was confirmed by analyzing the proportions of emotion categories at the mandatory points. When averaged over the 4 judges, the most common affective state was neutral (36.9%), followed by confusion (21.2%), flow (18.8%), and boredom (16.7%). The remaining states of delight, frustration and surprise constituted 6.5% of the observations. The more salient voluntary points had a rather different distribution. The most prominent affective state was confusion (37.7%), followed by delight (19.2%), and frustration (19.1%). The remaining affective states comprised 24.0% of the observations (boredom, surprise, flow, and neutral, in descending order). Therefore, the affective states that presumably accompany heightened arousal (confusion, delight, and frustration) are more prominent at the voluntary points.

One consequence of exclusively relying on voluntary judgments for facial expression coding is that the frequency of these points is substantially less than the mandatory observations. There were 2688 data points for the mandatory observations, but only 1133 points of voluntary observations. A subset of the voluntary points was identified by only a single judge (self, peer, or one of the trained judges). This sample size reduction problem was mitigated by an exhaustive voluntary affect judgment session in which the trained judges repeated the judgment procedure with the added requirement that they had to provide affect ratings on all 1133 voluntary emotion observations. We found that the two trained judges agreed 64% of the time (N = 720) yielding a kappa score of 0.49. This kappa score is higher than that achieved for the mandatory observations (0.31) but substantially lower when compared to the purely voluntary observations (0.71). However, it is on par with reliability scores reported by other researchers who have assessed identification of emotions by humans (e.g. Ang et al., 2002; Forbes-Riley & Litman, 2004).

Since facial action unit coding is a time consuming, labor intensive process, we sampled 212 emotion episodes out of the 720 data points obtained from the exhaustive voluntary affect judgment procedure. These points were sampled to approximate a uniform distribution of the different emotions, i.e., an approximately equal number of observations was obtained from each participant and for each of the affective states of boredom, confusion, delight, frustration, and neutral. Surprise was not included because this emotion was extremely rare. Flow was also excluded because it rarely appeared in the voluntary emotion samples.

Scoring Procedure. The expression of emotions tends to be short in duration, lasting only about 3 seconds (Ekman, 2003). Two judges independently coded the facial videos 3-4 seconds before the emotional episode, using the Facial Action Coding System (Ekman & Friesen, 1978). Specifically, 212 video clips 3-4 seconds in duration were prepared for action unit coding. The clips were not associated with any discernable emotion annotation, so the judges were unaware of the learner's emotion while viewing the clips. The two AU coders were not the trained judges used for the emotion judgment procedure. Each coder watched the clips and recorded the AUs present along with the time of each observation. It is important to note that we decided to focus on a subset (N = 33) of the action units that were most relevant to the emotions under exploration.

Results

We computed the proportion of AUs observed by each rater and discarded AUs that occurred relatively infrequently. We adopted an ad-hoc selection threshold of 3%, such that AUs that appeared in less than 3% of the samples were discarded. When averaged across both judges, we preserved 12 AUs that collectively represented 80% of the observations. The 21 less frequent AUs comprised the remaining 20% of the observations and were subsequently discarded.

Facial Action Unit			Prop.	Kappa
Upper Face	AU1	Inner Brow Raiser	.056	.642
	AU2	Outer Brow Raiser	.033	.534
	AU4	Brow Lowerer	.057	.779
	AU7	Lid Tightener	.079	.590
	AU43	Eye Closure	.047	.605
	AU45	Blink	.172	.681
Lower Face	AU12	Lip Corner Puller	.100	.707
Lip Parting & Jaw Opening	AU25	Lips Part	.089	.912
	AU26	Jaw Drop	.056	.851
Head Positions	AU55	Head Tilt Left	.035	.770
	AU56	Head Tilt Right	.035	.665
Eye Positions	AU64	Eyes Down	.037	.833
Other	-	-	.206	-

Table 1. Proportion of Action Units Observed

Table 1 presents the proportion of each of the AUs averaged across the two human coders. Kappa scores between the two coders for each of the AUs are also presented. We note that the majority of the activity of the facial features during emotional experiences occurred on the upper face, with the mouth area a close second. The kappa scores also indicate that the level of agreement achieved by the AU judges in coding the target action units ranged from fair to excellent (*M* = .721, *SD* = .117).

Relationship between Action Units and Emotions

Correlations were computed to determine the extent to which each of the AUs were diagnostic of the affective states of boredom, confusion, delight, frustration, and neutral. Two sets of correlations were computed in order to determine whether significant patterns emerged across both independent coders. These correlations are presented in Table 2. The analyses revealed that there was a good degree of concordance among the two judges. Barring a few anomalies that may be attributed to individual differences of the coders, the directions of the relationships are consistent between the two AU coders. In the subsequent discussion we only focus on the significant correlations presented in Table 2, when both coders achieved a consensus.

We found that confusion was manifested by a lowered brow (AU 4), the tightening of the eye lids (AU 7), and a notable lack of a lip corner puller (AU 12). Figure 1b presents an example of confusion in which AUs 4 and 7 are prominent (see furrowed brow). This pattern replicates D'Mello et al.'s (2006) results when learners verbally expressed their affective states while interacting with AutoTutor. However, one exception is that, in the D'Mello et al. (2006) study, the presence of a lip corner puller was associated with a state of confusion.

The present study revealed that a number of action units that span the entire face can distinguish delight from neutral. In particular, the presence of AU 7 (lid tightener), AU 12 (lip corner puller), AU 25 (lips part), and AU 26 (jaw drop) coupled with an absence of AU 45 (blink) segregate this emotion from neutral. These patterns are generally consistent with a smile, as illustrated in Figure 1(c) where AUs 7 and 12 are activated.

Frustration is a state that is typically associated with significant physiological arousal, yet the facial features we tracked were not very good at distinguishing this emotion from neutral. The only significant correlation with frustration was obtained for AU 12 (lip corner puller) – perhaps indicative of a half smile as evident in Figure 1d. This may be an attempt by the learner to disguise an emotion associated with negative connotations in society.

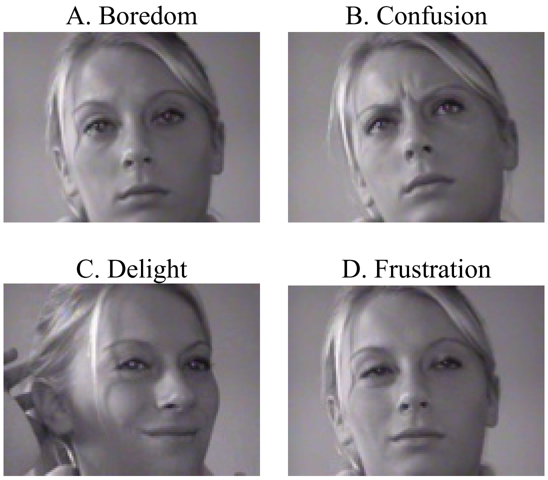
A. Boredom B. Confusion
C. Delight D. Frustration

Figure 1. Examples of Affective States

It appears that boredom is not easily distinguishable from neutral on the basis of the facial features. Indeed, boredom resembles an expressionless face (see Figure 1a). This result replicates an earlier finding by Craig et al. (2004), where no action unit was found to be associated with boredom.

Discriminating between Affective States

A discriminant analysis was performed with the affective state of the learner as the dependent variable and the various facial features as predictors. In this analysis only the AUs that significantly correlated with the affective states were included as predictors (AU 1, AU 4, AU 7, AU 12, AU 25, AU 26, and AU 45). Although four discriminant functions were calculated, the results indicated that only the first two functions were statistically significant ($\chi 2(28) = 200.8$, $p < .001$ for function 1 and $\chi 2(18) = 79.64$, $p < .001$ for

Facial Action Unit		**Affective State** Boredom (*N*=26)		Confusion (*N*=59)		Delight (*N*=43)		Frustration (*N*=47)		Neutral (*N*=37)	
		Judge1	Judge2	Judge1	Judge2	Judge1	Judge2	Judge1	Judge2	Judge1	Judge2
AU1	Inner Brow Raiser				.196						
AU4	Brow Lowerer	-.186		.458	.417		-.191			-.240	-.160
AU7	Lid Tightener	-.247	-.223	.157	.275	.240	.180			-.288	-.270
AU12	Lip Corner Puller	-.260	-.300	-.208	-.150	.522	.456	.167	.161	-.265	-.224
AU25	Lips Part					.342	.337			-.156	-.197
AU26	Jaw Drop					.363	.282			-.164	-.145
AU43	Eye Closure							.195		-.178	
AU45	Blink		.313			-.169	-.180				

Note. All listed correlations statistically significant at the p < .05 level.

Table 2 Statistically Significant Correlations between Action Units and Affective States function 2). These functions were able to account for 98% of the variance, with 63.5% of the variance explained by the first function and the remaining 34.5% of the variance attributed to function 2. Correlations between the predictor variables and the discriminant functions indicated that AU 12 (lip corner puller) was the best predictor for function 1 while AU 1 (inner brow raiser), AU4 (brow lowerer), and AU 7 (lid tightener) were the best predictors for function 2. This implies that function 1 utilizes information from the lower face (i.e. the mouth) to discriminate between the affective states. The second function relies on information from the upper face (i.e. the brow and eyelids) to disambiguate the affective states.

The discriminant function was also able to successfully predict the affective states with an accuracy of 49.1%, a significant improvement over the base rate of 0.21 (kappa = .35). This is a reasonable level of accuracy in inferring complex mental states. We also computed kappa scores for individually detecting each of the affective states. The results indicated that the discriminant function was most successful in detecting delight (kappa = .65) followed by confusion (kappa = .41). This was expected since both these states are typically accompanied by animated facial expressions. The reliability of detecting the more subtle affective states of boredom and neutral was lower than delight and confusion (kappa = 0.12 for both states). Our results also indicated that the discriminant analysis was unable to distinguish frustration from the other emotions. In fact, the kappa score for this emotion reflected an accuracy equal to random guessing (kappa = -0.07).

Discussion

This study examined the facial features that accompany the affective states that routinely occur during learning complex topics with AutoTutor. We have discovered several important patterns in the manner in which learners convey these emotions though their faces. The highly animated affective states of confusion and delight are easily detectable from facial expressions. It is tempting to speculate, from an evolutionary perspective, that learners use their face as a social cue to indicate that they are confused, which potentially recruits resources to alleviate their perplexity. Delight is also readily expressed on the face, perhaps because it is a positive emotion. However, it appears that learners do not readily display frustration, perhaps due to the negative connotations associated with this emotion. This finding is consistent with Ekman's theory of social display rules, in which social pressures may result in the disguising of negative emotions such as frustration.

The associations between the various facial features and the affective states of confusion and boredom generally replicate earlier findings from the emote-aloud study. For example, the raising of the inner brow coupled with the tightening of the lids appears to be the prototypical expression of confusion. However, for boredom, in neither study could any particular subset of action units be associated with this emotion. This suggests that the tracking of this emotion may have to use additional indicators, acoustic-prosodic features of speech and posture patterns. We have also had some success in separating boredom from neutral on the basis of dialogue features in the log file of interactions with AutoTutor (D'Mello & Graesser, 2006).

In the earlier study that utilized emote-aloud protocols (Craig et al., 2004), it was reported that frustration was associated with a raised inner and outer brow (AUs 1 and 2) and a dimpler (AU 14). However, these patterns were not replicated in the current study. This suggests that there might be occasional differences between our current offline methodology and our previous emote-aloud methodology, which was an on-line measure. A smaller sample of participants (N = 7) were run in the emote-aloud study whereas 28 participants were run in the current study.

The broader goal of this research is to transform AutoTutor into an affect-sensitive intelligent tutoring system (D'Mello et al., 2005). This endeavor involves the development of automatic affect detection systems and a reengineering of AutoTutor's pedagogical strategies to incorporate the learner's emotions. This research directly contributes to that effort by identifying the facial features that accompany certain emotions. There are existing systems that can automatically code AUs with reasonable accuracy (Cohn & Kanade, in press). We are currently investigating the possibility of integrating these computational systems to aid in inferring the affective states of the learner. We are also considering the use of additional sensors that track posture patterns, speech contours, and dialogue information. We hope that the combination of these methods will yield reliable estimates of the affective states of the learner.

Acknowledgments

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

Ang, J., Dhillon, R., Krupski, A., Shriberg, E., & Stolcke, A. (2002). Prosody-based automatic detection of annoyance and frustration in human-computer dialog. In J. H. L. Hansenet al (Eds.), Proceedings of the International Conference on Spoken Language Processing (ICSLP'02) (pp. 2037-2039). Denver.

Cohn, J. F. & Kanade, T. (In press). Use of automated facial image analysis for measurement of emotion expression. In J. A. Coan & J. B. Allen (Eds.), The handbook of emotion elicitation and assessment. Oxford University Press Series in Affective Science. New York: Oxford.

Craig S., D'Mello S., Gholson B., Witherspoon A., Sullins J., and Graesser A, (2004). Emotions during learning: The first steps toward an affect sensitive intelligent tutoring system. E-learn (in press). Association for the Advancement of Computing in Education, Norfolk, VA.

Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience. Harper-Row: NY.

D'Mello, S. K., Craig, S. D., Gholson, B.; Franklin, S., Picard, R., & Graesser, A. C. (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.

D'Mello, S. K., Craig, S. D., Sullins, J., & Graesser, A. C. (2006). Predicting affective states through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. International Journal of Artificial Intelligence in Education, 16, 3-28.

D'Mello, S., & Graesser, A.C. (2006). Affect detection from human-computer dialogue with an intelligent tutoring system. In J. Gratch, M. Young, R. Aylett, D. Ballin, and P. Oliver (Eds.), Lecture Notes in Computer Science: Intelligent Virtual Agents: 6th International Conference (pp. 54-67). Berlin, Heidelberg, Germany: Springer.

Elfenbein, H. A. & Ambady, N. (2002) On the universality and cultural specificity of emotion recognition: a metaanalysis, Psychological Bulletin, 128, 203–235

Ekman, P. 2003. Emotions Revealed. New York: Henry Holt and Company.

Ekman, P. & Friesen, W.V. (1969). Nonverbal leakage and clues to deception. Psychiatry, 32, 88-105.

Ekman, P, & W. V. Friesen. (1978). The facial action coding system: A technique for the measurement of facial movement. Palo Alto: Consulting Psychologists Press

Ekman, P., & Friesen, W. (1972). Constants across culture in the face and emotion. Journal of Personality and Social Psychology, 17, 124–129. Ericsson, K. A. & Simon, H. A. (1993). Protocol analysis: Verbal reports as data. Revised edition. Cambridge, MA: The MIT Press.

Forbes-Riley, K. & 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, & A. Olney. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. IEEE Transactions in Education, 48, 612-618.

Graesser, A.C., Jackson, G.T., & McDaniel, B. (2007). AutoTutor holds conversations with learners that are responsive to their cognitive and emotional states. Educational Technology, 47, 19-22.

Graesser, A.C., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. (2006). Detection of emotions during learning with AutoTutor. In R. Son (Ed.), Proceedings of the 28th Annual Meetings of the Cognitive Science Society (pp. 285-290). Mahwah, NJ: Erlbaum.

Graesser, A.C., & Olde, B.A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. Journal of Educational Psychology, 95, 524-536

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & the TRG (1999). AutoTutor: A simulation of a human tutor. Journal of Cognitive Systems Research, 1, 35-51.

Graesser, A.C., Witherspoon, A., McDaniel, B., D'Mello, S., Chipman, P., Gholson, B. (2006). Detection of emotions during learning with AutoTutor. In R. Son (Ed.), Proceedings of the 28th Annual Meetings of the Cognitive Science Society. (pp. 285-290). Mahwah, NJ: Erlbaum.

Harter, S. (1981). A model of mastery motivation in children: Individual differences and developmental change. In W. A. Collins (Ed.). The Minnesota symposium on child psychology. Aspects of the development of competence, 14, (pp. 215-255). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kapoor, A., Mota, S. and Picard, R. (2001). Towards a learning companion that recognizes affect. AAAI Fall Symposium 2001, North Falmouth MA, November 2001.

Kort, B., Reilly, R., & Picard, R. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy—building a learning companion. In T. Okamoto, R. Hartley, Kinshuk, & J. P. Klus (Eds.), Proceedings IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges (pp. 43-48). IEEE Computer Society.

Meyer, D. K., & Turner, J. C. (2002). Discovering Emotion in Classroom Motivation Research. Educational Psychologist, 37, 107-114. Miserandino, M. (1996) Children who do well in school: Individual differences in perceived competence and autonomy in above-average children, Journal of Educational Psychology, 88, 203–214.

Patrick B., Skinner, E. & Connell, J. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain, Journal of Personality and Social Psychology, 65, 781–791.

Picard, R. W. (1997). Affective computing. Cambridge, MA: MIT Press. Rozin, P. & Cohen, A. B. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of maturally occurring facial expressions of Americans, Emotion, 3, 68–75.

Snow, R. E., Corno, L., & Jackson, D. (1996). Individual differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (Eds.) Handbook of educational psychology (pp. 243-310). New York: Macmillan.

Stipek, D. (1998) Motivation to learn: from theory to practice (3rd ed.) Boston, MA: Allyn and Bacon.