

MINING THE LEARNING BEHAVIOR CHARACTERISTICS OF COLLEGE STUDENTS IN BLENDED LEARNING: AN EMPIRICAL STUDY BASED ON A DECADE OF LONGITUDINAL DATA

HongJun Zhang^{1*}, XinXin Kang²

¹*Office of Information Construction and Management, Shaanxi Normal University, Xi'an 710062, Shaanxi, China.*

²*Information Office, Xi'an Conservatory of Music, Xi'an 710061, Shaanxi, China.*

**Corresponding Author: HongJun Zhang*

Abstract: Blended learning has become the mainstream model in higher education. However, existing studies on learners' behavioral characteristics mostly rely on short-term and small sample data, and lack in-depth exploration of the intrinsic relationship between objective behavioral logs and subjective self-assessments. This study, based on ten-year longitudinal tracking data (N=3693), integrates platform objective learning logs and learners' subjective self-assessment scales, uses stepwise regression to screen key behavioral variables, and then employs K-means clustering to identify learners' behavioral types. The results show that there is a systematic separation between objective academic performance and subjective self-assessment results, and gender and academic background are important moderating variables; the impact of learning behaviors on learning outcomes exhibits a non-linear characteristic, and interaction quality and self-regulation ability are the core levers beyond "time accumulation"; self-assessed learning behaviors are generally high and have relatively small individual differences, which contrasts sharply with the significant individual differences in objective behavioral indicators. This study provides a methodological example for multi-source and longitudinal learning analysis and has important practical implications for the design of adaptive learning systems and the formulation of precise teaching intervention strategies.

Keywords: Blended learning; Learning behavior; Learning analysis

1 INTRODUCTION

Blended learning, which combines the flexibility of online learning with the interactivity of offline teaching, has become the mainstream model in higher education. Its core value lies in leveraging technology to enable precise insight into the learning process and personalized support. Learning behavior refers to the observable and explicit actions exhibited by learners during their learning activities. Behavioral science theory indicates that behavior is the external manifestation of an individual's thoughts, motivations, and needs. By statistically analyzing and interpreting the data of learning behaviors, it is possible to effectively infer the psychological state of learners and predict their learning behavior patterns [1]. Therefore, mining effective learning patterns based on behavioral data, optimizing the learning environment and teaching strategies is an important research direction in the current field of learning analysis.

Based on this, this study takes blended learning as the research scenario, integrates the two core dimensions of autonomous learning behavior and interactive learning behavior, and relies on the students' objective learning logs accumulated over the ten years from 2006 to 2015 on the national first-class blended course platform of Modern Educational Technology of Shaanxi Normal University. It combines the subjective self-assessment data of learners to systematically explore the behavioral characteristics of college students in blended learning and clarify the predictive path of learning behavior on learning outcomes. This study focuses on three core questions:

RQ1: Based on the ten-year longitudinal data, what are the overall characteristics of college students' blended learning behaviors? What are the differences in learning behaviors among different demographic groups?

RQ2: What is the relationship between objective learning behavior, subjective self-assessed learning behavior and blended learning effectiveness? Which behavioral variables are the key predictors of learning outcomes?

RQ3: Based on key behavioral variables, which groups of learners with different behavioral patterns can be identified? What are the differences in learning outcomes among different groups?

2 LITERATURE REVIEW

The classification and measurement of learning behaviors form the basis of learning analysis research. Early studies focused on constructing a universal online learning behavior classification framework. Peng Wenhui et al. (2006) proposed a multi-dimensional and multi-level online learning behavior model, dividing online learning behaviors into five dimensions: collecting information, organizing information, publishing information, exchanging information, and using information [2]. Yang Jinlai et al. (2008) based on the definitions of behavioral theory and behavioral science for the behavioral system, sorted out more than 20 common online learning behaviors [3]. Hannafin (2003) further, from the perspective of the functions of the learning process, classified online learning behaviors into four categories:

trajectory behavior, social behavior (including real-time communication and discussion-based communication), resource learning behavior, and assessment and reflection behavior[4]. The above classification framework provides theoretical support for understanding the structure of learning behaviors, but it mostly remains at the level of conceptual construction and does not clearly define specific measurable indicators.

In recent years, with the rapid development of learning analytics technology, researchers have further subdivided learning behaviors to the level of platform operations, enabling automated tracking and quantitative analysis of learning behaviors. Zarzour et al. (2020) focused on fine-grained behaviors such as page turning, returning, posting, liking, sharing, commenting, etc. [5]; Yoon et al. (2021) systematically reviewed 11 video learning behaviors including video playback, pausing, bookmarking, slide navigation, conditional judgment, evaluation, reply, view switching, filtering, information seeking, and annotation [6]. These studies significantly enhanced the collectability and computability of behavioral data, but the underlying assumption is that all recorded behaviors have equal analytical value. However, which behavioral indicators truly have stable and significant predictive power for learning outcomes remains a matter of consensus [7].

More importantly, the existing research has two common shortcomings in terms of methodology: Firstly, the data cycle is limited, mainly relying on cross-sectional or short-term tracking data, lacking longitudinal data across multiple years. The stability of learning behaviors across semesters and grades, as well as the long-term applicability of short-term research conclusions, have not been fully verified. Secondly, the data modality is single, overly relying on platform objective log data, while ignoring the driving role of internal psychological factors such as learning motivation and self-efficacy on behaviors. The separation phenomenon between objective behaviors and subjective self-assessment and its internal mechanism lack systematic exploration in blended learning scenarios.

In summary, the current learning behavior analysis framework has significant research gaps in three aspects: the validity of indicators, the longitudinal stability, and the integration of subjective and objective data. Therefore, this study is based on the blended learning scenario, with autonomous learning behavior and interactive learning behavior as the core constructs. It integrates the objective behavioral logs of the platform accumulated over ten years and the subjective self-evaluation data of learners to construct a learning behavior analysis framework that is predictive, robust, and psychologically interpretable, thereby addressing the shortcomings of existing research.

3 RESEARCH METHODS

3.1 Research Background and Data Sources

The data for this study were sourced from the self-developed "Modern Educational Technology" blended teaching course platform of Shaanxi Normal University (<http://edutech.snnu.edu.cn>). This course is a national-level first-class blended course. The platform was independently developed by the course team and has been continuously operated. The research subjects were full-time undergraduate and graduate students who participated in this course during the period from 2006 to 2015. During the data processing, multi-source data association and matching were achieved through student numbers. To ensure the completeness of data for the three core dimensions of autonomous learning behavior, interactive learning behavior, and self-assessment learning behavior, this study excluded samples with missing core indicators (if any core dimension was missing, the sample was deleted). As a result, 3,693 valid samples were obtained. The composition of the samples is as follows: 1,498 males (40.6%), 2,195 females (59.4%); 1,502 in the liberal arts (40.7%), 1,640 in science/engineering (44.4%), 551 in technical subjects (14.9%); 2,902 undergraduates (78.6%), 791 postgraduates (21.4%).

3.2 Variable Measurement

The main variables of this study and their descriptions are as follows:

3.2.1 Dependent variable

The effectiveness of blended learning is measured by two indicators: ① Academic performance, which is the comprehensive assessment score of the course (including homework, online learning, online exams, and final exams), divided into four grades: 4 = Excellent (≥ 85 points), 3 = Good (75 - 84 points), 2 = Average (60 - 74 points), 1 = Fail (< 60 points); ② Self-assessment of learning effectiveness, which uses a Likert 5-point scale to measure learners' overall evaluation of their own learning effectiveness, with binary classification: 2 = Better, 1 = Worse.

3.2.2 Independent variable

The independent variables are divided into three categories: ① Autonomous learning behavior (16 indicators): covering online study duration, study resource duration, effective study duration of courses, pass rate of key knowledge, pass rate of ordinary knowledge, normal withdrawal rate, average number of active days, and average duration of study intervals, etc.; ② Learning interaction behavior (16 indicators): including participation activity level, centrality, content contribution degree, interaction breadth, interaction depth, as well as interaction frequency and quality indicators of BLOG, BBS, MESSAGE systems; ③ Self-assessment of learning behavior, measured by a 5-point Likert scale to evaluate the overall evaluation of learners' own learning behavior (Beh_Avg).

3.2.3 Control variables

The control variables are demographic characteristics, including gender, major category (Liberal arts, science/engineering, and art disciplines), and academic stage (undergraduate, graduate student).

3.3 Data Analysis Process

This study adopts a two-stage analytical strategy:

Stage 1: Screening of key predictor variables. With academic performance as the dependent variable, all behavioral indicators and demographic control variables were included in stepwise regression analysis (significance level $\alpha=0.05$). The behavioral variables with significant predictive effects on learning outcomes were screened, and the data dimension was reduced to lay the foundation for clustering analysis.

Stage 2: Cluster analysis of learners. Based on the selected key behavioral variables, the K-means clustering algorithm was used to cluster students after Z-score standardization. The optimal number of clusters is determined by elbow method, and verified by contour coefficient. To enhance robustness, the clustering process was randomly initialized 50 times and the solution with the smallest total within-group sum of squares was selected.

3.4 Data Preprocessing

In order to eliminate the influence of dimensional differences, all continuous features were normalized by max-min and converted to the interval $[0,1]$. In order to weaken the systematic errors caused by the differences between annual assessment content and scoring standards, and ensure the comparability of ten-year data across years, the continuous data (scale scores, platform behavior indicators) were standardized by Z-score in different years. Handling missing values: A small number of missing values ($< 5\%$) were imputed by the annual mean. The samples missing in any of the three core dimensions were eliminated, and finally 3693 complete analysis samples were formed. IBM SPSS Statistics 24 and Python 3.10 were used for data analysis.

4 ANALYSIS OF RESEARCH RESULTS

4.1 Descriptive Statistics and Group Differences

4.1.1 Global behavioral characteristics

The statistical results of the standardized data of 3693 learners showed that: first, the individual differences of learners in the learning habit indicators such as time investment, learning efficiency, active days, and learning interval were extremely significant, and the Coefficient of Variation (CV) of most indicators was greater than 0.6. Secondly, the distribution of interaction activity in the learning community was extremely unbalanced, showing the characteristics of the coexistence of a small number of highly active contributors and a large number of low-participation learners. Third, the mean value of self-rated learning behavior (Beh_Avg) was 3.34 ($SD=0.43$, 5-point scale), with the overall positive and individual differences being very small, which was in sharp contrast with the significant differences of objective behavior indicators.

4.1.2 Demographic differences

Gender differences. Girls' academic performance was significantly better than boys' ($t=4.72$, $p<0.001$), but boys' self-evaluation learning effect was significantly higher ($\chi^2=8.34$, $p=0.004$). The contrast between objective performance and subjective evaluation was prominent in the gender dimension, which may be related to the influence of traditional gender concepts on self-assessment standards.

Section and subject differences. The academic performance and self-evaluation effect of graduate students were significantly better than those of undergraduate students ($p<0.01$). In the subject dimension, liberal arts students had the best performance, and technical students had the lowest performance, but technical students had the highest self-evaluation learning effect, which reflected the regulatory effect of subject evaluation standards on learners' self-cognition.

4.2 Identification of Key Behavioral Predictors

In order to identify the indicators with significant predictive power for blended learning effect, this study took academic performance and self-rated learning effect as the dependent variables. Firstly, the correlation analysis between 33 behavioral indicators and demographic control variables was carried out, and 20 behavioral indicators significantly correlated with learning effect were screened out. Then, considering that academic performance and self-rated learning effect were categorical variables, this study used Logistic regression analysis to investigate the predictive effect of covariates (gender, academic stage, discipline background) and the above 20 predictor variables on blended learning effect of college students. Stepwise Logistic regression results showed that 10 behavioral variables with significant predictive effects were retained in the final model ($F(10,3682)=127.45$, $p<0.001$, adjusted $R^2=0.312$). Among them, the normal dropout rate, extended registration time, average duration of course learning, average duration of online learning, active days, average duration of learning interval in autonomous learning behavior, and the proportion of total BBS topic reading and the proportion of BBS circles in interactive learning behavior were entered into the model. It is worth noting that "total online time" was not included in the final model, indicating that time investment alone has limited predictive power for learning outcomes.

4.3 Dimension Feature Mining of Learning Behavior

4.3.1 Learning behavior clustering: Six groups of learners

Based on the 10 key behavioral variables (excluding demographic variables), K-means clustering was performed on the standardized data. With 6,446 data points as input, using the elbow rule for evaluation, it was found that the DBI index reached its minimum value of 0.5466 when the number of clusters K was 6, thus determining the optimal number of clusters to be 6.

4.3.2 Description of behavioral characteristics

In order to reveal the detailed characteristics of different groups in terms of autonomous learning behavior, interactive learning behavior, and self-assessment learning behavior, this study conducted descriptive statistical analysis on 6 clusters across 3 dimensions and 10 behavioral characteristics. The analysis results are shown in Table 1.

Table 1 Analysis of Behavioral Characteristics of Different Clusters

Features	Cluster 1 (681)	Cluster 2 (1429)	Cluster 3 (1672)	Cluster 4 (1448)	Cluster 5 (918)	Cluster 6 (298)
Normal exit rate	0.555	0.361	0.612	0.258	0.470	0.651
Registration(CB1)	0.052	0.082	0.040	0.145	0.060	0.040
extension time(CB2)						
Average course study duration(CB3)	0.239	0.291	0.220	0.323	0.265	0.187
Average online duration(CB4)	0.269	0.314	0.254	0.347	0.293	0.229
Active days(CB5)	0.217	0.167	0.244	0.138	0.193	0.309
Average learning interval duration(CB6)	0.227	0.253	0.211	0.265	0.241	0.180
BBS topic reading total percentage(IB1)	0.561	0.463	0.596	0.393	0.518	0.639
BBS circle percentage(IB2)	0.005	0.003	0.011	0.003	0.004	0.031
Autonomous learning behavior(CB)	0.601	0.485	0.680	0.413	0.541	0.788
Interactive learning behavior(IB)	0.561	0.463	0.596	0.393	0.518	0.639
Self-assessment learning behavior(Beh)	0.005	0.003	0.011	0.003	0.004	0.031

From the descriptive statistics results, it can be seen that each cluster shows significant differences and common patterns in learning behaviors, mainly manifested as structural contradictions such as "duration $\neq$ quality" and "frequency $\neq$ depth". Cluster 6 has the highest comprehensive score in autonomous learning (0.788) and interactive learning (0.639), demonstrating a high completion rate, high activity level, and strong interactivity; Cluster 4 ranks first in "average course learning duration" (0.323) and "average online duration" (0.347), but has the lowest "active days" (0.138), and its interaction behavior score (0.393) and self-assessment behavior score (0.003) are both at the bottom, highlighting the contradictory characteristics of "high investment, low frequency, and low interaction".

Each cluster generally presents the problem of "unequal balance between autonomous learning and interactive learning": The scores for autonomous learning are generally higher than those for interactive learning, and the self-assessment behavior is overall low, reflecting the weak self-reflection and feedback mechanism of the learners. Among them, Cluster 2 and Cluster 5 have a medium learning frequency, but the interaction participation is insufficient; Cluster 3, although having a high activity level, still has a low self-assessment score.

4.3.3 The complex interrelationship among behavioral patterns, self-perception, and learning outcomes

In order to more clearly observe the differences in students' learning outcomes among different clusters, this study presents the data through line graphs, as shown in Figure 1. The various behavioral indicators present a complex and interwoven relationship. Cluster analysis clearly reveals the significant differences in students' learning behaviors. Based on the differences and characteristics of different types of learners in their learning behaviors, combined with the research in this paper, the six types of learners are named as: the Isolated Group (Cluster 1), the Passive Procrastinator Group (Cluster 2), the Shallow Participation Group (Cluster 3), the Stable and Active Group (Cluster 4), the Highly Engaged Group (Cluster 5), and the Deep Interaction Leadership Group (Cluster 6). The core characteristics of each group are as follows:

The behaviorally detached group (cluster 1): low autonomous participation, low interactive behavior, the longest learning interval, the lowest online time, weak willingness to participate in the learning platform, significant learning procrastination, and the lowest academic performance and self-evaluation effect.

Passive procrastination group (cluster 2): long learning interval, high online time, low interaction quality, mainly surface passive interaction such as browsing and liking, no deep interaction is formed, and the learning effect is not good.

Shallow participation group (cluster 3): the learning interval is moderate, the interaction frequency is slightly increased, and active behaviors such as posting for the first time appear, but shallow interaction is still the core, and the learning performance and self-evaluation effect are slightly improved.

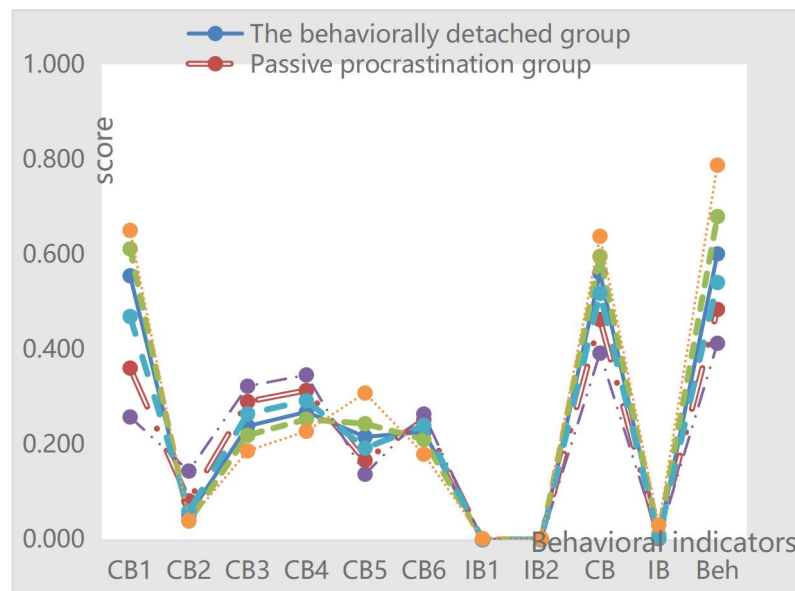


Figure 1 Line Graph Showing the Mean Values of Each Characteristic Behavior of Different Clusters

Stable and active group (cluster 4): the online time and active days are above the middle, the interactive behavior is stable and the quality is gradually improved, the regular learning rhythm is formed, and the learning performance and self-evaluation effect are above the middle level.

Highly engaged group (cluster 5): the shortest learning interval, the highest online time, frequent interaction behaviors with excellent quality, strong learning motivation, and high scores and self-evaluation effects.

Deep interaction leading group (cluster 6): the interaction quality was the best (the proportion of BBS topic replies and the proportion of circles were the first), the independent learning behavior performance was outstanding, not only the learning effect was excellent, but also played a leading role in the learning community.

The comparison results of the learning performance of each group showed that the learning behavior and learning effect were not a simple linear relationship of "the more investment, the better effect". The stable active group (cluster 4) had higher online time investment, but the academic performance (78.9 points) was significantly lower than that of the deep interaction leading group (cluster 6, 87.6 points). The deep interaction leading group achieved the best results with medium time investment, which confirmed that interaction quality and learning regularity were the core factors driving the nonlinear growth of learning effectiveness. The learning performance of the shallow participation group (cluster 3) (72.1 points) was better than that of some high time investment groups, which further verified that the "quality" of learning behavior had a stronger predictive power than the "quantity" of learning behavior.

5 DISCUSSION

Based on ten years of longitudinal multimodal data, this study reveals three core findings: the systematic separation of objective performance and subjective self-evaluation, the nonlinear association between learning behavior and learning effectiveness, and the six types of learner behavior classification in blended learning scenarios. The detailed discussion is as follows:

5.1 The Phenomenon of Subjective and Objective Separation: The Double Regulation of Gender and Discipline Culture

This study found that girls achieved significantly higher objective scores than boys, but their self-assessed learning outcomes were lower; students in technical subjects had the lowest objective scores, yet their self-assessed learning outcomes were the highest. This phenomenon is not a measurement error; rather, it is an external manifestation of the differences in the reference frameworks and success criteria used for self-evaluation among different groups.

Previous studies have provided a theoretical explanation for this finding. Pajares (2002) pointed out that females tend to adopt stricter standards or make upward comparisons in their academic self-evaluation, while males are more likely to adopt looser standards or make downward comparisons [8]. This gender difference is manifested in this study as follows: although female students have higher objective grades, their strict self-evaluation standards lead to a lower self-assessment. At the same time, the discipline culture has shaped the cognitive paradigm of "ability". The art disciplines (such as art and sports) focus on practical performance and creative expression, and their success criteria are different from standardized tests. Therefore, although art disciplines students have lower test scores, they may still give higher self-assessments based on the value scale of their own field. This finding is consistent with the existing studies on the moderating effect of disciplinary identity.

From the perspective of learning analysis methodology, this finding holds significant implications. Relying solely on objective behavior logs may underestimate the learning satisfaction of women and students of liberal arts, while relying

solely on self-assessment scales may overestimate the knowledge acquisition of students of technical subjects. The integration of multimodal data is a necessary approach to ensure the validity of research conclusions.

5.2 Nonlinear Behavior-Effectiveness Relationships: The Leverage of Interaction Quality and Self-Regulation

The clustering results clearly challenge the linear perception that "the more time invested, the better the learning outcome". The deeply interactive group (Cluster 6) achieved the best learning results with non-maximum online duration as its core feature, which lies in high-quality interaction and strong self-regulation ability. However, the stable and active group (Cluster 4) had sufficient time investment but failed to achieve the optimal learning outcome due to insufficient interaction depth. The underlying mechanism is that high-quality deep interaction enables learners to externalize knowledge, defend viewpoints, and reconstruct through reflection, thereby promoting deep understanding of concepts; while good self-regulation, such as a regular learning rhythm, avoids inefficient consumption of cognitive resources. When the learning duration exceeds a certain threshold, the marginal benefit decreases or even turns negative. Without strategic adjustment, simply accumulating time can easily lead to the "false activity" predicament. This explanation is in line with the consensus in the field of learning analytics that "quality of input is superior to quantity of input" [9], and also consistent with the "redundancy effect" in the cognitive load theory.

5.3 Behavior Typologies for Six Groups: From Descriptions to Guidelines for Differentiated Interventions

The six groups of learners identified in this study have distinct behavioral characteristics, providing an empirical basis for the transition from "uniform teaching" to "precise intervention" in higher education. There are fundamental differences in the learning difficulties and needs of different groups: the alienated group and the passive procrastinator group need to lower the participation threshold and rely on external impetus to integrate into the learning process; the shallow participation group has crossed the threshold, but lacks cognitive strategies to transform surface interactions into deep processing; the stable and active group and the highly engaged group have already demonstrated strong autonomy, but may encounter risks of saturation or burnout; the deep interaction leading group has formed an efficient learning model.

This classification framework corresponds to the three elements of autonomy, competence and belongingness in the Self-Determination Theory [10]: The alienation group lacks basic belongingness and competence; the shallow participation group needs to enhance competence to move beyond the surface level; the highly engaged group needs to maintain autonomy and challenge to prevent burnout; the deep interaction-leading group has achieved a high level of satisfaction of these three elements and can then shift to contribution and leadership. This classification framework indicates that effective teaching intervention is not a one-size-fits-all approach of increasing or reducing burdens, but rather a precise matching based on the behavioral characteristics of learners. In the future, intelligent teaching systems can rely on real-time behavioral data to automatically identify group types and push differentiated support strategies.

6 CONCLUSION AND IMPLICATIONS

This study is based on a ten-year longitudinal multimodal dataset and employs a two-stage analysis method consisting of stepwise regression screening and K-means clustering to reach the following conclusions.

First, there is a systematic separation between objective academic performance and subjective self-assessment results. Girls perform better academically than boys but rate themselves lower. Students in technical subjects have the lowest scores but rate themselves the highest. This finding confirms that relying solely on objective logs or self-assessment scales can lead to errors. The integration of multimodal data is a necessary methodological foundation for learning analysis.

Secondly, the regularity of learning, the habit of systematic withdrawal, and the depth of community participation are the core indicators for predicting learning outcomes. Among the 10 key variables selected through stepwise regression, "average duration of learning intervals (reverse)", "normal withdrawal rate", and "percentage of BBS circles" have the largest effect sizes, revealing a predictive path beyond "time accumulation".

Thirdly, the relationship between behavior and outcome exhibits a non-linear characteristic. The quality of interaction and the ability of self-regulation are the core levers that surpass the accumulation over time. Among the six types of learners, the group with deep interaction led achieved the best results with moderate time investment, while the stable and active group, which invested the most time, did not reach the optimal level.

The above conclusion has clear guiding significance for the practice of blended learning: The course design should break away from the "time tracking" paradigm and build a learning environment that supports deep interaction and self-regulation; The learning analysis system should integrate subjective and objective data to achieve multimodal learning warnings and personalized feedback; Teaching intervention should be implemented precisely based on behavioral classification, promoting low-threshold participation for the isolated group, building interaction scaffolds for the shallow participation group, providing challenging tasks and autonomous space for the highly engaged group, and exerting the value of peer tutoring for the deeply interactive leading group. Future research should conduct replication verification across courses and institutions, introduce dynamic clustering and sequence mining methods to reveal the time evolution patterns of learning behaviors, and develop an adaptive intervention system based on the six types of learner classification for experimental verification.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This study was supported by the general topic of "Research on the Cognitive and Emotional Interaction Mechanism and Intervention in College Students' Digital Learning Supported by Digital Intelligence and Evidence-Based Methods" (SGH25Y3650) of the "14th Five-Year Plan" for Education Science in Shaanxi Province for the year 2025;

This study was supported by the Philosophy and Social Sciences Research Project of Weinan City in 2026 under the Shaanxi Provincial Philosophy and Social Sciences Research Special Project (2026HZ0905);

This study was supported by the teacher development research project of Xi'an Music College for the year 2025, "Research on the Development of General Knowledge Courses for Music Art Students Using Artificial Intelligence" (xyjf2509).

REFERENCES

- [1] Kong Lili, Ma Zhiqiang, Yi Yuhuo, et al. Research on the Model of Influencing Factors of Online Learning Behavior Based on Behavioral Science Theory. *Open Learning Research*, 2017, 22(05): 46-53. DOI: 10.19605/j.cnki.kfxxyj.2017.05.007.
- [2] Peng Wenhui, Yang Zongkai, Huang Kebin. Analysis of Online Learning Behaviors and Its Model Research. *China Educational Technology*, 2006(10): 31-35.
- [3] Yang Jinlai, Zhang Yixiang, Ding Rongtao. Research on Learning Behavior Monitoring Based on Online Learning Platform. *Computer Education*, 2008(11): 65-68. DOI: 10.16512/j.cnki.jsjy.2008.11.019.
- [4] Hannafin M, Hill J R, Oliver K, et al. Cognitive and learning factors in web-based distance learning environments. *Handbook of distance education*, 2003: 245-260.
- [5] Zarzour H, Bendjaballah S, Harirche H. Exploring the behavioral patterns of students learning with a Facebook-based e-book approach. *Computers & Education*, 2020, 156: 103957.
- [6] Yoon M, Lee J, Jo I H. Video learning analytics: Investigating behavioral patterns and learner clusters in video-based online learning. *The Internet and Higher Education*, 2021, 50: 100806.
- [7] Amnueypornsakul B, Bhat S, Chinprutthiwong P. Predicting attrition along the way: the UIUC model. *Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs*. Doha, Qatar: Association for Computational Linguistics, 2014: 55-59.
- [8] Pajares F. Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice*, 2022, 41(2): 116-125.
- [9] Henrie C R, Halverson L R, Graham C R. Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 2015, 90: 36-53.
- [10] Ryan R M, Deci E L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 2000, 55(1): 68-78.