The Effects of Autonomy on Emotions and Learning in Game-Based Learning Environments

Amanda E. Bradbury, Michelle Taub, & Roger Azevedo
{aebradbu@ncsu.edu, mtaub@ncsu.edu, razeved@ncsu.edu}
North Carolina State University, Department of Psychology, 2310 Stinson Drive, 640 Poe Hall
Raleigh, NC 27695 USA

Abstract

The current study examined the impact of agency on college students' emotions and learning during gameplay with CRYSTAL ISLAND, a game-based learning environment designed to foster microbiology learning. 96 undergraduate students (59% female) from a large North American university participated in the study. Participants were randomly assigned to one of three experimental conditions (i.e., full agency, partial agency, no agency), based on the level of control granted during gameplay, and were asked to uncover the source, identity, and best treatment for a mysterious illness. Results revealed participants in the partial agency condition achieved the highest (pre- to post-test) proportional learning gain (PLG), even when controlling for session duration. Additionally, there was a positive correlation between evidence scores of four emotions (anger, fear, confusion, and frustration) and PLG within the partial agency condition—meaning the higher the evidence of the above emotions, the higher the PLG. Further, a stepwise multiple regression showed anger as the sole predictor of PLG. Results from this study have important implications for understanding the role of autonomy and emotions during learning and problem solving with GBLEs designed to foster scientific thinking in STEM. The current study suggests that although GBLEs offer significant learning benefits, they also induce several emotions that can facilitate or inhibit learning gains, requiring further examination.

Keywords: human agency; emotions; learning; game-based learning environments; science

Autonomy is a critical determinant in human learning, problem solving, and performance (Bandura, 2001). Despite its importance in cognitive science, there is a paucity of research that experimentally manipulates autonomy and explores its impact on learning and emotions, in STEM game-based learning environments (GBLEs). Various levels of autonomy likely affect learners' abilities to monitor and regulate their cognitive, affective, metacognitive, and motivational processes in dynamic, non-linear learning environments involving planning (e.g., coordinating multiple goals), learning activities (e.g., reading scientific texts), and scientific reasoning (e.g., collecting evidence and testing hypotheses) in different ways. Further, little is understood of how autonomy affects emotions in GBLEs, and in turn, how these emotions affect learning outcomes (Azevedo, Taub, Mudrick, Farnsworth, & Martin, 2016; D'Mello & Graesser, 2012). Our study focuses on the effects of autonomy on emotions and the impact of both on learning and problem solving within the GBLE, CRYSTAL ISLAND.

GBLEs offer powerful platforms to enhance student learning, problem solving, and performance. However, a majority of the research focuses on engagement and motivation and is often criticized for (1) a lack of theoretical framing, (2) questionable operationalizations of key constructs (e.g., engagement, motivation), (3) overreliance on self-report measures, and (4) dubious empirical support, based on a lack of experimental rigor, methodological shortcomings, and inappropriate analytical techniques (see Mayer, 2014). Additionally, much of this research fails to assess learning gains, choosing to take an "everything but learning" approach, such as measuring engagement or motivation alone while ignoring educational outcomes (Mayer, 2014). Further, GBLEs have been criticized for overshadowing educational content with game elements that are superfluous and distracting to learning goals, drawing learner attention away from important educational content (Mayer & Johnson, 2010). Interestingly, many of these distractors (e.g., game narratives, interesting characters) are the very elements thought to increase student motivation, engagement and positive emotions (Sabourin & Lester, 2014). Further, research has indicted that while distractors may present opportunities for off-task behaviors, leading to decreased learning gains (Rowe, McQuiggan, Robison, & Lester, 2009), off-task behaviors could in-fact be a strategy to alleviate frustration, allowing the student to reduce frustration and thereby increase learning gains (Sabourin, Rowe, Mott, & Lester, 2014).

Students experience a diverse range of emotions when learning, which likely influence cognitive processes and academic performance (see Calvo, D'Mello, Gratch, & Kappas, 2014). We address this issue by using online trace methods (e.g., facial expression detection software [FACET; Version 6.2], and logfiles), to assess the impact of autonomy on emotions and learning during gameplay (see Azevedo et al., 2016; Calvo et al., 2014), thereby increasing understanding of emotional monitoring and regulation in GBLEs (Rowe, Shores, Mott, & Lester 2011). This research can inform the design of future intelligent, adaptive GBLEs that not only teach complex instructional material effectively but also train the skills necessary to successfully monitor and regulate emotions during learning, leading to improved learning outcomes.

Theoretical Framework

D'Mello and Graesser's (2012) model of affective dynamics suggests certain emotional states arise as the result of an impasse during deep learning, creating cognitive disequilibrium. This model focuses on four learner-centered emotional states: flow/engagement, confusion, frustration and boredom. When learners reach a state of disequilibrium (e.g., during reading complex text), they are likely to experience confusion which if unresolved will likely transition to frustration, which if also left unresolved, will lead to boredom and disengagement from the activity (e.g., reading, inspecting diagrams). This model posits that students systematically shift between learning-centered states during complex learning and that these shifts are predictive of learning, problem solving, and scientific reasoning. For instance, frustration is much more likely to transition to boredom than to engagement/flow, as learners have not yet transitioned to confusion, where through effortful reasoning and problem solving they can resolve an impasse and return to equilibrium. However, this model has some drawbacks. For instance, it ignores other emotional states such as the seven basic emotions (e.g., anger; Ekman, 1973), assuming that other basic emotions are unimportant to learning. Lastly, this model has not been used to examine autonomy and extended learning with GBLEs such as CRYSTAL ISLAND.

Current Study

The goal of the current study was to examine the effects of autonomy on emotions and learning during gameplay with GBLEs such as CRYSTAL ISLAND. By experimentally manipulating autonomy, we could empirically observe how different levels of autonomy (e.g., agency conditions) affected learning gains as well as emotional states, and in turn, how these emotional states affected learning gains. Our research questions were as follows: 1) What are the effects of autonomy on proportional learning gains with CRYSTAL ISLAND, after controlling for session duration? 2) What are the effects of autonomy on learners' emotions throughout their interaction with CRYSTAL ISLAND? 3) Do evidence scores of emotional states predict PLG during gameplay with CRYSTAL ISLAND and are there differences in emotion evidence scores between high and low performers?

Our hypotheses were as follows. **(H1):** Participants in the partial agency condition will show significant PLG compared to the full agency and no agency conditions. **(H2):** The full agency condition will exhibit the highest evidence of positive emotions such as joy and the lowest evidence of negative emotions such as anger and frustration compared to the partial and no agency conditions. **(H3):** Higher evidence scores of negative emotions such as anger, confusion and frustration will lead to increased PLG in all conditions.

Method

Participants

96 undergraduate students (59% female) from a large North American university participated in the current study. Participants' ages ranged from 18 to 29 (M = 19.99, SD = 1.79) and were randomly assigned to one of three experimental condition: full agency, partial agency or no agency (see Experimental Procedure). Additionally, they were compensated $10/hour for participating.

Materials

At the start of the experimental session, participants read and completed the informed consent, a demographics questionnaire and a series of self-report questionnaires. These questionnaires probed participants' emotions and motivation (e.g., Emotions and Values; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011) as well as achievement goals (Elliot & Murayama, 2008). Participants also completed a pretest (M = 11.94, SD = 2.79; 57% correct) and post-test (M = 13.92, SD = 2.86; 66% correct) on microbiology knowledge: a 21-item, four-choice multiple-choice test, with 12 factual and 9 procedural questions. Participants also completed the Perceived Interest Questionnaire (Schraw, Bruning, & Svoboda, 1995), Intrinsic Motivation Inventory (Ryan, 1982), and Presence Questionnaire (Witmer & Singer, 1998).

CRYSTAL ISLAND

CRYSTAL ISLAND is a narrative-centered GBLE used to foster students' self-regulated learning, scientific reasoning, and problem-solving skills (Rowe et al., 2011). Participants experience the game in first person perspective, arriving on a tropical island where they discover a mysterious illness has infected the community. Taking a protagonist role, participants explore the island, seek clues by speaking to residents and patients, read content on microbiology and use lab equipment to scan for possible transmission sources, all to discover the source, identity, and best treatment for the infectious disease.

Buildings CRYSTAL ISLAND has five buildings, each embedded with a multitude of books, research papers, posters, food items, and non-player characters (NPCs). In the infirmary, participants interview sick patients and interact with the NPC, Kim the camp nurse, who provides the game narrative. Through this interaction, they gather pertinent information such as overall goals, background information, and clues pointing towards possible illness types and transmission sources. In the two living quarters (a dorm room and a microbiologist's home), participants converse with microbiology experts and another patient, and read books and posters on various microbiology topics. In the dining hall, participants meet Quentin the camp cook, who offers insight into what foods he had prepared and sick patients had eaten prior to the outbreak. Using information and clues gathered from these buildings, participants can infer which items are the likely transmission source and then test these hypotheses by scanning these food items in the laboratory.

Game Elements Participants complete concept matrices as they read about microbiology in books and research articles. For example, as they read about E. coli, they must fill in a diagram asking questions related to the reading (i.e., where E. coli is located, symptoms and common diagnostic tests). Additionally, by interacting with NPCs, participants receive valuable information (i.e., evidence), such as symptoms and food eaten. As participants collect evidence and begin making inferences, they can track and organize symptoms, test results, and make a final diagnosis via a diagnosis worksheet. This worksheet supports problem-solving processes by allowing participants to offload information as they interact with the game environment, later using this information to make a final diagnosis, identify the transmission source, and propose a treatment plan. For instance, they may read about influenza then check the diagnosis worksheet to find the symptoms match the current epidemic. Additionally, participants generate hypotheses regarding which food items are the likely transmission source as well as the type of pathogen they might carry. These hypotheses are tested by collecting and scanning food items, and testing for a virus, bacterium, mutagen, or carcinogen. If a test comes back positive for a pathogenic substance, the participant can confirm the transmission source and add their finding to the diagnosis worksheet. Once participants correctly identify the illness type, transmission source, and treatment plan, the mystery is solved and the game concludes.

Experimental Procedure

Conditions Participants were randomly assigned to one of three conditions (i.e., full agency, partial agency, no agency) prior to gameplay. These conditions varied in the level of autonomy assigned to each player, ranging from full autonomy (full agency), to some autonomy (partial agency), to no autonomy at all (no agency). In the *full condition*, participants were free to explore the game environment and its elements as much or as little as they wished, choosing what buildings to visit, what books to read, and with which NPCs to interact. Conversely, the *partial condition* contained strict game parameters with a pre-set order in which players visited buildings and a requirement that they interact with all game artifacts (e.g., read all books/posters, speak with all NPCs, etc.) before advancing to other buildings. In the *no condition*, participants did not play CRYSTAL ISLAND but instead watched a narrated video of an expert playing the game. This was an optimal instructional path designed to enhance learning without the opportunity to exercise autonomy as participants had no control over any aspect of the gameplay or content.

Experimental Procedure The experimental session lasted one to two and a half hours depending on condition (M = 89.64 min, SD = 18.37 min). Upon arrival, participants were greeted, directed to the workstation and asked to review and complete the informed consent. Next, they received an overview of the study, donned an electro dermal activity (EDA) bracelet (Empatica E4), and completed the microbiology pretest. Then, the SMI RED 250 eye tracker was calibrated using a 9-point calibration. Following successful calibration, a baseline for the facial recognition of emotion software (FACET) and EDA were established using Attention Tool (Version 6.2). Participants were then given instructions for the experimental session that included an overview of the game scenario covering their role as the protagonist, the importance of reading (i.e., books, articles, and posters), interacting with NPCs and scanning food items to solve the mystery. During gameplay, we collected logfiles, eye-tracking, facial expressions of emotions, and physiological data on all participants in the full and partial agency conditions only. Upon game conclusion, participants completed several self-report measures and the microbiology post-test, after which they were debriefed, thanked, and paid for their time.

Coding and Scoring

For the purposes of the current study, only logfiles and FACET data were used. Additionally, pre- and post-test scores (out of 21 possible points) of microbiology content knowledge were used to generate a PLG score (see below).

Logfiles Logfile data captured the sequence and timing of participants' movements and actions within the game (e.g., talking to NPCs, reading books). For this study, only session duration was analyzed. This variable was extracted from the trace data. Additionally, logfile data were only captured in the full and partial agency conditions as the no agency condition watched a video play-through (91 min) of CRYSTAL ISLAND rather than play, thus not generating any log-file data.

Facial Expression Data Each experimental session included a video of the participant, which was later analyzed using FACET, facial expression recognition software included with Attention Tool. We used FACET (sampling rate of 30Hz) to analyze the following nine basic and learning-centered emotions: joy, anger, contempt, frustration, confusion, surprise, fear, sadness and disgust (see, Dente, Küster, Skora, & Krumhuber, 2017, regarding the software's validity). Each emotion was given an evidence score automatically generated by FACET representing the likelihood of an expert human coder to similarly categorize the expression. This score was based on a logarithmic scale (base 10), meaning that a score of one indicated the likelihood of 10 human coders coding for that emotions while a score of two indicated the likelihood of 100 human coders coding for that emotion, and so forth. For the purposes of the current study, the mean evidence score for the entire session duration was used for each participant. The range of evidence scores for all emotions and across participants was 0 to 1.98, excluding negative values. Negative scores indicated the emotion was not likely present, and since we were interested in emotions present, all negative values were replaced with zero.

Proportional Learning Gain (PLG) PLG scores were calculated from pre- and post-test ratios scores of microbiology content knowledge, using Witherspoon, Azevedo, and D'Mello's (2008) formula. For example, if a participant scored an 11 out of 21 on the pre-test and a 15 out of 21 on the post-test then their PLG score was .40.

Median Split High versus low performers were determined through a median split of the PLG variable for the partial agency condition. The median for this condition was .40 (range: -0.17 to 0.70).

Results

Research Question 1: What are the effects of autonomy on proportional learning gains with CRYSTAL ISLAND, after controlling for session duration?

To investigate the effects of autonomy on PLG, we conducted an ANCOVA, using condition as the independent variable and session duration as a covariate, see Table 1 for mean session duration by condition. Results indicated a significant main effect for condition, $F(2, 88) = 3.35$, $p = .003$, $\eta_p 2 = .13$. Post hoc LSD analyses indicated that the partial agency condition ($M = .35$, $SD = .23$) showed significantly higher PLG than both the full ($M = .18$, $SD = .27$) and no agency conditions ($M = .11$, $SD = .28$); however, there was no difference between the full and no agency conditions.

Table 1. *Mean session duration (min) by condition.*

	Full Agency *M (SD)*	Partial Agency *M (SD)*	No Agency *M (SD)*
Session Duration	78.69 (21.92)	98.65 (18.43)	91.00 (0)

Research Question 2: What are the effects of autonomy on learners' emotions throughout their interaction with CRYSTAL ISLAND?

A MANCOVA was conducted using mean evidence scores of the basic and learner-centered emotions as the nine dependent variables and condition as the one independent variable. No significant main effect was found by condition; Wilk's $\lambda = .78$, $F(16, 164) = 1.39$, $\eta_p^2 = .12$. Univariate results revealed that disgust, $F(2, 89) = 4.15$, $p = .02$, $\eta_p 2 = .09$, anger, $F(2, 89) = 4.12$, $p = .02$, $\eta_p 2 = .02$, and joy $F(2, 92) = 3.48$, $p = .04$, $\eta_p 2 = .07$, showed statistically significant differences between conditions. No other emotions demonstrated significant differences. Post hoc LSD analyses indicated that those in the full agency condition exhibited higher levels of disgust ($M = .22$, $SD = .34$) and anger ($M = .55$, $SD = .62$) compared to those in the partial agency condition ($M = .14$, $SD = .24$; $M = .37$, $SD = .49$, respectively). Additionally, those in the full agency condition exhibited higher levels of joy ($M = .25$, $SD = .44$) compared to the partial agency condition ($M = .06$, $SD = .13$; see Table 2).

Table 2. *Mean emotion evidence scores by condition.*

	Experimental Conditions			F-test Results	
Emotional State	Full Agen	Part Agen	No Agen	*F-Stat*	Comparisons
	M (SD)	*M (SD)*	*M (SD)*	*F(p)*	
Disgust	.22 (.34)	.14 (.24)	.04 (.12)	4.15 (.02)	(P = F > N = P)
Anger	.55 (.62)	.37 (.49)	.18 (.35)	4.12 (.02)	(P = F > N = P)
Joy	.25 (.44)	.06 (.13)	.09 (.27)	3.48 (.04)	(F > P = N)
Frustration	.38 (.49)	.20 (.32)	.16 (.31)	2.88 (.06)	(P = F > N = P)
Surprise	.16 (.36)	.19 (.32)	.11 (.27)	.48 (.62)	(F = P = N)
Fear	.18 (.31)	.10 (.15)	.09 (.20)	1.45 (.24)	(F = P = N)
Contempt	.06 (.13)	.06 (.12)	.05 (.14)	.05 (.95)	(F = P = N)
Sadness	.23 (.28)	.23 (.29)	.18 (.31)	.32 (.73)	(F = P = N)
Confusion	.45 (.52)	.33 (.40)	.26 (.46)	1.30 (.28)	(F = P = N)

Note: F = full agency, P = no agency, N = no agency conditions

Research Question 3: Do evidence scores of emotional states predict PLG during gameplay with CRYSTAL ISLAND and are there differences in emotion evidence scores between high and low performers?

To assess the relationship between emotions and PLG while playing CRYSTAL ISLAND, four correlation matrices were created: overall (all conditions; $n = 92$), full agency ($n = 30$), partial agency ($n = 32$), and no agency ($n = 30$). The full and no agency conditions as well as all conditions combined showed no correlations between emotions and PLG; however, for the partial condition, four emotions were significantly positively correlated with PLG, anger, $r(30) = .39$, $p = .03$, fear, $r(30) = .36$, $p = .04$, confusion, $r(30) = .39$, $p = .03$, and frustration, $r(30) = .39$, meaning the higher the evidence of the above emotions, the higher the PLG.

To determine the predictive power of anger, fear, confusion, and frustration on PLG within the partial agency condition, a stepwise multiple regression analysis was conducted. Results indicated that anger ($\beta = .39$, $p = .03$, $R^2 = .15$) was the sole predictor of PLG, meaning that more evidence of anger predicted better PLG, accounting for 15% of the variability in PLG.

Given the regression results for the partial agency condition, we performed a median split on these participants' PLG to examine whether there were differences between high- and low-performers' experienced emotions. Result of an independent samples t-test revealed that high performers exhibited significantly more evidence of facially expressed frustration, $t(18) = -3.75$, $p < .002$, $d = -1.78$, anger, $t(19) = -3.47$, $p < .003$, $d = -1.58$, and confusion, $t(21) = -2.97$, $p < .007$, $d = -1.29$, compared to low performers.

Discussion

Results of the current study revealed that students achieved the highest PLG in the partial agency condition compared to the full and no agency conditions, even after controlling for sessions duration. These results support H1, demonstrating the positive impacts of seceding partial agency to improve learning outcomes in GBLEs. Previous research explains that while offering a high degree of user control allows learners to regulate their own learning, constructing knowledge based on the representations they find useful, this responsibility can lead to disorientation and negative learning outcomes when learners are unsure which path to follow (Greene, Bolick, & Robertson, 2010), suggesting there may be on optimal level of autonomy to improve learning outcomes in GBLE. Future research should empirically test different parametrization of autonomy on GBLEs to assess the optimal level of autonomy to foster learning across domains.

For research question two, participants in the full agency condition were more emotionally expressive than those in no and partial agency conditions. For instance, those in the full agency condition showed significantly higher evidence of joy than those in the partial and no agency conditions, as well as significantly higher evidence of anger and disgust compared to the no agency condition. These results run contrary to our original hypothesis (H2), expecting the full agency condition to experience the least negative emotions; however, the full agency condition did experience the highest evidence of joy, partially supporting H2. A plausible explanation could be that those in the full agency had a greater potential to express autonomy which led to more emotional expressivity throughout task performance (Azevedo et al., 2016). A next step involves a micro-level analysis mapping specific game events (e.g., reading books, testing evidence, etc.) with emotional expressivity (e.g., higher evidence scores) and emotional states.

As for research question three, no correlations between emotional states and PLG were found with 1) all conditions combined, 2) the full agency condition or 3) the no agency condition; however, this was not the case within the partial agency condition. The partial agency condition found significant positive correlations between PLG and evidence scores of facially expressed anger, fear, confusion and frustration, meaning the higher evidence of the above emotions, the higher a participant's PLG. After imputing the aforementioned emotions into a stepwise multiple regression conducted within the partial agency condition, anger was the sole predictor of PLG. Further, high performers in the partial agency condition exhibited significantly higher evidence of anger, frustration and confusion compared to low performers, demonstrating that negative emotions, typically thought as unconducive to learning (Sabourin & Lester, 2014), can have positive effects on learning outcomes. Previous work has reach similar conclusions, finding confusion, if appropriately regulated and resolved, as beneficial to learning (D'Mello, Lehman, Pekrun,& Graesser, 2014).

In the current study, fear, anger, frustration and confusion had a positive effect on PLG, but only when the participant seceded partial control of the learning environment (i.e., partial agency condition). One explanation for these results could be explained using the model of affective dynamics (D'Mello & Graesser, 2012). For instance, participants are likely to experience confusion and frustration when learning difficult subject matter and will hence experience cognitive disequilibrium (D'Mello & Graesser, 2012). Equilibrium (e.g., engagement/flow state) is regained through effortful reasoning, problem solving and reflection; however, when left unresolved, learners can digress from confusion to frustration and eventually disengage from the learning activity (D'Mello et al., 2014).

In the current study, participants were asked to learn new information in order to solve complex problems: what disease was infecting the community, what was the transmission source, and how to best treat patients. However, each condition offered different paths to learn this information (via varying levels of autonomy) and in turn affected emotions and learning differently. For instance, in the no agency condition, participants might have felt frustrated at not being able to play the game and this frustration may have led to boredom and disengagement, explaining poor PLG. In the full agency condition, participants could reduce confusion and frustration by simply avoiding books, research articles, or interactions with aspects of the game they found unappealing; however, even though they would return to equilibrium through these actions, they would have missed valuable educational content, thus reducing PLG. Conversely, the partial agency condition was forced to interact with all elements of the game before leaving a room. This stipulation may have forced participants to work through the confusion and frustration they experienced because they could not progress with the not step of the game until required actions (e.g., finishing a conversion with the NPC, filling in a concept matrix correctly) were completed. Therefore, these participants were more likely to engage in the effortful reasoning and problem solving necessary for both deep learning and a return to equilibrium.

Limitations

There were a number of limitations with the current study. First, the operationalization of autonomy in the partial and full agency condition as this was a first attempt to parameterize key assumptions of autonomy in a GBLE. Also, because we were looking at autonomy, there may have been other metacognitive processes (e.g., motivation) affecting learning gains that we did not control for or measure, we only used log-files, FACET, and learning outcomes data. Converging these data along with EDA and eye-tracking data would further elucidate the role of autonomy and emotions during learning. For instance, eye-tracking data could be used to examine what activity a participant was engaged in prior, during and after the onset of a certain emotion. Additionally, EDA data could be used to validate the presence and relevance of emotions. For instance, spikes in EDA data could be mapped onto emotion evidence scores to determine when spikes and high emotion evidence scores co-occur revealing the quality of appraisals mechanisms (Gross, 2015).

Implications and Future Directions

These results have important implications for understanding the role of autonomy and emotions during learning and problem solving with GBLEs designed to foster scientific thinking in STEM. The current study suggests GBLEs induce several basic and learning–centered emotions depending on the level of autonomy granted to a learner and that autonomy and emotions can either facilitate or inhibit learning. However, further empirical examination is required. Future research should design and test additional experimental manipulations that operationalize key assumptions of autonomy (Bandura, 2001). Further, our results revealed a need to extend models and theories of affect to include basic emotions when considering transitions between emotional states in learning environments (D'Mello & Graesser, 2012) and would benefit by including Gross's process model of emotional regulation along with emotion regulation strategies (Gross, 2015).

Methodologically, converging the multimodal multichannel data will allow researchers to examine the impact of autonomy on emotions and their impact on learning, problem solving, and reasoning. For example, how do emotions fluctuate during different activities during learning with GBLEs? What is their specific behavioral signature in terms of onset/trigger event, intensity, duration, evidence of emotion regulation strategy, and so forth? How do these emotions related to specific GBLE activities (e.g., reading books and posters, interviewing patients, interpreting results, deriving hypotheses)? Such questions can be addressed by traditional statistics as well as data mining and machine learning techniques and lead to the design of intelligent GBLEs capable of detecting, tracking, modeling, and fostering adaptive, real-time scaffolding to learners, depending on their individual needs, thus ensuring optimal learning.

Acknowledgments

This research has been supported by funding from the Social Sciences and Humanities Research Council of Canada (SSHRC 895–2011–1006) and members of the SMART Lab and the IntelliMedia Group.

References

Attention Tool (Version 6.2) [Computer software]. Boston, MA: IMotions.

Azevedo, R., Taub, M., Mudrick, N., Farnsworth, J., & Martin, S. (2016). Using research methods to investigate emotions in computer-based learning environments. In P. Schutz & M. Zembylas (Eds.), *Methodological advances in research on emotion and education*. Amsterdam, The Netherlands: Springer.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.

Calvo, R. A., D'Mello, S., Gratch, J., & Kappas, A. (Eds.). (2014). *The Oxford handbook of affective computing*. Oxford, England: Oxford University Press.

Dente, P., Küster, D., Skora, L., & Krumhuber, E. (2017). Measures and metrics for automatic emotion classification via FACET. *Proceedings of the Convention of the Society for the Study of Artificial Intelligence and the Simulation of Behaviour*.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction, 22*, 145-157.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153-170.

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology, 100*, 613-628.

Ekman P. (ed) (1973). *Darwin and facial expression*. New York, NY: Academic Press.

Greene, J. A., Bolick, C. M., & Robertson, J. (2010). Fostering historical knowledge and thinking skills using hypermedia learning environments: The role of self-regulated learning. *Computers & Education, 54*, 230-243.

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*, 1-26.

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. Cambridge, MA: MIT Press.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research, 42*, 241-265.

Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48.

Rowe, J., McQuiggan, S., Robison, J., & Lester, J. (2009). Off-task behavior in narrative-centered learning environments. In S. Craig & D. Dicheva (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education.* Berlin, Germany: Springer-Verlag.

Rowe, J., Shores, L., Mott, B., & Lester, J. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education, 21*, 115-133.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450.

Sabourin, J. L., & Lester, J. C. (2014). Affect and engagement in game-based learning environments. *IEEE Transactions on Affective Computing, 5*, 45-56.

Sabourin, J. L., Rowe, J. P., Mott, B. W., & Lester, J. C. (2013). Considering alternate features to classify off-task behavior as emotion self-regulation: A supervised learning approach. *Journal of Educational Data Mining, 5*, 9-38.

Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. *Journal of Literacy Research, 27*, 1-17.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*, 225-240.

Witherspoon, A., Azevedo, R., & D'Mello, S. (2008). The dynamics of self-regulatory processes within self- and externally regulated learning episodes during complex science learning with hypermedia. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent Tutoring Systems 2008*. Berlin, Germany: Springer.