Toward a Feature-driven Understanding of Students' Emotions during Interactions with Agent-based Learning Environments: A Selective Review

Jason M. Harley, PhD
[a]University of Montréal, Computer Science and Operations, 2920 Chemin de la Tour, Pavillon André-Aisenstadt 2194, Montréal, QC, CAN, H3C 3J7
[b]McGill University, Department of Educational and Counselling Psychology, 3700 McTavish Street 614, Montréal, QC, CAN, H3A 1Y2
jason.harley@umontreal.ca
*Corresponding author

Biography: Dr. Jason M. Harley is a postdoctoral fellow in the Computer Science and Operations Department at the University of Montréal and research associate at McGill University in the Department of Educational and Counselling Psychology. He completed his PhD in Educational Psychology at McGill in the summer of 2014. His research examines emotions, self and co-regulated learning, and advanced learning technologies, including intelligent tutoring systems, serious games, and augmented reality. He has published fifteen handbook chapters, journal articles, and conference proceedings and presented or co-authored thirty conference papers in peer-reviewed, international venues in education, psychology, and computer science. He has contributed to numerous research projects that have been funded by American and Canadian funding agencies and has been awarded doctoral and postdoctoral fellowships from the Fonds Québécois de la recherche sur la Société et la culture (FQRSC) and the Social Science and Humanities Research Council of Canada (SSHRC).

Roger Azevedo, PhD
North Carolina State University, Psychology 2310 Stinson Drive, Poe Hall 703, Raleigh, NC, USA 2765-7650
razeved@ncsu.edu

Biography: Dr. Roger Azevedo joined North Carolina State University in August 2013 as a Chancellor's Faculty Excellence Program cluster hire in the Digital Transformation of Education. He is a Professor in the Department of Psychology in the area of Human Factors and Ergonomics. His main research area includes examining the role of cognitive, metacognitive, affective, and motivational self-regulatory processes during learning with computer-based learning environments. He has published over 200 peer-reviewed papers, chapters, and refereed conference proceedings in the areas of educational, learning, cognitive, and computational sciences. He received his doctorate from McGill University in 1998 and completed his postdoctoral training in Cognitive Psychology at Carnegie Mellon University. He is the editor of the Metacognition and Learning journal and his research is funded by the NSF, NIH, IES, SSHRC, NSERC, and CFI. He is a fellow of the American Psychological Association and the recipient of the NSF Early Faculty Career Award.

Note. This paper is an extension of a selective review published in the following conference proceeding: Harley, J. M., & Azevedo, R. (2014). Understanding students' emotions during interactions with advanced agent-based learning environments: A selective review. In S. Trausan-Matu., K. Boyer., M. Crosby., K. Panourgia (Eds.), *Lecture Notes in Computer Science: Vol. 8474. Intelligent Tutoring Systems (pp. 629-631)*. Berlin, Heidelberg: Springer-Verlag.

The last few decades have witnessed an explosion in the study of the complex role of emotions in a multitude of learning contexts (Azevedo & Aleven, 2013; D'Mello, 2013; Calvo & D'Mello, 2010, 2011, 2012; Harley, in press; Pekrun & Linnenbrink-Garcia, 2014). Amidst the application of traditional and cutting-edge methods to computer-based learning environments (CBLEs), however, many fundamental questions remain unanswered: How do students feel about interacting with specific types of CBLEs? Does the incidence of discrete emotions vary between similar types of environments? What features support or hinder learners' experience of adaptive emotions within these environments? This review is unique in attempting to answer these questions as they relate to a type of CBLE, namely agent-based learning environments (ABLEs; which feature pedagogical agents) and in so doing address these gaps in the research literature (Calvo & D'Mello, 2010; D'Mello, 2013). Understanding how and why learners' feel they way they do toward ABLEs is important because of the relationship between learning and emotions in which emotions can both support (e.g., enjoyment, hope) and hinder learning (e.g., boredom, anxiety; Pekrun, 2011; Pekrun, Daniel, Perry, Goetz, Stupinsky, 2010; Pekrun, Goetz, Frenzel, Petra, & Perry, 2011).

Agent-based Learning Environments

ABLEs are a type of computer-based learning environment (e.g., multi-agent systems, intelligent tutoring systems, serious games) used to help students learn various educational and professional topics such as science, math, computer literacy, and cultural sensitivity. These environments use artificial intelligence (AI) to respond and adapt to learners' individual needs as they interact with the system and, in some cases, train and foster students' cognitive, emotional, and metacognitive self-regulatory skills (Arroyo, Burleson, Tai, Muldner, & Woolf, 2013; Azevedo & Aleven, 2013; Azevedo et al., 2013; Conati & Maclaren 2009; D'Mello & Graesser, 2013; D'Mello, Lehman, Pekrun, & Graesser, 2014; Kinnebrew, Biswas, Sulcer, Taylor, 2013; Sabourin, Mott, & Lester, 2011). ABLEs can provide students with the opportunity to interact with educational material in a variety of forms including text, diagrams, videos, and interactive simulations. Many also contain a variety of embedded digital learning tools such as structured or unstructured note-taking features and open learner models that provide learners with opportunities to structure and monitor their learning progress (Azevedo et al., 2013; Feyzi-Behnagh et al., 2013).

ABLEs are unique from other computer-based learning environments because of their use of pedagogical agents (PAs) represented by an animated character (usually a 3D human head and torso) that serve several functions such as providing immediate and tailored prompts and feedback (e.g., hints, summaries, encouragement, and evaluations) to support student learning. Pedagogical agent-learner interactions are handled by one or several embodied PAs that vary along numerous dimensions including presence or absence of natural language processing, speech, facial expressions, gestures, gross motor behaviors, gender, and race. Several of these dimensions have been shown to influence learners' interactions, reaction to, and their opinions of PAs (Arroyo et al., 2013; Baylor & Kim, 2009). Furthermore, relationships between learners' individual differences (e.g., personality traits, gender) and PA characteristics have been found (e.g., PAs with different pedagogical interventions, gender; Arroyo et al., 2013; Harley et al., under review). For example, Arroyo et al. (2013) found that female students expressed positively valenced emotions most often and engaged in more productive behaviors when they interacted with female PAs as opposed to male students who had better learning outcomes when no PAs were present and the worst performance and affective outcomes when they interacted with female PAs.

Research has found that the visual presence of PAs does not distract students (i.e., produce a *split-attention effect*) from learning (Craig, Gholson, & Driscoll, 2002; Moreno, Mayer, Spires, & Lester, 2001). Moreover, in addition to the advantages that computer-based learning environments provide learners with (e.g., individualized feedback and low-stakes assessment) the presence of a PA has been found to have a positive effect on learners' perception of their learning experience as well as their performance outcomes with these environments (Andre, Rist, & Muller, 1999; Moreno et al., 2001; Lester, Convers, Stone, Kahler, & Barlow, 1997).

This selective review drew on results from six different ABLE environments in order to better understand the relationship between learners' emotions and these systems (see below for description of inclusion/ exclusion criteria): AutoTutor, Crytal Island, Operation ARIES!, MetaTutor, Prime Climb, and Wayang Outpost. A brief summary of each environment is provided below that includes their learning content, educational level, and distinguishing pedagogical features. A short overview of the programs of research that inform their development and design is also provided. ABLEs are presented alphabetically.

AutoTutor. AutoTutor is an ABLE designed to help college students learn about Newtonian physics, computer literacy, and critical thinking (Graesser & D'Mello, 2012; Graesser, et al., 2004; Graesser, Chipman, Haynes, & Olney, 2005; D'Mello & Graesser, 2010, 2011, 2013). AutoTutor scaffolds students' learning through a dialogue system with a PA that uses natural language processing and is organized into questions and problems that require students to engage in reasoning in providing explanations to the PA. AutoTutor deploys several tutorial approaches to facilitate learning, including providing feedback on learner's answers (e.g., "good job", "not quite"), pumping the learner for more information (e.g., "X is a type of what?"), correcting misconceptions, answering questions, and summarizing answers (D'Mello & Graesser, 2013). A newer version of AutoTutor that detects and responds to students' emotional states through empathetic and motivational responses as well as PA-embodied facial expressions and gross motor behaviors exists but is not included in the review of emotional states due to the inclusion/exclusion criteria in this selective review (D'Mello et al., 2010; D'Mello, Lehman, & Graesser, 2011). AutoTutor have been most notably used to investigate human-PA interactions, deep learning, and multi-method approaches to measuring, predicting, and responding to students' emotional states. AutoTutor's effectiveness to support learning has been examined in over twenty experiments to-date and produced effect sizes (improvements in learning from a pretest baseline) with an average of 0.8 sigma (D'Mello & Graesser, 2013).

Crystal Island. Crystal Island is a multi-agent, computer-based learning environment that deploys a rich, narrative-centered, inquiry-based approach to teaching eighth-grade microbiology and genetics (Huff & Nietfeld, 2009; McQuiggan, Lee, & Lester, 2007; McQuiggan & Lester, 2009; McQuiggan, Rowe, Lee, & Lester, 2008a; McQuiggan, Goth, Ha, Rowe, & Lester, 2008b; McQuiggan, Hoffman, Nietfeld, Robinson, & Lester, 2008c; McQuiggan, Robinson, & Lester, 2010; Mott & Lester, 2006; Robinson, McQuiggan, & Lester, 2009; Sabourin, Mott, & Lester, 2011; Sabourin, Shores, Mott, Bradford, & Lester, 2012; Lester, McQuiggan, & Sabourin, 2011). An additional distinguishing feature of Crystal Island is its game-like environment involving a developed storyline that puts the learner at the center of the narrative as a playable agent (i.e., avatar) and uses high-fidelity game engines including Valve Software's Source™ engine and more recently the Unity™ engine. This technology helps learners immerse themselves within the 3D ABLE that further incorporates scaffolds for effective learning such as metacognitive prompts for students' to monitor their progress toward their goal and embedded note-taking tools to assess and update in-game hypotheses. Crystal Island has an affect-aware system that delivers short, text-based empathetic responses (one or two sentences) to students based on their previously reported emotional state. Responses involve the PA either mirroring (parallel empathy) the learners' emotional state or drawing on contextual information (e.g., task success; reactive empathy) in order to elicit a more positive emotional state (McQuiggan & Lester, 2007; McQuiggan, Robison, & Lester, 2010). Subsequent research noted, however, that agents' responses to students' negatively-valenced emotional states (e.g., frustration) were rated significantly worse than responses to positively-valenced emotional states (e.g., excitement; Robison, McQuiggan, & Lester, 2009). Crystal Island has also been used to investigate inquiry-based learning, affect, engagement, perception of control, and a number of metacognitive and self-regulated learning strategies. A paired samples t-test comparing pre to post-test learning revealed that students scored significantly higher on a microbiology knowledge test following their interaction with Crystal Island (Rowe, Shores, Mott, &Lester, 2011).

Operation ARIES! Operation ARIES! (Lehman, et al., 2011; D'Mello, Lehman, Pekrun, & Graesser, 2014, Millis, Forsyth, Butler, Graesser, & Halpern, 2011) is a multi-agent ABLE and serious game designed to teach students about research methods through the presentation of case studies that are flawed and require the learner to evaluate their scientific merit. One of the novel features of this ABLE is its use of trialogues: conversations between the learner and two PAs; a peer agent (Chris) and a tutor/expert agent (Dr. Williams). The game-like elements of Operation ARIES! consist of a narrative and science fiction context in which aliens are seeking to conquer earth through the publication of bad research. Operation ARIES! has been most prominently used to induce confusion (through incorrect and / or conflicting messages from the two PAs) and to investigate this state's use as a pedagogical tool for deep learning (Lehman et al., 2011; D'Mello et al., 2014). Research has found that when contradictions were successful in inducing confusion they led better performance on posttest scores and transfer tasks than when learners were not presented with wrong or contradictory information from the two pedagogical agents (effects varied by experimental condition; Lehman et al., 2011).

MetaTutor. MetaTutor is a multi-PA, hypermedia learning environment that facilitates students learning of biological science content knowledge (e.g., the human circulatory system) and development of self-regulatory skills (Azevedo 2009; Azevedo, Behnagh, Duffy, Harley, & Trevors, 2012; Azevedo, & Chauncey-Strain, 2011; Azevedo, Moos, Johnson, & Chauncey, 2010; Azevedo, Behnagh, Duffy, Harley, & Trevors, 2012a; Azevedo et al., 2012b, 2013; Bouchet, Harley, Trevors, & Azevedo, 2013a; Duffy, Azevedo, Karabachian, & Dhillon, 2013; Feyzi-Beghnagh et al., 2013; Harley, Bouchet, & Azevedo, 2011, 2012, 2013a; Harley et al., 2013b; Harley, Taub, Bouchet, Henchey, & Azevedo, 2013c; Jaques, Conati, Harley, & Azevedo, 2014; Taub, Azevedo, Bouchet, & Khosravifar, 2014; Trevors, Duffy, & Azevedo, 2014; Trevors et al., 2013). Four distinct PAs are displayed asynchronously in the upper right-hand corner of the environment and provide different types of scaffolding to students during the learning session. More specifically: *Mary* supports students' metacognitive monitoring; *Sam* fosters students' use of self-regulated learning strategies (e.g., effective summarization, note-taking); *Pam* helps students set subgoals for the learning session and prompts them to activate their prior knowledge before the session; and *Gavin* guides students through the features of MetaTutor as well as administers self-report assessments of learning (e.g., subgoal and page quizzes, pre and post tests). MetaTutor has two different conditions that vary in the degree of scaffolding provided by the PAs. In the prompt and feedback condition, the PAs actively assist learners with the previously described behaviors, while the learner is left to regulate their learning in the control condition on their own (with the exception of a minimized version of Pam's role and Gavin normal role).

MetaTutor has been used to investigate students' use of self and co-regulated learning strategies, experience of emotions, gaze behavior, and the influence of trait emotions, goal orientations, and epistemic beliefs on learning (using proportion learning gains). MetaTutor has proven to be effective in improving students' scores on counter-balanced multiple-choice pre and post-tests, although results across studies suggest that PA behavior (experimentally manipulated by condition) does not have a significant effect on learning (Azevedo et al., 2012b; Bouchet, Harley, & Azevedo, 2013b; Harley et al., 2014).

Prime Climb. Prime Climb (Conati 2011; Conati & Maclaren 2009a,b; Conati & Manske, 2009; Conati & Zhao, 2004; Muir & Conati, 2012) is an educational game developed by Conati and Maclaren to help sixth and seventh grade students practice number factorization. The objective of the game is for students to work in pairs to climb a series of mountains that are divided into numbered sectors. The tools available to students in Prime Climb include a magnifying glass that allows students to see a number's factorization and a 3D model of a fully bodied cartoon PA (Merlin). The PA is autonomous and provides individual support both on demand (from a help box) and unsolicited when the student, according to the system's programming, does not seem to be learning.

The agent relies on probabilistic models of player's factorization knowledge, which is continuously updated throughout the game to determine when to make pedagogical interventions. This model allows the PA to intervene when it recognizes that the student is missing key pieces of information required to make a correct move or when after making a correct move, the success of the move was based on luck rather than knowledge. Conati and Maclaren (2009a) describe three different levels of hints that the agent provides: (1) reminders, for example, to think about number factorization or common factors; (2) a magnifying glass to see a number's factorization; and (3) the provision of common factors between two numbers and of number factorizations. In addition to these hints, the agent will attempt to stimulate reasoning if it suspects that a correct answer was a lucky guess (e.g., "Great, you are right this time. Do you know why?") and will occasionally encourage the student by congratulating him/her when he/she is successful. PrimeClimb has been used as an environment to detect emotions, examine gaze behavior, and build and test probabilistic models. Learning gains have been observed for different versions of PrimeClimb, the largest from student interactions with more sophisticated versions of the ABLE where hints are provided to students; PA hints have been shown in other studies to be related (when attended to) to better performance (Conati & Zhao, 2004; Muir & Conati, 2012).

Wayang Outpost. Wayang Outpost is an ITS environment that features storylines, animated characters, instructional videos, and problem solving hints, situated in a fictional island research station setting (Arroyo et al., 2007, 2009a; Arroyo, Woolf, Royer, & Tai, 2009b; Dragon et al., 2008; Woolf, 2010; Woolf et al., 2009, 2010). An animated PA, Jake or Jane, offer problem-solving advice and attempt to facilitate learners' experiences of positive emotions through empathetic strategies, such as mirroring participants' emotions and deploying emotionally supportive dialogue. An important characteristic of the PAs that facilitates their deployment of an empathetic strategy is that they are intended to be perceived as "study partners" or "learning companions" (LCs) rather than tutors or teachers. Accordingly, Jack and Jane are programmed to be non-invasive (i.e., minimally disruptive to the learner) and "work" on their own computers (PAs are animated, sitting in desks in the bottom right hand corner of the system interface) to solve the same problems as the student and to react only when the student finishes his/her problem. These reactions include embodied actions (such as body posture) and facial expressions, which allow the PA to display a range of emotions including confidence, excitement, boredom, etc. As LCs, the PAs mirror the learners' emotions using these embodied reactions as well as by deploying adaptive, context-dependent messages.

Wayang Outpost has been used to help prepare 12-16 year old students for standardized state exams in mathematics (geometry, statistics) with a 92% pass rate as compared to a 72% pass rate with students who did not use the ABLE (Woolf et al., 2009). Moreover, Wayang Outpost has been used by thousands of students and led to increased learning gains (M = 0.12 from pre-test to post-test) after two class periods (Woolf et al., 2009). Research with Wayang Outpost has been done in the classroom (computer lab setting) and examined the effectiveness of the ABLE's emotional interventions, approaches for detecting and modeling students emotions through the use of different methods (e.g., facial features from video data and physiological data), and the influence of PA scaffolding and characteristics (e.g., gender) on different demographics of students (e.g., low achieving students).

Emotions

Debate over how many different emotions exist and how they should be organized in relation to one another remains a popular topic in contemporary research on emotions (D'Mello, 2013; Pekrun, 2011). As a result, a variety of different discrete emotions are included in this review. A review of the theories that they stem from is beyond the scope of this paper, but the following sets of discrete emotions were used by researchers: universal, basic emotions (Ekman, 1992); OCC theory emotions (named after: Ortony, Clore, & Collins, 1998), academic achievement emotions (Pekrun, 2011), and learning-related emotions (D'Mello & Graesser, 2013; D'Mello et al., 2014). D'Mello and Graesser distinguish several of the learning-related emotions as cognitive-affective hybrids because of their strong relationship to cognitive features, such as comprehension. In this review, cognitive-affective emotions refer to these emotions (e.g., engagement, confusion). Adaptive emotions refer to those emotions (regardless of their membership to the aforementioned sets) that facilitate students' learning, such as engagement and curiosity. Conversely, non-adaptive emotions tend to interfere with learning (anger, boredom).

In this paper, we define emotions based on Pekrun's theoretical framework in which emotions are defined as multi-faceted phenomena involving coordinated psychological and physiological processes (affective, cognitive, psychological, motivational, and expressive components; Pekrun, 2006, 2011). Emotions are responses to situations (e.g., academic achievement situations such as a test or studying with an ABLE) that are perceived as relevant to an individual's current goals. Pekrun's control-value theory of achievement emotions has empirically demonstrated the influence of different appraisal mechanisms, namely appraisals of control and value, on the types of emotions experienced by students. Pekrun (2006, 2011) defines the term subjective value in the context of achievement emotions as the value of an activity and its outcome(s), and more broadly, as the perception that an action or outcome is positive or negative in nature. Individuals can make intrinsic appraisals of value, which refers to the inherent value of an activity, or extrinsic appraisals of value, which deal with the instrumental usefulness of actions or outcomes with regard to the attainment of a goal. Pekrun defines subjective control in the context of achievement emotions as one's control over achievement activities and their outcomes, or more generally, as one's perception of the causal influence they exert over their actions and outcomes (Pekrun, 2006, 2011).

Research Objectives

The primary objective of this review is to provide a critical analysis of the effectiveness of ABLEs in facilitating students' experience of adaptive emotions. This review differs from other reviews, meta-analyses, and surveys of the literature that examine emotions and computer-based learning environments (CBLEs) in the following ways. First, unlike a recent selective meta-analysis conducted by D'Mello (2013), this review is more selective by focusing only on CBLEs with PAs, thus excluding many intelligent tutoring systems and serious games that measure students' emotions, such as the Cognitive Tutor Algebra I and the Incredible Machine (Baker, D'Mello, Rodrigo, & Graesser, 2010). As such, comparisons are drawn between more similar learning environments. Second, this review is nonetheless more inclusive by examining any study that measured emotions using one or more methods so long as they met the criteria described below. As a result, this paper reviews a larger number of ABLEs (six rather than three; D'Mello, 2013). Third, this paper compares and contrasts learners' incidence of each of the discrete emotions reported for six ABLEs rather than creating aggregate scores across the environments. This approach facilitates a discussion of different features of these ABLEs and allows inferences to be made regarding why differences exist in learners' emotions between environments.

Methods

Inclusion and Exclusion Criteria

Seven studies were selected for analysis on the basis of the following inclusion criteria. First, studies had to measure more than one discrete emotional state using a forced-choice measure. Second, they had to report the incidence of emotions as either proportions or frequency. Third, in the case of multiple published articles based on the same or part of a common data set, the study with the larger sample size was taken. Studies that measured emotions using a dimensional framework (e.g., valence/ arousal) or only a single discrete emotion were excluded.

Representation of students' emotions across similar studies. To avoid redundancy and over-representation of some ABLEs, multiple studies were identified that reported learners' emotions by taking an average of learners' proportions of emotional states, as advocated by D'Mello and Graesser (2013). This approach was only taken when the emotions reported were similar. The described criteria allowed this review to extend the number of ABLEs examined from three in D'Mello's (2013) selective meta-analyses (AutoTutor, Operation ARIES!, and Crystal Island) to six (including MetaTutor, PrimeClimb, and Wayang Outpost). The emotions reported in Table 1 were primarily collected from self-report data or variations of self-report data (i.e., retrospective protocol) with three exceptions: Harley, Bouchet, and Azevedo (2013) used automatic facial expression coding, while Dragon et al., (2008) and D'Mello and Graesser (2013) used judges to code facial expressions and postures.

Results

Preliminary results revealed that a wide array of discrete emotional labels (25) were used to describe learners' emotional experiences with ABLEs, including basic, OCC, learner-centered, academic achievement, and other psychological states (e.g., neutral). The greatest amount of overlap in terms of the emotional states reported came from AutoTutor, Operation ARIES!, Crystal Island, and Wayang Outpost. A wider array of learner-centered emotions were reported by Harley et al. (2013a) with MetaTutor, but these states were excluded from this review because they were collected using Likert-type scales.

Table 1 was created to eliminate the redundancy of the large number of emotional labels by reducing the number of emotions to a set that could: (1) be operationalized as different emotional states and (2) that reduced the number of emotional labels while maintaining as much meaningful variation in learners' emotions as possible. This synthesis was completed using a set of definitions based on the research and operationalization of emotions by Pekrun (2011) and D'Mello, Lehman, and Person (2010). Emotions were therefore associated within the dimensions of valence and activation. Positively-valenced, activating emotions that were specifically related to learning and characterized as cognitive-affective states (e.g., curiosity, engagement) were grouped together because they represent ideal emotional states where the learner is not just feeling "good and energized" (e.g., delighted, happy), but in an emotional state where they are prepared to learn effectively.

Table 1.

Proportions of grouped discrete emotions experienced with ABLEs

			ABLE						
			AutoTutor (D'Mello & Graesser 2013)	Operation ARIES! (D'Mello, et al., 2014)	Crystal Island (McQuiggan, et al., 2010; Sabourin et al., 2011)		MetaTutor (Harley et al., 2013a)	Prime Climb (Conati & Maclaren, 2009a)	Wayang Outpost (Dragon, et al., 2008)
Val.	Act.	Emotion							
+	Act.	Happy/Joy/ Delight /Excitement	.06	.02	.25	.14	.09	.92	.34
+	Act.	Eng./Flow/ Focus/ Curiosity	.24	.24	.42	.41			
+	Act.	Admiration						.82	
+	De-Act.	Concentrated / Satisfied							.58

-	Act.	Anger/ Frustration	.13	.06	.07	.16	.03		.06
-	Act.	Fear/ Anxiety/ Distress		.01	.09	.05	.00	.08	
-	Act.	Disgust/ Contempt/ Reproach					.00	.18	
-	De-Act.	Boredom/ Tired	.18	.33	.03	.09			.02
-	De-Act.	Sadness			.02		.03		
+/-	Act.	Confusion	.17	.09	.13	.16			
+/-	Act.	Surprise	.03	.01			.03		
NA	Baseline	Neutral	.19	.26			.77		
+	Act.	-	.30	.26	.67	.55	.09	.92/.82	.34
+	De-Act.	-							.58
-	Act.	-	.13	.07	.16	.21	.03	.08/.18	.06
-	De-Act.	-	.18	.33	.05	.09	.03	-	.02
+/-	Act.?	-	.20	.10	.13	.16	-	-	
NA	Baseline	-	.19	.26	-	-	.77	-	

Note. Numbers represent the proportion of learners who reported experiencing an emotion. The proportions of reported discrete emotions are collapsed along the dimensions of activation and valence in the lower portion of the table. Emotions could add up to more than 100% if they possessed different object-focuses (e.g., PA [admiration/reproach] vs. event outcome [joy/distress]) (Conati & Maclaren, 2009a). Some emotions were pre-grouped by the authors including: concentrated/satisfied, excited (joyful, actively engaged), frustrated (angry), and bored (tired). Not all emotions could be consistently plotted within a dimensional framework that includes valence and arousal. For these emotions (confusion and admiration) the authors exercised their conceptualization of these emotions.

Discussion

Table 1 reveals that more than half of the emotions learners' experienced in Crystal Island and Prime Climb were positively-valenced and activating, while AutoTutor, Operation ARIES!, MetaTutor, and Wayang Outpost were associated with substantially lower levels of positively-valenced and activating affect (< 34%). In Crystal Island, the majority of these emotions were cognitive-affective states (e.g., engagement), which suggests that this environment facilitates the most adaptive emotional responses in learners. While it is better that students are joyful than distressed, the emotional reporting dichotomies in Prime Climb (admiration/reproach and joy/distress), on the other hand, unfortunately do not provide much information regarding the relevance of these emotional states to learning (when compared to cognitive-affective states). For example, students could be joyful (e.g., report that they feel "good" or "very good" about playing Prime Climb), but this only tells us that they are having fun, not whether they are feeling good because they are being educationally stimulated. For example, they might be having fun because they like the illustrations or gameplay rather than being engaged by the learning material. The situation is similar for learners' experience of happiness with

MetaTutor, although only a small proportion of learners' emotions fall into the positively-valenced, activating category. In contrast, AutoTutor and Operation ARIES!'s breakdown of emotions is similar to that of Crystal Island where most of the positive emotions are cognitive-affective states. Wayang Outpost was associated with a mixture of purely affective (e.g., joy, excited) and cognitive-affective states (concentrated, satisfied) that fit into this category, although the cognitive-affective states were classified as positively-valenced and de-activating rather than activating as in the other ABLEs.

Gamification and emotions in ABLEs

Some results may be explained by the game-like features of Crystal Island and Prime Climb, which were implemented with the purpose of making the games immersive and engaging in addition to educational. The high proportions of positively-valenced and activating emotional and cognitive-emotional states suggest that these features were implemented effectively with regard to their affective goals. In the other systems that do not have such features, it is therefore reasonable to expect a lower incidence of non-cognitive-affective states (e.g., delight, joy) because there are few features embedded within these systems to be "delighted" or "joyful" about from an appraisal point of view (e.g., students have limited control and are not likely to make high appraisals of value in these experimental studies; Harley et al., 2013a; Pekrun, 2011). One interesting caveat to this note is that Operation ARIES! is identified primarily as a serious game because of its narrative features about an alien invasion. These results suggest that narrative may not be enough to trigger positive emotional responses to an otherwise traditional (non-game feature enhanced) ABLE.

Wayang Outpost is unique in its measurement of positively-valenced, de-activating emotions (i.e., concentration/ satisfaction) which made up more than half of the emotions students reported, conforming with the previous explanation that lower activating positive emotional states are more expected in a non-serious game environment. Similar to Operation ARIES!, Wayang Outpost makes use of a fictional contextualization for the learning session (a research outpost), but whether this has a positive impact on students' emotional states is unknown because no studies have experimentally manipulated it's presence or absence.

It is important to note that when students were not experiencing positively valenced emotions, they were not necessarily experiencing maladaptive ones. For example, neutral states made up a considerable portion of the reported emotions in AutoTutor and Operation ARIES! and the majority of emotional states in MetaTutor. When interpreting what it means to have students in a neutral state while learning, it is important to remember that, while positive emotions are known to facilitate learning, their presence is not a requirement for students to learn (D'Mello et al., 2014; Harley et al., 2013). Therefore it is not practical to consider low levels or absence of positive emotions as evidence that something is amiss. Neutral emotions therefore may represent a state where students are not distracted by negative, activating emotions (e.g., frustration, anxiety), nor disengaged by negative, deactivating emotions (e.g., boredom), but are able to do what they need to do: learn.

Another interesting finding from Table 1 is that learners across ABLEs experienced small proportions of activating, negatively-valenced emotions such as anger or anxiety (<16%). Rather, the most disruptive emotional state students experienced was boredom. In Operation ARIES!, boredom was experienced more often than any other emotional state (24%) and was the third most frequently experienced state in AutoTutor (18%). In contrast, few instances of boredom were detected during learning sessions with Crystal

Island and Wayang Outpost. These results pertaining to boredom are likely related to the inclusion of effectively implemented game features in Crystal Island (e.g., 3D environment exploration) and exclusion of them in AutoTutor and Operation ARIES!. It is possible that learners in Wayang Outpost found the environment stimulating enough to avoid experiencing boredom, but not stimulating enough to be engaged or excited. Since other ABLEs did not measure the deactivating and positively valenced emotions, it is possible that some were coded as boredom or neutral, thus exaggerating the percentages. AutoTutor, Operation ARIES!, and Crystal Island reported fairly similar incidence rates of confusion, but research has shown that confusion can be beneficial and is indeed part of deep learning (D'Mello et al., 2014). Therefore, a state of confusion is not necessary a negative one, as per traditional perspectives, but simply an expected part of learning.

Appraisals and emotions in ABLEs

Aside from game features, another salient and differentiating characteristic is academic relevance (i.e., how relevant the learning material in the ABLE is to the student's academic performance and achievement). This feature is important because of its relationship to students' appraisals of task value (Pekrun, 2006, 2011). Specifically, the more directly related the content of an ABLE is to a student's academics (e.g., course, standardized examination), the more likely they are to value the task. Crystal Island, Wayang Outpost, and Prime Climb have content that is directly related to either a course learners are taking or to a required standardized exam. MetaTutor, and AutoTutor, on the other hand, do not and are more flexible with the participants they recruit. Therefore, with the exception of Operation ARIES!, the emotions experienced in these environments conform with expectations: learners' who interacted with ABLEs that were more directly tied to academic content experienced larger proportions of positive emotions. For example, mean levels of interest and excitement were higher in Wayang Outpost (M = 2.90 – 3.50; Arroyo, et al., 2009a) than the positively-valenced, activating self-reported emotions measured in MetaTutor (Ms = 2.53 – 2.93; Harley et al., 2013a).

Perceived control is a second appraisal dimension (Pekrun, 2006, 2011) that differs between these environments. Control is typically difficult to evaluate in ABLEs because it requires a nuanced understanding of the context, mechanics, programming rules, level of adaptivity, and levels of scaffolding of the different systems. In this section, perceived control will be discussed by considering the variations in choices that students have in each of the six ABLEs. Choice is not unique to a specific ABLE because each has at least some degree of choice (e.g., the order in which content is navigated). One ABLE that stands out from the others is Crystal Island which provides students with more varied and meaningful choices that influence their ability to complete the problem they are trying to solve (to identify the disease plaguing the island). These choices include moving their character to different settings (e.g., laboratory, residence), initiating conversations with multiple PAs, and selecting when to run tests and which tests to run in the laboratory. Students can also choose to ignore recommendations from the PAs, which is not always possible in the other ABLEs. The proportions and types of positively-valenced emotions, which characterize the emotions that students experience in this environment relative to AutoTutor, Operation ARIES!, and MetaTutor, conform with this explanation.

Prime Climb presents choice to students in terms of which square they select to move to, although this is a constrained set of options. It may be, however, that this is enough to foster a sufficient appraisal of control to not negatively impact learners' emotions or, alternatively, that that their appraisals of value concerning the ABLE due to its game-like (e.g., board game) features is sufficiently compensatory. Other hypermedia

ABLEs such as Wayang Outpost and MetaTutor provide students with a degree of choice regarding the learning tools and content they interact with. For example, students can chose to take notes or change pages or topics that they are learning about with MetaTutor (Azevedo et al., 2013). This more general level of choice may not be sufficient to elicit higher appraisals of choice and, subsequently, more positive emotional states though a controlled experimental manipulation would be needed to determine its impact.

Conclusion

In summary, it would seem that game-like elements, when implemented in a sufficient quantity (e.g., more than a narrative context) and with sufficient quality to make the environment truly game-like, is related to learners' experience of positive, activating emotions (Conati & Maclaren, 2009a; McQuiggan, Robison, & Lester, 2010; Sabourin, et al., 2011). Similarly, the relevance of content to students' academics and the affordance of choice in an ABLE also appear to be related to learners' experience of positive emotions (Conati & Maclaren, 2009a; Dragon et al., 2008; McQuiggan, Robison, & Lester, 2010; Sabourin, et al., 2011).

This selective review of ABLE environments also suggests that there is a range amongst them in terms of the incidence of desired emotions they induce in students (positively-valenced, activating). However, few negatively-valenced, activating emotions are observed by students while interacting with ABLEs which is encouraging. Instead, the greatest challenge for researchers to target in emotional interventions appears to be feelings of boredom.

Several of the ABLEs described in this selective review have versions that include affective feedback in which PAs respond to a learners' emotional state (e.g., frustration) with a prompt or feedback that is intended to improve how they feel. Unfortunately, the only study that was part of the comparative review of emotions between ABLEs was one of the two empirical papers on Crystal Island (Robison et al., 2009) and the proportions of emotions were similar to those of the second that did not contain affective feedback (McQuiggan, et al., 2010). Therefore, the affective feedback provided by PAs in Crystal Island in Robison et al.'s (2009) study did not seem to have a strong influence on learners' feelings while interacting with the ABLE. Other research with ABLEs has found evidence that affective feedback from PAs can be effective in supporting students' experience of positively valenced emotions, though some students benefit more than others (e.g., female students, low-prior knowledge students; Arroyo et al., 2013; D'Mello & Graesser, 2013).

Another finding of this selective review was that neutrality was found to be one of the most commonly appearing states in those environments that measured learners' experiences of neutral emotionality (and the most common in MetaTutor; D'Mello & Graesser, 2013; D'Mello et al., 2014; Harley, et al., 2013a). Therefore, future research to investigate neutral emotions is warranted because it is important to capture the range of students' emotional states, including those that may be considered non-emotional or baseline in nature. Overall, the present review suggests that it is better to be neutral than bored while interacting with ABLEs, and as such, our measurements of emotions should reflect these nuances.

Future directions: Toward an emotional assessment focus

The observations made in this review are the product of a qualitatively different type of review of ABLEs research than that found in other publications (e.g., D'Mello, 2013) in that it can be characterized as an *emotional assessment* approach. Such an approach is concerned with evaluating how learners are reacting to ABLEs and the different features within them. Whereas preliminary reviews have examined emotions across learning environments, the corresponding discussions typically do not go into sufficient detail and/or are based on aggregate emotional scores across systems (Baker, et al., 2010; D'Mello, 2013). More studies with forced-choice emotional labels and their incidence are needed, however, to validate and expand upon the number of environments reviewed and the samples they draw upon. Therefore, this review and its observations should be seen as recommendations for the development, evaluation, and refinement of current and future ABLEs. Moreover, the discussions in this review are based on observations of differences and similarities regarding how learners felt interacting with six ABLEs and the features these environments possessed. Further research where these features are experimentally controlled and manipulated would be needed to make causal claims between ABLE features (e.g., game-like elements) and learner emotions.

As research with ABLEs continues to advance, emotional assessment approaches will also become necessary for developing and validating frameworks to design emotional interventions that take contextual information into consideration, including: (1) the emotional state of a learner at the time the system decides to intervene; (2) learners' profile characteristics (e.g., personality, trait emotions, prior domain knowledge, age, ability to accurately monitor and regulate their self-regulatory processes); (3) learners' current and emerging understanding of learning material; (4) the role and function of the PA (e.g., modeling knowledge and skills, providing scaffolding, competing for resources, cooperating on task); and, (5) learners' disposition toward the PA(s) (e.g., contempt, curiosity).

Acknowledgements. This research was supported by a Joseph-Armand Bombardier Canada Graduate Scholarship for Doctoral research from the Social Science and Humanities Research Council (SSHRC) of Canada awarded to the first author. The authors would like to thank Nathan C. Hall, Susanne P. Lajoie, and Alenoush Saroyan for their feedback on this manuscript.

References

Andre, E., Rist, T., & Muller, J. (1999). Employing AI methods to control the behavior of animated interface agents. *Applied Artificial Intelligence, 13*, 415–448.

Arroyo, I., Burleson, W., Tai, M., Muldner, K., & Woolf, B (2013). Gender differences in the use and benefit of advanced learning technologies for mathematics. *Journal of Educational Psychology, 105*(4), 957-969.

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D.,…Wolf, P. B. (2007). Repairing disengagement with non-invasive interventions. In R. Rosemary, K. R. Koedinger, & J. Greer (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 195-202). Amsterdam, Netherlands: IOS Press.

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

Arroyo, I., Woolf, B. P., Royer, J. M., & Tai, M. (2009a). Affective gendered learning companions. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 41-48). Amsterdam, Netherlands: IOS Press.

Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning, 4*(1), 87-95.

Azevedo, R., & Aleven, V. (Eds.). (2013). *International handbook of metacognition and learning technologies*. Amsterdam, The Netherlands: Springer.

Azevedo, R., Behnagh, R., Duffy, M., Harley, J., & Trevors, G. (2012a). Metacognition and self-regulated learning with advanced learning technologies. In D. Jonassen &

S. Land (Eds.), *Theoretical foundations of learning environments* (2nd ed.) (pp. 171-197). Mahwah, NJ: Erlbaum.

Azevedo, R., & Chauncey-Strain, A., (2011). Integrating cognitive, metacognitive, and affective regulatory processes with MetaTutor. In R. A. Calvo & S. K. D'Mello (Eds.), *New Perspectives on Affect and Learning Technologies* (Vol. 3, pp. 141-154). New York, NY: Springer-Verlag.

Azevedo, R., Harley, J., Trevors, G., Feyzi-Behnagh, R., Duffy, M., Bouchet, F., & Landis, R.S. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (Vol. 26, pp. 427-449). New York, NY: Springer.

Azevedo, R., Landis, R.S., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J…Hossain, G. (2012b). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. In S. A. Cerri, W. J. Clancey, G. Papadourakis, and K. Panourgia (Eds.), *Lecture Notes in Computer Science: Vol. 7315. Intelligent Tutoring Systems (pp. 212-221).* Berlin, Heidelberg: Springer-Verlag.

Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist*, *45*, 210 - 223.

Baker, R. S. D., D'Mello, S. K., Rodrigo, M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223-241.

Baylor, A. L., & Kim, Y. (2009). Designing non-verbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, *25*, 450-457.

Bouchet, F., Harley, J. M., & Azevedo, R. (2013a). The impact of different pedagogical agents' adaptive self-regulated prompting strategies with MetaTutor. In C. H. Lane, K. Yacef, J. Mostow, P. Pavik (Eds.), *Lecture Notes in Computer Science: Vol. 7926. Artificial Intelligence in Education* (pp. 815-819). Berlin, Heidelberg: Springer-Verlag.

Bouchet, F., Harley, J.M., Trevors, G., & Azevedo, R. (2013b). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining, 5*(1), 104-146.

Calvo, R.A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1,*18-37.

Calvo R. A., & D'Mello, S. (Eds.) (2011). *New Perspectives on Affect and Learning Technologies*. New York: Springer.

Calvo, R. A., & D'Mello, A. C. (2012). Frontiers of affect-aware learning technologies. *Intelligent Systems*, *27(6)*, 86-89.

Conati, C. (2011). Combining cognitive appraisal and sensors for affect detection in a framework for modeling user affect new perspectives on affect and learning technologies. In R. A. Calvo & S. D'Mello (Eds.), *New perspectives on affect and learning technologies* (pp. 71-84). New York: Springer-Verlag.

Conati, C., & Maclaren, H. (2009a). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction, 19*, 267-303.
Conati, C., & Maclaren, H. (2009b). Modeling user affect from causes and effects. In G.J. Houben, G. McCalla, F. Pianesi, & M. Zancanaro (Eds.), *Lecture Notes in Artificial Intelligence: Vol. 5535. User Modeling, Adaptation, and Personalization* (pp. 4-15). Berlin, Germany: Springer-Verlag.
Conati, C. & Manske, M. (2009). Evaluating adaptive feedback in an educational computer game. In Z. Ruttkay, M. Kipp, A. Nijholt & H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents*: Vol. 5773 (pp. 146-158). Berlin, Heidelberg: Springer-Verlag.
Conati, C., & Zhao, X. (2004). Building and evaluating an intelligent pedagogical agent to improve the effectiveness of an educational game. *Proceedings of the 9th International Conference on Intelligent User Interfaces* (pp. 6-13). New York, NY: ACM.
Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*(2), 428.
D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology. 105(4),* 1082-1099.
D'Mello, S., & Graesser, A. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction, 20,* 147-187.
D'Mello, S., & Graesser, A.C. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion, 25*, 1299-1308.
D'Mello, S. K., & Graesser, A. C. (2013). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems, 2*. Advance online publication. doi: 10.1145/2395123.2395128
D'Mello, S. K., Lehman, B., & Graesser, A. (2011). A motivationally supportive affect-sensitive AutoTutor. In R. A. Calvo & S. K. D'Mello (Eds.), *New Perspectives on Affect and Learning Technologies* (pp. 113-126). New York: Springer.
D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction,* 29, 153-170.
D'Mello, S. K., Lehman, B., & Person, N. (2010). Monitoring affective states during effortful problem solving activities. *International Journal of Artificial Intelligence in Education, 20,* 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 V. Aleven, J. Kay & J. Mostow (Eds.), *Lecture Notes in Computer Science: Vol. 6094. Intelligent Tutoring Systems* (pp. 245-254). Berlin, Heidelberg: Springer-Verlag.
Duffy, M., Azevedo, R., Karabachian, L., & Dhillon, I. (2013, April). *Motivation matters: Interactions between achievement goals and scaffolding for self-regulated learning within a multi-agent intelligent tutoring system*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Dragon, T., Arroyo, I., Woolf, B., Burleson, W., El Kaliouby, R., & Eydgahi, H. (2008). Viewing student affect and learning through classroom observation and physical sensors. In B. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Lecture Notes in Computer Science: Vol 5091. Intelligent Tutoring Systems* (pp. 29-39). Berlin, Germany: Springer-Verlag.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*, 169-200.

Feyzi-Behnagh, R., Azevedo, R., Legowski, E., Reitmeyer, K., Tseytlin, E., & Crowley, R.S. (2013). Metacognitive scaffolds improve self-judgments of accuracy in a medical intelligent tutoring system. *Instructional Science, 42,* 159-181.

Feyzi-Behnagh, R., Trevors, G., Azevedo, R., Yang, W., Bélanger-Cantara, V., Henchey, J.,Bouchet, F., Pacampara, N., & Wang, G. (April, 2013). *Understanding multimedia learning by converging process and product data*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Graesser, A., & D'Mello, S. K. (2012). Emotions during the learning of difficult material. In B. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 57: pp. 183-226). San Diego, CA: Elsevier.

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

Graesser, A, Lu, S., Jackson, G., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods*, *36*(2), 180-192.

Harley, J. M. (in press.). Measuring emotions: A survey of cutting-edge methodologies used in advanced agent-based learning environment research. In S. Tettegah (Ed.). *Emotions and technology: Communication of feeling for, with and through digital media*. Elsevier.

Harley, J.M, Bouchet, F., Azevedo, R. (2011). Examining learner's emotional responses to pedagogical agents' tutoring strategies. In H. Högni Vilhjálmsson, S. Koop, S. Marsella, and K. R. Thórisson (Eds.), *Lecture Notes in Artificial Intelligence: Subseries of Lecture Notes in Computer Science: Vol. 6895. Intelligent Virtual Agents* (pp. 449-450). Berlin, Heidelberg: Springer-Verlag

Harley, J.M, Bouchet, F., & Azevedo, R. (2012). Measuring learners' co-occurring emotional responses during their interaction with a pedagogical agent in MetaTutor. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.). *Lecture Notes in Computer Science: Vol. 7315*. *Intelligent Tutoring Systems* (pp. 40-45). Berlin, Heidelberg: Springer-Verlag.

Harley, J. M., Bouchet, F., & Azevedo, R. (2013a). Aligning and comparing data on learners' emotions experienced with MetaTutor. In C. H. Lane, K. Yacef, J. Mostow, P. Pavik (Eds.), *Lecture Notes in Computer Science: Vol. 7926. Artificial Intelligence in Education* (pp. 61-70). Berlin, Heidelberg: Springer-Verlag.

Harley, J. M., Bouchet, F., Papaionnou, N., Carter, C., Azevedo, R., & Landis, R. (2014, April). Assessing learning with MetaTutor, a multi-agent hypermedia learning environment. In M.W. Chu, & J. M. Harley (Chairs), *Innovative practices for assessment in computer based learning environments.* Symposium conducted at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Harley, J. M., Carter, C., Papaioannou, N., Bouchet, F., Azevedo, R., & Landis, R. (2013b, April). *Examining learners' academic achievement emotions during science learning with an intelligent, hypermedia multi-agent system*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Harley, J. M., Carter, C.K., Papaionnou, N., Bouchet, F., Azevedo, R., Landis, R. L., & Karabachian, L.R. (under review). Examining the predictive relationship between personality and emotion traits and learners' agent-directed emotions.

Harley, J. M., Taub, M., Bouchet, F., Henchey, J., & Azevedo, R. (2013c, April). *Profiling learners' co-regulation patterns with a pedagogical agent in an intelligent tutoring system for learning about human biology*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Huff, J., & Nietfeld, J. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*(2), 161-176.

Jaques, N., Conati, C., Harley, J. M., & Azevedo, R. (2014). Predicting affect from gaze behavior data during interactions with an intelligent tutoring system. In S. Trausan-Matu., K. Boyer., M. Crosby., K. Panourgia (Eds.), *Lecture Notes in Computer Science: Vol. 8474. Intelligent Tutoring Systems (pp. 629-631)*. Berlin, Heidelberg: Springer-Verlag.

Kinnebrew, J. S, Biswas, G., Sulcer, B., & Taylor, R. S. (2013). Investigating self-regulated learning in teachable agent environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (Vol. 26, pp. 451-470). New York, NY: Springer.

Lehman, B., D'Mello, S. K., Strain, A. C., Gross, M., Dobbins, A., Wallace, P., . . . Graesser, A. C. (2011). Inducing and tracking confusion with contradictions during critical thinking and scientific reasoning. In G. Biswas, S. Bull, J. Kay & A. Mitrovic (Eds.), *Artificial Intelligence in Education:* Vol. 6738. (pp. 171-178). Berlin Heidelberg: Springer-Verlag.

Lester, J., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. (1997). The persona effect: Affective impact of animated pedagogical agents. In C. Ware & D. Wixon (Eds.). Proceedings of the Conference on Human Factors in Computing Systems (pp. 359–366). New York, NY: ACM.

Lester, J. C., McQuiggan, S. W., & Sabourin, J. L. (2011). Affect recognition and expression in narrative-centered learning environments. In R. A. Calvo & S. D'Mello (Eds.), *New Perspectives On affect and Learning Technologies* (pp. 85-96). New York: Springer-Verlag.

McQuiggan, S., Goth, J., Ha, E., Rowe, J., & Lester, J. (2008b). Student note-taking in narrative-centered learning environments: Individual differences and learning effects. In B. Woolf, E. Aïmeur, R. Nkambou & S. Lajoie (Eds.), *Lecture Notes in Computer Science: Vol. 5091. Intelligent Tutoring Systems* (pp. 510-519). Berlin, Germany: Springer-Verlag.

McQuiggan, S. W., Hoffman, K. L., Nietfeld, J. L., Robison, J. L., & Lester, J. C. (2008c). *Examining self-regulated learning in a narrative-centered learning environment: An inductive approach to modeling metacognitive monitoring.* Paper presented at the World Conference on Educational Multimedia, Hypermedia, and Telecommunications, Vienna, Austria.

McQuiggan, S., Lee, S., & Lester, J.. (2007). Early prediction of student frustration. In A. Paiva, R. Prada & R. Picard (Eds.), *Lecture Notes in Computer Science*: *Vol. 4738. Affective Computing and Intelligent Interaction* (pp. 698-709): Berlin, Heidelberg: Springer-Verlag.

McQuiggan, S., Rowe, J., Lee, S., & Lester, J. (2008a). Story-based learning: The impact of narrative on learning experiences and outcomes. In B. Woolf, E. Aïmeur, R. Nkambou & S. Lajoie (Eds.), *Lecture Notes in Computer Science: Vol. 5091. Intelligent Tutoring Systems* (pp. 530-539). Berlin, Germany: Springer-Verlag.

McQuiggan, S. W., & Lester, J. C. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies, 65*, 348-360.

McQuiggan, S. W., & Lester, J. C. (2009). Modeling affect expression and recognition in an interactive learning environment. *International Journal of Learning Technology 4*, 216-233.

McQuiggan, S.W., Robison, J. L., & Lester, J. C. (2010). Affective transitions in narrative-centered learning environments. *Journal of Educational Technology & Society*, *13*(1), 40-53.

Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A., & Halpern, D. (2011). Operation ARIES!: A serious game for teaching scientific inquiry serious games and edutainment applications. In M. Minhua, A. Oikonomou, & L. C. Jain (Eds.), *Serious Games and Edutainment Applications* (pp. 169-195). London, UK: Springer.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, *19*(2), 177-213.

Mott, B., & Lester, J. (2006). Narrative-centered tutorial planning for inquiry-based learning environments. In M. Ikeda, K. Ashley & T.-W. Chan (Eds.), *Lecture Notes in Computer Science: Vol.* 4053. *Intelligent Tutoring Systems* (pp. 675-684). Berlin, Heidelberg: Springer-Verlag.

Muir, M., & Conati, C. (2012). An analysis of attention to student-adaptive hints in an educational game. In S. A. Cerri, W. J. Clancey, G. Papadourakis, and K. Panourgia (Eds.), *Lecture Notes in Computer Science: Vol: 7315. Intelligent Tutoring Systems (pp.* 112-122*).* Berlin, Heidelberg: Springer-Verlag.

Ortony, A., Clore, G.L., Collins, A. (1988). *The cognitive structure of emotions.* Cambridge, UK: Cambridge University Press.

Pekrun, R. (2011). Emotions as drivers of learning and cognitive development. In R. A. Calvo & S. D'Mello (Eds.), *New Perspectives on Affect and Learning Technologies* (pp. 23-39). New York: Springer.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18,* 315-341.

Pekrun, R., Daniels, L. M., Perry, R. P., Goetza, T., & Stupnisky, R. H. (2010). Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology, 102,* 531-549.

Pekrun, R., Goetz, T., Frenzel, A.C., Petra, B., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 34-48.

Pekrun, R. & Linnenbrink-Garcia, L. (2014). *International Handbook of Emotions in Education.* Hoboken: Taylor and Francis.

Robison, J., McGuiggan, S. W., & Lester, J. (2009). Evaluating the consequences of affective feedback in intelligent tutoring systems. In J. Cohn, A. Nijholt, & M. Pantic (Eds.). *Proceedings of the International Conference on Affective Computing & Intelligent Interaction* (pp. 37-42). Amsterdam, The Netherlands: IEEE Press.

Sabourin, J., Mott, B., & Lester, J. C. (2011). Modeling learner affect with theoretically grounded dynamic Bayesian networks. In S. D'Mello, A. Graesser, B. Schuller & J.-C. Martin (Eds.), *Affective Computing and Intelligent Interaction*: *Vol. 6974* (pp. 286-295). Berlin Heidelberg: Springer-Verlag.

Sabourin, J., Shores, L. R., Mott, B. W., & Lester, J.C. (2012). Predicting student self-regulation strategies in game-based learning environments. In S. Cerri, W. Clancey, G. Papadourakis & K. Panourgia (Eds.), *Lecture Notes in Computer Science: Vol. 7315. Intelligent Tutoring Systems* (pp. 141-150). Berlin, Heidelberg: Springer-Verlag.

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014) Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior*, *39,* 356-367.

Trevors, G., Duffy, M., Azevedo, R. (2014). Note-taking within MetaTutor: interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. *Educational Technology Research and Development*, 1-22.

Trevors, G., Feyzi-Behnagh, R., Azevedo, R., Yang, W., Henchey, J., Bélanger-Cantara, V., Bouchet, F., Wang, G., & Pacampara, N. (April, 2013). *Eye-movement patterns in science multimedia as a function of epistemic beliefs and learning task conditions*. Paper presented at a roundtable at the annual meeting of the American Educational Research Association, San Fransico, USA.

Woolf, B. P. (2010). Social and caring tutors. In V. Aleven, J. Kay, & J. Mostow (Eds.). *Lecture Notes in Computer Science: Vol. 6094. Intelligent Tutoring Systems* (pp. 5-14). Berlin, Heidelberg: Springer-Verlag.

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,* 129-164.

Woolf, B., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., & Christopherson, R. M. (2010). The effect of motivational learning companions on low-achieving students and students with learning disabilities. In V. Aleven, J. Kay, & J. Mostow (Eds.). *Lecture Notes in Computer Science: Vol. 6094. Intelligent Tutoring Systems* (pp. 327-337). Berlin, Heidelberg: Springer-Verlag.