Studies in Social Science Research
ISSN 2690-0793 (Print) ISSN 2690-0785 (Online)
Vol. 7, No. 1, 2026
www.scholink.org/ojs/index.php/sssr

Original Paper

Analysis of Learning Behaviors in SPOC-Based Blended Teaching of Linear Algebra

Kaiyi Fan[1*]

[1] Xihua University, Chengdu, Sichuan Province, 610039, China

[*] Kaiyi Fan (1999–), male, Han ethnicity, from Zizhong County, Neijiang City, Sichuan Province. Undergraduate student at the School of Science, Xihua University, Student ID: 3120192109119, is the first author and corresponding author of this paper.

Abstract

SPOC (Small Private Online Course), as a teaching model that deeply integrates online learning with classroom instruction, has been widely applied in the field of higher education in recent years. This study takes the Linear Algebra course at a certain university as an example and systematically analyzes the characteristics of students' learning behaviors based on log data from the SPOC blended teaching platform. The findings reveal: First, students' video-watching behavior exhibits a distinct "beginning-and-end concentration" pattern, with a negative correlation between video length and completion rate. Second, online test performance shows a significant positive correlation with final exam scores, but the submission times of chapter tests are considerably dispersed. Third, forum interaction displays a phenomenon of "marginal participation," with active posters accounting for less than 15% of students. Fourth, cluster analysis of learning behaviors categorizes students into three types: "balanced," "cramming," and "passive," with significant differences in academic performance among the different types. Fifth, overall satisfaction with blended teaching is high, but individual differences in student learning engagement are substantial. Based on these findings, this paper proposes improvement suggestions such as optimizing video resource design, strengthening process-oriented evaluation incentives, promoting deep forum interaction, and implementing stratified teaching interventions. The research conclusions provide data support and practical references for the effective implementation of SPOC blended teaching in mathematics courses.

Keywords

SPOC, Blended Teaching, Linear Algebra, Learning Behavior Analysis, Learning Engagement

1. Introduction

1.1 Research Background

1.1.1 The Deep Integration of Information Technology and Higher Education

In recent years, information technologies represented by big data, cloud computing, and artificial intelligence are profoundly changing the form and methods of education. The state attaches great importance to the construction of educational informationization. The *Education Informatization 2.0 Action Plan* clearly proposes to promote the deep integration of information technology and education and teaching, and to build a new "Internet + Education" ecosystem. In this context, university classroom teaching is undergoing a profound transformation from traditional lectures to online-offline blending, from teacher-centered to student-centered, and from standardized to personalized. How to effectively use information technology to improve teaching quality has become a core issue in university teaching reform.

1.1.2 The Development and Characteristics of the SPOC Teaching Model

SPOC (Small Private Online Course) is a concept proposed by Professor Armando Fox at the University of California, Berkeley, in 2013. It can be understood as the "classroom version of MOOC." Unlike the large scale and openness of MOOCs, SPOC has typical characteristics such as small scale, restrictiveness, and blending. SPOC is generally targeted at registered students of a university. It combines high-quality online course resources with offline classroom teaching. Before class, students engage in self-directed learning through videos and quizzes. Class time is then used for in-depth interactive activities such as discussions, Q&A sessions, and exercises. This model fully utilizes the richness and convenience of online resources while retaining the emotional connection and immediate feedback advantages of face-to-face teaching. It is considered a more sustainable teaching model than MOOCs.

1.1.3 The Teaching Difficulties of the *Linear Algebra* Course

Linear Algebra is an important foundational course for majors in science, engineering, economics, and management. It plays an irreplaceable role in cultivating students' abstract thinking, spatial imagination, and logical reasoning abilities. However, traditional *Linear Algebra* teaching faces many difficulties: the course content is abstract (concepts such as vector spaces, linear transformations, and eigenvalues are difficult to understand), leading to a lack of intuitive feeling among students; class time is limited, making it difficult for teachers to balance concept explanation and example exercises; in large-class teaching environments, individual differences among students are significant (varying greatly in mathematical foundation, learning habits, and acceptance speed), and a unified pace cannot cater to the needs of all students. These difficulties have long made *Linear Algebra* one of the courses with a high failure rate. SPOC blended teaching provides a new idea for solving this dilemma.

1.2 Research Purpose and Significance

1.2.1 Research Purpose

This study aims to systematically analyze the application effect of SPOC-based blended teaching in the *Linear Algebra* course. Specific goals include: First, to describe the typical learning behavior characteristics of students in the SPOC environment (video watching, online testing, forum interaction, etc.); Second, to reveal the relationship between learning behaviors and academic performance; Third, to identify differences in behavior patterns among different types of learners; Fourth, to propose targeted suggestions for optimizing SPOC blended teaching.

1.2.2 Theoretical Significance

From a theoretical perspective, this study helps deepen the understanding of learning behavior patterns in blended learning environments. Existing research on learning behavior mostly focuses on MOOC platforms, with relatively few studies on SPOC environments, especially for mathematics courses. As an abstract foundational mathematics course, the learning behavior characteristics of *Linear Algebra* may differ from other courses. This study can enrich the application of blended learning theory in specific subject areas.

1.2.3 Practical Significance

From a practical perspective, this study can provide data support for frontline teachers to optimize SPOC instructional design. By analyzing which learning behaviors are positively correlated with academic performance and which student groups need focused attention, teachers can adjust their teaching strategies more effectively and improve teaching outcomes. At the same time, the research conclusions can also serve as a reference for the blended teaching reform of other mathematics courses (such as Calculus, Probability Theory).

1.3 Domestic and International Research Status

1.3.1 International Research on the SPOC Teaching Model

Internationally, SPOC research mainly focuses on three areas: first, comparative studies on the effectiveness of SPOC versus traditional teaching, with most studies confirming that SPOC is superior to pure classroom teaching in terms of knowledge acquisition and learning motivation; second, SPOC learning analytics research, using platform log data to reveal learning behavior patterns; third, research on the design and optimization of SPOC courses, exploring the impact of elements such as video length, interaction frequency, and assessment methods on learning outcomes. Overall, international research is characterized by diverse methods and data-driven approaches.

1.3.2 Application of SPOC in Mathematics Teaching in China

Domestic research on SPOC started around 2014 and has shown rapid growth in recent years. In the field of mathematics teaching, existing research mainly focuses on the application effect of the SPOC model in courses such as *Calculus*, *Linear Algebra*, and *Probability Theory*. Most studies show that SPOC blended teaching can improve student performance, enhance learning attitudes, and strengthen autonomous learning abilities. However, existing research mostly focuses on effect comparison (SPOC class vs. control class), with relatively weak in-depth analysis of the learning process. There is a lack of detailed analysis based on behavioral log data regarding process-oriented questions such as "how students learn in the SPOC environment" and "which behaviors predict learning success."

1.3.3 Limitations of Existing Research

Overall, existing research has the following shortcomings: First, the granularity of behavior analysis is relatively coarse, mostly general descriptions (e.g., average video watching duration), lacking in-depth mining of behavior patterns and sequences; Second, there are few SPOC studies specifically targeting the *Linear Algebra* course. The abstractness and logical nature of Linear Algebra may produce unique learning behavior characteristics; Third, there is a lack of systematic analysis of student type classification and behavioral differences. This study aims to explore these areas.

1.4 Research Methods and Content Framework

1.4.1 Research Methods

This study employs a combination of quantitative and qualitative methods. Quantitatively, based on learning behavior log data from the SPOC platform (including video watching records, test scores, forum posts, login frequency, etc.), descriptive statistics, correlation analysis, cluster analysis and other methods are used for data analysis. Qualitatively, semi-structured interviews are conducted with some students to understand their learning experiences, difficulties, and needs, explaining the deep-seated reasons behind the quantitative data.

1.4.2 Research Subjects

The subjects are 180 students from a certain university who took the *Linear Algebra* course in the first semester of the 2023-2024 academic year, covering majors in engineering, science, economics, and management. The course adopted a SPOC-based blended teaching model. The online learning platform was the Chaoxing Learning Platform. Course resources included instructional videos (48 videos across 12 chapters, totalling approximately 480 minutes), chapter tests, and a discussion forum.

1.4.3 Content Framework

This paper is divided into five parts. Part 1 is the introduction, presenting the research background, purpose, significance, current state, and methods. Part 2 describes the implementation framework of SPOC blended teaching in *Linear Algebra*. Part 3 conducts a multi-dimensional analysis of learning behaviors. Part 4 explores the relationship between learning behaviors and learning outcomes. Part 5 proposes teaching optimization suggestions and summarizes the research conclusions.

2. Implementation Framework of SPOC Blended Teaching in *Linear Algebra*

2.1 Overall Course Design

2.1.1 Blended Online-Offline Teaching Structure

The SPOC design of this course follows a three-stage structure: "pre-class online learning—in-class deep interaction—post-class expansion and consolidation." In the pre-class stage, students watch instructional videos and complete pre-class quizzes on the platform to gain a preliminary understanding of basic concepts and methods. In the in-class stage, the teacher no longer systematically lectures on all content but instead focuses on problems revealed in students' previews, organizing group discussions and in-class exercises. In the post-class stage, students complete chapter tests and expansion tasks, and engage in Q&A and discussions via the forum.

The core logic of this design is the reasonable allocation of cognitive load: lower-order learning tasks (memorization, understanding) are moved forward to pre-class online completion, while higher-order learning tasks (application, analysis, evaluation) are concentrated in high-quality face-to-face interactions in the classroom. Online and offline components are not simply additive but are organic wholes with different focuses, complementing each other.

2.1.2 Reorganization and Adjustment of Teaching Content

The SPOC model requires the re-combing and reorganization of teaching content. The course team divided the *Linear Algebra* content into several clean learning units, each unit corresponding to a core knowledge point (e.g., definition and properties of determinants, elementary row operations of matrices, linear dependence of vector groups, etc.). Each unit is configured with 1-2 instructional videos, with video duration controlled between 8-15 minutes, striving to maintain content integrity while reducing cognitive fatigue.

The presentation format of the video content was also designed specifically: abstract concepts are accompanied by geometric intuition as much as possible (e.g., displaying the geometric meaning of linear combinations of vectors on a plane); complex calculation processes use step-by-step demonstrations and color coding; key conclusions and formulas are emphasized through pop-up prompts. This design aims to enhance the comprehensibility and appeal of online self-learning.

2.1.3 Composition of the Evaluation System

The total course grade consists of a regular grade (50%) and a final exam grade (50%). The regular grade is further broken down into: video viewing completion (10%), chapter tests (20%), forum participation (5%), classroom performance (10%), and group assignments (5%). This evaluation system aims to guide students to maintain continuous engagement throughout the course, avoiding the phenomenon of "not studying regularly, cramming before the exam." At the same time, clear time windows are set for video viewing and chapter tests, with no make-up for missed deadlines, to reinforce time management awareness.

2.2 Construction of Online Learning Resources

2.2.1 Production of Micro-Video Resources

Instructional videos are the core resource for SPOC online learning. This course produced 48 instructional videos, covering six major modules: determinants, matrices, vectors, linear equation systems, eigenvalues and eigenvectors, and quadratic forms. The production of each video followed these principles: a "Lesson Objectives" section at the beginning (1 minute), clearly informing students what they should master after learning; the middle section (6-12 minutes) divided into three parts: "concept introduction—theorem derivation—example demonstration"; a "Key Points Summary" at the end (1 minute), reviewing the core content. The videos were produced using screen recording and a writing tablet, with the teacher annotating and calculating directly on the slides, simulating the feel of a blackboard.

It is worth noting that after analyzing data from the first batch of videos, the course team found that the average completion rate (proportion of students watching fully) for videos exceeding 12 minutes was significantly lower than for videos within 12 minutes (58% vs. 81%). Based on this finding, subsequent long videos were split, ensuring most videos were controlled to around 10 minutes.

2.2.2 Online Test Questions and Resource Library

Chapter tests are important tools for assessing online learning effectiveness. Each learning unit is equipped with 5-10 online test questions, including multiple-choice, fill-in-the-blank, and simple calculation questions. The design of the test questions follows the principle of "foundation-first, reasonable gradient": 70% of the questions test basic concepts and routine calculations, 20% test variant applications, and 10% are set with certain difficulty to challenge advanced students.

The immediate feedback function of the testing system is fully utilized: after students submit their answers, they immediately see whether they were correct or incorrect, along with reference answers. Some questions are accompanied by explanatory videos. For questions answered incorrectly twice in a row, the system automatically pushes similar practice questions for reinforcement training. This immediate feedback and adaptive pushes help students correct errors and consolidate understanding in a timely manner during self-study.

2.2.3 Construction of Forum and Interactive Community

The forum is an important space for student-student and teacher-student interaction in the SPOC environment. This course's forum is divided into four sections: first, a "Q&A" section where students post problems encountered in learning, with teachers and teaching assistants responding within 24 hours; second, an "Example Discussion" section encouraging students to exchange different solutions to typical problems; third, an "Extended Reading" section posting application cases of Linear Algebra in fields like machine learning, cryptography, and network analysis; fourth, a "Learning Experience" section where students share their learning methods and insights.

To incentivize forum participation, the teaching team regularly selects "Best Question" and "Best Answer," awarding bonus points to the regular grade and praising them in class. At the same time, typical forum questions are compiled weekly and discussed centrally in class, forming a closed loop of "online questions—offline answers."

2.3 Design of Classroom Teaching Activities

2.3.1 Analysis of Pre-class Preview Situations

Before each class, the teacher logs into the platform to view students' preview data: video watching progress, chapter test average scores, hot topics in the discussion forum. Based on these data, the teacher can accurately identify common difficulties and typical errors of the students, adjusting the focus of classroom teaching accordingly.

For example, in the unit on "Linear Dependence of Vector Groups," the data showed that over 60% of students made errors on questions about "determining linear dependence/independence," and discussions were exceptionally active in the forum around "whether there exist non-zero coefficients making the linear combination zero." This indicated that students were encountering difficulties understanding the abstract definition of dependence. Consequently, the teacher decided to reduce time spent on theorem proofs in class and increase geometric intuition explanations and categorized discussion exercises.

2.3.2 Implementation of the Flipped Classroom

Class time primarily uses the flipped classroom format. The specific process is as follows: First, the teacher conducts a "pre-test" and Q&A session for 5-10 minutes, quickly checking the effectiveness of the preview. Then, a "case discussion" is organized for 15-20 minutes, where a comprehensive problem is given (e.g., determining the rank of a vector group, solving linear equation systems with parameters). Students discuss in small groups and report their ideas. Finally, the teacher provides a "concise lecture and summary," supplementing extended content and emphasizing error-prone points and key ideas.

This classroom organization format demands more from the teacher. The teacher needs to transform from a "sage on the stage" to a "guide on the side" – no longer unidirectionally pouring knowledge but organizing discussions, guiding thinking, intervening appropriately, and summarizing and elevating. It also requires students to invest sufficient time in previewing before class; otherwise, effective classroom discussion will be difficult.

2.3.3 In-class Exercises and Immediate Feedback

The classroom is equipped with an immediate response system (clicker or mobile app). The teacher can pose multiple-choice or true/false questions at key points, with all students answering in real-time, and the results displayed instantly on the big screen. This immediate feedback has multiple functions: helping the teacher quickly judge students' level of understanding and decide whether further explanation is needed; helping students accurately position their own comprehension; increasing classroom participation through whole-class responses, reducing off-task behavior. Data show that in classrooms using immediate response systems, the frequency of students actively raising their hands to ask questions is about 40% higher than in traditional classrooms.

2.4 Platform Data Collection Mechanism

2.4.1 Learning Behavior Log Data

The Chaoxing Learning Platform automatically records every learning action of students, generating detailed behavior logs. The recorded fields include: user ID, timestamp, behavior type (video playback, test submission, forum post, resource download, etc.), behavior object (which specific video, which question), behavior duration (video viewing duration), behavior outcome (test score), etc.

This course collected approximately 120,000 behavior logs, covering the complete learning cycle of all 180 students. These granular behavior data are the core analytical objects of this study.

2.4.2 Extraction of Key Behavior Indicators

Based on the research objectives, we extracted the following key behavior indicators from the raw logs:

- **Video Viewing**: total videos viewed, total duration, duration per video, video completion rate, number of replays.
- **Online Testing**: submission time for each chapter test per student, score, number of attempts, average correctness rate.
- **Forum Participation**: number of posts, number of replies, number of times a post was replied

to, number of featured posts, login frequency.

- **Learning Regularity**: average daily logins, time distribution of learning (peak hours), regularity of learning intervals.

2.4.3 Data Privacy and Ethical Protection

During data collection and analysis, this study strictly adhered to academic ethical norms. All student data were anonymized upon export, with personally identifiable information such as student names and IDs replaced by random numbers. The results of data analysis are presented only at the group level and do not disclose identifiable information of individual students. Students were informed at the beginning of the semester that their learning behavior data would be used for teaching research, with data usage limited to course improvement and academic publication. The research protocol was approved by the university's Academic Ethics Committee.

3. Multi-dimensional Analysis of Learning Behavior Characteristics

3.1 Analysis of Video-Watching Behavior

3.1.1 Overall Completion Rate

Statistical analysis of viewing data for 48 instructional videos showed that the average video completion rate (proportion of students watching fully) was 72.3%. The completion rate exhibited a distinct "U-shaped" distribution: the completion rate for the first video (Definition of Determinant) was highest at 89.4%; completion rates for videos in the middle chapters decreased slightly to about 65%-70%; the completion rate for the last video (Standardization of Quadratic Forms) rebounded to 78.2%. This phenomenon aligns with existing research, reflecting students' high enthusiasm at the beginning of the course and increased focus during final review.

Notably, completion rates varied across different modules. The matrix operations module had the highest video completion rate (average 81.3%), as this content is heavily computational, has clear application value, and students perceive it as moderately difficult. The eigenvalues and eigenvectors module followed (76.8%). The vector group linear dependence module had the lowest video completion rate (62.5%), as this content is highly abstract, making students prone to fear-induced difficulty and giving up midway.

3.1.2 Relationship between Video Length and Completion Rate

A significant negative correlation existed between video length and completion rate. Videos were divided into three groups by length: short videos (≤8 minutes, total 15), medium videos (8-12 minutes, total 22), long videos (≥12 minutes, total 11). The average completion rates were 85.6%, 73.2%, and 54.7%, respectively. This finding has practical instructional implications: while ensuring the integrity of teaching points, long videos should be split wherever possible, controlling each video within 8 minutes, or 9-10 minutes as an acceptable compromise.

Further analysis revealed that the attrition rate in the first 3 minutes of a video was only 5%, but it gradually increased from the 4th minute, with a noticeable attrition peak appearing around the 7th minute (cumulative attrition 18%). This suggests that teachers need to set "reminder points" or "suspense questions" in the middle to latter part of videos to maintain student attention.

3.1.3 Repeated Viewing and Learning Engagement

About 35% of students engaged in repeated viewing of videos, of which about 12% showed high-frequency repeated viewing (watching the same video 3 times or more). Repeated viewing was concentrated on knowledge points considered difficult, such as linear dependence and eigenvector calculation. Correlation analysis showed a moderate positive correlation between the number of replays and chapter test scores ($r=0.42$, $p<0.01$), indicating that actively rewatching videos is an effective remedial learning strategy.

From an individual difference perspective, students with weaker learning foundations had a higher average number of replays (2.3 times), but their test scores were still lower than those of students with better foundations (even though the latter had fewer replays). This indicates that repeated viewing is only one dimension of learning engagement; differences in learning strategies (whether they take notes or try to derive independently) may explain some performance differences.

3.2 Analysis of Online Testing Behavior

3.2.1 Distribution of Test Completion Times

A one-week completion window was set for chapter tests. The data showed that about 40% of students completed the test within the first two days of the window, 25% in the middle two days, 20% in the last two days, and 15% concentrated their submissions in the last 6 hours before the window closed. This "beginning-and-end concentration" submission pattern reflects students' time management habits: some students tend to "finish early for peace of mind," while others are accustomed to "meeting deadlines at the last minute."

Interestingly, there was a weak but significant negative correlation between submission time and test score ($r=-0.18$, $p<0.05$): the closer the submission to the deadline, the slightly lower the average score. This might be because students rushing to meet the deadline are less prepared, or these students have lower learning engagement overall. Teachers can give appropriate praise to students who complete tests early, cultivating a good time management culture.

3.2.2 Test Scores and Knowledge Mastery

Overall, the average correctness rate for chapter tests was 78.6%, with a standard deviation of 16.3%. Differences existed between chapters: the correctness rate for matrix multiplication and inverse matrices was highest (86.2%), while the correctness rate for vector group linear dependence was lowest (67.5%). This result is consistent with the chapter-wise distribution of video completion rates, suggesting that students "learn well the chapters they are willing to learn, achieving high scores," forming a virtuous cycle.

Correlation analysis between the average test score and final exam score showed a moderate to strong positive correlation ($r=0.58$, $p<0.001$), indicating that online tests reflect students' true mastery level relatively well. However, it is worth noting that about 8% of students had "high test scores but low final scores," and another 6% had "low test scores but high final scores." The former may reflect issues of short-term memory and test anxiety, while the latter suggests that some students made up for deficiencies in chapter tests through intensive final exam preparation.

3.2.3 Multiple Attempts and Learning Improvement

The testing system allowed students 3 attempts, taking the highest score as the result for that test. The data showed that about 62% of students used 2 or 3 attempts. The average correctness rate on the first attempt was 68.3%, which increased to 78.9% on the last attempt, an average improvement of 10.6 percentage points. This margin of improvement indicates that allowing retries is not only an assessment mechanism but also an effective learning mechanism – after seeing error feedback, students actively relearn relevant content, correct misunderstandings, and thus improve their accuracy.

However, further analysis revealed significant differences among students in the extent of benefit from "retry opportunities." Higher-performing students had a smaller improvement margin (average improvement of 5.2 percentage points) because their initial correctness rate was already high; while average-performing students had the largest improvement margin (average improvement of 14.3 percentage points), for whom retries are an important remedial opportunity.

3.3 Analysis of Forum Interaction Behavior

3.3.1 Overall Level of Forum Participation

Overall forum activity was relatively low. Throughout the semester, total posts were 312, total replies were 586, averaging 1.7 posts and 3.3 replies per student. Considering that every student was encouraged to participate in forum discussions at least once per week, actual participation levels were far below expectations.

Further analysis of the participation structure revealed a typical "Pareto distribution" in forum interaction: about 12% of students contributed over 60% of posts, about 78% of students had posts between 0-2, and about 10% of students never posted or replied. This means forum discussions were primarily dominated by a few active students, with most students in a state of "lurking and marginal participation." Marginal participation may stem from various reasons: students' introverted personalities, fear of appearing ignorant by asking questions, questions already raised and answered by others, or a belief that forum interaction offers limited help in improving grades.

3.3.2 Quality Analysis of Questions and Answers

Qualitative analysis of post content revealed that questions could be divided into three types: concept understanding (45%, e.g., "What is the geometric meaning of linear transformation?"), problem-solving methods (38%, e.g., "Is there a simpler algorithm for this problem?"), expansion and extension (17%, e.g., "How are eigenvalues used in image compression?"). The quality of answers varied considerably: some posts received detailed answers and in-depth discussion from teachers or peers; more posts received only brief responses (e.g., "See page XX of the textbook," "Use Cramer's rule"), lacking full explanation and elaboration.

Notably, high-quality questions tended to receive high-quality answers. This indicates that cultivating students' questioning ability is equally important – a good question can stimulate in-depth discussion. Teachers can demonstrate in class how to ask good academic questions or incorporate "question quality" into the regular grade.

3.3.3 Interaction Patterns and Learning Outcomes

The average final exam score of active forum students (with posts+replies >10) was 84.6, significantly higher than that of inactive students (posts+replies <3) at 71.3 ($t=4.82$, $p<0.001$). However, whether this correlation is causal or due to selection bias (i.e., students who are better at learning are more willing to post on the forum) needs careful judgment.

Interview data provided some clues. Several high-achieving students indicated that they consolidated their understanding by answering others' questions, as "teaching others is the best way to learn." In contrast, lower-achieving students indicated they tended to "browse" the forum rather than "participate" because they "didn't know what to ask" or "feared being ridiculed for asking stupid questions." These findings suggest that forum interaction does indeed promote student learning, but a safe and inclusive interaction atmosphere is needed to lower the participation threshold for disadvantaged students.

3.4 Cluster Analysis of Learning Behaviors

3.4.1 Classification of Student Types

Based on behavior indicators such as video viewing duration, online test scores, forum posts, and login frequency, the K-means clustering algorithm was used to divide students into three types:

Type 1: Balanced (47% of students). These students performed consistently and positively on various behavior indicators: high video viewing completion rate (average 91%), good chapter test scores (average 82 points), moderate forum participation (average 2.5 posts), and regular logins (average 6.2 times per week). They can progress steadily according to the course rhythm, with stable learning engagement.

Type 2: Cramming (32% of students). The behavioral characteristic of these students is "loose in daily routine, concentrated before exams": video viewing is intermittent, with a large number of videos watched intensively just before deadlines; test submission times are generally late; forum participation is low; login frequency surges just before the final exam. They represent a typical "exam-oriented" learning model.

Type 3: Passive (21% of students). These students score low on all indicators: low video viewing completion rate (average below 45%), repeatedly missing test submission windows, almost zero forum participation, long login intervals, and even no learning records for some chapters. They are the main group experiencing learning difficulties and the key target for teaching intervention.

3.4.2 Differences in Academic Performance among Different Student Types

Significant differences in final exam scores existed among the three student types. Balanced averaged 83.7 points, Cramming averaged 74.2 points, Passive averaged 58.6 points ($F=31.6$, $p<0.001$). This difference reveals the long-term effects of different learning models: "sustained engagement" is superior to "temporary cramming," which is significantly superior to "low engagement." Although cramming students can achieve passing or even moderate scores through intensive pre-exam review, there is still a clear gap compared to balanced students, and this learning model is less effective for long-term knowledge retention.

It is worth noting that there is internal differentiation within the cramming type: about 40% of cramming students achieved final scores above 80. They have strong self-learning ability and test-taking skills and can achieve good results even with an uneven engagement pattern; however, another 30% of cramming students scored below 65, indicating that not everyone is suitable for temporary cramming.

3.4.3 Intervention Strategies for Different Student Types

Differentiated intervention strategies should be adopted for different student types. For balanced students, give them sufficient autonomy and challenging tasks, such as recommending extended reading, inviting them to act as forum "little tutors," etc., to prevent learning burnout. For cramming students, set more "process incentives" during the teaching process, such as adding small quiz nodes, issuing learning progress reminders, to encourage them to distribute learning engagement more evenly throughout the course. For passive students, early identification and early intervention are needed: students with abnormal login behavior in the first two weeks of the semester should be flagged as "warning objects," with teaching assistants conducting individual conversations to understand their difficulties and formulate remediation plans.

The value of cluster analysis lies in moving beyond the assumption of the "average student," acknowledging learner heterogeneity, and implementing precision teaching support accordingly. This is an important direction for data-driven teaching decision-making.

4. Exploration of the Relationship between Learning Behaviors and Learning Outcomes

4.1 Learning Engagement and Academic Performance

4.1.1 Combined Effect of Multi-dimensional Engagement

Learning engagement is a multi-dimensional concept, including behavioral engagement (video watching, assignment completion), emotional engagement (learning interest, perceived value), and cognitive engagement (learning strategies, metacognition). The behavioral data in this study primarily reflect the behavioral engagement dimension. Multiple regression analysis showed that the three variables of video viewing completion rate, average test score, and forum participation frequency jointly explained 52.3% of the variance in final exam scores, with average test score contributing the most (standardized regression coefficient $\beta=0.47$).

This finding has two practical implications: First, behavioral engagement is indeed an important predictor of academic performance, and platform data can be used to issue learning warnings to students. Second, test scores have the strongest predictive power, indicating that "learning by practicing" is particularly crucial in mathematics learning – just watching videos without sufficient practice makes it difficult to internalize knowledge.

4.1.2 Sensitivity Analysis of Engagement Timing

The timing of learning engagement is equally important. The semester was divided into three periods: early (weeks 1-6), middle (weeks 7-12), and late (weeks 13-16). The correlation between learning engagement indicators in each period and final exam scores was calculated separately. The results showed that early engagement had the strongest correlation with final scores ($r=0.51$), followed by middle engagement ($r=0.38$), and late engagement was weakest ($r=0.19$). This finding reveals the importance of "early engagement": students who build a solid foundation and avoid falling behind early on have more confidence and efficiency in subsequent learning, forming a virtuous cycle. In contrast, students who fall behind early may find it difficult to fully catch up even with double effort later.

Teachers should pay special attention to students' engagement at the very beginning of the course, promptly reminding those who miss video viewing in the first week or score low on the first test. A small amount of engagement within the first two weeks of the semester may contribute more to the final score than a large amount of engagement just before the final exam.

4.1.3 Relationship between Learning Regularity and Outcomes

Students' login and study time distribution is also noteworthy. Students were divided into "regular learners" (log in at fixed times daily, evenly distributed study duration) and "random learners" (irregular logins, fluctuating study duration). The performance difference between the two groups was significant: regular learners averaged 79.4 points, random learners averaged 70.6 points ($t=3.45$, $p<0.01$). Regular learners tend to have better time management and self-control skills. These "non-cognitive abilities" are particularly important in online learning environments.

Interestingly, there was no significant difference in performance between "night owl" students (whose main study time was concentrated after 11:00 PM) and "morning person" students, but the former self-reported higher levels of study pressure. This suggests that as long as sufficient engagement and regularity are ensured, the specific time period chosen for studying may not be a determining factor.

4.2 Contribution of Different Learning Segments

4.2.1 Marginal Benefit of Video Watching

A regression model with interaction terms was used to examine the marginal benefit of video watching for students at different performance levels. The results showed: For lower-performing students (final score below 65), improving video viewing completion had the most significant positive impact on scores. For middle-performing students (65-80), the effect of video watching diminished, while the effect of test practice became more prominent. For high-performing students (above 80), the marginal benefit of video watching approached zero; they benefited more from extended reading and in-depth discussions.

This quantitative finding supports the necessity of differentiated instructional design: weaker students need to ensure they complete core video resources; average students should focus more on practice and error correction; excellent students need higher-level challenges and expansion content.

4.2.2 Diagnostic and Promotive Functions of Test Practice

Chapter tests have dual functions: a diagnostic function (exposing problems in students' understanding) and a promotive function (deepening understanding and memory through practice). A path analysis model showed that the average test score directly affected the final exam score (direct effect 0.37) and also indirectly affected the final exam score (indirect effect 0.11) by influencing subsequent video-watching behavior. This indicates that students who perform well on tests have more confidence and motivation to watch subsequent videos, forming a positive cycle of "good test performance → receptive to videos → better learning."

It is worth noting that the "corrective feedback" function of tests is underestimated by many students. In interviews, several students mentioned that they only focused on their test scores and did not carefully read the answer explanations for incorrect questions. Teachers should explicitly emphasize that learning from mistakes is more important than getting a high score and consider incorporating the analysis of incorrect answers into the assessment.

4.2.3 Moderating Role of Forum Interaction

The impact of forum interaction on learning outcomes is not direct but more often plays a moderating role. Hierarchical regression analysis showed that for students with frequent forum participation, the effects of video watching and test practice on scores were stronger. In other words, forum interaction may act as a "catalyst," enhancing the benefits of other learning behaviors.

Analysis of forum content revealed that effective forum interaction usually meets three conditions: the question has depth of thought (not simply "what is the answer"), the responder provides a clear explanation (not simply "the answer is X"), and there are follow-ups and further discussion. Shallow Q&A interactions have a limited effect on promoting learning benefits.

4.3 Prediction of Performance by Learning Behavior Patterns

4.3.1 Early Behavior Warning Indicators

Based on learning behavior data from the first 4 weeks, a logistic regression model was constructed to predict students likely to struggle academically (failing the final exam). The results showed three indicators had significant predictive power: video viewing completion rate below 30% in the first two weeks, first chapter test score below 50%, and less than 10 login days in the first 4 weeks. The combination of these three indicators could improve the identification accuracy of students at risk of failing to 82%. This means teachers can identify about 80% of potential failing students early in the semester (by the 5th week) and intervene in a timely manner.

The value of an early warning system lies in shifting from "post-hoc remediation" to "preventive intervention." Conversing with a student only after they have already failed is far less effective than intervening when problems are budding.

4.3.2 Pattern Recognition of Learning Trajectories

Longitudinal clustering of students' week-by-week learning engagement indicators identified four typical learning trajectories: "sustained high engagement" (about 25%, engagement curve stably high), "sustained low engagement" (about 15%, consistently low), "low-to-high" (about 35%, starting to exert effort mid-term), and "fluctuating" (about 25%, engagement ebbing and flowing). The final exam scores of these four groups were significantly different, in order: sustained high engagement (86.4 points), low-to-high (76.8 points), fluctuating (70.2 points), sustained low engagement (52.7 points).

The "low-to-high" group accounted for a large proportion and scored in the middle, indicating that many students go through a process of "adaptation-adjustment-catch-up." This finding provides insight for teachers: do not give up on students just because of low early engagement; they may well catch up in the middle of the term. The key is to provide timely reminders and encouragement.

4.3.3 Variability of Behavior Patterns and Intervention Windows

Are learning behavior patterns variable? Analysis of the low-to-high group revealed that their "turning point" usually occurred after a significant feedback event: possibly the low score warning from the first chapter test or an individual conversation with the teacher. This suggests the existence of an "intervention window period" – providing appropriate encouragement or warnings to students at critical time points (e.g., after the first test, around the mid-term) can effectively change their subsequent learning behavior.

For the sustained low engagement group, simple reminders often have limited effect. Their problems may be deeper, involving multiple factors such as learning motivation, self-efficacy, or external distractions, requiring more systematic and sustained support.

4.4 Exploration of Factors Influencing Learning Behaviors

4.4.1 Influence of Individual Student Characteristics

Individual student characteristics significantly influence learning behaviors. There was no significant difference between male and female students in video viewing completion rates, but male students had slightly higher posts (average 2.1 vs. 1.4, $p<0.05$). Differences existed among students from different majors: engineering students had the highest video viewing completion rate, followed by economics and management students, and science students had the lowest (possibly because science students had a better foundation in Linear Algebra and felt they did not need to watch videos). Type of learning motivation (intrinsic vs. extrinsic) was an important predictor distinguishing balanced and cramming students.

These findings suggest that when analyzing learning behavior data, teachers should consider students' background characteristics and avoid "one-size-fits-all" judgments.

4.4.2 Influence of Course Design Factors

Many elements of course design influence learning behavior. The presentation format of videos (teacher on screen vs. pure screen recording) had no significant impact on completion rates, but interactive questions embedded in videos significantly reduced attrition rates (decrease of about 12 percentage points). The difficulty level of tests affected retry behavior: too difficult leads to student abandonment, too easy provides no practice value. Incentive mechanisms in the forum (such as bonus points, in-class praise) had a short-term boosting effect on participation, but their long-term effect diminished.

Course design is a process of continuous improvement, requiring teachers to conduct A/B testing based on behavioral data, constantly adjusting and improving.

4.4.3 External Environment and Time Management

External environmental factors were also frequently mentioned by students. About 35% of students indicated that "club activities or student work take up a lot of time, causing video watching to be intermittent"; about 28% reported "dormitory environment is noisy, making it difficult to concentrate on watching videos"; about 20% said that "tasks from multiple online courses overlap, making it difficult to balance time." These problems are beyond the control of a single course teacher, but they can be mitigated in the following ways: providing mobile learning support (facilitating fragmented learning time), setting flexible but clear final deadlines, and announcing the full semester's task timeline in advance to help students plan.

Time management ability is an important prerequisite for online learning. Teachers can hold a "time management workshop" at the beginning of the semester to help students master task planning and prioritization methods.

5. Conclusion and Recommendations

5.1 Main Research Conclusions

This study systematically analyzed the characteristics of students' learning behaviors in the *Linear Algebra* course based on learning behavior data from the SPOC blended teaching platform, drawing the following main conclusions:

First, video-watching behavior exhibits a "beginning-and-end concentration" pattern, with a significant negative correlation between video length and completion rate. Videos within 10 minutes can achieve a completion rate of over 80%, while videos exceeding 12 minutes have a completion rate below 60%.

Second, online test performance is highly correlated with final exam scores, but there is a serious lag in test submission times, with nearly 15% of students submitting within the last 6 hours before the deadline, and their scores are lower.

Third, forum interaction exhibits a phenomenon of "marginal participation," with approximately 78% of students in a state of shallow participation or onlooking. There is a positive correlation between forum activity and academic performance.

Fourth, cluster analysis identified three types of students: balanced (47%), cramming (32%), and passive (21%). Their academic performance differs significantly, requiring categorized interventions.

Fifth, the timing of learning engagement is crucial. Early engagement is much more predictive of final scores than late engagement, making early warning and timely intervention very important.

Sixth, the marginal benefits of different learning segments vary for students at different performance levels. Differentiated teaching and precision support are effective ways to improve overall teaching effectiveness.

5.2 Suggestions for Teaching Improvement

5.2.1 Optimizing the Design and Restructuring of Video Resources

Based on the finding of a negative correlation between video length and completion rate, it is suggested to split existing videos exceeding 12 minutes into 2-3 shorter videos, each focusing on one sub-knowledge point. Embed interactive questions in the videos (popping up one multiple-choice question every 3-5 minutes) to maintain student attention and provide immediate feedback. At the same time, provide video transcripts and key point summaries for students to quickly review and retrieve information.

5.2.2 Strengthening Process-Oriented Evaluation and Incentive Mechanisms

For cramming and passive students, strengthen the binding force and attractiveness of process-oriented evaluation. Specific measures include: setting up more checkpoints (e.g., a small test every two weeks) rather than only chapter tests; linking test "retries" to "incorrect answer analysis reports" to encourage students to truly learn from their mistakes; introducing "learning progress visualization and leaderboards" to add healthy competition and incentives; incorporating learning behavior data (e.g., learning regularity indicators) into the regular grade.

5.2.3 Promoting Deep Forum Interaction and Knowledge Construction

To address the problem of "marginal participation" in forums, the following improvement measures are suggested: lower the participation threshold by setting up a "Novice Question Area" to create a safe and inclusive questioning atmosphere; shift the focus of forum participation from "quantity-oriented" to "quality-oriented," selecting "Best Question" and "Best Answer"; ensure timely responses and follow-up questions from teachers and teaching assistants to guide discussions deeper; compile typical excellent posts in the forum into a "Best Posts Collection" for students' reference.

5.2.4 Implementing Stratified Teaching and Individualized Interventions

Based on student type classification, implement differentiated support strategies. Establish an early warning system to identify high-risk students based on behavior data from the first 4 weeks, with teaching assistants providing one-on-one help and planning guidance. Offer two levels of resources and tasks – "Basic Consolidation" and "Expansion Challenge" – to meet the needs of students at different levels. Regularly analyze learning behavior data, identify individuals requiring attention, and send targeted reminders and encouragement via system messages or emails.

5.3 Research Limitations and Future Directions

This study has the following limitations: First, the research sample was only 180 students from a single university, so the external validity of the conclusions needs to be tested in other institutions and courses. Second, the learning behavior analysis is primarily based on quantitative data, with insufficient depth in exploring students' emotional experiences, cognitive strategies, and other aspects during the learning process. Third, the research period was only one semester, failing to track students' knowledge retention effects and performance in subsequent courses. Fourth, the cluster analysis and causal inference are still relatively preliminary, and the influence of selection bias and confounding variables could not be completely ruled out.

Future research can be deepened in the following directions: First, conduct large-scale cross-institutional and multi-course studies to test the generalizability of the conclusions. Second, adopt methods such as Experience Sampling Method (ESM) and eye-tracking to finely characterize learners' real-time cognitive states. Third, conduct randomized controlled experiments to test the causal effects of specific teaching interventions. Fourth, combine Qualitative Comparative Analysis (QCA) to explore the multiple pathways to success for different types of students. Fifth, extend the tracking period to investigate the long-term effects of blended teaching on learning outcomes.

References

Chen, L., & Guo, Y. J. (2020). A Review of Research on Online Learning Behavior Analysis. *Distance Education in China*, *2020*(2), 36-45.

Fox, A. (2013). From MOOCs to SPOCs. *Communications of the ACM*, *56*(12), 38-40.

Gasevic, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64-71.

He, B., & Cao, Y. (2015). SPOC: Teaching Process Innovation Based on MOOC. *China Educational Technology*, *2015*(3), 22-29.

Kang, Y. Q. (2014). The "Post-MOOC Era" of Online Education – An Analysis of SPOC. *Tsinghua Journal of Education*, *35*(1), 85-93.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. *Proceedings of the Third International Conference on Learning Analytics and Knowledge ACM*, *2013*, 170-179.

Li, H. D., & Wang, C. H. (2019). Construction of a Quality Evaluation System for SPOC Blended Teaching. *Higher Education Exploration*, *2019*(2), 67-72.

Li, S., & He, Z. J. (2018). Research on Factors Influencing Online Student Academic Achievement Based on Learning Analytics. *Open Education Research*, *24*(3), 92-101.

Liu, Q. T., & Ye, Y. (2019). Research on an Online Learning Early Warning Model Based on Learning Behavior Data. *Distance Education in China*, *2019*(8), 33-41.

Liu, Z., & Liu, S. N. Y. (2021). Research on the Relationship Between Learning Behavior Patterns and Learning Outcomes in Blended Learning Environments. *Journal of Educational Technology*, *42*(2), 92-99.

Ministry of Education, China. *Education Informatization 2.0 Action Plan*. 2018.

Wang, N., & Zhang, H. (2018). Practice of SPOC-Based Blended Teaching Model in Advanced Mathematics. *Journal of Mathematics Education*, 2018, 27(4), 73-78.

Wang, Q., & Wang, Y. (2014). SPOC: The Transition from MOOC to Campus Teaching. *China Educational Informatization*, *2014*(11), 12-15.

Wang, S. Q., & Huang, R. H. (2016). Research on SPOC Instructional Design and Its Application. *Modern Educational Technology*, *26*(2), 98-104.

Zhang, Q., & Ma, Z. Q. (2017). Research on SPOC Learning Behavior from the Perspective of Learning Analytics. *Journal of Educational Technology*, *38*(5), 78-84.