Examining the Relationship Between Performance Feedback and Emotions in Diagnostic Reasoning: Toward a Predictive Framework for Emotional Support

Amanda Jarrell[1(✉)], Jason M. Harley[1,2], Susanne Lajoie[1], and Laura Naismith[3]

[1] Department of Educational and Counselling Psychology, McGill University,
Montreal, Canada
`amanda.jarrell@mail.mcgill.ca`
[2] Computer Science and Operations Research, Université de Montréal, Montreal, Canada
[3] Toronto Western Hospital, University Health Network, Toronto, Canada

Abstract. The purpose of this research is to understand achievement emotions resulting from performance feedback in a medical education context where 30 first and second year medical students learned to diagnose virtual patients in an intelligent tutoring system (ITS), BioWorld. We found that students could be organized into groups using cluster analyses based on the emotions they reported after receiving performance feedback: a positive emotion cluster, negative emotion cluster, and low overall emotion cluster. Medical students in the positive achievement emotion cluster had the highest performance on the diagnostic reasoning cases; those in the negative achievement emotion cluster had the lowest performance; and students categorized as belonging to the low overall achievement emotion cluster had mean performance levels that fell between the two. From the results we propose critical performance thresholds that can be used to predict emotions following performance feedback.

Keywords: Emotions · Affect · Intelligent tutoring systems

1 Introduction: Success, Failure and Emotions

Failure and success generate emotional responses that have a strong impact on future learning, educational decisions, and even health [1]. Given that emotional responses influence motivation and behaviors following successes and failures, it is critical to understand achievement emotions. Achievement emotions are affective arousal tied to achievement activities or outcomes [1]. We focus on retrospective outcome achievement emotions. These emotions arise when learners reflect back on the outcome of an achievement task, and include feelings of pride, anger, shame, joy, and relief [1, 2]. Retrospective outcome emotions are important because they contribute to the formation of future appraisals, which can influence prospective and concurrent state emotions (e.g., anxiety) as well as motivation-related learning behaviors (e.g., drop-out) [1]. Supporting emotional resilience is especially important in educational contexts that train students for high-stakes, real-life situations, such as medical education [3].

The purpose of this research is to understand achievement emotions resulting from performance feedback, as medical students learn to diagnose virtual patients in an intelligent tutoring system (ITS), BioWorld [4]. In this study we answered two research questions: (1) Do retrospective outcome emotions cluster in a meaningful way? And (2) is there a significant performance difference between emotion cluster groups?

2   Methods

Participants. This study consisted of 30 first and second year medical students from a North American university. The sample consisted of 11 men and 19 women, with an average age of 23 ($SD = 2.60$). The data analyzed in this paper were collected as part of a larger study in BioWorld [5].

BioWorld. BioWorld is an ITS that scaffolds students as they learn medical problem solving. Each case begins with a patient history where medical students gather evidence and based on this evidence propose initial hypotheses. Medical students 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, they receive individualized feedback on their solution based on an aggregated expert solution.

Academic Achievement Emotions (AEQ). The emotion questionnaire used in this study was adapted from the Achievement Emotion Questionnaire (Retrospective) (AEQ; [6]) to measure emotions experienced after receiving performance feedback (i.e. pride, joy, relief, anger and shame).

Performance. Performance in BioWorld was measured by examining the accuracy of students' diagnoses. Accuracy was operationalized in this study based on the percentage of evidence matches with the expert solution.

Procedure. Participants were asked to solve three cases of varying levels of difficulty: easy, moderate and difficult. The order of the cases was counterbalanced to mitigate practice effects. Feedback on performance was provided to the learner at the end of each case. After finishing each case, learners completed the AEQ.

3   Results

3.1   Emotion Clusters

The results from the 3-cluster k-means cluster analyses portrayed a robust pattern of emotion clusters. For each analysis, participants in cluster 1 appear to experience low overall affect; cases in cluster 2 appear to experience high negative affect; and cases in cluster 3 appear to experience high positive affect. The final cluster centers and the number of cases in each cluster are shown in Table 1.

Table 1. Final cluster z-score means on the retrospective outcome emotion varibles

Analysis	Variable	Cluster 1 Low Affect	Cluster 2 Negative Affect	Cluster 3 Positive Affect
		$n = 17$	$n = 4$	$n = 9$
	Shame	-.46	**1.56**	-.61
	Anger	-.36	**1.95**	-.58
Easy	Joy	-.49	-.02	**1.85**
	Pride	.02	-.54	**1.60**
	Relief	-.54	.39	**1.37**
		$n = 13$	$n = 7$	$n = 10$
	Shame	-.68	**1.26**	-.03
	Anger	-.60	**1.49**	-.14
Moderate	Joy	-.71	.20	.61
	Pride	-.49	-.25	.72
	Relief	-.83	.58	.46
		$n = 17$	$n = 8$	$n = 5$
	Shame	-.28	**1.52**	-.14
	Anger	-.45	**1.43**	-.31
Difficult	Joy	-.86	.23	**1.24**
	Pride	-.74	-.60	**1.10**
	Relief	-.60	.33	.10
		$n = 47$	$n = 22$	$n = 21$
	Shame	-.42	**1.24**	-.35
	Anger	-.54	**1.47**	-.33
All cases	Joy	-.67	.19	**1.30**
	Pride	-.39	-.39	**1.27**
	Relief	-.69	.45	**1.08**

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 [7].

3.2   Emotion Clusters and Performance

Four ANOVAs were conducted to determine if significant differences existed between emotion clusters and performance for the: (1) moderate, (2) difficult and (3) easy cases as well as (4) across all cases. The results of the ANOVAs revealed a significant difference in performance between emotion clusters for all analyses except for the moderate case (see Table 2).

Table 2. Results of the ANOVAs Conducted for Emotion Clusters and Performance

Analysis	df	MS	F	p	η^2
Moderate	2	82.24	.56	.580	.04
Difficult	2	1857.73	7.10	.003*	.35
Easy	2	988.11	4.16	.027*	.24
All Cases	2	2374.23	6.08	.003*	.12

The descriptive statistics suggest that participants in the positive affect cluster had the highest performance; those in the negative affect cluster had the lowest performance; and participants in the low overall affect cluster had mean performance that fell between the two (Table 3). Tukey LSD post-hoc comparisons were conducted for all significant ANOVAs. These results revealed that several of the emotion clusters differed significantly on performance (Table 3).

Table 3. Tukey Post-hoc Comparisons Conducted for Emotion Clusters and Performance

Analysis	Cluster	Performance		Comparisons				p
		M	SE					
	Cluster 1: Low Affect	41.07	3.37	Low.	>	Neg.		N.A.
Moderate	Cluster 2: Negative Affect	35.14	4.60	Neg.	<	Pos.		N.A.
	Cluster 3: Positive Affect	38.20	3.85	Pos.	<	Low		N.A.
	Cluster 1: Low Affect	48.94	17.34	Low.	>	Neg.		.838
Difficult	Cluster 2: Negative Affect	45.00	16.13	Neg.	<	Pos.		.005*
	Cluster 3: Positive Affect	77.20	10.40	Pos.	>	Low.		.005*
	Cluster 1: Low Affect	73.53	3.74	Low.	>	Neg.		.028*
Easy	Cluster 2: Negative Affect	50.00	7.70	Neg.	<	Pos.		.035*
	Cluster 3: Positive Affect	74.44	5.14	Pos.	>	Low.		.989
	Cluster 1: Low Affect	54.53	2.88	Low.	>	Neg.		.073
All cases	Cluster 2: Negative Affect	43.18	4.21	Neg.	<	Pos.		.002*
	Cluster 3: Positive Affect	64.14	4.31	Pos.	>	Low.		.159

4   Conclusion and Future Directions

We interpret these results to suggest that two critical performance thresholds that must be met for a participant to experience negative, positive or low affect (i.e. a low performance threshold and a high performance threshold). If a participant's performance falls (1) below the low performance threshold, then the participant will experience negative affect, (2) above the high performance threshold, then the participant will experience positive affect and (3) above the low performance threshold but below high performance threshold, then the participant will experience low overall affect. Future research could apply these thresholds to predict emotional responses without needing to ask students how they feel. This would allow ITSs to swiftly and appropriately debrief students to prevent the formation of harmful future appraisals and down-regulate the experience of negative emotions following poor performance and failure.

References

1. Pekrun, R., Perry P.P.: Control-value theory of achievement emotions. In: International Handbook of Emotions in Education, pp. 120–141. Routledge, New York (2014)
2. Pekrun, R.: The control-value theory of achievement emotions. Educational Psychology Review **18**(4), 315–341 (2006)
3. Artino Jr, A.R., Holmboe, E.S., Durning, S.J.: Can achievement emotions be used to better understand motivation, learning, and performance in medical education? Medical Teacher **34**(3), 240–244 (2012)
4. Lajoie, S.: Developing professional expertise with a cognitive apprenticeship mode. In: Development of Professional Expertise, Cambridge, UK, pp. 61–83 (2009)
5. Naismith, L.M.: Examining motivational and emotional influences on medical students' attention to feedback in a TRE for learning clinical reasoning. McGill University (2013)
6. Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., Perry, R.P.: Measuring emotions in students' learning and performance. Contemporary Ed. Psychology **36**, 36–48 (2011)
7. Meyers, L., Gamst, G., Guarino, A.: Applied Multivariate Research: Design and Interpretation. SAGE, Thousand Oaks (2013)