Chapter 10: Emotional Learning Analytics

Sidney K. D'Mello

Departments of Psychology and Computer Science & Engineering, University of Notre Dame, USA
DOI: 10.18608/hla17.010

ABSTRACT

This chapter discusses the ubiquity and importance of emotion to learning. It argues that substantial progress can be made by coupling the discovery-oriented, data-driven, analytic methods of learning analytics (LA) and educational data mining (EDM) with theoretical advances and methodologies from the affective and learning sciences. Core, emerging, and future themes of research at the intersection of these areas are discussed.

Keywords: Affect, affective science, affective computing, educational data mining

At the "recommendation" of a reviewer of one of my papers (D'Mello, 2016), I recently sought to learn a new (for me) statistical method called generalized additive mixed models (GAMMs; McKeown & Sneddon, 2014). GAMMs aim to model a response variable with an additive combination of parametric and nonparametric smooth functions of predictor variables, while addressing autocorrelations among residuals (for time series data). At first, I was mildly *displeased* at the thought of having to do more work on this paper. *Anxiety* resulted from the thought that I might not have the time to learn and implement a new method to meet the revision deadline. So I did nothing. The *anxiety* transitioned into mild *panic* as the deadline approached. I finally decided to look into GAMMs by downloading a recommended paper. The paper had some eye-catching graphics, which piqued my *curiosity* and motivated me to explore further. The *curiosity* quickly turned into *interest* as I read more about the method, and eventually into *excitement* when I realized the power of the approach. This motivated me to slog through the technical details, which led to some intense emotions - *confusion* and *frustration* when things did not make sense, *despair* when I almost gave up, *hope* when I thought I was making progress, and eventually, *delight* and *happiness* when I actually did make progress. I then attempted to apply the method to my data by modifying some R-syntax. More *confusion*, *frustration*, and *despair* interspersed with *hope*, *delight*, and *happiness*. I eventually got it all working and wrote up the results. Some more emotions occurred during the writing and revision cycles. Finally, I was done. I felt *contentment*, *relief*, and a bit of *pride*.

As this example illustrates, there is an undercurrent of emotion throughout the learning process. This is not unique to learning as all "cognition" is tinged with "emotion". The emotions may not always be consciously experienced (Ohman & Soares, 1994), but they exist and influence cognition nonetheless. Also, emotions do not occur in a vacuum; they are deeply intertwined within the social fabric of learning. It does not take much to imagine the range of emotions experienced by the typical student whose principle occupation is learning. Pekrun and Stephens (2011) call these "academic emotions" and group them into four categories. *Achievement* emotions (contentment, anxiety, and frustration) are linked to learning activities (homework, taking a test) and outcomes (success, failure). *Topic* emotions are aligned with the learning content (empathy for a protagonist while reading classic literature). *Social* emotions such as pride, shame, and jealousy occur because education is situated in social contexts. Finally, *epistemic* emotions arise from cognitive processing, such as surprise when novelty is encountered or confusion in the face of an impasse.

Emotions are not merely incidental; they are functional or they would not have evolved (Darwin, 1872; Tracy, 2014). Emotions perform *signalling functions* (Schwarz, 2012) by highlighting problems with knowledge (confusion), problems with stimulation (boredom), concerns with impending performance (anxiety), and challenges that cannot be easily surpassed (frustration). They perform *evaluative functions* by serving as the currency by which people appraise an event in terms of its value, goal relevance, and goal congruence (Izard, 2010). Emotions perform *modulation functions* by constraining or expanding cognitive focus with negative emotions engendering narrow, bottom-up, and focused modes of processing (constrained focus) (Barth & Funke, 2010; Schwarz, 2012) in comparison to positive emotions, which facilitate broader, top-down, generative processing (expanded focus) (Fredrickson & Branigan, 2005; Isen, 2008). Indeed, emotions pervade thought as is evident by their effects on memory, problem solving, decision making, and other facets of cognition (see Clore & Huntsinger, 2007, for a review).

So what exactly is an "emotion"? Truth be told, we do not really know or at least we do not fully agree (Izard, 2010). This can be readily inferred from the most recent debates on the psychological underpinnings of emotion - a debate sometimes referred to as the "100 year old emotion war" (Lench, Bench, & Flores, 2013; Lindquist, Siegel, Quigley, & Barrett, 2013). Fortunately, there is general agreement on the following key points. Emotions are conceptual entities that arise from brain–body–environment interactions. But you will not find them by looking in the brain, the body, or the environment. Instead, emotions *emerge* (Lewis, 2005) when organism–environment interactions trigger changes across multiple time scales and at multiple levels - neurobiological, physiological, behaviourally expressive, action-oriented, and cognitive/metacognitive/subjective. The "emotion" is reflected in these changes in a manner modulated by the ongoing situational context. The same emotional category (e.g., anxiety) will manifest differently based on a triggering event (Tracy, 2014), the specific biological/cognitive/metacognitive processes involved (Gross, 2008; Moors, 2014), and sociocultural influences (Mesquita & Boiger, 2014; Parkinson, Fischer, & Manstead, 2004). For example, an anxiety-inducing event will trigger distinct "episodes" of anxiety depending on the specific circumstance (public speaking, test taking), the temporal context (one day vs. one minute before the speech), the neurobiological system (baseline arousal), and the social context (speaking in front of colleagues vs. strangers). This level of variability and ambiguity is expected because humans and their emotions are dynamic and adaptive. Rigid emotions have little evolutionary value.

Where do learning analytics (LA) and educational data mining (EDM) fit in? On one hand, given the central role of emotions in learning, attempts to analyze (or data mine) learning without considering emotion will be incomplete. On the other hand, giving the ambiguity and complexity of emotional phenomena, attempts to study emotions during learning without the methods of LA and EDM will only yield shallow insights. Fortunately, there is a body of work that adopts a data-driven analytic approach to study the incidence and influence of emotions on the processes and products of learning. In this chapter, I highlight some of the core, emerging, and future themes in this interdisciplinary research area.

Let us begin with a note on terminology. Emotion is related, but not equivalent to motivation, attitudes, preferences, physiology, arousal, and a host of other constructs often used to refer to it. Emotions are also distinct from moods and affective traits (Rosenberg, 1998). Emotion is not the same as a feeling. Hunger is a feeling but is not an emotion. Neither is pain. There is also some contention as to what constitutes an emotion. Anger is certainly an emotion, but what about confusion? Confusion has affective components (feelings of being confused, characteristic facial expressions; D'Mello & Graesser, 2014b), but there is some debate as to whether it is an emotion (Hess, 2003; Rozin & Cohen, 2003). Thus, in the remainder of this chapter, I use the more inclusive term "affective state" rather than the more restrictive term "emotion".

CORE THEMES

I selected the following four themes to highlight the use of LA/EDM methods to study affect during learning. I also review one or two exemplary studies within each theme in some depth rather than cursorily reviewing many studies. This means that many excellent studies go unmentioned, but I leave it to the reader to explore the body of work within each theme. I recommend review papers, when available, to facilitate this process.

Affect Analysis from Click-Stream Data

One of the most basic uses of LA/EDM techniques is to use the rich stream of data generated during interactions with learning technologies in order to understand learners' cognitive processes (Corbett & Anderson, 1995; Sinha, Jermann, Li, & Dillenbourg, 2014). A complementary set of insights can be gleaned when affect is included in the mix, as illustrated in the study below.

Bosch and D'Mello (in press) conducted a lab study on the affective experience of students during their first programming session. Novice students (N = 99) were asked to learn the fundamentals of computer programming in the Python language using a self-paced computerized learning environment involving a 25-minute scaffolded learning phase and a 10-minute non-scaffolded fadeout phase. All instructional activities (coding, reading text, testing code, receiving errors, etc.) were logged and videos of students' faces and computer screens were recorded. Students provided affect judgments at approximately 100 points (every 15 seconds) over the course of viewing these videos immediately after the learning session via a retrospective affect judgment protocol (Porayska-Pomsta, Mavrikis, D'Mello, Conati, & Baker, 2013). The affective states of interest were anger, anxiety, boredom, confusion, curiosity, disgust, fear, frustration, flow/engagement, happiness, sadness, and surprise. Only engagement, confusion, frustration, boredom, and curiosity occurred with sufficient frequency to warrant further analysis.

The authors examined how interaction events give rise to affective states, and how affective states trigger various behaviours. They constructed time series that interspersed interaction events (from clickstream data) and affective states (self-reports) for each student during the scaffolded learning phase. Time series modelling techniques (D'Mello, Taylor, & Graesser, 2007) were used to identify significant transitions between affective states and interaction events. The resultant model is shown as a directed graph in Figure 10.1. There were some transitions between interaction events that did not include an affective state (dashed lines). This was due to the infrequency of affect sampling (every 15 seconds) relative to other interaction events (as frequent as 1 second).

The more interesting transitions include affective states. In particular, confusion and frustration were both preceded by an incorrect solution submission (SubmitError); these affective states were then followed by a hint request (ShowHint), or by constructing code (Coding), which itself triggered confusion and frustration. Reading instructional texts (including problem descriptions) was a precursor of engagement, curiosity, boredom, and confusion but not frustration. In other words, all the key affective states were related to knowledge assimilation (reading) and construction (coding) activities. However, only confusion and frustration accompanied failure (Submit Error) and subsequent help-seeking behaviours (ShowHint), which are presumably learning opportunities. Taken together, the transition model emphasizes the key role of impasses and the resultant negative activating states of confusion and frustration to learning (D'Mello & Graesser, 2012b; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003). It also illustrates how affect is interspersed throughout the learning process.

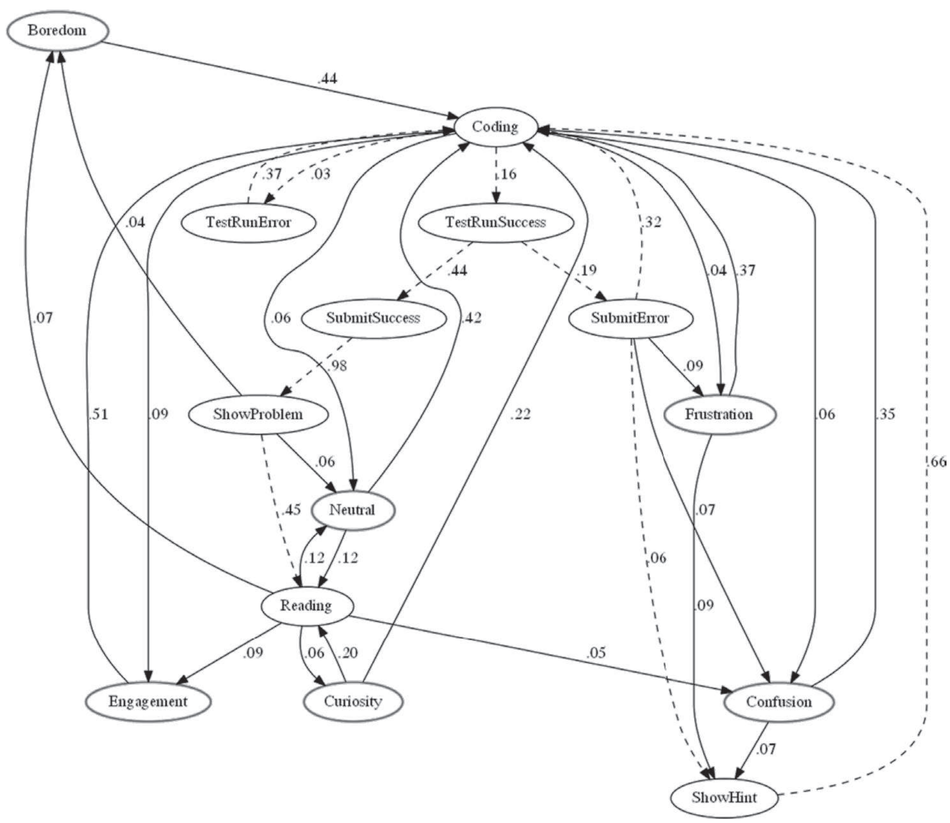

Figure 10.1. Significant transitions between affective states and interaction events during scaffolded learning of computer programming. Solid lines indicate transitions including affect. Dashed lines indicate transitions not involving affective states. ShowProblem: starting a new exercise; Reading: viewing the instructions and/or problem statement; Coding: editing or viewing the current code; ShowHint: viewing a hint; TestRunError: code was run and encountered a syntax or runtime error; TestRunSuccess: code run without syntax or runtime errors (but was not checked for correctness); SubmitError: code submitted and produced an error or incorrect answer; SubmitSuccess: code submitted and was correct.

Affect Detection from Interaction Patterns

Affective states cannot be directly measured because they are conceptual entities (constructs). However, they emerge from environment–person interactions (context) and influence action by modulating cognition. Therefore, it should be possible to "infer" affect by analyzing the unfolding context and learner actions. This line of work, referred to as "interaction-based", "log-file based", or "sensor-free" affect detection was started more than a decade ago (Ai et al., 2006; D'Mello, Craig, Sullins, & Graesser, 2006) and was recently reviewed by Baker and Ocumpaugh (2015).

As an example, consider Pardos, Baker, San Pedro, and Gowda (2013), who developed affect detectors for ASSISTments, an intelligent tutoring system (ITS) for middle- and high-school mathematics, used by approximately 50,000 students in the US as part of their regular mathematics instruction (Razzaq et al., 2005). The authors adopted a supervised learning approach to build automated affect detectors. They collected training data from 229 students while they used ASSISTments in school computer labs. Human observers provided online observations (annotations) of affect as students interacted with ASSISTments using the Baker-Rodrigo Observation Method Protocol (BROMP) (Ocumpaugh, Baker, & Rodrigo, 2012). According to this protocol, trained observers provide live annotations of affect based on observable behaviour, including explicit actions towards the interface, interactions with peers and teachers, body movements, gestures, and facial expressions. The observers coded four affective states (boredom, frustration, engaged concentration, and confusion) and two behaviours (going off-task and gaming the system). Supervised learning techniques were used to discriminate each affective state from other states (e.g., bored vs. others) using features extracted from ASSISTments log files (performance on problems, hint requests, response times, etc.). Affect detection accuracies ranged from .632 to .678 (measured with the A-prime metric [similar to area under the receiver operating characteristic curve – AUC or AUROC]) for affect and .802 to .819 for behaviours. The classifier was validated in a manner that ensured generalizability to new students from the same population by enforcing strict independence among training and testing data.

Pardos et al. (2013) also provided preliminary evidence on the predictive validity of their detectors. This was done by applying the detectors on log files from a different set of 1,393 students who interacted with ASSISTments during the 2004–2006 school years - several years before the measure was developed. Automatically measured affect and behaviour moderately correlated with standardized test scores. Further, San Pedro, Baker, Bowers, and Heffernan (2013) attempted to predict college enrollment based on the automatic detectors. They applied the detectors to existing log files from 3,707 students who interacted with ASSISTments from 2004 to 2009. College enrollment information for these students was obtained from the National Student Clearinghouse. Automatically measured affective states were significant predictors of college enrollment several years later, which is a rather impressive finding.

Affect Detection from Bodily Signals

Affect is an embodied phenomenon in that it activates bodily response systems for action. This should make it possible to infer learner affect (a latent variable) from machine-readable bodily signals (observables). There is a rich body of work on the use of bodily signals to detect affect as discussed in a number of reviews (Calvo & D'Mello, 2010; D'Mello & Kory, 2015; Zeng, Pantic, Roisman, & Huang, 2009). The research has historically focused on interactions in controlled environments, but researchers have begun to take this work into the real world, notably computer-enabled classrooms. The study reviewed below reflects one such effort by our research group and collaborators, but the reader is directed to Arroyo et al. (2009) for their pioneering work on affect detection in computer-enabled classrooms.

Bosch, D'Mello, Baker, Ocumpaugh, and Shute (2016) studied automated detection of affect from facial features in a noisy real-world setting of a computer-enabled classroom. In this study, 137 middle and high school students played a conceptual physics educational game called Physics Playground (Shute, Ventura, & Kim, 2013) in small groups for 1.5 to 2 hours across two days as part of their regular physics/physical science classes. Trained observers performed live annotations of boredom, confusion, frustration, engaged-concentration, and delight using the BROMP field observation protocol as in the ASSISTments study discussed above (Pardos et al., 2013). The observers also noted when students went off-task.

Videos of students' faces and upper bodies were recorded during game-play and synchronized with the affect annotations. The videos were processed using the FACET computer-vision program (Emotient, 2014), which provides estimates of the likelihood of 19 facial action units (Ekman & Friesen, 1978) (e.g., raised brow, tightened lips), head pose (orientation), and position (see Figure 10.2 for screenshot). Body movement was also estimated from the videos using motion filtering algorithms (Kory, D'Mello, & Olney, 2015) (see Figure 10.3). Supervised learning methods were used to build detectors of each affective state (e.g., bored vs. other states) using both facial expressions and bodily move-

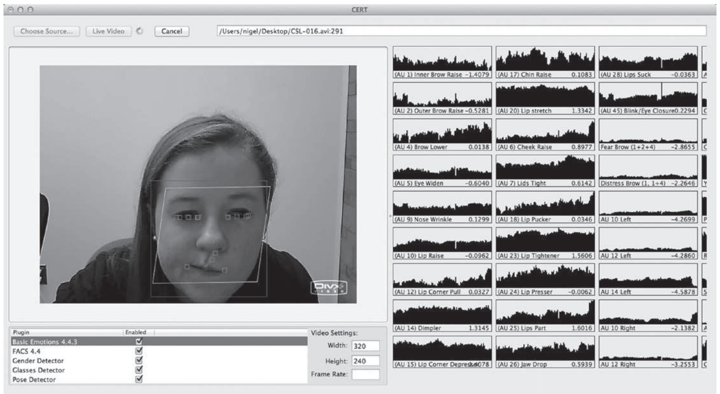

Figure 10.2. Automatic tracking of facial features using the Computer Expression Recognition Toolbox. The graphs on the right show likelihoods of activation of various facial features (e.g., brow lowered, eyelids tightening).

Figure 10.3. Automatic tracking of body movement from video using motion silhouettes. The image on the right displays the areas of movement from the video playing on the left. The graph on the bottom shows the amount of movement over time.

ments. The detectors were moderately successful with accuracies (quantified with the AUC metric as noted above) ranging from .610 to .867 for affect and .816 for off-task behaviours. Follow-up analyses confirmed that the affect detectors generalized across students, multiple days, class periods, and across different genders and ethnicities (as perceived by humans).

One limitation of the face-based affect detectors is that they are only applicable when the face can be automatically detected in the video stream. This is not always the case due to excessive movement, occlusion, poor lighting, and other factors. In fact, the face-based affect detectors were only applicable to 65% of the cases. To address this, Bosch, Chen, Baker, Shute, and D'Mello (2015) used multimodal fusion techniques to combine interaction-based (similar to previous section) and face-based detection. The interaction-based detectors were less accurate than the face-based detectors (Kai et al., 2015), but were applicable to almost all of the cases. By combining the two, the applicability of detectors increased to 98% of the cases, with only a small reduction (<5% difference) in accuracy compared to face-based detection.

Integrating Affect Models in Affect-Aware Learning Technologies

The interaction- and bodily-based affect detectors discussed above are tangible artifacts that can be instrumented to provide real-time assessments of student affect during interactions with a learning technology. This affords the exciting possibility of closing the loop by dynamically responding to the sensed affect. The aim of such affect-aware learning technologies is to expand the bandwidth of adaptivity of current learning technologies by responding

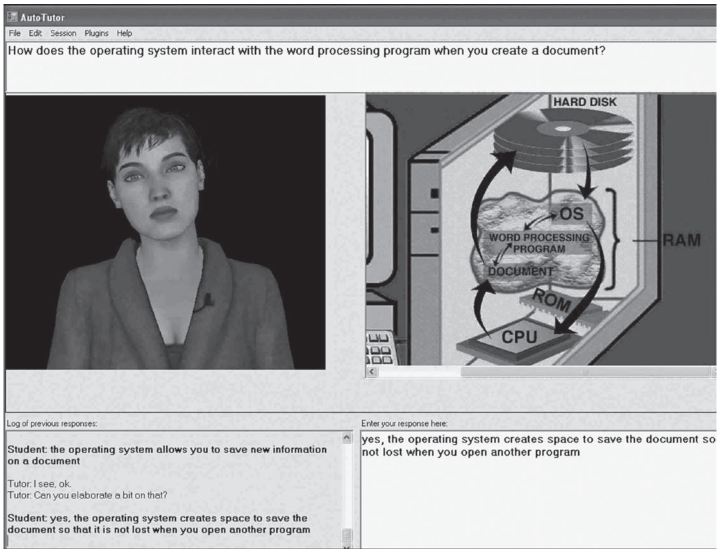

Figure 10.4. Affective AutoTutor: an intelligent tutoring system (ITS) with conversational dialogs that automatically detects and responds to learners' boredom, confusion, and frustration.

to what students feel in addition to what they think and do (see D'Mello, Blanchard, Baker, Ocumpaugh, & Brawner, 2014, for a review). Here, I highlight two such systems, the Affective AutoTutor (D'Mello & Graesser, 2012a) and UNC-ITSPOKE (Forbes-Riley & Litman, 2011).

Affective AutoTutor (see Figure 10.4) is a modified version of AutoTutor - a conversational ITS that helps students develop mastery on difficult topics in Newtonian physics, computer literacy, and scientific reasoning by holding a mixed-initiative dialog in natural language (Graesser, Chipman, Haynes, & Olney, 2005). The original AutoTutor system has a set of fuzzy production rules that are sensitive to the cognitive states of the learner. The Affective AutoTutor augments these rules to be sensitive to dynamic assessments of learners' affective states, specifically boredom, confusion, and frustration. The affective states are sensed by automatically monitoring interaction patterns, gross body movements, and facial features (D'Mello & Graesser, 2012a). The Affective AutoTutor responds with empathetic, encouraging, and motivational dialog-moves along with emotional displays. For example, the tutor might respond to mild boredom with, "This stuff can be kind of dull sometimes, so I'm gonna try and help you get through it. Let's go". The affective responses are accompanied by appropriate emotional facial expressions and emotionally modulated speech (e.g., synthesized empathy or encouragement).

The effectiveness of Affective AutoTutor over the original non-affective AutoTutor was tested in a between-subjects experiment where 84 learners were randomly assigned to two 30-minute learning sessions with either tutor (D'Mello, Lehman, Sullins et al., 2010). The results indicated that the affective tutor helped learning for low-domain knowledge learners during the second 30-minute learning session. The affective tutor was less effective at promoting learning for high-domain knowledge learners during the first 30-minute session. Importantly, learning gains increased from Session 1 to Session 2 with the affective tutor whereas they plateaued with the non-affective tutor. Learners who interacted with the affective tutor also demonstrated improved performance on a subsequent transfer test. A follow-up analysis indicated that learners' perceptions of how closely the computer tutors resembled human tutors increased across learning sessions, was related to the quality of tutor feedback, and was a powerful predictor of learning (D'Mello & Graesser, 2012c). The positive change in perceptions was greater for the affective tutor.

As a second example, consider UNC-ITSPOKE (Forbes-Riley & Litman, 2011), a speech-enabled ITS for physics with the capability to automatically detect and respond to learners' certainty/uncertainty in addition to the correctness/incorrectness of their spoken responses. Uncertainty detection was performed by extracting and analyzing the acoustic-prosodic features in learners' spoken responses along with lexical and dialog-based features. UNC-ITSPOKE responded to uncertainty when the learner was correct but uncertain about the response. This was taken to signal an impasse because the learner is unsure about the state of their knowledge, despite being correct. The actual response strategy involved launching explanation-based sub-dialogs that provided added instruction to remediate the uncertainty. This could involve additional follow-up questions (for more difficult content) or simply the assertion of the correct information with elaborated explanations (for easier content).

Forbes-Riley and Litman (2011) compared learning outcomes of 72 learners who were randomly assigned to receive adaptive responses to uncertainty (adaptive condition), no responses to uncertainty (non-adaptive control condition), or random responses to uncertainty (random control condition). In this later condition, the added tutorial content from the sub-dialogs was given for a random set of turns in order to control for the additional tutoring. The results indicated that the adaptive condition achieved slightly (but not significantly) higher learning outcomes than the random and non-adaptive control conditions. The findings revealed that it was perhaps not the presence or absence of adaptive responses to uncertainty, but the number of adaptive responses that correlated with learning outcomes.

EMERGING THEMES

Research at the intersection of emotions, learning, LA, and EDM, has typically focused on one-on-one learning with intelligent tutoring systems (Forbes-Riley & Litman, 2011; Woolf et al., 2009), educational games (Conati & Maclaren, 2009; Sabourin, Mott, & Lester, 2011), or interfaces that support basic competencies like reading, writing, text-diagram integration, and problem solving (D'Mello & Graesser, 2014a; D'Mello, Lehman, & Person, 2010; D'Mello & Mills, 2014). Although these basic lines of research are quite active, recent work has focused on analyzing affect across more expansive interaction contexts that more closely capture the broader sociocultural context surrounding learning. I briefly describe four themes of research to illustrate a few of the exciting developments.

Affect-Based Predictors of Attrition and Dropout

Early indicators of risk and early intervention systems are some of the "killer apps" of LA and EDM (Jayaprakash, Moody, Lauría, Regan, & Baron, 2014). Most fielded systems focus on academic performance data, demographics, and availability of financial assistance. These factors are undoubtedly important, but there are likely alternate factors that come into play. With this in mind, Aguiar, Ambrose, Chawla, Goodrich, and Brockman (2014) compared the predictive power of traditional academic and demographic features with features indicative of behavioural engagement in predicting dropout from an Introduction to Engineering Course. Their key finding was that behaviourally engaging with e-portfolios, measured by number of logins, number of artifacts submitted, and number of page hits, was a better predictor of dropout than models constructed from academic performance and demographics alone. Although affect was not directly measured in this study, behaviourally engaging with e-portfolios can be considered a sign of interest, which is a powerful motivating emotion.

Sentiment Analysis of Discussion Forums

Language communicates feelings. Hence, sentiment analysis and opinion mining techniques (Pang & Lee, 2008) have considerable potential to study how students' thoughts (expressed in written language) about a learning experience predicts relevant behaviours (most importantly attrition). In line with this, Wen, Yang, and Rosé (2014) applied sentiment analysis techniques on student posts on three Massive Open Online Courses (MOOCs). They observed a negative correlation between the ratio of positive to negative terms and dropout across time. More recently, Yang, Wen, Howley, Kraut, and Rosé (2015) developed methods to automatically identify discussion posts that were indicative of student confusion. They showed that confusion reduced the likelihood of retention, but this could be mitigated with confusion resolution and other supportive interventions.

Classroom Learning Analytics

Recent advances in sensing and signal processing technologies have made it possible to automatically model aspects of students' classroom experience that could previously only be obtained from self-reports and cumbersome human observations. For example, second-generation Kinects can detect whether the eyes or mouth are open, if a person is looking away, and if the mouth has moved, for up to six people at a time (Microsoft, 2015). In one pioneering study, Raca, Kidzinski, and Dillenbourg (2015) tracked students in a classroom using multiple cameras affixed around the blackboard area. Computer vision techniques were used for head detection and head-pose estimation, which were then used to train a detector of student attention (validated via self-reports). This emerging area, related to the field of multimodal learning analytics (Blikstein, 2013), is poised for considerable progress in years to come.

Teacher Analytics

Teachers should not be left out of the loop since teacher practices are known to influence student affect and engagement. Unfortunately, quantifying teacher instructional practices relies on live observations in classrooms (e.g., Nystrand, 1997), which makes the research difficult to scale. To address this, researchers have begun to develop methods for automatic analysis of teacher instructional practices. In a pioneering study, Wang, Miller, and Cortina (2013) recorded classroom audio in 1st to 3rd grade math classes and developed automatic methods to predict the level of discussions in these classes. This work was recently expanded to analyze several additional instructional activities (lecturing, small group work, supervised seatwork, question/answer, and procedures and directions) in larger samples of middle-school literature and language-arts classes using teacher audio alone (Donnelly et al., 2016a) or a combination of teacher and classroom audio (Donnelly et al., 2016b). Blanchard et al. (2016) used teacher audio to automatically detect teacher questions, achieving a .85 correlation with the proportion of human-coded questions. The next step in this line of work is to use information on what teachers are doing to contextualize how students are feeling, which in turn influences what they think, do, and learn.

FUTURE THEMES

Let me end by briefly highlighting some potential future themes of research. One promising area of research involves a detailed analysis of the emotional experience of learners and communities of learners across the extended time scale of a traditional course, a flipped-course, or a MOOC (Dillon et al., 2016). A second involves the study of emotion regulation during learning, especially how LA/EDM methods can be used to identify different regulatory strategies (Gross, 2008), so that the more beneficial ones can be engendered (e.g., Strain & D'Mello, 2014). A third would jointly consider emotion alongside attentional states of mindfulness, mind wandering, and how emotion-attention blends like the "flow experience" (Csikszentmihalyi, 1990) emerge and manifest in the body and in behaviour. A fourth addresses how the so-called "non-cognitive" (Farrington et al., 2012) traits like grit, self-control, and diligence modulate learner emotions and efforts to regulate them (e.g., Galla et al., 2014). A fifth would monitor emotions of groups of learners during collaborative learning and collaborative problem solving (Ringeval, Sonderegger, Sauer, & Lalanne, 2013) given the importance of collaboration as a critical 21st century skill (OECD, 2015).

Finally, quoting William James's classic 1884 treatise on emotion: "The most important part of my environment is my fellow-man. The consciousness of his attitude towards me is the perception that normally unlocks most of my shames and indignations and fears" (p. 195). Research to date has mainly focused on the achievement, epistemic, and topic emotions. However, an analysis of learning in the sociocultural context in which it is situated must adequately address the social emotions of pride, shame, guilt, jealousy, envy, and so on. This is both a future theme and a grand research challenge.

CONCLUSION

Learning is not a cold intellectual activity; it is punctuated with emotion. The emotions are not merely decorative, they have agency. But emotion is a complex phenomenon with multiple components that dynamically unfold across multiple time scales. And despite great strides in the fields of affective sciences and affective neuroscience, we know little about emotions, and even less on emotions during learning. This is certainly not to imply that we should refrain from modelling emotion until there is more theoretical clarity. Quite the opposite. It simply means that we need to be mindful of what we are modelling when we say we are modelling emotion. We also need to embrace, rather than dilute, the complexity and ambiguity inherent in emotion. If anything, the discovery-oriented, data-driven, analytic methods of LA and EDM, along with an emphasis on real-world data collection, has the unique potential to advance both the science of learning and the science of emotion. It all begins by incorporating emotion into the analysis of learning.

ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation (NSF) (DRL 1108845 and IIS 1523091), the Bill & Melinda Gates Foundation, and the Institute for Educational Sciences (R305A130030). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the funding agencies.

REFERENCES

Aguiar, E., Ambrose, G. A. A., Chawla, N. V., Goodrich, V., & Brockman, J. (2014). Engagement vs. performance: Using electronic portfolios to predict first semester engineering student persistence. *Journal of Learning Analytics*, *1*(3), 7–33.

Ai, H., Litman, D. J., Forbes-Riley, K., Rotaru, M., Tetreault, J. R., & Pur, A. (2006). Using system and user performance features to improve emotion detection in spoken tutoring dialogs. *Proceedings of the 9th International Conference on Spoken Language Processing* (Interspeech 2006).

Arroyo, I., Woolf, B., Cooper, D., Burleson, W., Muldner, K., & Christopherson, R. (2009). Emotion sensors go to school. In V. Dimitrova, R. Mizoguchi, B. Du Boulay, & A. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 17–24). Amsterdam: IOS Press.

Baker, R., & Ocumpaugh, J. (2015). Interaction-based affect detection in educational software. In R. Calvo, S. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford handbook of affective computing* (pp. 233–245). New York: Oxford University Press.

Barth, C. M., & Funke, J. (2010). Negative affective environments improve complex solving performance. *Cognition and Emotion*, *24*(7), 1259–1268. doi:10.1080/02699930903223766

Blanchard, N., Donnelly, P., Olney, A. M., Samei, B., Ward, B., Sun, X., . . . D'Mello, S. K. (2016). Identifying teacher questions using automatic speech recognition in live classrooms. *Proceedings of the 17th Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL 2016) (pp. 191–201). Association for Computational Linguistics.

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (LAK '13), 8–12 April 2013, Leuven, Belgium. New York: ACM.

Bosch, N., Chen, H., Baker, R., Shute, V., & D'Mello, S. K. (2015). Accuracy vs. availability heuristic in multimodal affect detection in the wild. *Proceedings of the 17th ACM International Conference on Multimodal Interaction* (ICMI 2015). New York: ACM.

Bosch, N., & D'Mello, S. K. (in press). The affective experience of novice computer programmers. *International Journal of Artificial Intelligence in Education.*

Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., & Shute, V. (2016). Using video to automatically detect learner affect in computer-enabled classrooms. *ACM Transactions on Interactive Intelligent Systems*, 6(2), doi:10.1145/2946837.

Calvo, R. A., & D'Mello, S. K. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. doi:10.1109/T-AFFC.2010.1

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393–399. doi:10.1016/j.tics.2007.08.005

Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, *19*(3), 267–303.

Corbett, A., & Anderson, J. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 4(4), 253–278.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

D'Mello, S. K. (2016). On the influence of an iterative affect annotation approach on inter-observer and self-observer reliability. *IEEE Transactions on Affective Computing*, *7*(2), 136–149.

D'Mello, S. K., Blanchard, N., Baker, R., Ocumpaugh, J., & Brawner, K. (2014). I feel your pain: A selective review of affect-sensitive instructional strategies. In R. Sottilare, A. Graesser, X. Hu, & B. Goldberg (Eds.), *Design recommendations for adaptive intelligent tutoring systems: Adaptive instructional strategies* (Vol. 2, pp. 35–48). Orlando, FL: US Army Research Laboratory.

D'Mello, S., Craig, S., Sullins, J., & Graesser, A. (2006). Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education*, *16*(1), 3–28.

D'Mello, S., & Graesser, A. (2012a). AutoTutor and Affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. ACM *Transactions on Interactive Intelligent Systems*, 2(4), 23:22–23:39.

D'Mello, S., & Graesser, A. (2012b). Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2), 145–157. doi:10.1016/j.learninstruc.2011.10.001

D'Mello, S., & Graesser, A. (2012c). Malleability of students' perceptions of an affect-sensitive tutor and its influence on learning. In G. Youngblood & P. McCarthy (Eds.), *Proceedings of the 25th Florida Artificial Intelligence Research Society Conference* (pp. 432–437). Menlo Park, CA: AAAI Press.

D'Mello, S., & Graesser, A. (2014a). Inducing and tracking confusion and cognitive disequilibrium with breakdown scenarios. *Acta Psychologica*, 151, 106–116.

D'Mello, S. K., & Graesser, A. C. (2014b). Confusion. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 289–310). New York: Routledge.

D'Mello, S. K., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. ACM *Computing Surveys*, 47(3), 43:41–43:46.

D'Mello, S., Lehman, B., & Person, N. (2010). Monitoring affect states during effortful problem solving activities. *International Journal of Artificial Intelligence in Education*, 20(4), 361–389.

D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., . . . Graesser, A. (2010). A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning. In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 245–254). Berlin/Heidelberg: Springer.

D'Mello, S., & Mills, C. (2014). Emotions while writing about emotional and non-emotional topics. *Motivation and Emotion*, 38(1), 140–156.

D'Mello, S., Taylor, R. S., & Graesser, A. (2007). Monitoring affective trajectories during complex learning. In D. McNamara & J. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 203–208). Austin, TX: Cognitive Science Society.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Dillon, J., Bosch, N., Chetlur, M., Wanigasekara, N., Ambrose, G. A., Sengupta, B., & D'Mello, S. K. (2016). Student emotion, co-occurrence, and dropout in a MOOC context. In T. Barnes, M. Chi, & M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (EDM 2016) (pp. 353–357). International Educational Data Mining Society.

Donnelly, P., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., . . . D'Mello, S. K. (2016a). Automatic teacher modeling from live classroom audio. In L. Aroyo, S. D'Mello, J. Vassileva, & J. Blustein (Eds.), *Proceedings of the 2016 ACM International Conference on User Modeling, Adaptation, & Personalization* (ACM UMAP 2016) (pp. 45–53). New York: ACM.

Donnelly, P., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., . . . D'Mello, S. K. (2016b). Multi-sensor modeling of teacher instructional segments in live classrooms. *Proceedings of the 18th ACM International Conference on Multimodal Interaction* (ICMI 2016). New York: ACM.

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

Emotient. (2014). FACET: Facial expression recognition software.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago, IL: University of Chicago Consortium on Chicago School Research.

Forbes-Riley, K., & Litman, D. J. (2011). Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor. *Speech Communication*, 53(9–10), 1115–1136. doi:10.1016/j.specom.2011.02.006

Fredrickson, B., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, 19(3), 313–332. doi:10.1080/02699930441000238

Galla, B. M., Plummer, B. D., White, R. E., Meketon, D., D'Mello, S. K., & Duckworth, A. L. (2014). The academic diligence task (ADT): Assessing individual differences in effort on tedious but important schoolwork. *Contemporary Educational Psychology*, 39(4), 314–325.

Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612–618. doi:10.1109/TE.2005.856149

Gross, J. (2008). Emotion regulation. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 497–512). New York: Guilford.

Hess, U. (2003). Now you see it, now you don't - the confusing case of confusion as an emotion: Commentary on Rozin and Cohen (2003). *Emotion*, 3(1), 76–80.

Isen, A. (2008). Some ways in which positive affect influences decision making and problem solving. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 548–573). New York: Guilford.

Izard, C. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review*, 2(4), 363–370. doi:10.1177/1754073910374661

James, W. (1884). What is an emotion? *Mind*, 9, 188–205.

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47.

Kai, S., Paquette, L., Baker, R., Bosch, N., D'Mello, S., Ocumpaugh, J., . . . Ventura, M. (2015). Comparison of face-based and interaction-based affect detectors in physics playground. In C. Romero, M. Pechenizkiy, J. Boticario, & O. Santos (Eds.), *Proceedings of the 8th International Conference on Educational Data Mining* (EDM 2015) (pp. 77–84). International Educational Data Mining Society.

Kory, J., D'Mello, S. K., & Olney, A. (2015). Motion tracker: Camera-based monitoring of bodily movements using motion silhouettes. *PloS ONE*, 10(6), doi:10.1371/journal.pone.0130293.

Lench, H. C., Bench, S. W., & Flores, S. A. (2013). Searching for evidence, not a war: Reply to Lindquist, Siegel, Quigley, and Barrett (2013). *Psychological Bulletin*, 113(1), 264–268.

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences*, 28(2), 169–245.

Lindquist, K. A., Siegel, E. H., Quigley, K. S., & Barrett, L. F. (2013). The hundred-year emotion war: Are emotions natural kinds or psychological constructions? Comment on Lench, Flores, and Bench (2011). *Psychological Bulletin*, 139(1), 264–268.

McKeown, G. J., & Sneddon, I. (2014). Modeling continuous self-report measures of perceived emotion using generalized additive mixed models. *Psychological Methods*, 19(1), 155–174.

Mesquita, B., & Boiger, M. (2014). Emotions in context: A sociodynamic model of emotions. *Emotion Review*, 6(4), 298–302.

Microsoft. (2015). Kinect for Windows SDK MSDN. Retrieved 5 August 2015 from https://msdn.microsoft.com/en-us/library/dn799271.aspx

Moors, A. (2014). Flavors of appraisal theories of emotion. *Emotion Review*, 6(4), 303–307.

Nystrand, M. (1997). Opening dialogue: Understanding the dynamics of language and learning in the English classroom. *Language and Literacy Series*. New York: Teachers College Press.

Ocumpaugh, J., Baker, R. S., & Rodrigo, M. M. T. (2012). Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0. New York.

OECD (Organisation for Economic Co-operation and Development). (2015). PISA 2015 collaborative problem solving framework.

Ohman, A., & Soares, J. (1994). Unconscious anxiety: Phobic responses to masked stimuli. *Journal of Abnormal Psychology*, *103*(2), 231–240.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135.

Pardos, Z., Baker, R. S. J. d., San Pedro, M. O. C. Z., & Gowda, S. M. (2013). Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (LAK '13), 8–12 April 2013, Leuven, Belgium (pp. 117–124). New York: ACM.

Parkinson, B., Fischer, A. H., & Manstead, A. S. (2004). *Emotion in social relations: Cultural, group, and interpersonal processes*. Psychology Press.

Pekrun, R., & Stephens, E. J. (2011). Academic emotions. In K. Harris, S. Graham, T. Urdan, S. Graham, J. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (pp. 3–31). Washington, DC: American Psychological Association.

Porayska-Pomsta, K., Mavrikis, M., D'Mello, S. K., Conati, C., & Baker, R. (2013). Knowledge elicitation methods for affect modelling in education. *International Journal of Artificial Intelligence in Education*, 22, 107–140.

Raca, M., Kidzinski, L., & Dillenbourg, P. (2015). Translating head motion into attention: Towards processing of student's body language. *Proceedings of the 8th International Conference on Educational Data Mining*. International Educational Data Mining Society.

Razzaq, L., Feng, M., Nuzzo-Jones, G., Heffernan, N. T., Koedinger, K. R., Junker, B., . . . Choksey, S. (2005). The assistment project: Blending assessment and assisting. In C. Loi & G. McCalla (Eds.), *Proceedings of the 12th International Conference on Artificial Intelligence in Education* (pp. 555–562). Amsterdam: IOS Press.

Ringeval, F., Sonderegger, A., Sauer, J., & Lalanne, D. (2013). Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. *Proceedings of the 2nd International Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EmoSPACE) in conjunction with the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*. Washington, DC: IEEE.

Rosenberg, E. (1998). Levels of analysis and the organization of affect. *Review of General Psychology*, 2(3), 247–270. doi:10.1037//1089-2680.2.3.247

Rozin, P., & Cohen, A. (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.

Sabourin, J., Mott, B., & Lester, J. (2011). Modeling learner affect with theoretically grounded dynamic Bayesian networks. In S. D'Mello, A. Graesser, B. Schuller, & J. Martin (Eds.), *Proceedings of the 4th International Conference on Affective Computing and Intelligent Interaction* (pp. 286–295). Berlin/Heidelberg: Springer-Verlag.

San Pedro, M., Baker, R. S., Bowers, A. J., & Heffernan, N. T. (2013). Predicting college enrollment from student interaction with an intelligent tutoring system in middle school. In S. D'Mello, R. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining* (EDM 2013) (pp. 177–184). International Educational Data Mining Society.

Schwarz, N. (2012). Feelings-as-information theory. In P. Van Lange, A. Kruglanski, & T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 289–308). Thousand Oaks, CA: Sage.

Shute, V. J., Ventura, M., & Kim, Y. J. (2013). Assessment and learning of qualitative physics in Newton's playground. *The Journal of Educational Research*, 106(6), 423–430.

Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). Your click decides your fate: Inferring information processing and attrition behavior from MOOC video clickstream interactions. Paper presented at the Empirical Methods in Natural Language Processing: Workshop on Modeling Large Scale Social Interaction in Massively Open Online Courses. http://www.aclweb.org/anthology/W14-4102

Strain, A., & D'Mello, S. (2014). Affect regulation during learning: The enhancing effect of cognitive reappraisal. *Applied Cognitive Psychology*, 29(1), 1–19. doi:10.1002/acp.3049

Tracy, J. L. (2014). An evolutionary approach to understanding distinct emotions. *Emotion Review*, 6(4), 308–312.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction, 21*(3), 209–249. doi:10.1207/S1532690XCI2103_01

Wang, Z., Miller, K., & Cortina, K. (2013). Using the LENA in teacher training: Promoting student involvement through automated feedback. *Unterrichtswissenschaft*, 4, 290–305.

Wen, M., Yang, D., & Rosé, C. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us? In J. Stamper, S. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (EDM 2014), 4–7 July, London, UK (pp. 130–137): International Educational Data Mining Society. http://www.cs.cmu.edu/~mwen/papers/edm2014-camera-ready.pdf

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: Recognizing and responding to student affect. *International Journal of Learning Technology*, 4(3/4), 129–163.

Yang, D., Wen, M., Howley, I., Kraut, R., & Rosé, C. (2015). Exploring the effect of confusion in discussion forums of massive open online courses. *Proceedings of the 2nd ACM conference on Learning@Scale* (L@S 2015), 14–18 March 2015, Vancouver, BC, Canada (pp. 121–130). New York: ACM. doi:10.1145/2724660.2724677

Zeng, Z., Pantic, M., Roisman, G., & Huang, T. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 39–58.