Article

Journal of Educational Computing Research
2021, Vol. 59(3) 429–452

DOI: 10.1177/0735633120967318
journals.sagepub.com/home/jec
SAGE

Quantifying the Influence of Achievement Emotions for Student Learning in MOOCs

Bowen Liu[1], Wanli Xing[2], Yifang Zeng[3] and Yonghe Wu[1]

Abstract

Massive Open Online Courses (MOOCs) have become a popular tool for worldwide learners. However, a lack of emotional interaction and support is an important reason for learners to abandon their learning and eventually results in poor learning performance. This study applied an integrative framework of achievement emotions to uncover their holistic influence on students' learning by analyzing more than 400,000 forum posts from 13 MOOCs. Six machine-learning models were first built to automatically identify achievement emotions, including K-Nearest Neighbor, Logistic Regression, Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machines. Results showed that Random Forest performed the best with a kappa of 0.83 and an ROC_AUC of 0.97. Then, multilevel modeling with the "Stepwise Build-up" strategy was used to quantify the effect of achievement emotions on students' academic performance. Results showed that different achievement emotions influenced students' learning differently. These findings allow MOOC platforms and instructors to provide relevant emotional feedback to students automatically or manually, thereby improving their learning in MOOCs.

[1]Faculty of Education, East China Normal University, Shanghai, China
[2]School of Teaching and Learning, University of Florida, Gainesville, Florida, United States
[3]Department of Educational Psychology and Leadership, Texas Tech University, Lubbock, Texas, United States

Corresponding Author:
Wanli Xing, College of Education, University of Florida, 2-215, Normal Hall, Gainesville, FL 32611, United States.
Email: wanli.xing@coe.ufl.edu

Keywords
achievement emotions, learning performance, sentiment analysis, multilevel modeling, MOOCs

Massive Open Online Courses (MOOCs), which aim to provide open learning opportunities to thousands of people for free or at a low cost, and are becoming one of the fastest growing online learning environments (Buhr et al., 2019). The ability for anyone in the world with reliable Internet to enroll in MOOC courses at any time attracts thousands of learners for self-improvement (Yang et al., 2015). Learners can increase their knowledge and benefit from MOOCs through the course content and interactions with peers from a diverse range of places and cultures (Henderikx et al., 2019). MOOCs have become an important part of students' daily learning (Dmoshinskaia, 2016), and have received increasing attention from the public.

The learning and interaction process usually involves emotional experience, which may affect students' learning in MOOCs. The separation of time and space in the context of MOOCs hinders the transmission of students' emotional information (Liu et al., 2018). Compared to traditional face-to-face learning, learners are also less likely to interact immediately with instructors or other students because of the large number of students in MOOCs (Yang et al., 2015). Lack of emotional interaction and support causes learners to abandon their learning and eventually results in dropout or poor learning performance in MOOCs (Arguel et al., 2017). Thus, supporting the emotional experience of thousands of learners is crucial for MOOCs to improve their learning and outcomes (Buhr et al., 2019).

Most previous studies have focused on positive and negative emotions and their effect on student dropout in MOOCs (e.g., Dmoshinskaia, 2016; Dillon et al., 2016). However, as emotions are complex constructs, it would be overly simplistic to characterize them as simply being positive and negative. Particularly, achievement emotions are specific emotions in academic settings related to learning processes and learning outcomes (Pekrun, 2017). Achievement emotions involve goal-directed and appraisal-driven psychological processes (Jarrell et al., 2016), which are triggered by psychological subsystems, including affective, cognitive, motivational, and expressive behavior (Pekrun, 2017). The relationship between achievement emotions and students' learning is complex. Positive achievement emotions do not always produce positive effects on learning achievement and negative achievement emotions do not always exert negative effects on learning (Pekrun, 2006).

While a few studies have examined how achievement emotions impact student dropout in one or two MOOC courses (e.g., Wen et al., 2014; Yang et al., 2015), these studies only randomly selected a few emotions and did not examine the achievement emotions holistically. Moreover, no study so far has holistically quantified the effect of achievement emotions on students' learning in MOOCs. Therefore, this study focuses on quantifying the holistic effect of students' achievement emotions on their learning in MOOCs based on an integrated framework of achievement emotions. Compared with previous studies on emotions in MOOCs, this study fills research gaps by making four contributions to the field. First, we focus on quantifying the impact of learners' achievement emotions on their learning performance rather than dropout/retention in MOOCs. Second, we rely on control-value theory to reveal the holistic effect of achievement emotions by classifying achievement emotions into four categories. Third, machine-learning methods are used to automatically identify student achievement emotions, which may be further used in other educational research studies and beyond. Fourth, 13 samples of MOOCs are analyzed to quantify the influence of achievement emotions to increase the generalizability of our findings.

Literature Review

Theoretical Framework

Emotions are attitudes and experiences triggered by psychological subsystems, including affective, cognitive, motivational, expressive, physiological processes, and expressive behavior (Pekrun, 2017), which influence learning and achievement in academic settings (Pekrun, 2006), as shown in Figure 1. The influence of emotions on students' learning and achievement is mediated by cognitive and motivational mechanisms (Pekrun, 1992). For cognitive mechanisms, emotions

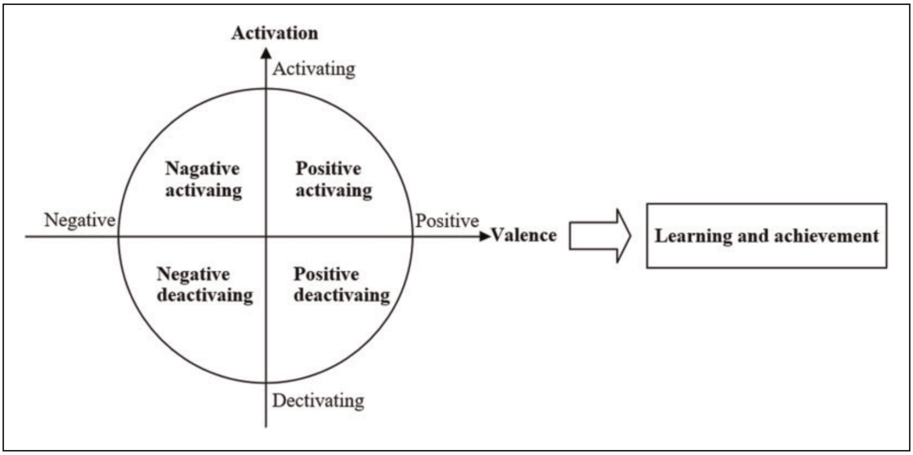

Figure 1. Link Between Achievement Emotions and Learning.

have a fundamental influence on cognitive and learning processes that contribute to learning, including storage, processing, and retrieval of information, memory, attention, perception, use of learning strategies, and self-regulation of learning (Pekrun, 2011; Phelps, 2006). For motivational mechanisms, emotions can affect students' intrinsic and extrinsic learning motivation (Pekrun, 1992), which ultimately influences students' learning and achievement (Hascher, 2010).

Achievement emotions are the academic emotions directly linked to learning activities or learning outcomes, which means that achievement emotions relate to not only the process of learning but also the outcomes of learning (Pekrun et al., 2011). The control-value theory of achievement emotions provides an integrative framework for analyzing the antecedents and effects of achievement emotions in learning (Pekrun, 2006). The control-value theory postulates that two important proximal antecedents of achievement emotions are subjective control over learning and the subjective value of learning. Subjective control refers to the perceived controllability of achievement-related actions and outcomes implied by causal expectations, causal attributions of achievement, and competence appraisals, while subjective value refers to the subjective importance of these activities and outcomes (Pekrun & Stephens, 2010).

According to control-value theory, achievement emotions can be described by two dimensions: valence and activation. The dimension of valence can be divided into positive and negative emotions. The dimension of activation can be divided into physiologically activating and deactivating emotions (Pekrun et al., 2011). The valence and activation of achievement emotions are independent of each other. Thus, four categories of achievement emotions can be generated by combining valence and activation: positive activating, positive deactivating, negative activating, and negative deactivating. Positive activating emotions, including enjoyment, hope, joy, and pride, relate to positive appraisal of learning activities and outcomes, which are thought to support interest, intrinsic motivation, and deep learning (Pekrun et al., 2017). Negative deactivating emotions, including boredom and disappointment, relate to negative appraisal of learning activities and outcomes without any intention of making effort, which are thought to undermine motivation and information processing (Pekrun et al., 2017). Positive deactivating emotions, including relaxation, relief, and contentment, relate to positive attitudes toward learning processes and outcomes, but do not invest much effort and time. Negative activating emotions, including shame, anxiety, and frustration, relate to negative attitudes toward learning processes and outcomes, but focus on investing effort and time (Xing et al., 2019).

Achievement Emotions in MOOCs

Studies on achievement emotions in MOOCs mainly focus on student dropout and analyze these emotions from a discrete perspective. Based on a description of achievement emotions (Pekrun et al., 2002), Dillon et al. (2016) developed a "Pick-Two" list of emotions to allow learners to report their emotions in one MOOC. They found that hope and enjoyment were the most frequently reported emotions. Anxiety, confusion, frustration, and hope were significantly positively related to learners' dropout in the MOOC. Wen et al. (2014) found a significant correlation between emotions and dropout within a specific MOOC context. In a study on two MOOCs, Yang et al. (2015) found that learners' confusion in forums was significantly positively related with dropping out of the learning community. Dmoshinskaia (2016) conducted sentiment analysis on learners' comments in two MOOC discussion forums with a dictionary-based approach. The results indicated that both learners who did not express any negative emotions and learners who expressed negative emotions frequently were likely to drop the course. In contrast, learners who occasionally expressed some negative emotions were likely to complete the course. Xing et al. (2019) found that learners' achievement emotions had a significant impact on their survival duration in one MOOC. They concluded that learners' negative activating emotions significantly influence student dropout and being exposed to deactivating emotions significantly affects their survival in the course. However, positive activating emotions did not show any impact on learner dropout in the MOOC. These studies focused on students' dropout and only investigated a few discrete emotions rather than holistically analyzing achievement emotions.

A few studies focused on the role of achievement emotions in learning performance in MOOCs, but did not quantify the effect of these emotions on students' learning. In a study on an art MOOC, Tucker et al. (2014) found that learners' sentiments were slightly positively related to their quiz scores, but negatively related to their homework assignment scores. In a 45-minute MOOC, Gong et al. (2019) found that learners' positive confusion had a significant positive predictive effect on their learning outcomes, while boredom had a significant negative predictive effect on their learning achievements. By using lag sequential analysis, they further found that the high-achieving group (learners with scores ranking in the top 50%) experienced more sequences from negative to positive confusion, but fewer sequences from frustration to boredom and from negative confusion to surprise, compared with the low-achievement group.

Sentiment Analysis of MOOC Learners

While many studies of emotions in MOOCs use self-report (e.g., Henderikx et al., 2019; Pekrun et al., 2011) or manual analysis of interviews or forums (e.g., Dillon et al., 2016), sentiment analysis, an automatic approach for identifying and classifying emotions by analyzing large amounts of opinion data, has become more popular over the years. Sentiment analysis can provide understanding of learners' attitudes, satisfaction, and possible sentiment predictors for dropping out or academic performance (Dmoshinskaia, 2016) and consists of four main phases: data acquisition, data preparation, review analysis, and sentiment classification. MOOCs forums are the most used resource for performing sentiment analysis (Mite-Baidal et al., 2018). In order to identify sentiment from forum posts in MOOCs, Wen et al. (2014) defined a sentiment ratio; the ratio of positive versus negative words used in the post set. Kechaou et al. (2011) built a combined supervised machine-learning model based on Hidden Model Markov (HMM) and Support Vector Machine (SVM) to detect learners' emotions (positive or negative) from their reviews on an e-learning platform. Cobos et al. (2019) developed a tool that supports emotion analysis for various textual contents (learning goals, video transcription, test questions and answers, and forum posts) in MOOCs by using natural language processing techniques. Hew et al. (2020) used TextBlob3, an open-sourced text sentiment-computing engine, to compute the learner sentiment scores of MOOC reviews ranging from -1.0 to +1.0, with higher scores indicating more positive sentiment.

Summary

Given the importance of achievement emotions in student learning, several studies have begun to explore these emotions in MOOCs. However, there are four limitations in the documented studies. First, most current studies investigated how discrete emotions influenced the MOOC dropout/retention rate, and few have investigated how such emotions influence students' learning performance. Second, most previous studies limited achievement emotions to two (positive or negative) without examining achievement emotions holistically (Rani & Kumar, 2017). Third, most research studied learners' achievement emotions through student self-report with surveys or interviews. There are still limited studies that systematically and automatically identify student achievement emotions. Fourth, the existing studies on achievement emotions usually relied on one MOOC course and a few analyzed a couple of MOOC courses, which might make it difficult to generalize their findings.

Methodology

Research Purpose

Beyond positive and negative emotions, this study aims to uncover the influence of achievement emotions on students' learning in MOOCs holistically. The central research question is: How do achievement emotions influence students' learning performance in MOOCs? To answer this, relying on the control-value theory of achievement emotions, we analyzed 13 MOOC courses with more than 400,000 posts to quantify the effect of achievement emotions on students' learning performance using machine learning and multilevel modeling.

Research Context and Dataset

The study collected students' forum posts from 13 MOOCs on a popular MOOC platform. We chose this MOOC platform because it brings together courses from many well-known universities, which is similar to most other MOOC platforms. Therefore, the results were likely to be generalized to other MOOCs platforms. The 13 MOOCs were selected by convenience sampling, including six STEM courses and seven liberal arts courses. The 13 MOOCs on different subjects are helpful in generalizing our findings. These courses were selected because their course design was exemplary for current MOOCs, which were all comprised of video lectures, discussions forums, quizzes, assignments, and exams (Bonafini et al., 2017). Moreover, these courses were involved in a large number of students and were highly valued by students. The 13 courses proceeded through one to eight terms from 2014 to 2019, with each term lasting for five to seventeen weeks. Each course offered discussion forums to students to support them with voluntary participation interactions. There were usually three sub-forums in the 13 courses, including the Teacher Q & A sub-forum, for asking teachers questions about the assignments, tests, and courseware contents; the Classroom Communication sub-forum for discussing learning contents among peers; and the General Discussion sub-forum, for sharing experiences and ideas about general topics such as courses, study, work, and life. In total, there were 21,192 active learners who posted a total of 406,726 posts in the discussion forums of the 13 courses. Figure 2A shows the number of learners who remained active in discussion forums during the 13 courses, and Figure 2B provides an overview of the number of posts over the 13 courses.

Analysis Variables

Dependent Variable. Learning performance: students' learning performance was calculated by accumulating their scores on quizzes, assignments, and exams. Students' learning performance ranged from 0 to 100.

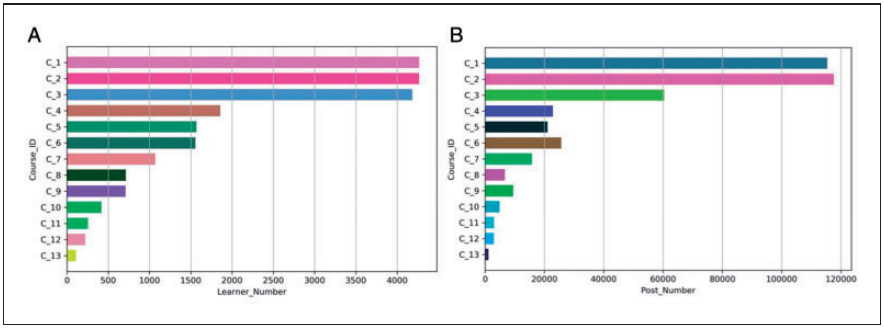

Figure 2. Descriptive Statistics in the MOOC Forum. A: The number of learners over the 13 courses. B: The number of posts over the 13 courses.

Independent Variable. Four variables were built for achievement emotions: positive activating emotions (PA), positive deactivating emotions (PD), negative activating emotions (NA), and negative deactivating emotions (ND), according to the control-value theory of achievement emotions. Each type of achievement emotion was calculated by the number of posts falling into one of the four achievement emotions divided by the total number of posts that the student posted in discussion forums. For example, a student's PA was calculated by the number of posts falling into positive activating emotions divided by the total number of posts of the student.

Control Variable. Total_posts: the number of posts students contribute to forums is a strong indicator of their commitment (Yang et al., 2015). This variable was built with the total number of posts a student contributed to associated forums in a course.

Achievement Emotions Analysis

It is a daunting challenge to manually read, code, and analyze these data because of the sheer volume of posts in MOOC forums. Therefore, this study built a machine-learning model to detect the achievement emotions of each post in discussion forums automatically.

Sampling and Coding. The hand-coding of a sample of posts into certain achievement emotions is the first step in building a machine-learning model. In order to reduce data variance, a stratified sampling method was used to sample a subset of posts randomly from the total posts (406,726 posts) of the 13 MOOCs. A total of 4,098 posts were sampled when the sampling fraction was set to 0.01 for each term. After that, the entire sample of 4,089 posts was coded into achievement emotions by two graduate researchers independently. According to the control-value theory, positive activating emotions were coded with the indicators of enjoyment, hope, and pride. Positive deactivating emotions were coded with the relaxation and relief indicators. Negative activating emotions were coded with the shame, anxiety, and frustration indicators. Negative deactivating emotions were coded with the boredom and disappointment indicators. The Cohen's Kappa for the four achievement emotions was 0.876, which indicated a high inter-rater reliability of coding between the two coders (Cohen, 1960). For inconsistent codes, the two coders discussed together and finally jointly determined a consistent code of achievement emotions. The coded posts were then used as training data for the machine-learning model. Table 1 shows the samples of the post-coding.

Feature Extraction. Textual features extraction is the second step for achieving emotion detection. During this step, all coded sample posts were extracted as

Table 1. Samples of the post coding.

Post 1	Thank you, teacher! I have mastered a lot of new knowledge through doing homework these days. Homework is too important for consolidating knowledge. I am very happy to learn this course and hope that I can gain something new every day.	PA
Post 2	In general, it is not difficult for me to learn this course. Therefore, I do not have much pressure on completing the course. I think this course is still useful for my work.	PD
Post 3	I feel that I have learned something from the course. However, I didn't know how to participate in the discussion due to my poor knowledge base. Even though I wanted to give up a few times, I insisted on completing the course. But I was very disappointed with the courses grades I got. Thanks anyway!	NA
Post 4	I didn't gain anything after participating in two classes. Now I want to quit the course immediately!	ND
Post 5	If failing to uploading homework, you can try to change your browser. I just encountered this problem, and I successfully submitted the homework after changing my browser.	None

Table 2. The Features Set for Achievement Emotion Detection.

Summary language variables	
word count, words per sentence, dictionary words, words > 6, number of emotion words	
Linguistic features	
Grammar	common verbs, quantifiers, numbers
Function words	adverbs, prepositions, conjunctions, negations
Punctuation	period, comma, colon, semicolon, question mark, exclamation mark
Affect features	affective processes, positive emotion, negative emotion, anxiety, anger, sadness, love
Topic features	relativity, motion, space, time, work, achievement, leisure, home, assent, swear words

a set of features. The study extracted textual features based on the Linguistic Inquiry and Word Count (LIWC) library (Pennebaker et al., 2015). The LIWC measures the psychological meaning of words such as emotions expressed in the text (Almatrafi et al., 2018; Gao et al., 2013), which has been proven effective in identifying achievement emotions in MOOC discussion forums (Yang et al., 2015). As shown in Table 2, four categories of textual features were extracted by using the LIWC, including language summary, linguistic, affect, and topic features. These extracted features were then used as the input for machine learning algorithms.

Model Building and Application. Building a machine-learning model based on the extracted features is the third step for constructing a machine-learning model to detect achievement emotions automatically. The study employed six classic supervised machine-learning algorithms to optimize the performance of the machine-learning model, including K-Nearest Neighbor (KNN), Logistic Regression, Nave Bayes, Decision Tree, Random Forest, and Support Vector Machines (SVM). The 10-fold cross validation was used to obtain an averaged performance for the evaluation of the model. The study used six metrics, including accuracy, precision, recall, Kappa, F1_score, and area under curve (ROC_AUC), to show the actual performance of these six machine learning models. The model with the best performance was applied to the remaining forum posts to automatically classify them into different achievement emotions.

Multilevel Analysis

We further utilized a multilevel modeling technique that entails modeling data at two different levels of analysis (course level and student level) to examine the role of achievement emotions (specifically, PA, PD, NA, and ND) in influencing students' learning performance. A "Stepwise Build-up" strategy was adopted to balance between the deductive and inductive approach to achieve a final model of selection (Hox et al., 2010). Several models were estimated in a hierarchical fashion by conducting a baseline analysis (M0; null model with random intercept-only) with which we could compare later models. At this point, we also calculate the Intraclass Correlation (ICC) to ensure that multilevel modeling is appropriate for the purpose of this study. An ICC coefficient of 0.06 was considered as a sufficient level of ICC for utilizing multilevel modeling (Cheung & Au, 2005; Duncan et al., 1998; Stapleton, 2006). We presented the M0 (equation (1)) as follows:

$$\begin{gathered}\text{Level 1}: \quad \mathrm{Score_{ij}} = \beta_{0j} + \mathrm{e_{ij}} \\ \text{Level 2}: \quad \beta_{0j} = \gamma_{00} + \mathrm{u_{0j}} \\ \mathrm{e} \sim \mathrm{N}(0, \sigma_e^2) \\ \mathrm{u} \sim \mathrm{N}(0, \tau_{00})\end{gathered} \tag{1}$$

Then, based on this hypothetical framework, we added five level-1 predictors, including four types of achievement emotions and total_posts (control variable) with a fixed slope (M1; Equation (2)). This model indicated that different students may have different starting scores after controlling for the multilevel structure, but the slopes were assumed to be fixed for each level-1 predictor.

$$\begin{aligned}
\text{Level 1}: \quad \text{Score}_{ij} &= \beta_{0j} + \beta_{1j}{*}\text{PA}_{ij} + \beta_{2j}{*}\text{PD}_{ij} + \beta_{3j}{*}\text{NA}_{ij} + \beta_{4j}{*}\text{ND}_{ij} \\
&\quad + \beta_{5j}{*}\text{total_posts} + e_{ij} \\
\text{Level 2}: \quad \beta_{0j} &= \gamma_{00} + u_{0j} \\
\beta_{1j} &= \gamma_{10} \\
\beta_{2j} &= \gamma_{20} \\
\beta_{3j} &= \gamma_{30} \\
\beta_{4j} &= \gamma_{40} \\
\beta_{5j} &= \gamma_{50} \\
e &\sim N(0, \sigma_e^2) \\
u &\sim N(0, \tau_{00})
\end{aligned} \tag{2}$$

Subsequently, a third model (M2; Equation (3)) was conducted by adding a level-2 predictor (the length of the course). The slopes remained fixed as in M1. Because the level-2 predictor is non-significant in the model, it was excluded from the following models. Then, in the fourth and fifth models (M3 and M4), we examined the random effects in model building based on M1. Specifically, in M3 (Equation (4)), we added slope variance to all predictors in a variable-by-variable manner and fixed the intercept in M4 (Equation (5)).

$$\begin{aligned}
\text{Level 1}: \quad \text{Score}_{ij} &= \beta_{0j} + \beta_{1j}{*}\text{PA}_{ij} + \beta_{2j}{*}\text{PD}_{ij} + \beta_{3j}{*}\text{NA}_{ij} + \beta_{4j}{*}\text{ND}_{ij} \\
&\quad + \beta_{5j}{*}\text{total_posts} + e_{ij} \\
\text{Level 2}: \quad \beta_{0j} &= \gamma_{00} + \gamma_{01}{*}\text{length}_{0j} + u_{0j} \\
\beta_{1j} &= \gamma_{10} \\
\beta_{2j} &= \gamma_{20} \\
\beta_{3j} &= \gamma_{30} \\
\beta_{4j} &= \gamma_{40} \\
\beta_{5j} &= \gamma_{50} \\
e &\sim N(0, \sigma_e^2)
\end{aligned} \tag{3}$$

$$u \sim N(0, \tau_{00})$$

$$\begin{aligned}
\beta_{0j} &= \gamma_{00} + u_{0j} \\
\beta_{1j} &= \gamma_{10} + u_{1j} \\
\beta_{2j} &= \gamma_{20} \\
\beta_{3j} &= \gamma_{30} \\
\beta_{4j} &= \gamma_{40} \\
\beta_{5j} &= \gamma_{50} \\
e &\sim N(0, \sigma_e^2) \\
\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} &\sim N\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \tau_{00} & \\ \tau_{10} & \tau_{11} \end{bmatrix}
\end{aligned} \tag{4}$$

$$\begin{aligned}
\beta_{0j} &= \gamma_{00} \\
\beta_{1j} &= \gamma_{10} + u_{1j} \\
\beta_{2j} &= \gamma_{20} \\
\beta_{3j} &= \gamma_{30} \\
\beta_{4j} &= \gamma_{40} \\
\beta_{5j} &= \gamma_{50} \\
e &\sim N(0, \sigma_e^2) \\
u &\sim N(\tau_{10}, \tau_{11})
\end{aligned} \tag{5}$$

Finally, we tested the interaction effects between achievement emotions and total_posts in predicting students' learning performance (M5). In order to find the optimal choice of model, a Likelihood Ration Test (LRT) was conducted. The last equation (M5; Equation (6)) in this research can be expressed as:

$$\text{Level 1}: \quad \text{Score}_{ij} = \beta_{0j} + \beta_{1j} * PA_{ij} + \beta_{2j} * PD_{ij} + \beta_{3j} * NA_{ij} + \beta_{4j} * ND_{ij} + \beta_{5j} * \text{total\_post}_{ij} + e_{ij}$$

$$\text{Level 2}: \quad \beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21} * \mathrm{total\_post_j} \tag{6}$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31} * \mathrm{total\_post_j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50} \mathrm{e} \sim \mathrm{N}(0, \sigma_e^2)$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \mathrm{N} \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \tau_{00} & \\ \tau_{10} & \tau_{11} \end{bmatrix}$$

The Akaike Information Criterion (AIC) (Akaike, 1974), the Bayesian Information Criterion (BIC) (Schwarz, 1978), and the deviance were employed to assess each model mentioned above. In addition, a t-test was performed on every fixed effect as a significance test, 95% of confidence intervals were calculated for the standard deviation of each random effect, and the variance reduction of the two levels was evaluated in the final model (Hox et al., 2010). Last, the effect size measured related to the variance explained for the overall model f2 was calculated, where 0.02 is a small effect, 0.15 is a medium effect, and 0.35 is a large effect (Cohen, 1992).

Results

Achievement Emotions of MOOC Learners

Based on the extracted features, six machine-learning models were built to automatically detect the achievement emotions of posts in MOOC discussion forums. The metrics results for the performance of the six machine learning models are shown in Table 3. Figure 3 shows the ROC curve of the six machine learning models. Random Forest performed the best in terms of all six metrics (accuracy, precision, recall, Kappa, F1_score, and ROC_AUC) among the six machine learning models. Thus, the Random Forest model was applied to

Table 3. The Performance of the Machine Learning Models.

Model	Accuracy	Kappa	ROC_AUC	Precision	Recall	F1_score
KNN	0.79	0.73	0.92	0.78	0.79	0.79
Logistic Regression	0.49	0.36	0.77	0.48	0.49	0.47
Naïve Bayes	0.37	0.21	0.69	0.39	0.37	0.32
Decision Tree	0.76	0.70	0.86	0.76	0.76	0.76
Random Forest	0.86	0.83	0.97	0.86	0.86	0.86
SVM	0.41	0.27	0.83	0.44	0.41	0.39

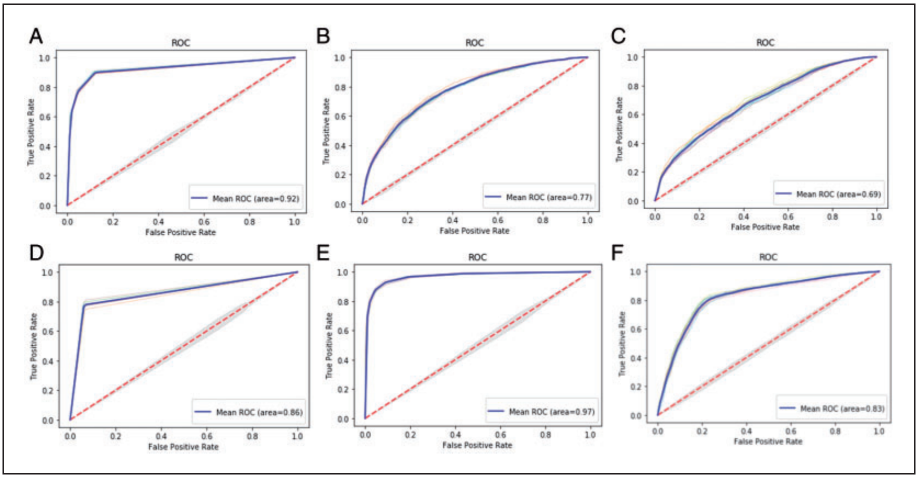

Figure 3. ROC Curve of the Machine Learning Models. A: ROC curve of KNN. B: ROC curve of Logistic Regression. C: ROC curve of Gaussian NB. D: ROC curve of Decision Tree. E: ROC curve of Random Forest. F: ROC curve of SVM.

Table 4. Descriptive Statistics for the Multilevel Analysis.

	Mean	SD	Min	Max
PA	0.231	0.172	0.000	1.000
PD	0.189	0.142	0.000	1.000
NA	0.087	0.104	0.000	1.000
ND	0.024	0.051	0.000	0.600
Total_posts	19.192	24.986	4.000	750.000
Learning performance	63.408	32.420	0.070	100.00

identify the achievement emotions for the remaining posts. Table 4 shows the descriptive statistics of the four kinds of achievement emotions contained within all the forum posts in 13 MOOCs. The descriptive statistics of the control variable and the learning performance of students are also included in Table 4 for the multilevel analysis in the next section.

Impact of Achievement Emotions on Learning Performance

First, we fit the data to the null model (M0) through Maximum Likelihood (ML) estimation to determine if multilevel modeling is the appropriate technique for analysis. The results of the estimated fixed and random effects for M0 are reported in Table 5. The ICC coefficient of the null model was 0.1585, indicating that 15.85% of the total variance of students' performance score

Table 5. Models Estimate for the Two-Level Analysis on Students' Learning.

	M0	M1	M3	M5
Fixed Effects (SD, t)				
Intercept ($\hat{\gamma}_{00}$)	58.72(3.58, 16.39) ***	58.02(3.65, 15.88) ***	58.26(3.42, 16.89) ***	56.26(3.41, 16.51) ***
PA ($\hat{\gamma}_{10}$)		−4.45(1.51, −2.95) **	−12.36(5.82, −2.13) *	−11.67(5.72, −2.04) *
PD ($\hat{\gamma}_{20}$)		3.65(1.68, 2.18) *	5.24(1.68, 3.12) **	−2.95(2.00, −1.48)
NA ($\hat{\gamma}_{30}$)		−21.66(2.06, −10.51) ***	−20.54(2.06, −9.96) ***	7.48 (2.58, 2.90) **
ND ($\hat{\gamma}_{40}$)		−27.02(3.93, −6.88) ***	−25.75(3.92, −6.58) ***	−22.71(3.89, −5.84) ***
Total_posts ($\hat{\gamma}_{50}$)		.28(.01, 33.68) ***	.28(.01, 33.60) ***	.43(.02, 18.81) ***
PD:*Total_posts* ($\hat{\gamma}_{21}$)				.83(.10, 8.40) ***
NA:*Total_posts* ($\hat{\gamma}_{31}$)				−2.98(.16, −18.45) ***
Random Effects (95% CI)				
Intercept ($\hat{\tau}_{00}$)	1165.1(9.10, 20.01)	166.2(9.13, 29.08)	146.5(8.50, 18.95)	141.7(8.36, 18.64)
PA ($\hat{\tau}_{11}$)			387 (12.91, 32.31)	373.5(12.66, 31.78)
Residual ($\hat{\sigma}_e^2$)	876.6(29.33, 29.89)	823.8(28.43, 28.98)	817(28.31, 28.86)	802.0(28.05, 28.59)
Model Fit				
Deviance	203806	202491	202346	201954
AIC	203812	202507	202366	201978
BIC	203836	202571	202445	202073
χ^2		1315.2	145.24	391.88
df	3	8	10	12
p		<.001	<.001	<.001
Reference		M0	M1	M3

Note:

1. Estimates, standard errors, and t were reported for fixed effects.
2. Estimates and 95% Confidence interval were reported for random effects.
3. * significant at the .05 level; ** significant at the .01 level; *** significant at the .001 level.
4. M0: null model; M1: adding level-1 predictors; M3: random slope; M5: adding interactions.

was embedded in the course-level, which means that multilevel modeling is the proper statistical model to utilize.

The following procedures were also conducted through ML estimation. Results of estimated fixed and random effects for each model (M1-M5), as well as model fit indexes, are presented in Table 5. Based on the results, M1, which added five level-1 predictors, fit the data significantly better than M0 with $\chi^2(5) = 1315.2$, $p < .001$. However, after adding the level-2 predictor (M2), the result of LRT was non-significant, as its p value was larger than .05. For the sake of parsimony, we kept M1 as the baseline mode for the following one to compare. When fitting M3, only adding PA to the random effects achieved convergency. The results of the LRT test showed that M3 was significantly better than M1 with $\chi^2(1) = 145.24$, $p < .001$.

With regard to fixing the intercept in M4, the result of the LRT test indicated that M3 is significantly better than M4 with $\chi^2(2) = 1048.2$, $p < .001$. Thus, we added interactions between four achievement emotions and total_posts on M5 based on M3. The convergency could only be achieved in the model by adding interactions between PD and total_posts, NA, and total_posts. The LRT yielded a $\chi^2(2) = 391.88$, $p < .001$. Thus, M5 is the final model.

As shown in Table 5, $\hat{\gamma}_{00} = 56.26$ (SD = 3.41, $p < .001$) indicates that the average of students' academic performance is 56.26 across all students and courses after controlling for other variables. We found that students' PA and ND scores can negatively predict their learning performance ($\hat{\gamma}_{10} = -11.67$, $\hat{\gamma}_{40} = -22.71$, $p < .001$); significantly, NA score and total_posts can significantly positively predict students' learning performance ($\hat{\gamma}_{30} = 7.48$, $\hat{\gamma}_{50} = 43$, $p < .01$), while their PD score cannot significantly predict their learning achievement ($\hat{\gamma}_{30} = -2.95$, $p > .05$). We also found that the influence of students' PD score on their academic performance became non-significantly negative from significantly positive after adding interactions, specifically, $\hat{\gamma}_{20}$ decreased from 5.24, $p < .01$ in M3 to -2.95, $p > .05$ in M5. Although NA still has a significant effect on students' learning performance $\hat{\gamma}_{30}$, it increased from -20.50 point in M3 to 7.48 in M5.

For the effects of interaction on students' learning performance, the interaction between students' PD score and total_posts significantly influenced their learning performance ($\hat{\gamma}_{21} = .83$, SD = .10, $p < .001$), while the interaction between NA and total_posts influenced their learning performance negatively ($\hat{\gamma}_{31} = -2.98$, SD = .16, $p < .001$). We plotted the moderator effects of total_posts on the relationship between PD and NA and total_posts, respectively, at high (mean +1 SD), medium (mean), and low (mean – 1 SD) levels of total_posts (see Figure 4). The results indicated that the indirect effect of two types of achievement emotions on students' learning performance via total_posts differed. For students who have a higher level of total_posts, their academic performance is usually higher than lower total_posts with the same level of PD (see A in Figure 4). Specifically, for the students with high or medium levels of total_posts, their

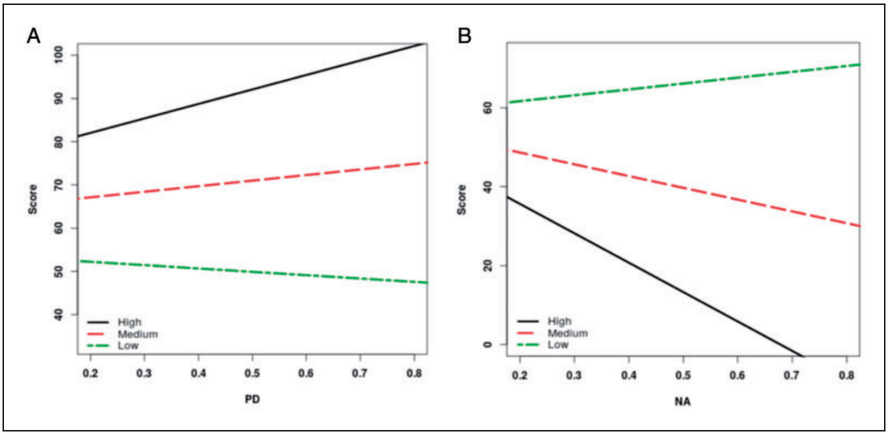

Figure 4. The Moderating Effect of total_posts on the Relationship Between PD and Score, NA and Score.

academic performance score increases as PD increases; however, for those with low levels of total_posts, their academic performance score decreases as PD increases. Differing from the interaction effect between PD and total_posts, students with a higher level of total_posts usually have a lower level of academic performance score with the same level of NA (see B in Figure 4). More specifically, for the students with high or medium levels of total_posts, their academic performance decreases as NA increases; however, for those with a low level of total_posts, their academic performance score increases as NA increases.

In addition, when compared to the null model, the final model has a reduction in residual variance of 8.5% and a reduction in intercept variance of 14.17%. These results indicated that the final model was able to explain a small to medium portion of level-1 variance and a rather medium portion of level-2 variance. The effect size related to the variance explained for the overall model is $f^2 = \frac{R^2}{1-R^2} = 0.21$, which means that the overall model has a medium to large effect size in explaining students' learning performance in this two-level multilevel model.

Discussion

Understanding the role of emotional experience in learning helps improve students' learning processes and outcomes in MOOCs (Gong et al., 2019). The results of this study are summarized as follows: (1) Negative activating emotions have a significant positive influence on learners' academic performance.

However, the interaction of negative activating emotions and total posts is negatively related to learners' learning performance. (2) Negative deactivating emotions are negatively related to learners' academic performance. (3) Positive deactivating emotions have no significant influence on learners' learning achievement. However, the interaction of positive deactivating emotions and total posts is positively related to learners' academic performance. (4) Positive activating emotions are negatively related to learners' academic performance, but this negative influence is much smaller compared to that of negative deactivating emotions.

Negative activating emotions positively contribute to academic performance possibly because they can motivate students to invest effort to avoid failure or solve problems, especially for those students who aim to succeed in MOOCs (D'Mello et al., 2014; Pekrun, 2017). Negative activating emotions are usually induced when the failure is judged to be due to controllable causes such as lack of effort (Pekrun, 2006). Thus, students may be motivated by negative activating emotions to adjust their learning behavior or learning strategies to avoid failure. For example, negative activating emotions such as anxiety can facilitate students' use of rigid learning strategies such as rehearsal and memorization of learning materials (Pekrun et al., 2011), which may be helpful for improving their learning performance. This finding is consistent with Gong et al. (2019), who found that learners' negative activating emotions, such as confusion, had a significant positive effect on their academic performance.

Negative deactivating emotions negatively contribute to students' learning, likely because they generally reduce learners' motivation to invest effort in learning (Pekrun et al., 2002). Negative deactivating emotions are usually induced when the learning activity is not sufficiently controllable and lacks any incentive value (Pekrun, 2006). Students will experience negative deactivating emotions if they have a low expectancy of success as well as a high expectancy of failure. Thus, students' attention, motivation, and strategy use would be undermined by their deactivating negative emotions (Pekrun, 2017). For example, negative deactivating emotions, such as boredom, reduce students' systematic use of learning strategies and lead to superficial processing of information (Pekrun et al., 2011), which may impair their achievement outcomes in MOOCs. This finding is consistent with the finding of Gong et al. (2019), who determined that learners' negative deactivating emotions, such as boredom, had a significant negative effect on their academic achievement.

Positive deactivating emotions are not related to learners' academic performance, likely because they reduce learners' motivation to continue making efforts to succeed (Sweeny & Vohs, 2012). Students experience positive deactivating emotions when the value of success or the non-occurrence of failure reaches its maximum (Pekrun, 2006). For example, learners will experience positive deactivating emotions, such as relief and relaxation, if they anticipate that they will be able to pass the exam. However, these emotions could reduce students' efforts and systematic use of learning strategies (Pekrun, 2017). Thus, positive deactivating emotions have a complex impact on students' academic performance, which, as shown in this study, may not always be positive and significant.

Surprisingly, positive activating emotions have a negative effect on academic performance. The reason may be that positive activating emotions distract students' attention away from the task they are performing (Pekrun, 2017). Positive activating emotions are usually experienced when the learning activity and success are positively valued and perceived as resulting from subjective internal control (Pekrun, 2006). Positive activating emotions, such as enjoyment and pride, can increase interest and motivation and facilitate the systematic use of learning strategies (Fredrickson, 2001). However, these emotions reduce students' academic learning performance, which requires sustained attention (Pekrun, 2017). Thus, positive activating emotions also have a complex impact on students' academic performance/achievement, which was further reflected as having no effect on learner dropout/survival in MOOCs in the study by Xing et al. (2019). This finding corresponds to the study of Meinhardt and Pekrun (2003), who concluded that positive emotions reduced students' performance by distracting their attention.

Moreover, the total number of posts had a moderating effect on the relationship between positive deactivating emotions and students' learning, as well as between negative activating emotions and students' learning. Positive deactivating emotions are significantly positively related to learners' academic performance after adding interaction with total posts, which indicates that learners' academic performance increases as positive deactivating emotions increase for those students with more posts. This may be because although positive deactivating emotions reduce motivation in the short term, they reinforce motivation to reengage with learning in the long run (Sweeny & Vohs, 2012). Therefore, when students engage in long-term forum discussions, which results in a large number of posts, positive deactivating emotions will maintain their motivation to learn, thereby improving their academic performance. Negative activating emotions negatively contribute to students' learning after adding interaction with total posts, which indicates that learners' academic performance decreases as negative activating emotions increase for those students with more posts. Our hypothesis is that students build an increasingly negative learning climate when they post too many texts expressing negative emotions, even though these emotions are activating, thereby decreasing their academic performance. This finding is similar to the research of Dmoshinskaia (2016), who found that learners who experienced many negative emotions were likely to drop the course. However, learners who experienced some appropriate negative emotions were likely to complete the course.

This study has significant implications for effective MOOC building. New mechanisms are required to help learners regulate and manage their emotions in order to move towards more successful MOOCs. This study built an effective and high-performance machine-learning model with easy features extraction to automatically detect students' achievement emotions, which could be employed to analyze students' emotional states in MOOC forums and beyond. Moreover, we identified the role of achievement emotions in student academic performance, yielding a new direction for the adaptation and personalization of MOOCs from an affective perspective. It will be possible to improve students' learning by recommending more appropriate learning materials and strategies for students to regulate their emotions if learners' emotional states can be recognized automatically by the MOOC platform (Cobos et al., 2019). Understanding the achievement emotions of learners in MOOCs can also help instructors understand students better and improve students' learning by providing emotional support (Hew et al., 2020). For example, teachers should try to minimize negative deactivating emotions because of their negative impact on students' learning. Teachers can also introduce some level of confusion in the learning topics and forums to ignite students' curiosity, since negative activating emotions can positively affect students' learning performance. However, they should pay careful attention to students who post a fair amount of negative deactivating emotions consistently because of the negative impact on students' learning of such emotions when interacting with total posts.

There are some limitations to this study. First, the findings in the study were relatively correlational, even though we selected a large MOOC sample and controlled some related variables. Significant relationships between achievement emotions and students' learning were identified. However, the results could also be interpreted as students' individual differences. More control experiments could be conducted in the future to examine how achievement emotions influence student academic performance in MOOCs. Second, the study only considered learners who had posted on MOOCs forum, which ignored learners who are active readers of forums but did not actually post on them.

Conclusion

This study employed an integrative framework of achievement emotions to uncover the influencing mechanism of achievement emotions on students' learning from more than a dozen MOOC courses. First, a machine-learning model with high prediction performance was built to automatically identify the achievement emotions in forum posts. Then, multilevel modeling was used to quantify the effect of learners' achievement emotions on their learning performance. The novelty of this study comes from the theoretical approach used in the study and the new research direction for quantifying the connection between achievement emotions and student learning relying on large-scale data analytics. Tracking and monitoring student achievement emotions can help MOOC platforms and instructors provide related emotional feedback to students automatically or manually.

There are a few future research directions. First, even though our findings indicate that different achievement emotions influenced students' learning differently, the reasons behind them are unclear. Interviews and surveys could be conducted in future studies to further examine the relationship between student achievement emotions and learning performance. Second, we provide MOOC designers and instructors with interventions to support students' emotional experiences and interactions. Future studies can apply these interventions to MOOCs to test their effectiveness in improving students' learning performance. Third, beyond examining students' emotions from textual posts, multimodal emotion detection can be used in future studies to enhance the detection performance of students' achievement emotions from multiple sources such as behavioral observation, facial expressions, and physiological reactions. More studies are also encouraged to replicate this research in other MOOC platforms and courses.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Bowen Liu https://orcid.org/0000-0003-3151-578X
Wanli Xing https://orcid.org/0000-0002-1446-889X

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/tac.1974.1100705

Almatrafi, O., Johri, A., & Rangwala, H. (2018). Needle in a haystack: Identifying learner posts that require urgent response in MOOC discussion forums. *Computers & Education*, *118*, 1–9. https://doi.org/10.1016/j.compedu.2017.11.002

Arguel, A., Lockyer, L., Lipp, O. V., Lodge, J. M., & Kennedy, G. (2017). Inside out: Detecting learners' confusion to improve interactive digital learning environments. *Journal of Educational Computing Research*, *55*(4), 526–551. https://doi.org/10.1177/0735633116674732

Bonafini, F. C., Chae, C., Park, E., & Jablokow, K. W. (2017). How much does student engagement with videos and forums in a MOOC affect their achievement? *Online Learning*, *21*(4), 223–240. https://doi.org/10.24059/olj.v21i4.1270

Buhr, E. E., Daniels, L. M., & Goegan, L. D. (2019). Cognitive appraisals mediate relationships between two basic psychological needs and emotions in a massive open online course. *Computers in Human Behavior*, *96*, 85–94. https://doi.org/10.1016/j.chb.2019.02.009

Cheung, M. W. L., & Au, K. (2005). Applications of multilevel structural equation modeling to cross-cultural research. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(4), 598–619. https://doi.org/10.1207/s15328007sem1204_5

Cobos, R., Jurado, F., & Villén, Á. (2019). *Moods in MOOCs: Analyzing emotions in the content of online courses with edX-CAS* [Paper presentation]. 2019 IEEE Global Engineering Education Conference (EDUCON) (pp. 1467–1474). IEEE. https://doi.org/10.1109/educon.2019.8725107

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. https://doi.org/10.1016/j.learninstruc.2012.05.003

Dillon, J., Bosch, N., Chetlur, M., Wanigasekara, N., Ambrose, G. A., Sengupta, B., & D'Mello, S. K. (2016, June). *Student emotion, co-occurrence, and dropout in a MOOC context* [Paper presentation]. 9th International Conference on Educational Data Mining (pp. 353–357). https://files.eric.ed.gov/fulltext/ED592723.pdf

Dmoshinskaia, N. (2016). *Dropout prediction in MOOCs: Using sentiment analysis of users' comments to predict engagement* [Unpublished master's thesis]. University of Twente. http://essay.utwente.nl/70461/1/Dmoshinskaia-MA-BMS.pdf

Duncan, T. E., Alpert, A., & Duncan, S. C. (1998). Multilevel covariance structure analysis of sibling antisocial behavior. *Structural Equation Modeling: A Multidisciplinary Journal*, *5*(3), 211–228. https://doi.org/10.1080/10705519809540102

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218–226. https://doi.org/10.1037/0003-066x.56.3.218

Gao, R., Hao, B., Li, H., Gao, Y., & Zhu, T. (2013). Developing simplified Chinese psychological linguistic analysis dictionary for microblog. *Brain and Health Informatics*, *8211*, 359–368. https://doi.org/10.1007/978-3-319-02753-1_36

Gong, L., Liu, Y., & Zhao, W. (2019, March). *Dynamics of emotional states and their relationship with learning outcomes during learning python with MOOC* [Paper presentation]. 2019 7th International Conference on Information and Education Technology (pp. 71–76). https://doi.org/10.1145/3323771.3323821

Hascher, T. (2010). Learning and emotion: Perspectives for theory and research. *European Educational Research Journal*, *9*(1), 13–28. https://doi.org/10.2304/eerj.2010.9.1.13

Henderikx, M., Lohr, A., & Kalz, M. (2019). *Enjoyed or bored? A study into achievement emotions and the association with barriers to learning in MOOCs* [Paper presentation]. Transforming Learning With Meaningful Technologies (pp. 15–27). Springer. https://doi.org/10.1007/978-3-030-29736-7_2

Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computers & Education*, *145*, 103724. https://doi.org/10.1016/j.compedu.2019.103724

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Jarrell, A., Harley, J. M., & Lajoie, S. P. (2016). The link between achievement emotions, appraisals, and task performance: Pedagogical considerations for emotions in CBLEs. *Journal of Computers in Education*, *3*(3), 289–307. https://doi.org/10.1007/s40692-016-0064-3

Kechaou, Z., Ammar, M. B., & Alimi, A. M. (2011, April). *Improving e-learning with sentiment analysis of users' opinions* [Paper presentation]. 2011 IEEE Global Engineering Education Conference (EDUCON) (pp. 1032–1038). https://doi.org/10.1109/educon.2011.5773275

Liu, Z., Pinkwart, N., Liu, H., Liu, S., & Zhang, G. (2018). Exploring students' engagement patterns in SPOC forums and their association with course performance. *EURASIA Journal of Mathematics, Science and Technology Education*, *14*(7), 3143–3158. https://doi.org/10.29333/ejmste/91667

Meinhardt, J., & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition & Emotion*, *17*(3), 477–500. https://doi.org/10.1080/02699930244000039

Mite-Baidal, K., Delgado-Vera, C., Solís-Avilés, E., Espinoza, A. H., Ortiz-Zambrano, J., & Varela-Tapia, E. (2018). *Sentiment analysis in education domain: A systematic literature review* [Paper presentation]. International Conference on Technologies and Innovation (pp. 285–297). https://doi.org/10.1007/978-3-030-00940-3_21

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, *41*(4), 359–376. https://doi.org/10.1111/j.1464-0597.1992.tb00712.x

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. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2011). Emotions as drivers of learning and cognitive development. In R. A. Calvo & S. K. D'Mello (Eds.), *New perspectives on affect and learning technologies* (pp. 23–39). Springer. https://doi.org/10.1007/978-1-4419-9625-1_3

Pekrun, R. (2017). Emotion and achievement during adolescence. *Child Development Perspectives*, *11*(3), 215–221. https://doi.org/10.1111/cdep.12237

Pekrun, R., & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass*, *4*(4), 238–255. https://doi.org/10.1111/j.1751-9004.2010.00259.x

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. https://doi.org/10.1207/s15326985ep3702_4

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC*. University of Texas at Austin.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, *57*(1), 27–53. https://doi.org/10.1146/annurev.psych.56.091103.070234
Rani, S., & Kumar, P. (2017). A sentiment analysis system to improve teaching and learning. *Computer*, *50*(5), 36–43. https://doi.org/10.1109/mc.2017.133
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136
Stapleton, L. M. (2006). Using multilevel structural equation modeling techniques with complex sample data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 345–383). Information Age Publishing.
Sweeny, K., & Vohs, K. D. (2012). On near misses and completed tasks: The nature of relief. *Psychological Science*, *23*(5), 464–468. https://doi.org/10.1177/0956797611434590
Tucker, C., Pursel, B. K., & Divinsky, A. (2014). Mining student-generated textual data in MOOCs and quantifying their effects on student performance and learning outcomes. *The ASEE Computers in Education (CoED) Journal*, *5*(4), 84–95. https://doi.org/10.18260/1-2-22840
Wen, M., Yang, D., & Rose, C. (2014, July). *Sentiment analysis in MOOC discussion forums: What does it tell us?* [Paper presentation]. 7th International Conference on Educational Data Mining (pp. 130–137). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.660.5804&rep=rep1&type=pdf
Xing, W., Tang, H., & Pei, B. (2019). Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs. *The Internet and Higher Education*, 43, 100690. https://doi.org/10.1016/j.iheduc.2019.100690
Yang, D., Wen, M., Howley, I., Kraut, R., & Rose, C. (2015). *Exploring the effect of confusion in discussion forums of massive open online courses* [Paper presentation]. The Second (2015) ACM Conference on Learning @ Scale (pp. 121–130). ACM. https://doi.org/10.1145/2724660.2724677

Author Biographies

Bowen Liu is a PhD student in Educational Technology, Faculty of Education, East China Normal University. His research interests are educational data mining and learning analytics.

Wanli Xing is an assistant professor of Educational Technology at University of Florida. His research interests are artificial intelligence, learning analytics, STEM education and online learning.

Yifang Zeng is a PhD student in quantitative educational psychology at Texas Tech University. Her research interests are quantitative methods and their applications in education and psychology.

Yonghe Wu is a professor in Educational Technology, Faculty of Education, East China Normal University. His research interests are artificial intelligence, learning analytics and STEM education.