Exploring the Pattern and Influence Factors in the Correlation Between Learning Engagement and Academic Achievement in Online Learning

Yaoyuan Zhou [1], Yuhan Gai[1], Jiamin Ren[1],Qingyang Liu[1]

[1] School of Educational Technology in Beijing Normal University

Abstract**:** *This study investigates the complex patterns linking learning engagement to academic achievement in online learning environments. By analyzing data from 217 university students using regression, K-means clustering, ANOVA, and Epistemic Network Analysis (ENA), the research reveals that simple positive correlations obscure distinct student profiles. Cluster analysis identified three groups(high engagement & achievement, low engagement & achievement, low engagement/high achievement), including a notable "low-engagement/high-achievement" profile. Results indicate that behavioral engagement serves as the primary differentiator between high and low academic achievement. Furthermore, ENA findings suggest that low-engagement/high-achievement students compensate for lower behavioral frequency with superior cognitive quality, characterized by integrated knowledge generation, whereas low-performing students rely on superficial referencing. These findings challenge linear engagement–achievement models, demonstrating that academic success depends less on the magnitude of engagement than on the structural quality of cognitive processing and behavioral execution.*

Keywords: Online learning, Learning engagement, Learning analysis, Discourse analysis

1. Introduction

Originating in the late 20th century, online learning was initially introduced into higher education to provide flexible learning modalities for students—thereby expanding the educational market—and to afford greater geographical flexibility for faculty and staff (McGaughey, 2022; Xie et al., 2020). Over the past two decades, the significance of online learning has increasingly surged, establishing it as a pivotal educational format characterized by its unique environment and conditions (National Center for Educational Statistics [NCES]). However, research indicates that the lack of effective teacher–student interaction within online environments impedes communication and the generation of personalized feedback, resulting in challenges related to low learning engagement and high attrition rates (Freidhoff, 2021; Trust, 2020). While online learning engagement is recognized as a critical factor influencing student achievement and instructional quality, with numerous studies demonstrating a positive correlation between engagement and learning outcomes (Lee, 2022), this relationship is moderated by various factors. Nevertheless, simple correlations are insufficient to accurately characterize student dynamics in online learning. Consequently, it is imperative to investigate the specific patterns underlying the relationship between learning engagement and academic achievement, as well as the mechanisms driving the formation of these patterns.

2. Literature Review

Learning engagement's definition can be dated back to Moore (1989), who conceptualized engagement through the lens of three distinct interaction types: learner–content, learner–instructor, and learner–learner interactions. In the specific context of online learning, Martin (2022) synthesized these perspectives to propose a refined definition: "online learning engagement is the productive cognitive, affective, and behavioral energy that a learner exerts interacting with others and learning materials and/or through learning activities and experiences in online learning environments."

Martin's framework delineates learning engagement into three dimensions: cognitive, affective, and behavioral. Cognitive engagement involves the psychological energy invested in learning activities, emphasizing mental exertion and interactions with materials or peers, for example, learning how to conduct online learning appropriately and how to engage effectively in collaborative interaction and learning with materials. Affective engagement captures students' emotional reactions, reflecting both individual responses and technological influences within the online context. It mainly focuses on building relationships and fostering a sense of community and belonging in the process of online learning. Behavioral engagement represents the physical manifestation of these cognitive and affective states through tangible actions. This includes participation in both academic and non-academic activities, accompanied by the demonstration of positive behaviors.

The association between learning engagement and academic achievement has been repeatedly substantiated across numerous studies. Puri (2025) indicated that in the context of mobile-assisted language learning, mobile devices effectively reduce cognitive load and anxiety, thereby serving as a mediating effect that explains the positive correlation between learning engagement and learning achievement. Furthermore, researchers (Hatahet, 2025; Oliver, 2025) have confirmed the positive impact of multidimensional engagement on learning achievement from the perspectives of cognitive engagement, affective engagement, and behavioral engagement, respectively. Collectively, these studies establish a solid foundation for conceptualizing the relationship between learning engagement and academic performance.

Current research predominantly focuses on the associations between learning engagement and other constructs, such as self-achievement (Wang, 2022) and teacher–student interaction (Ong, 2023), in relation to academic outcomes. However, scholars have noted that such studies typically employ cross-sectional designs, necessitating caution when interpreting causal relationships (Miao, 2022). Only a minority of scholars have examined the internal dynamics of learning engagement itself, proposing that engagement is multidimensional and that its distinct types interact with one another (Hatahet, 2025). This suggests the existence of more complex, underexplored relationships; indeed, the complex dynamics exhibited by students extend beyond simple linear associations, and latent patterns between these variables remain obscured. Moreover, while students of different profiles exhibit variations across the three dimensions of engagement, few studies have deeply investigated the underlying causes driving these multidimensional disparities. Accordingly, this study proposes the following two research questions:

1. What is the relationship between students’ online learning engagement and their academic achievement, and what characteristics does the student distribution exhibit?
2. What features characterize the epistemic networks of different student profiles, and what are the distinctions and commonalities among them?

3. Methods

3.1 Participant and Procedures

This study utilized course data from the three separate class sessions of the "Educational Technology" curriculum—comprising 14 classes sessions—offered during the 2024 spring for 15weeks at a university in Beijing. The course was delivered online via a self-developed platform, where students were required to receive course notifications, engage in online learning, post classroom comments, and complete tasks assigned within each session. The platform systematically collected students' post-class assignment scores, course completion rates, and comment content, while also capturing all operational log data generated by student users during the learning process. Figure 1 shows the shortcut of platform applied in online learning.

A total of 309 students provided informed consent, confirming their voluntary and anonymous participation in the study. Following data cleaning and verification procedures, 217 valid samples were obtained. The final sample consisted of 149 females (approximately 68.6%) and 68 males (approximately 31.4%), representing a diverse range of majors including law, chemistry, mathematics, education, and literature. We collected 432 comments in the platform overall.

Fig.1 Shortcut of the self-developed online learning platform

3.2 Instrument and Data Analysis

To quantify students' learning engagement, this study adopted the conceptual framework proposed by Martin (2022) and measured engagement across three distinct dimensions. First, cognitive engagement was assessed using students' participation assessment scores, which were calculated as a composite of content completion (34%), chapter quizzes (50%), and attendance (16%). Second, regarding affective engagement was assessed by questions from the Online Self-regulated Learning Questionnaire (OSLQ) developed by Barnard et al. (2009); the overall Cronbach's alpha coefficient for the scale was 0.90, supporting the instrument's validity. Finally, behavioral engagement was measured by collecting interaction log data from the online platform across five dimensions: page views, key-point clicks, comments, AI utilization, and pause frequency. To measure the behaviour engagement, we calculated the frequency of interaction and converts values respectively into its percentile rank as the final score.Ultimately, the mean score of these three dimensions was computed to determine the students' overall level of learning engagement. At the end of the semester, students were required to complete an examination, and their exam scores were used as the measure of academic achievement.

The analytical procedure consisted of three steps. Initially, we performed a linear regression to assess the predictive validity of the overall learning engagement index on academic achievement. Following this, K-means clustering was utilized to uncover latent profiles characterizing different engagement-achievement patterns. In the final stage, a Mixed designed ANOVA were applied to decompose the specific impacts of the three engagement dimensions (cognitive, affective, and behavioral) on academic achievement, and to characterize the heterogeneity of engagement profiles across the identified clusters.

To characterize the features of different student profiles, all comment data from the online instructional platform were collected and encoded using the asynchronous online discussion coding framework established by Biasutti (2017). The students' comment texts were coded across seven dimensions: inferencing, producing, developing, evaluating, summarizing, organizing, and supporting. Two researchers participated in the coding process. After brief training to ensure conceptual alignment in the coding scheme, both researchers independently conducted the coding based on the agreed concepts, and the final Cohen's kappa reached 0.91, indicating a high level of agreement and strong inter-coder reliability.Subsequently, Epistemic Network Analysis (ENA) was conducted using the web-based ENA tool (https://app.epistemicnetwork.org/) to model the cognitive characteristics of the different student profiles and to visualize the divergences in cognitive networks between groups.

4. Result

4.1 The Correlation and Pattern Between Learning Engagement and Academic Achievement

The descriptive statistics of learning engagement and academic achievement for all students are shown in Table 1. It can be seen that the average score for learning engagement is lower than the average score for academic achievement, but the data stability for both is similar. Among the three types of learning engagement, the average scores for cognitive engagement, affective engagement, and behavioral engagement decrease sequentially, and the data stability also gradually declines. This indicates that students generally have a high level of cognitive engagement, while affective and behavioral engagement vary greatly from person to person, with significant differences between samples.

Table 1. Descriptive statistics on learning engagement and academic achievement

Descriptive statistics	Learning engagement	Cognitive engagement	Affective engagement	Behavior engagement	Academic achievement
Mean	72.778	86.677	77.917	53.741	89.065
Stand Deviation	10.971	4.313	15.982	29.331	8.965

The results from both simple and multiple linear regression analyses indicated that overall learning engagement significantly and positively predicted academic achievement (β=0.179, t=3.28, p=0.001). However, the model exhibited weak explanatory power (R^2=0.048) In the multiple linear regression analysis of the three engagement dimensions, only cognitive engagement emerged as a significant predictor; nonetheless, the explanatory power remained limited (t=3.868, p<0.001, R^2 = 0.120) These findings suggest that while a statistical association exists at the aggregate level, a single composite engagement metric may obscure more complex underlying mechanisms. Given the potential existence of distinct patterns within the sample, it is necessary to employ cluster analysis to further explore whether specific engagement profiles exist that could better elucidate the variance in academic achievement.

Subsequently, K-means clustering was performed on the standardized total learning engagement scores and academic achievement to identify potential "engagement–achievement" learning patterns within the sample. To determine the optimal number of clusters, we experimented with various clustering configurations; the computational results are presented in Table 2. Balancing model fit with analytical interpretability, we selected K=3 as the final clustering solution, as it yielded the highest Silhouette Coefficient and a relatively high Calinski-Harabasz Index. The algorithm parameters were set to random_state=42 and n_init=10.

Table 2. Different results of K-means clustering

K	Silhouette Coefficient	Calinski-Harabasz	Sum of Squared Errors
2	0.3822	147.75	257.23
3	0.3849	167.17	169.37
4	0.3718	168.2	128.82
5	0.3681	170.82	102.77

As illustrated in Table 3 and Figure 2, the clustering results yielded three distinct student groups. In terms of the two indicators, Cluster 1 exhibited high levels of both engagement and achievement. Cluster 2 was characterized by generally lower academic achievement and a dispersed but overall low distribution of engagement. Cluster 3 presented the most distinct profile, characterized by low engagement but high achievement.

Table 3. The details of three Clusters' student

Cluster	Number	Learning Engagement		Academic Achievement	
		M	SD	M	SD
Cluster 1	102	81.57	6.06	93.32	5.39
Cluster 2	43	69.26	9.07	75.26	6.65
Cluster 3	72	64.23	6.17	91.28	5.53

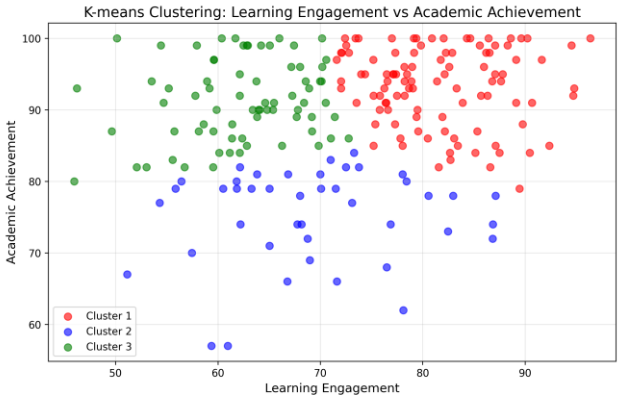


Fig. 2 K-means clustering results under the dimensions of learning engagement and academic performance

4.2 The Differences of Three Dimensions of Learning Engagement in Three Patterns

To further investigate the differences among the three student profiles across the three dimensions of learning engagement, a mixed-design Analysis of Variance (ANOVA) was conducted. In this analysis, the cluster category served as the between-subjects factor, while the dimension of learning engagement served as the within-subjects factor. The results are presented in Table 4. Regarding the between-subjects effects, significant differences in overall learning engagement levels were observed among the different cluster groups (F (2,214) =175.0, p<.001), verifying that the cluster analysis effectively discriminated participants based on engagement dimensions. Regarding the within-subjects' effects, significant differences were found across the three dimensions of learning engagement for the overall sample (F (2, 428) = 272.0, p < .001), indicating distinct levels of student engagement across cognitive, affective, and behavioral domains. Furthermore, the significant interaction effect between the engagement dimensions and cluster category suggests that the variations in learning engagement among the different clusters are not merely attributable to simple differences in overall magnitude (F (4, 428) = 51.0, p < .001); rather, they manifest as distinct patterns within the structural composition of engagement.

Table 4. The result of Analysis of Variance (ANOVA)

Effect type		Sum of squares	df	Mean square	F	P
Within subject's effect	Engagement type	143822	2	71911	272.0	<.001**
	Engagement type * Cluster	52879	4	13470	51.0	<.001**
Between subject's effect	Cluster	48384	2	24192	175.0	<.001**

To elucidate these patterns, this study proceeded with an analysis of estimated marginal means for the interaction effect (as shown in Figure 3), which clearly delineates three distinct engagement profiles. As indicated by the figure, the cognitive dimension exhibited the highest mean scores, with all three student clusters displaying a high degree of convergence. This suggests that most students possess a solid foundation in cognitive volition, indicating that cognitive factors are not the primary bottleneck distinguishing academic achievement. In contrast, mean scores for affective engagement began to decline, showing moderate differentiation between groups; notably, Cluster 3 (characterized by low engagement but high achievement) exhibited the most pronounced decrease. In the behavioral engagement dimension, a precipitous divergence emerged among the three patterns. Students in the high-engagement/high-achievement profile maintained elevated levels, whereas the other two clusters—characterized by lower overall engagement—experienced a sharp decline in scores.

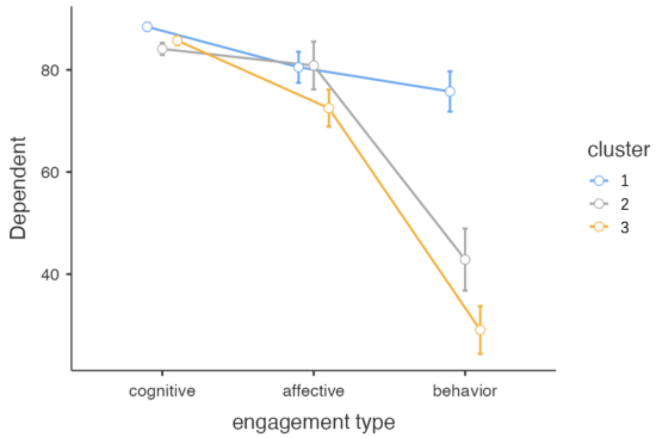
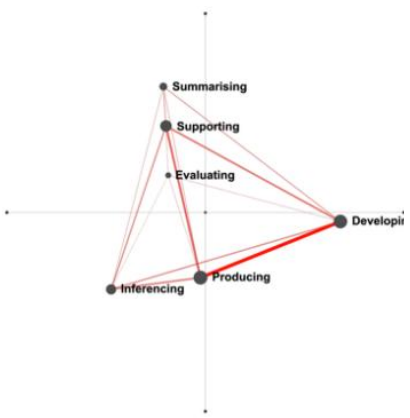


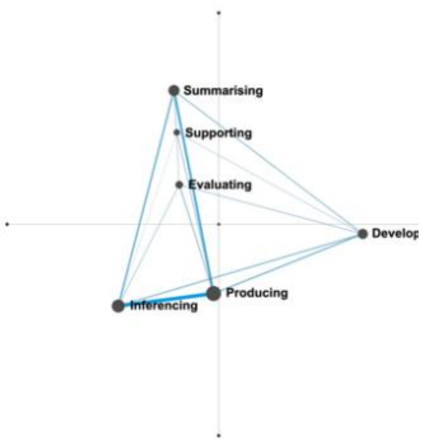
Fig.3 Estimated marginal means of three type of students

4.3 The Epistemic Network Analysis Result of Three Pattern Students

However, current results remain insufficient to explain why Cluster 3 students achieved high academic achievement despite lower levels of learning engagement. Consequently, this study encoded the comments posted by the three student clusters on the online learning platform and employed Epistemic Network Analysis (ENA) to delineate and compare the cognitive trajectories of different students, thereby uncovering the fundamental sources of these disparities. A moving window of 4 lines was used to capture connections between the current code and codes in the preceding 3 lines. The model achieved strong goodness of fit, with co-registration correlations of 0.92 (Pearson) for the first dimension, and 0.95 (Pearson) for the second dimension, ensuring the visualization accurately reflects the underlying network structure. The final analytical results are illustrated in Figures 4, 5 and 6.

Figure 4 presents the Epistemic Network Analysis results for Cluster 1, revealing that producing, developing, and supporting are the most prominent nodes. Students in this cluster tended to generate and offer ideas for constructing the project during forum interactions (producing), complemented by a high frequency of cognitive deepening (developing) and community support (supporting). In contrast, for Cluster 2, characterized by both lower learning engagement and academic achievement in figure 5, the centroid of the cognitive network shifted toward the left. While similarly centered on content generation (producing), this was primarily linked with content summarization (summarizing) and referencing (inferencing). Compared to Cluster 1, Cluster 2 demonstrated a more mechanical understanding of content, lacking deep reflective behaviors. Figure 6 illustrates Cluster 3 (low engagement, high achievement), which featured numerous high-frequency nodes with dense inter-nodal connections. It can be inferred that although these students exhibited a deficit in overall engagement magnitude, they maintained active cognitive processes and critical understanding; their engagement was characterized by lower frequency but higher quality, indicative of superior learning efficiency.





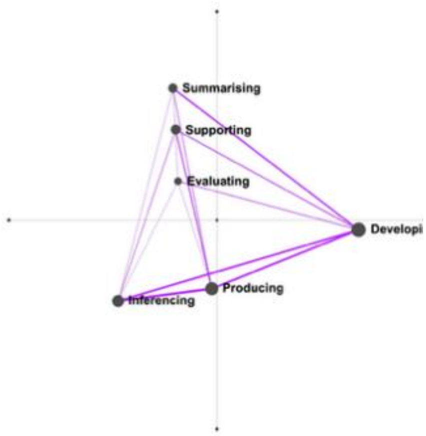


Fig.4 Epistemic Network of Cluster 1 Students | Fig.5 Epistemic Network of Cluster 2 Students | Fig.6 Epistemic Network of Cluster 3 Students

5. Conclusion

By employing regression analysis, K-means clustering, Analysis of Variance (ANOVA), and Epistemic Network Analysis (ENA), this study comprehensively addressed the relationship and distributional characteristics between students' online learning engagement and academic achievement. Furthermore, through the integration of multiple statistical methodologies, the research elucidated the underlying causes driving the emergence of distinct student profiles.

First, regression analysis revealed a positive correlation between students' learning engagement and academic achievement, a finding that aligns with previous research; however, the effect size was limited and the correlation proved weak. Moreover, in the multiple regression analysis, only cognitive engagement demonstrated a significant association with academic achievement. These results indicate that student dynamics cannot be adequately generalized by a simple positive correlation. Consequently, K-means clustering results delineated the student population into three primary groups: high engagement/high achievement, lower engagement/low achievement, and low engagement/high achievement.

Subsequently, ANOVA was employed to assess the differences among the three student clusters across the three dimensions of learning engagement. The estimated marginal means indicated that the disparities in engagement among the three groups were primarily concentrated in behavioral engagement. Specifically, students in the high-achievement/high-engagement cluster maintained elevated levels across all three engagement dimensions. In contrast, the other two clusters exhibited a substantial decline in behavioral engagement, underscoring the critical importance of behavioral participation in the context of online learning.

Finally, we utilized Epistemic Network Analysis (ENA) to further investigate the cognitive characteristics manifested in the comments of different student profiles on the online learning platform. A comparison of the resulting network visualizations reveals that high-engagement/high-achievement students concentrated on knowledge generation (producing), support (supporting), and deepening (developing). Conversely, lower-engagement/low-achievement students remained fixated on knowledge referencing (inferencing) and summarization (summarizing). Notably, students with low engagement but high achievement exhibited more comprehensive and integrated cognitive features, characterized by greater overall intellectual vitality. It can be inferred that within the low-engagement population, divergences in academic performance may be attributable to variations in the depth and diversity of cognition.

Through this study, we make both theoretical and practical contributions. Theoretically, this research refines the correlational patterns between learning engagement and academic achievement, extending beyond the singular positive correlation conclusions found in prior literature. Practically, it offers valuable insights for instructors delivering online learning to analyze student dynamics and enhance academic performance. Specifically, educators are advised to encourage students to actively utilize online platforms to bolster behavioral engagement, and to foster the posting of comments characterized by greater cognitive depth and breadth.

6. Limitation

First, the study collected data from only three online class sessions, and this relatively short data collection window may have limited the representativeness of the dataset. Second, the study did not examine specific behaviors within each subdimension of learning engagement, resulting in limited dimensional granularity; these issues may partly explain the relatively low coefficient of determination in the regression analysis between learning engagement and academic achievement. Finally, in the cluster analysis of student types, the selected clustering solution yielded a relatively low silhouette coefficient, and the scatterplot did not show clear separation between clusters, suggesting that the clustering results lacked sufficient explanatory power.

References

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. The internet and higher education, 12(1), 1-6. https://doi.org/10.1016/j.iheduc.2008.10.005

Biasutti, M. (2017). A coding scheme to analyse the online asynchronous discussion forums of university students. Technology, Pedagogy and Education, 26(5), 601-615. https://doi.org/10.1080/1475939X.2017.1365753

Freidhoff, J. R. (2021). Michigan's k-12 virtual learning effectiveness report 2019-20. Michigan Virtual University. https://michiganvirtual.org/research/publications/michigans-k-12-virtual-learning-effectiveness-report-2019-20/

Hatahet, T. (2025). Supporting active learning opportunities with feedback enhances student achievement in pharmaceutical innovation module: Implication of behavioural and cognitive engagement. Social Sciences & Humanities Open, 11, 101347. https://doi.org/10.1016/j.ssaho.2025.101347

Lee, J., Park, T., & Davis, R. O. (2022). What affects learner engagement in flipped learning and what predicts its outcomes?. British Journal of Educational Technology, 53(2), 211-228. https://doi.org/10.1111/bjet.12717

Martin, F., & Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. Educational Psychologist, 57(3), 162-177. https://doi.org/10.1080/00461520.2022.2089147

McGaughey, F., Watermeyer, R., Shankar, K., Suri, V. R., Knight, C., Crick, T., … Chung, R. (2022). 'This can't be the new norm': academics' perspectives on the COVID-19 crisis for the Australian university sector. Higher Education Research & Development, 41(7), 2231–2246. https://doi.org/10.1080/07294360.2021.1973384

Miao, J., & Ma, L. (2022). Students' online interaction, self-regulation, and learning engagement in higher education: The importance of social presence to online learning. Frontiers in psychology, 13, 815220. https://doi.org/10.3389/fpsyg.2022.815220

Moore, M. G. (1989). Three types of interaction. https://doi.org/10.1080/08923648909526659

National Center for Educational Statistics (NCES). (2021). Condition of education 2021. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021144

Olivier, E., Galand, B., Belanger, J., & Morin, A. J. (2025). Multilevel student-perceived teaching practices profiles: Associations with competence beliefs, task value, behavioral engagement, and academic achievement. Learning and Individual Differences, 118, 102631. https://doi.org/10.1016/j.lindif.2025.102631

Ong, S. G. T., & Quek, G. C. L. (2023). Enhancing teacher–student interactions and student online engagement in an online learning environment. Learning environments research, 26(3), 681-707. https://doi.org/10.1007/s10984-022-09447-5

Puri, B., Mushtaque, I., Fang, S., Chenhe, G., & Younas, A. (2025). The impact of mobile-based language learning on speaking and learning anxiety, engagement and achievement in Chinese language learning: The mediating role of cognitive load. Acta Psychologica, 259, 105400. https://doi.org/10.1016/j.actpsy.2025.105400

Trust, T., & Whalen, J. (2020). Should Teachers be Trained in Emergency Remote Teaching? Lessons Learned from the COVID-19 Pandemic. Journal of Technology and Teacher Education. https://doi.org/10.70725/307718pkpjuu

Wang, Y., Cao, Y., Gong, S., Wang, Z., Li, N., & Ai, L. (2022). Interaction and learning engagement in online learning: The mediating roles of online learning self-efficacy and academic emotions. Learning and Individual Differences, 94, 102128. https://doi.org/10.1016/j.lindif.2022.102128

Xie, X., Siau, K., & Nah, F. F. H. (2020). COVID-19 pandemic – online education in the new normal and the next normal. Journal of Information Technology Case and Application Research, 22(3), 175–187. https://doi.org/10.1080/15228053.2020.1824884