The Link Between Achievement Emotions, Appraisals and Task Performance: Pedagogical Considerations for Emotions in CBLEs

Amanda Jarrell[1]
Jason M. Harley[2]
Susanne P. Lajoie[1]

[1] Department of Educational and Counselling Psychology, McGill University, 3700 McTavish Street 614, Montre´al, QC H3A 1Y2, Canada
[2] Department of Educational Psychology, University of Alberta, 6-102 Education North, Edmonton, AB T6G 2G5, Canada

Amanda Jarrell
amanda.jarrell@mail.mcgill.ca
Jason M. Harley
jharley1@ulaberta.ca
Susanne P. Lajoie
susanne.lajoie@mcgill.ca

Abstract. Achievement emotions have a powerful influence on how students interact with current and future learning and performance tasks. As such, pedagogical practices that support adaptive student emotions are critical for teaching and learning in computer-based learning environments (CBLEs). This research investigates the relationship between during-task achievement emotions and participants' appraisals of task control, value, perceived performance and actual performance outcomes on a diagnostic reasoning task with a CBLE, XX. Based on the emotions participants reported experiencing during the task, we found that participants could be organized into three groups using a *k*-means cluster analysis: a positive, negative, and low emotion group. Participants assigned to the positive emotion group had the highest subjective appraisals of task value, task control, and the highest perceived performance; however, these participants had lower levels of actual performance when compared to learners assigned to the low emotion cluster, and had actual performance levels comparable to learners in the negative emotion cluster. These results provide preliminary evidence for fostering low emotionality rather than positive emotionality with pedagogical interventions in order to support better performance outcomes while learners engage in academic achievement tasks in CBLEs.

Emotions, affect, performance, computer-based learning environments, problem solving, medical decision making

Introduction

Achievement emotions profoundly influence how students interact with learning and performance activities including their goals, strategy use and persistence (Pekrun & Perry, 2014; Linnenbrink-Garcia & Barger, 2014; Pekrun, 2006). In addition, emotions experienced during educational endeavors shape subsequent behaviors, goals and emotions (Pekrun & Perry, 2014). Given the immediate and long-term implications emotions have on learning and achievement, it is important that instructors and system developers support adaptive achievement emotions during learning and performance tasks. For this to be possible, a better understanding of the antecedent factors that lead to specific achievement emotions and how these emotions interact with learning and performance is needed. The purpose of this study was to examine the emotions students experienced during authentic problem solving in a computer-based learning environment (CBLE), XX (Authors, 2009; Authors, 2015c), in relation to their subjective appraisals of task control, value, perceived performance and actual performance. Our work expands upon previous research on achievement emotions in CBLEs in two important ways: (1) We analyze achievement emotions using a person-centered approach to address the paucity of research using such analytic practices when studying emotions, and (2) we examine how multiple achievement emotions and their relative intensity relate to learner appraisals and task performance in a CBLE.

Theoretical Framework: The Control-Value Theory of Achievement Emotions

The control-value theory of achievement emotions (Pekrun & Perry, 2014; Pekrun, 2006) was used as the guiding theoretical framework for this study. According to this theory, achievement emotions are goal-directed and appraisal-driven multi-componential psychological processes that are tied to achievement activities and mediate effective learning (Pekrun & Perry, 2014). Emotions can be organized according to valence and arousal where valence refers to the intrinsic pleasantness or unpleasantness of an emotional experience and arousal refers to physiological activation or deactivation (Pekrun & Perry, 2014; Russell, Weis, & Mendelsohn, 1989). Specific emotions are initiated from individual appraisal or evaluations of a situation, in other words emotions have a particular object-focus. In the context of achievement activities, students will appraise the achievement activity and outcomes according to subjective appraisals of control and value. Subjective control refers to individual evaluations of agency over the achievement activity and its outcomes and subjective value is the individual evaluation of the importance of the activity and its outcomes (Pekrun, 2006). A combination of the object-focus and appraisal processes will elicit different types of achievement emotions, including prospective outcome emotions, activity emotions and retrospective outcome emotions. The focus of this paper is on prospective outcome and activity emotions. These emotions can be elicited during the achievement task depending on whether the learner is evaluating the outcome (e.g., hope) or the activity (e.g., enjoyment). The intensity of these emotions can also vary as a function of appraisal processes and individual differences such as culture (Fenzel, Thrash, Pekrun, & Goetz, 2007; Pekrun, 2006; Pekrun & Perry, 2014). Thus the valence, activation and intensity of an emotional experience is caused by the interaction between subjective appraisals of control and value. For example, a student who positively values a learning task and feels that they are in control of their learning outcome will experience enjoyment, whereas a student who does not value the learning task will experience boredom (Pekrun, 2006; Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010).

Emotions During Achievement Activities

The optimal emotional conditions to increase learning gains and improve performance outcomes stand to vary partly because of the different cognitive processes involved in learning and performance activities. Simply put, learning relies on encoding new information into memory and performance involves retrieving information from memory (Mayer, 2012). Emotions have the potential to impact these underlying processes and thereby learning and performance. With regards to learning, the control value theory of achievement emotions proposes that emotions differentially impact learning strategies and self-regulation of learning (Pekrun, 2006; Zimmerman & Labuhn, 2012). Activating positive emotions are considered adaptive because they facilitate the use of flexible learning strategies and meta-cognitive strategies for adapting learning to individual goals, whereas negative deactivating emotions can lead to learning strategies that result in surface cognitive processing and a reliance on external guidance (Pekrun, 2006). For example, enjoyment during learning has been associated with better academic outcomes because these positive emotions have been found to maintain cognitive resources during an achievement task (Pekrun, Goetz, Titz, & Perry, 2002), whereas boredom is associated with poorer outcomes because this negative emotion detracts from the learning task (Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010; Pekrun, Goetz, Titz, & Perry, 2002). Similarly, emotions influence performance in meaningful ways because positive emotions can facilitate information retrieval and preserve cognitive resources, while negative emotions can inhibit information retrieval and direct attention to task irrelevant thoughts (Pekrun & Perry, 2014). The interaction between emotions and cognition can result in better performance outcomes in the case of positive activating emotions, such as enjoyment and hope (Pekrun, Goetz, Titz, & Perry, 2002) and poorer performance outcomes for negative emotions, such as anxiety (Zeidner, 1998; Cassady, 2004; Zeidner, 2014).

Over the course of a single achievement task, such as when solving a difficult problem (D'Mello & Graesser, 2012) or learning complex material (Authors, 2015a, 2016a), a single student will experience many different emotions, often at different intensity levels. During these activities a single student might therefore experience both positive and negative emotions at different points. Thus, asking questions pertaining to how a single emotion (e.g., anxiety) relates to outcomes might have little utility when in reality a student will experience multiple emotions (e.g., anxiety, frustration, boredom, and enjoyment). It is therefore valuable to ask questions relating to how the profile of a students' emotional experience compares to the profile of another students' emotional experience in order to determine optimal conditions for learning and achievement.

Emotions and Instructional Interventions in CBLEs

The functionalities of CBLEs can be leveraged to support individual student's emotional needs to optimize learning and performance outcomes. This is because CBLEs have the potential to foster adaptive emotionality by targeting the antecedent factors of emotional responses, such as individual appraisals of control and value. This can be accomplished by, for instance, environmental design, scaffolding and adaptive feedback. Appraisals of task value and control can be bolstered by designing meaningful and authentic tasks with the capability of providing learners with autonomy. For example, CBLEs can be designed to target a particular group of students (e.g. medical students) and construct learning tasks that are both meaningful for their learning (e.g. diagnostic reasoning) and authentic by reflecting real-world applications (e.g. medical decision making in a hospital setting). Perceptions of control over achievement outcomes can be reinforced by providing students the ability to control their navigation through the environment, access additional information (e.g. a glossary of medical diseases and diagnostic tests) and receive assistance (e.g. a help option). However, there is a paucity of empirical work exploring the direct link between emotions, appraisal processes and outcomes, and as such empirically grounded instructional recommendations are tentative.

Research on emotions in CBLEs is not without challenges. In these environments emotions are typically reported with low intensities, particularly negative emotions (D'Mello 2013; Authors, 2014; Authors 2015a). These low intensities can restrict statistical variance, especially through floor effects, thus making it difficult to understand the nuanced relationship between emotions and their impact on learning and performance. This is a unique challenge because even moderate levels of negative emotions stand to impact achievement outcomes detrimentally.

Person-Centered Analyses of Emotions

Current examinations of emotions in achievement contexts have generally relied on variable-centered approaches whereas person-centered (i.e. intra-individual) investigations have been underrepresented (Pekrun, 2006). This is potentially problematic because the average emotional responses of a group of students might not represent a single student within that group (Pekrun, 2006). Since one of the goals for studying emotions in achievement contexts is to understand how to support *individual* students, an important direction for this area of research is to study the emotions using person-centered approaches.

Person-centered analyses place the individual rather than the variable at the focus of the analysis allowing researchers to investigate how psychological constructs are manifested within individuals rather than between variables (Linnenbrink-Garcia & Barger, 2014). In other words, rather than examining the distribution across or between groups of learners for a single emotion (e.g., anxiety), person-centered analyses allow us to identify trends (i.e., profiles) in the constellation of state emotions and their experience at varying levels of intensity (e.g., low anxiety and low enjoyment).

Person-centered analyses (e.g. cluster analyses) are gaining interest in emotion research (Pekrun & Hofmann, 1996; Lazarus, 2006; Martinent, Nicolas, Gaudreau, & Campo, 2013; Authors, 2015). These types of analyses are also particularly well-suited for addressing the unique challenges associated with studying emotions in CBLEs (i.e. floor effects and low variance). The *k*-mean cluster analysis is one person-centered approach that might enable researchers to overcome these challenges. This analysis requires that all variables are standardized prior to analysis, permitting a relative comparison between participants with similar emotion profiles. The current research employs this method to uncover the emotional profiles of learners who solved authentic science problems in XX (Authors, 2009; Authors, 2015c).

Current Study

This study sought to answer three research questions. (1) *Can learners be grouped according to the emotions experienced during problem solving?* (2) *Is there a relationship between emotion groups and task control and value appraisals?* And (3), *Is* *there a relationship between emotion groups and performance?* With regard to our first research question we hypothesized, based on previous work (Authors, 2015b), that participants' self-reported emotions would form three clusters: a positive, negative, and a low emotion cluster. For our second research question, we hypothesized that participants with overall positive emotions would have significantly higher task control and value appraisals than participants with overall negative emotions based on the the control-value theory of achievement emotions (Pekrun, 2006). For our third research question we developed two hypotheses based on participants' perceived performance and actual performance. First, and with regard to *perceived* performance, we hypothesized that participants with overall positive emotions would perceive their performance to be significantly better than participants with overall negative emotions based on higher appraisals of goal attainment (D'Mello, Lehman, Pekrun & Graesser, 2012; Carver & Scheier, 2014).We also hypothesized, based on prior research (Authors, 2015b) and the net effects of emotions on achievement (Pekrun & Perry, 2014; Pekrun, 2006) that learners with overall positive emotions would outperform learners with overall negative emotions and that learners with overall low emotions would have performances that fell between the two.

Methods

Participants

Participants (*N* = 26) from one large North American public university were pre-medical (*N* = 1) and medical students (*N* = 25) with a mean age of 24.40 (*SD* = 3.43).

Learning Environment

XX is a CBLE designed to help medical students effectively diagnose a patient through a diagnostic simulation. Medical students apply what they have learned in medical school to authentic diagnostic problems. Each problem begins with a patient history, which provides details on the case including relevant symptoms and relevant patient details (e.g. age; see Figure 1). Students propose initial hypotheses based on the evidence gathered in the patient history. Students can obtain further evidence by ordering laboratory tests that confirm or disconfirm a particular hypothesis, search for information using the online library, and request help using the consultation tool. After the final diagnosis is submitted, students receive individualized feedback on their solution based on an aggregated expert solution (see Figure 2). Participants are also provided an efficiency score which represents numerically how well their solution matched with the expert solution.

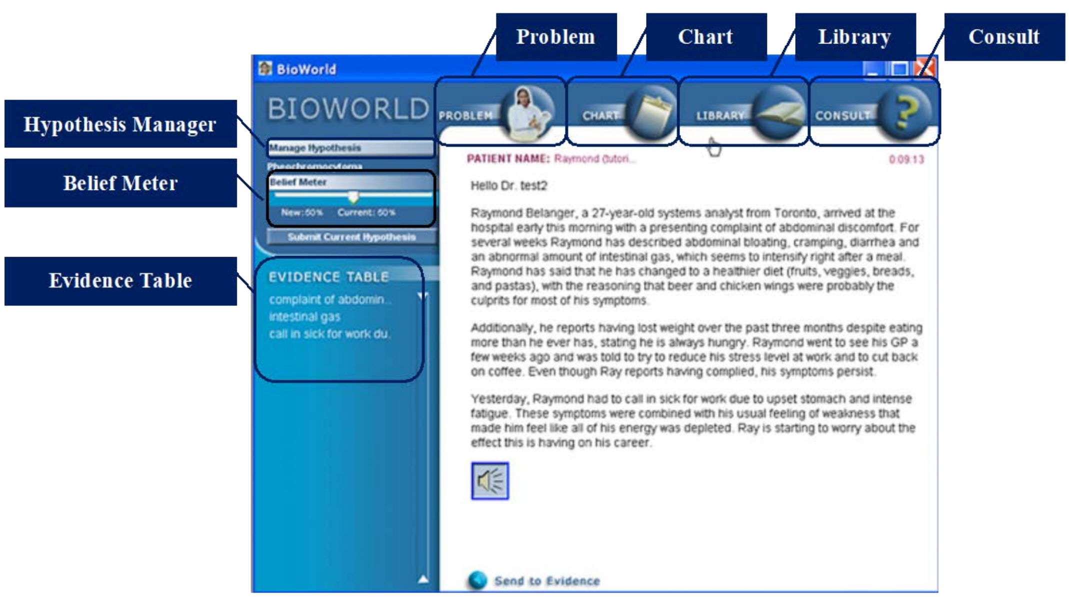

Figure 1. Screenshot of Primary Interface: Case Summary Tab.

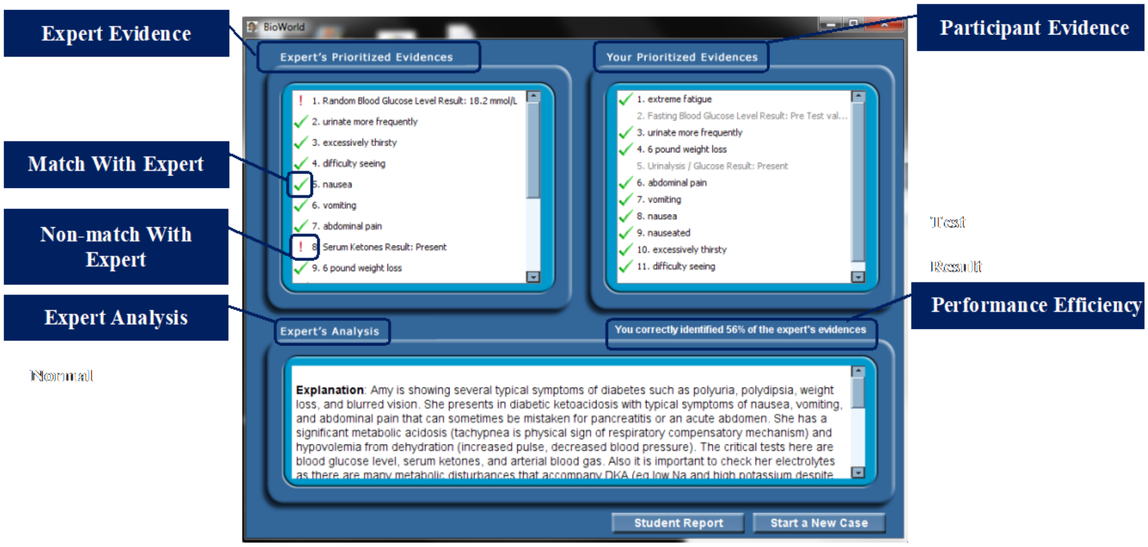

Figure 2. Screenshot of XX Interface: Expert Feedback Interface

Measures

Academic Achievement Emotion Questionnaire (AEQ). Due to the performance nature of solving diagnostic problems in XX, the academic achievement emotions questionnaire (AEQ) concurrent state emotions subscale for *test* emotions (AEQ; Pekrun, Goetz, Titz, & Perry, 2002) was used to measure the emotions learners' experienced *while* solving the diagnostic reasoning task (as opposed to concurrent studying or class-related emotions). The AEQ concurrent state test-taking emotions subscale consists of 27 items and measures enjoyment (3 items), pride (2 items), hope (2 items), anxiety (7 items), hopelessness (6 items), shame (5 items), and anger (2 items). The AEQ was adapted, according to the instructional manual (Pekrun et al., 2002) and previous studies (Naismith, 2013; Authors, 2015b) to measure learners' test-related concurrent emotions with XX. An example item used to measure enjoyment asked medical students to respond to the statement, "For me the task was a challenge that was enjoyable." Responses were measured on a 5-point Likert scale where a rating of 1 indicated that the participant strongly disagreed with the statement and a rating of 5 the participant strongly agreed with the statement.

Control and Value Appraisals. To measure appraisals of control participants asked to respond to the statement, “I felt in control of my performance on the task.” To measure value appraisals participants responded to the statement, “I valued the task.” For both items, responses were measured on a 5-point Likert scale where a rating of 1 indicated that the participant strongly disagreed with the statement and a rating of 5 the participant strongly agreed with the statement. Previous research has used similar single item measures to assess appraisals of subjective task control and task value (Tong et al, 2007; Goetz et al. 2010).

Performance. Participants’ perceived performance and actual performance were extracted from the XX logfiles. In this study, perceived performance was inferred from participants’ self-reported confidence in their final diagnosis reported as a percentage (0% - 100%) on the Belief Meter (left-hand of Figure 1). Participants’ actual performance was measured using solution efficiency (percent match with the expert solution, Figure 2).

Experimental Procedure. The data analyzed in this study were collected as part of a larger project that examined emotions in the context of diagnostic reasoning, which comprised of a demographics questionnaire, measures of participant traits and several measures of emotions, including participants’ emotional experience (i.e. emotion questionnaire and retrospective interview), expression (i.e. video analysis of facial expression) and physiological arousal (i.e. electrodermal activation). Only the measures relevant to the analyses in this study will be discussed in the subsequent sections. After reading and signing the informed consent form, participants completed a researcher-guided practice case to receive training on using and navigating XX. During this training case participants learned how to interact with XX to solve authentic diagnostic problems. Participants solved either two short diagnostic problems or one long problem. In either case, the session took approximately 2.5 hours to complete. After solving the last problem, participants were asked to report their *during-task* appraisals of control and value and their *during-task* (i.e., concurrent) emotions by completing a post questionnaire which contained the above mentioned control, value and AEQ items.

Data Analysis

Data Cleaning. The emotion data for one participant was missing due to a technical error when reporting concurrent emotions. Consequently, this participant was not included in the analyses. The performance efficiency data for two participants was missing due a technical problem during data collection. Consequently, these participants were not included during the analysis related to performance efficiency (but were for the others). A box plot analysis was conducted in IMB SPSS to detect univariate outliers for each of the nine continuous variables included in the analyses (i.e. enjoyment, pride, hope, anxiety, hopelessness, shame, anger, control, value, perceived performance and performance efficiency). Only one outlier was detected for the variable anger. This outlying score was replaced with the next most extreme non-outlying score (Meyers, Gamst, & Guarino, 2013).

Emotion Cluster Extraction. A *k*-means cluster analysis was performed on the five during-task emotions to identify groups of participants that were highly similar within groups and highly distinct between groups. Participant's mean emotion scores measured from the AEQ were converted to *z*-scores and analyzed using the *k* means clustering algorithm in IBM SPSS. A total of 125 data points were entered into the cluster model as each participant had a total of 5 unique emotion scores. The selection of the number of clusters is determined by the researcher and it is based on previous empirical work and theory (Daniels, Haynes, Stupnisky, Perry, Newall, & Pekrun, 2008). We selected a 3-cluster model as previous work demonstrated that a 3-cluster model was optimal for categorizing similar self-reported emotions (Authors, 2015b).

Results

RQ1: Can learners be grouped according to the emotions experienced during problem solving?

Three iterations were run in order to generate convergence (i.e., automatic validation). Follow-up univariate ANOVAs indicated that the clustered groups differed significantly on all seven AEQ variables (see Table 1 for concurrent emotion descriptive statistics).

The final cluster centers, together with the number of cases in each cluster are shown in Table 2. The cluster membership ranged from 5 to 17 learners. Cases in Cluster 1 tended to experience relatively high levels of positive emotions (enjoyment and hope); cases in Cluster 2 tended to experience relatively high levels of negative emotions (anger, hopelessness, shame and anger); and cases in Cluster 3 tended to experience relatively low overall affect. Final cluster memberships were used to define participant groups for the subsequent analysis.

Table 1

Descriptive Statistics of Concurrent Emotion Responses

Sate Emotion	*M*	*SD*	Min	Max
Enjoyment	3.49	.75	1.67	5.00
Pride	2.30	.72	1.00	4.00
Hope	2.88	.71	2.00	4.50
Anxiety	1.95	.70	1.00	3.57
Hopelessness	2.00	.99	1.00	4.33
Shame	1.82	.94	1.00	4.00
Anger	1.70	.72	1.00	3.00

Table 2

Final cluster z-score means on the concurrent state emotion variables

State Emotion	Cluster 1: Positive Emotion $n = 7$	Cluster 2: Negative Emotion $n = 5$	Cluster 3: Low Emotion $n = 13$
Enjoyment	**1.00**	-.66	-.28
Pride	.67	.69	-.63
Hope	**1.27**	-.39	-.53
Anxiety	-.19	**1.21**	-.36
Hopelessness	-.39	**1.38**	-.32
Shame	-.24	**1.55**	-.47
Anger	.22	**1.25**	-.60

Note: Clusters were interpreted by the *z*-scores. Z-scores above 1 were interpreted as high (bolded) and scores approaching zero were interpreted as low (Meyers, Gamst, & Guarino, 2013).

RQ 2: Is there a relationship between emotion clusters and task control and value appraisals?

Appraisal of Control. An ANOVA was conducted to determine if there was a significant difference between emotion clusters (IV; cluster membership) and appraisal of task control (DV; subjective task control; see table 3 for subjective task control descriptive statistics). Results indicated a trend towards a statistically significant relationship between emotion clusters and appraisals of task control with a large effect size, F (2, 22) = 2.97, p = .07, η^2 = .21. Taking into consideration the alpha and partial eta squared values of this omnibus, we conducted follow-up pairwise comparisons. The results from the pairwise comparisons indicated that the positive emotion group reported significantly higher levels of task control than the negative emotion group (*M* = 3.86, *SD* = .90 and *M* = 2.20, *SD* = 1.09, respectively). No other pairwise comparisons were significant (see table 4).

Table 3

Descriptive Statistics for Control and Value Appraisals

Appraisal	*M*	*SD*	Min	Max
Control	3.28	1.27	1.00	5.00
Value	4.28	.61	3.00	5.00

Table 4

Descriptive Statistics and Pairwise Comparisons for Emotion Clusters and Appraisals

Appraisal	Emotion Cluster						Pairwise Comparisons			
	Positive		Negative		Low					
	M	*SD*	*M*	*SD*	*M*	*SD*				*p*
Control	3.86	.90	2.20	1.10	3.38	1.33	Positive	>	Negative	.026*
							Positive	>	Low	.403
							Low	>	Negative	.070
Value	4.57	.53	3.80	.45	4.31	.63	Positive	>	Negative	.032*
							Positive	>	Low	.339
							Low	>	Negative	.108

Note: The *p*-values on the left correspond to significant and non-significant alphas related to the pairwise comparison. Values with an * = $p < .05$.

Appraisal of Value. An ANOVA was conducted to determine if there was a significant difference between emotion clusters (IV; cluster membership) and appraisal of task value (DV; subjective task value) (see table 3 for subjective task control descriptive statistics). Results suggest a trend towards a statistically significant relationship between emotion clusters and appraisals of task value with a large effect size, F (2, 22) = 2.65, p = .09, η^2 = .19. Consistent with the rational presented above, we have also chosen to conduct the follow-up pairwise comparisons. The results from the pairwise comparisons indicate that the positive emotion group reported significantly higher levels of task value than the intense negative emotion group (M = 4.57, SD = .53 and M = 3.80, SD = .45, respectively). No other pairwise comparisons were significant (see table 4).

RQ 3: Is there a relationship between emotion clusters and performance?

Perceived Performance. An ANOVA was conducted to determine if there was a significant difference between emotion clusters (IV; cluster membership) and perceived performance (DV; confidence; see table 5 for performance descriptive statistics). Results from the ANOVA suggest that there was no statistically significant relationship between emotion clusters and perceived performance although a medium effect size was found, F (2, 22) = .66, p = .54, η^2 = .06. Interpretations of the descriptive statistics revealed that learners in the positive emotion cluster had the highest perceived performance followed by the low emotion and negative emotion clusters (see table 6).

Actual Performance. An ANOVA was conducted to determine if there was a significant difference between emotion clusters (IV; cluster membership) and actual performance (DV; percent match with expert; see table 5 for performance descriptive statistics). Results from the second ANOVA suggest that there was no statistically significant relationship between emotion clusters and actual performance, although a medium effect size was found, F (2, 20) = 1.47, p = .25, η^2 = .13. Descriptive statistics indicate that learners in the low emotion cluster outperformed learners in both the positive and negative emotion clusters, which had comparable mean levels (see table 6).

Table 5

Descriptive Statistics for Performance

Performance	*M*	*SD*	Min	Max
Perceived Performance	73.73	21.69	24	100
Actual Performance	44.25	24.35	7	100

Table 6

Descriptive Statistics for Emotion Clusters and Performance

Performance Type	Emotion Cluster					
	Positive		Negative		Low	
	M	*SD*	*M*	*SD*	*M*	*SD*
Perceived Performance	81.00	10.71	66.40	27.73	73.69	24.45
Actual Performance	35.00	20.29	32.75	29.24	50.92	23.35

Discussion

In summary, we found that: (1) Based on the emotions learners experienced while solving diagnostic problems participants could be organized into three meaningful groups: Positive emotions, negative emotions, and low overall emotions. (2) Learners in the positive emotion cluster had the highest subjective control and value appraisals in comparison to learners in the low and negative emotion clusters. (3) Learners in the positive emotion cluster had the highest perceived performance when compared to learners in the low and negative emotion clusters; however, (4) learners in the positive emotion cluster had lower levels of actual performance when compared to learners assigned to the low emotion cluster, and had actual performance levels comparable to learners in the negative emotion cluster.

Results related to our first research question supported our hypothesis that the emotions learners experienced while solving diagnostic problems would cluster in meaningful ways. This finding is in line with prior research on clustering self-reported emotion data and adds further support to theories and studies of emotion that group emotions into high-level categories according to the dimensions of valence and arousal (Pekrun & Perry, 2014; Authors, 2015a; Authors, 2015b). Participants in the low emotion group experienced low levels of emotions across all measured emotions irrespective of valence. As such, it might be appropriate to conceptualize this cluster as being similar to self-reports of neutral, given its center point on the affective grid (Russell, Weiss, & Mendelsohn, 1989). Alternatively, this group might represent a profile similar to the state of relaxation, which is posited in the control value theory of achievement emotions to be experienced when students engage in pleasant but routine activities (Pekrun, 2006; Pekrun & Perry, 2014). Thus the profile of low emotionality supports the perspective that students will not necessarily experience intense emotions during particular learning tasks which is in-line with current conceptions of emotions during achievement tasks. One unexpected finding, however, was that only moderate levels of pride were found in the positive emotion cluster (in comparison to relatively high levels of enjoyment and hopefulness). One possible explanation for this finding is that participants were asked to reflect back on how they were feeling *while* they were interacting with XX at which point they would not yet have received feedback on their performance and therefore had less to feel proud about regarding their performance.

Results also support our second hypothesis that participants assigned to the positive emotion cluster would have higher task control and value appraisals than participants in the negative emotion cluster. These results are inline with several of the emotion-appraisal relationships outlined by the control-value theory of achievement emotions (Pekrun, 2006). Particularly, participants with positive value appraisals and high levels of subjective control are associated with enjoyment and hope as found in the intense positive emotion cluster. On the other hand, negative value appraisals along with low levels of perceived control are consistently related to negative emotions such as hopelessness or anxiety, as found in the intense negative emotion cluster (Pekrun & Perry, 2014). Although no specific hypotheses were made regarding the appraisals of participants in the low emotion cluster, the results showed that this group did not differ significantly in their appraisals from participants in either the positive or negative emotion clusters. However, the descriptive statistics suggest that the low emotion group experienced moderate levels of control and value. These findings suggest that moderate levels of control and value appraisals might lead participants to experience emotions at low intensities.

The results did not, however, support the fourth hypothesis that participants in the positive emotion cluster would outperform participants in the negative and low emotion clusters. In fact, participants in the low emotion cluster outperformed participants in the positive and negative emotion clusters, which performed similarly. These unexpected findings can be explained by previous work. For example, one study found that both positive and negative emotions can lead to decreased task-related cognitive resources (Meinhardt & Pekrun, 2003). Similarly, other work has found that positive emotions can lead to task irrelevant thinking (Seibert & Ellis, 1991), and strong positive emotions such as enjoyment can be negatively associated with explicit learning (Rieber, & Noah, 2008). For example, high levels of enjoyment might have been elicited through exploratory learning (reading unrelated content) or from engaging in off-task behaviors such as experimenting with XX's software features (e.g., seeing how many lab tests can be run at once; gaming the system; Baker, D'Mello, Rodrigo, & Graesser, 2010). Thus, this study supports findings that enjoyment and high positive emotionality are not always conducive for learning and performance.

Indeed, related research has shown that positive emotions have also been associated with academic overconfidence (Hall et al., 2006). The results from the current study suggest that learners higher in positive emotionality might have been overconfident in their task performance given their greater confidence in their solution *and* lower performance on diagnostic accuracy. The gap between perceived and actual performance was, on the other hand, much lower for low emotionality learners. It is therefore possible that the ideal emotional state for individuals completing an achievement task may be overall low intensity irrespective of emotional valence.

This research emphasizes the influence of student emotions in performance tasks and highlights important pedagogical decisions for supporting achievement in these contexts. Current instructional recommendations concerning emotion focus on linking an emotional state, either discrete or dimensional to a particular context. This study provides evidence that we also need to consider the intensity of during task emotional experiences. This finding is consistent with conceptualizations of achievement emotions when considering the intensity of emotions experienced by an *individual* student. For example, a student experiencing minor levels of hopelessness during a task might still be able to maintain task-focus and do well, whereas a student experiencing acute levels of hopelessness might become occupied by task irrelevant thoughts and disengage from the aversive task. This is also supported by empirical work. One study found that agreement between physiological data ad self-reported emotion increased when only higher-intensity states were considered (Authors, 2015a). Another study found that high levels of negative emotions were uniformly detrimental to performance outcomes, however participants performed better when they had low levels of negative emotions and high levels of perceived control (Ruthig et al., 2008). Therefore it is possible that different levels of an emotion, such as, intermediate and high intensity should be associated with greater and lesser degrees of learning behaviors, cognitions, and other outcomes. In other words, instead of asking "what emotions do we need to worry about?" we need to also consider "at what level of emotional intensity should an intervention be triggered?" In the context of CBLEs, systems might be designed to disregard low levels of frustration experienced while problem solving; however, try to intervene after a participant has spent a certain amount of time in a moderate state of frustration as to avoid transitioning into a more intense and distracting state. Detecting and responding to student emotions is at the forefront of cutting edge affect-aware learning technologies (Calvo & D'Mello, 2012; Bosch, 2015).

A clear limitation of the current work is the sample size. Although our sample size was small, results were supported by large and medium effect sizes, suggesting that with a larger sample, significant relationships could be detected. A larger sample size will also make it possible to calculate the reliability of the AEQ measure used in this study.

That said, the reliability of this scale is well-established over a large number of empirical studies (xx).

Moreover, the participants recruited for this study came from a highly-limited, expert population, unlike other studies that sample from larger populations (e.g., undergraduate students, or undergraduate students in a specific faculty of department). As such, the sample is representative amongst studies with similar expert participants (Authors, 2015c; Authors, 2015d; Duffy et al., 2015; Naismith, 2013). A second limitation of the current work was with regard to how emotions were measured. Students were asked to report how they remembered feeling during the task after they had received feedback on their performance. It is possible that this might have biased participant responses such that participants who performed favorably recalled experiencing more positive emotions, valuing the task more and being more in control. The authors recognize the need to use multi-modal measurement tools during cognitively demanding tasks (Authors, 2015a; Duffy et al., 2015). To overcome this limitation, future research should integrate multiple assessments of emotions, particularly non-invasive measurements, such as physiological and behavioral measures (Authors, 2015a; Calvo & D'Mello, 2010). Future research will also aim to replicate the findings of this study with larger sample sizes and in other CBLEs in order to examine the robustness and generalizability of the results, in particular, between improved learning outcomes and lower levels of emotionality. In addition, future research should determine the specific causes of an emotion during learning with a CBLE at a finer level of granularity for the purposes of system development and during task interventions.

The significant relationships and the trends supported by medium effect sizes identified in this study provide critical support to the notion that during-task emotions have important implications for learning and performance in CBLEs. In particular, the results suggest that CBLEs could support learners by promoting low emotionality *rather* than intense positive emotionality while learners engage in a performance oriented achievement task. As such, it identifies an important area of future research: examining instructional intervention in CBLEs that can assist learners to control their emotional experiences, irrespective of valence. Finally, this study highlights the importance of person-centered analyses of emotions, in particular, to furthering an understanding of the nuanced relationship between achievement emotions and learning that are fundamental for instructional design of CBLEs.

Acknowledgments The research presented in this paper has been supported by a Doctoral fellowship from Richard H. Tomlinson (through McGill University) and a Doctoral Canadian Graduate Scholarship from the Social Science and Humanities Research Council (SSHRC) awarded to the first author. This research has also been supported by a Canadian Fund for Innovation and a Canada Research Chair grant awarded to Dr. Susanne P. Lajoie and a SSHRC partnership grant, headed by Dr. Susanne P. Lajoie. The authors would also like to thank Maher Chaouachi, Tenzin Doleck, Melissa Duffy, Maedeh Kazemi, Philippe Latour, Lila Lee, Laura Pipe, and Tara Tressel for assistance with data collection.

Compliance with ethical standards

Conflict of interest The authors declare that they have no known conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the McGill Research Ethics Board and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Authors, 2009
Authors, 2014
Authors, 2015a
Authors, 2015b
Authors, 2016c
Authors, 2016d

Baker, R.S.J., D'Mello, S.K., Mercedes, 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*, 223-241.

Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., ... & Zhao, W. (2015, March). Automatic detection of learning-centered affective states in the wild. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 379-388). ACM.

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

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

Cassady, J. C. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and Instruction*, *14*(6), 569-592.

Daniels, L. M., Haynes, T. L., Stupnisky, R. H., Perry, R. P., Newall, N. E., & Pekrun, R. (2008). Individual differences in achievement goals: A longitudinal study of cognitive, emotional, and achievement outcomes. *Contemporary Educational Psychology*, *33*(4), 584-608.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145-157.

D'Mello, S. K. (2013). A selectice meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*, 105, 1082-1099.

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

Doleck, T., Basnet, R. B., Poitras, E. G., & Lajoie, S. P. (2015). Mining learner–system interaction data: implications for modeling learner behaviors and improving overlay models. *Journal of Computers in Education*, *2*(4), 421-447.

Duffy, M., Lajoie, S., Jarrell, A., Pekrun, R., Azevedo, R., & Lachapelle, K. (2015, April). Emotions in medical education: Developing and testing a scale of emotions across medical learning environments. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Goetz, T., Frenzel, A. C., Stoeger, H., & Hall, N. C. (2010). Antecedents of everyday positive emotions: An experience sampling analysis. *Motivation and Emotion*, *34*(1), 49-62.

Hall, N. C., Perry, R. P., Ruthig, J. C., Hladkyj, S., & Chipperfield, J. G. (2006). Primary and secondary control in achievement settings: A longitudinal field study of academic motivation, emotions, and performance1. *Journal of Applied Social Psychology*, *36*(6), 1430-1470.

Harley, J. M., & Azevedo, R. (2014). Toward a feature-driven understanding of students' emotions during interactions with agent-based learning environments: A selective review. International Journal of Gaming and Computer-Mediated Simulation, 6(3), 17–34.

Harley, J. M., Bouchet, F., Hussain, S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. Computers in Human Behavior, 48, 615–625.

Harley, J. M., Carter, C. K., Papaionnou, N., Bouchet, F., Azevedo, R., Landis, R. L., & Karabachian, L. (2016). Examining the predictive relationship between personality and emotion traits and students' agent-directed emotions: Towards emotionally-adaptive agent-based learning environments. User Modeling and User-Adapted Interaction, 26, 177–219.

Jarrell, A., Harley, J. M., Lajoie, S. P., & Naismith, L. (2015). Examining the relationship between performance feedback and emotions in diagnostic reasoning: Toward a predictive framework for emotional support. In C. Conati & N. Heffernan (Eds.), Lectures notes in artificial intelligence (Vol. 9112, pp. 657–660). Artificial intelligence in education Switzerland: Springer.

Jarrell, A., Harely, J. M., & Lajoie, S. P. (2016). How do emotions experienced during problem solving relate to students' perceived and actual performance? Paper presented at the American Educational Research Association, Washington, DC.

Lajoie, S. P. (2009). Developing professional expertise with a cognitive apprenticeship model: Examples from avionics and medicine. In K. A. Ericsson (Ed.), Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments (pp. 61–83). Cambridge: Cambridge University Press.

Lajoie, S. P., Lee, L., Poitras, E., Bassiri, M., Kazemitabar, M., Cruz-Panesso, I., et al. (2015). The role of regulation in medical student learning in small groups: Regulating oneself and others' learning and emotions. Computers in Human Behavior, 52, 601–616.

Lazarus, R. S. (2006). Emotions and interpersonal relationships: Toward a person-centered conceptualization of emotions and coping. *Journal of personality*, *74*(1), 9-46.Martinent, G., Nicolas, M., Gaudreau, P., & Campo, M. (2013). A cluster analysis of affective states before and during competition. *Journal of sport & exercise psychology*, *35*(6), 600-611.

Linnenbrink-Garcia, L., & Barger, M. M. (2014). Achievement goals and emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp.142-161). New York: Routledge.

Mayer, R. (2011). Information processing. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook: Theories, constructs, and critical issues* (Vol. 1, pp. 85-100). Washington, DC: APA.

Meyers, L., Gamst, G., & Guarino, A. (2013). Applied Multivariate Research: Design and Interpretation. SAGE, Thousand Oaks, California.

Meinhardt, J., & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. Cognition and Emotion, 17, 477–500.

Naismith, L. M. (2013). Examining motivational and emotional influences on medical students' attention to feedback in a technology-rich environment for learning clinical reasoning. Unpublished doctoral dissertation. Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada.

Pekrun, R., & Hofmann, H. (1996, April). Affective and motivational processes: Contrasting interindividual and intraindividual perspectives. Paper presented at the annual meeting of the American Educational Research Association, New York.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, *18*(4), 315-341.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology, 102*(3), 531-549. doi:10.1037/a0019243

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist, 37,* 91-106.

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120-141). New York: Routledge.

Rieber, L. P., & Noah, D. (2008). Games, simulations, and visual metaphors in education: antagonism between enjoyment and learning. *Educational Media International*, *45*(2), 77-92.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57*(3), 493–502.

Ruthig, J. C., Perry, R. P., Hladkyj, S., Hall, N. C., Pekrun, R., & Chipperfield, J. G. (2008). Perceived control and emotions: Interactive effects on performance in achievement settings. *Social Psychology of Education*, *11*(2), 161-180.

Seibert, P. S., & Ellis, H. C. (1991). Irrelevant thoughts, emotional mood states, and cognitive task performance. *Memory & Cognition*, *19*(5), 507-513.

Tong, E. M., Bishop, G. D., Enkelmann, H. C., Why, Y. P., Diong, S. M., Khader, M., & Ang, J. (2007). Emotion and appraisal: A study using ecological momentary assessment. *Cognition and Emotion, 21*(7), 1361-1381

Zimmerman, B. J., & Labuhn, A. S. (2012). Self-regulation of learning: Process approaches to personal development. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook: Theories, constructs, and critical issues* (Vol. 1, pp. 399-425). Washington, DC: APA.

Zeidner, M. (1998). *Test anxiety: The state of the art*. Springer Science & Business Media.

Zeidner, M. (2014). Anxiety in education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), International handbook of emotions in education (pp. 265–288). New York, NY: Taylor & Francis.