Semerikov *et al. Discover Education* (2026) 5:15
https://doi.org/10.1007/s44217-025-01024-1

Discover Education

RESEARCH

Open Access

Multivariate analysis of Moodle components and grade distribution patterns for adaptive learning environments

Serhiy O. Semerikov[1,2,3,4*], Pavlo P. Nechypurenko[1,4], Tetiana A. Vakaliuk[3,2,1,4], Iryna S. Mintii[1,2,3,4,5,6] and Liliia O. Fadieieva[1]

*Correspondence:
Serhiy O. Semerikov
semerikov@gmail.com
[1]Kryvyi Rih State Pedagogical University, 54 Universytetskyi Ave., Kryvyi Rih 50086, Ukraine
[2]Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Street, Kyiv 04060, Ukraine
[3]Zhytomyr Polytechnic State University, 103 Chudnivsyka Street, Zhytomyr 10005, Ukraine
[4]Academy of Cognitive and Natural Sciences, 54 Universytetskyi Ave., Kryvyi Rih 50086, Ukraine
[5]Lviv Polytechnic National University, 12 Stepana Bandery Street, Lviv 79000, Ukraine
[6]University of Łódź, 68 Gabriela Narutowicza Street, Łódź 90-136, Poland

Abstract
The increasing sophistication of Learning Management Systems (LMS) has generated vast amounts of educational data with untapped potential for enhancing teaching and learning. While previous research has examined relationships between LMS components and overall student performance, less attention has been paid to how these components influence specific grade distributions. This study employs multivariate analysis techniques to investigate how Moodle components relate to distinct grade categories (A-F) across 985 courses at a Ukrainian pedagogical university. Using multiple regression models, canonical correlation analysis, and cluster analysis, we identify significant relationships between specific Moodle components and grade outcomes. Information components (particularly teacher presence) strongly predict excellence, while Activity x Communication interactions support satisfactory performance. Resource diversity combined with structured activities correlates with reduced failure rates. We further identify four distinct course design clusters with characteristic grade distribution patterns. The findings demonstrate how learning analytics can move beyond descriptive insights toward prescriptive guidance. This nuanced understanding of grade distribution patterns provides practical guidance for tailoring course design to diverse pedagogical objectives and student needs, thus contributing to the emerging field of personalized digital learning environments.

Keywords Learning management systems, Moodle, Grade distributions, Learning analytics, Adaptive learning, Context-aware learning environments

1 Introduction

The rapid advancement of digital learning technologies has transformed educational practices worldwide, with Learning Management Systems (LMS) becoming central to higher education pedagogy. These platforms generate substantial amounts of data that, when properly analyzed, can provide valuable insights into the complex relationship between learning design and student outcomes [1]. However, as Alnasyan et al. [2]

observe, the field is still developing approaches to translate these data-driven insights into actionable strategies for personalized learning experiences.

While LMS generate vast amounts of data about course design and student outcomes, most research has focused on predicting overall academic performance rather than understanding how specific LMS components influence different levels of achievement [3]. This gap is particularly relevant for developing adaptive learning environments, which require nuanced understanding of how course design elements affect diverse learning outcomes. By examining grade distributions rather than aggregate performance measures, we can identify which LMS components support excellence, ensure satisfactory performance, or prevent failure, thus enabling more targeted adaptations [4].

Moodle, as one of the most widely adopted LMS platforms globally, offers a rich environment for examining how digital learning components influence educational outcomes. Previous research has predominantly examined the relationship between LMS components and student performance as a unidimensional construct, typically measuring overall achievement through aggregated metrics [5, 6]. This approach, while valuable, overlooks the nuanced ways in which LMS components might differentially impact various levels of academic achievement.

A recent study by Fadieieva and Semerikov [7] and Semerikov et al. [8] investigated the interconnectedness of Moodle constructs and their relationship with student assessment using Structural Equation Modeling – Partial Least Squares (SEM-PLS). Their findings revealed significant relationships between Information and Assessment constructs, and between Activities and Communication constructs. However, the Assessment construct was treated as a single latent variable rather than examining how Moodle components might uniquely influence different grade categories.

This research gap presents a significant opportunity to enhance our understanding of how LMS design influences diverse learning outcomes. By analyzing grade distributions as separate dependent variables, we can identify which Moodle components are most effective for promoting excellence, ensuring satisfactory performance, or preventing failure. Such insights would allow instructors and instructional designers to tailor course designs to specific educational objectives, thus contributing to the development of more sophisticated adaptive learning environments.

The present study addresses this gap by employing multivariate analysis techniques to examine how different Moodle components relate to specific grade distributions across 985 courses at Kryvyi Rih State Pedagogical University (KSPU) in Ukraine. Building upon previous work, we adopt a novel analytical approach focused on grade distribution patterns rather than aggregate assessment outcomes, aligning with the growing research interest in developing more nuanced learning analytics models that can support context-aware personalization [9].

Our primary aims are to: (1) identify specific Moodle components associated with different grade outcomes, (2) examine multivariate relationships between component combinations and grade distribution patterns, and (3) identify course design archetypes associated with distinctive grade profiles.

This approach extends current learning analytics research by moving beyond descriptive and diagnostic analyses toward predictive and prescriptive insights that can guide instructional design decisions [10].

2 Theoretical framework and literature review

2.1 Social constructionist pedagogy and adaptive learning in moodle

Moodle's architecture embodies social constructionist pedagogy, which synthesizes three complementary learning theories to create a comprehensive framework for digital learning [11]. Constructivism, as articulated by Piaget [12], posits that learners actively build knowledge by connecting new experiences to existing mental frameworks. This principle manifests in Moodle through features that allow learners to organize and structure their own learning paths. Constructionism, developed by Papert [13], extends constructivist theory by emphasizing that learning becomes most effective when learners create tangible artifacts to share with others. Moodle operationalizes this through assignments, glossaries, and databases where students produce content for peer review and collaboration. Social constructivism, following Vygotsky [14], highlights that knowledge construction occurs through social interaction within cultural contexts, implemented in Moodle through forums, wikis, and workshops that facilitate collaborative learning within the Zone of Proximal Development.

These theoretical foundations translate into four core design principles that shape Moodle's functionality. First, the platform recognizes all participants as potential teachers and learners, enabling peer teaching through forums, wikis, and workshops where students can take instructional roles. Second, learning occurs through creating artifacts for others, with assignments, glossaries, and databases allowing learners to produce shareable content that benefits the entire learning community. Third, social interaction drives cognitive development, with communication tools facilitating dialogue that supports learners operating within their Zone of Proximal Development. Fourth, context shapes learning, allowing course structures to be adapted to disciplinary and cultural needs through flexible configuration options.

The platform operationalizes these principles through specific architectural decisions that support different aspects of learning. Collaborative knowledge construction is enabled through forums supporting threaded discussions for reflective dialogue, wikis facilitating collaborative content creation that shows the co-evolution of individual and collective knowledge [15], and workshops implementing peer assessment processes that embody the teacher-learner duality [16]. Scaffolded learning support is provided through conditional activities and completion tracking that enable instructors to create scaffolded learning paths guiding students through their Zone of Proximal Development. The gradebook and rubrics provide detailed feedback mechanisms essential for formative assessment within a constructivist framework, while features like adaptive quiz questions and lesson branching allow content to adjust based on learner responses [17].

Multiple representation formats support different learning styles through resource diversity including pages, files, URLs, and books, enabling learners to encounter concepts through various representations and aligning with constructivist emphasis on multiple perspectives [18]. Active learning through engagement is promoted via interactive components such as quizzes, assignments, and SCORM packages that move beyond information transmission to engage learners in applying, analyzing, and creating knowledge, which are core constructionist activities [19].

Within this pedagogical architecture, Moodle components serve distinct but interconnected functions that align with the theoretical framework. Information components provide course metadata and contextual information that helps learners understand course expectations and structure, including course descriptions, teacher information, educational level, semester timing, and course status. These components establish the learning context and support learner orientation and self-regulation, creating the foundation for constructivist learning by helping students understand the framework within which they will build knowledge.

Resource components deliver static learning materials in various formats for knowledge acquisition. Pages provide web-based content, URLs link to external resources, Files offer downloadable documents, Books present multi-page resources, Folders organize collections, and Labels display text or multimedia directly on the course page. These support diverse learning preferences through multiple content representations, aligning with constructivist principles that recognize different pathways to understanding.

Activity components engage learners in interactive tasks that promote active learning and skill development. Assignments allow submitted work, Quizzes provide assessments, Forums enable discussions, SCORM packages deliver interactive content, Choice activities create polls, and Workshops facilitate peer assessment. These components operationalize the constructionist principle of learning through doing, providing opportunities for learners to create and share artifacts that demonstrate their understanding.

Communication components facilitate interaction between learners and instructors to support collaborative knowledge construction. Forums enable asynchronous discussions while Chat supports synchronous conversations. These components embody social constructivism by enabling dialogue, peer learning, and community building essential for knowledge construction within cultural contexts.

Recent research extends this foundation by exploring how these theoretically-grounded components can support adaptive learning. Mirata and Bergamin [20] argue that effective adaptation must preserve the social and constructive aspects of learning rather than merely personalizing content delivery. This perspective aligns with the framework by Hilali et al. [21], which identifies the need to integrate learner modeling, content design, and adaptive technologies while maintaining pedagogical coherence.

Within this expanded framework, adaptive learning in Moodle involves dynamic scaffolding through conditional release of activities based on learner progress, personalized feedback via rubrics and competency frameworks that guide individual development, collaborative adaptation where peer interactions inform learning pathways, and context-aware resource selection based on learner characteristics and performance patterns. This integration of social constructionist principles with adaptive learning capabilities positions Moodle as more than a content delivery system. It becomes an environment for contextualized, collaborative knowledge construction that can respond to diverse learner needs while maintaining its pedagogical foundations [9].

2.2 Learning analytics and grade distributions

The field of learning analytics has evolved from descriptive analysis of learning behaviors to predictive modeling of student outcomes. Early studies focused primarily on predicting overall academic performance using aggregated metrics [22]. However, recent research recognizes the limitations of treating academic achievement as a unidimensional construct. Lópes-Pernas and Saqr [23] demonstrate that temporal dynamics of learning behaviors create complex trajectories that cannot be captured by single performance metrics. Their VaSSTra method reveals how students transition between different performance states over time, suggesting that grade categories represent qualitatively different learning states rather than points on a continuous scale.

Building on this trajectory-based understanding, Uwimpuhwe et al. [24] propose moving beyond standardized effect sizes toward a "gain index" that measures the percentage of learners making positive gains. This approach aligns with our focus on analyzing grade distributions, as it recognizes that interventions may differentially affect students at different performance levels. Their latent class analysis reveals distinct subpopulations of learners who respond differently to educational interventions, supporting our hypothesis that different Moodle components may have varying effects across grade categories.

Recent empirical studies provide evidence that students at different achievement levels require different types of support. Van der Velde et al. [25] conducted a large-scale evaluation of adaptive learning systems involving 100 million trials from 140,000 learners. They found that cold-start mitigation strategies were most effective when tailored to specific difficulty levels, with content-specific adaptations ("what") generally outperforming learner-specific adaptations ("who"). This finding suggests that the relationship between learning resources and outcomes may vary based on the cognitive demands of different achievement levels.

Similarly, Lawrence et al. [26] examined submission limits and regression penalties in automatic assessment systems, finding that different intervention strategies had varying effects on high-performing versus struggling students. High-achieving students benefited from unlimited submissions that allowed exploration and refinement, while struggling students performed better with structured limits that prevented trial-and-error approaches. These findings support our hypothesis that excellence, satisfactory performance, and failure prevention may require different combinations of Moodle components.

Several studies have examined how specific LMS components influence learning outcomes, though most have not differentiated between grade categories. Rienties et al. [27] analyzed ten years of implementation data from the Open University UK, identifying distinct patterns of LMS usage associated with different learning outcomes. They found that assimilative activities (like reading resources) showed stronger relationships with performance in knowledge-based assessments, while productive activities (like assignments) better predicted performance in skill-based assessments.

Wang et al. [28] investigated the impact of dashboard feedback types on learning effectiveness, finding significant interactions between feedback format and learner characteristics. Visual learners showed improved performance with graphical dashboards, while verbal learners benefited more from text-based feedback. This research suggests that resource diversity in Moodle may be particularly important for accommodating different learning preferences, potentially explaining differential effects across grade categories.

Research by Muñoz et al. [29] in their systematic review of adaptive learning technology identified patterns in how different technological features support learning. They found that scaffolding features (similar to Moodle's conditional activities) were most effective for intermediate learners, while advanced learners benefited more from open-ended exploration tools. This aligns with our conceptual model suggesting that Activity components may show stronger relationships with satisfactory performance than with excellence.

Despite these advances, significant gaps remain in understanding how LMS components differentially affect grade distributions. First, most studies treat performance as a continuous variable or binary outcome (pass/fail), missing the nuanced differences between grade categories. Second, few studies examine interaction effects between different types of LMS components, despite theoretical predictions from social constructivist pedagogy. Third, limited research connects learning analytics findings to practical course design strategies for achieving specific grade distribution objectives.

Our study addresses these gaps by examining how Moodle components relate to specific grade categories (A, B, C-D, F), investigating interaction effects between component types, and identifying course design patterns associated with characteristic grade distributions. This approach extends the current literature by providing empirical evidence for the theoretical predictions of differentiated support needs across achievement levels.

2.3 Context-aware adaptive learning environments

The development of context-aware adaptive learning environments represents a convergence of learning analytics, adaptive technologies, and pedagogical theory. Hasanov et al. [3] define adaptive context-aware learning environments (ACALEs) as systems that detect the learner's context and adapt learning materials accordingly, making learning contextually relevant. Their comprehensive survey identifies three key components of successful ACALEs: context detection mechanisms that gather information about learner state and environment, adaptation rules that determine how to modify the learning experience, and feedback loops that assess the effectiveness of adaptations. Recent implementations of context-aware systems demonstrate the importance of considering multiple contextual dimensions. Sihabudin Sahid et al. [30] propose a model of personalized context-aware e-learning that incorporates psychological experiences alongside traditional performance metrics. Their system adapts not only to cognitive factors but also to affective states, recognizing that emotional context significantly influences learning effectiveness. This multidimensional approach aligns with our conceptualization of grade categories as representing different cognitive, motivational, and behavioral states rather than merely different levels of knowledge.

Salazar et al. [31] demonstrate how context-awareness services can be integrated within multi-agent system architectures for instructional planning and resource recommendation. Their system monitors student progress, identifies contexts requiring intervention, and automatically adjusts resource presentation and activity sequencing. The success of their approach depends on accurate context modeling that captures not just current performance but also learning trajectories and patterns of engagement. This supports our approach of examining grade distributions as contextual indicators for adaptation rather than relying solely on aggregate performance metrics.

The operationalization of context-awareness in LMS settings faces several challenges identified by Aziz and Khoukhi [9]. They propose an ontological approach for service adaptation that addresses the complexity of educational contexts through structured knowledge representation. Their framework distinguishes between static context (learner characteristics, course properties) and dynamic context (current performance, engagement patterns), suggesting that effective adaptation requires considering both dimensions. Our identification of course design patterns and their associated grade distributions provides a bridge between these static and dynamic contexts, offering a practical framework for implementing context-aware adaptations.

2.4 Theoretical synthesis and conceptual model

Building on the theoretical foundations and empirical evidence reviewed above, we propose a conceptual framework that examines the relationships between Moodle components and specific grade categories. This framework integrates insights from social constructionist pedagogy, learning analytics research, and context-aware adaptive learning to hypothesize differential relationships between course design elements and achievement levels.

Our conceptual model synthesizes these theoretical perspectives to predict how different Moodle components support different types of learning outcomes. The model recognizes that excellence, good performance, satisfactory performance, and failure prevention represent qualitatively different learning states requiring different types of support. This differentiated view extends beyond traditional approaches that treat performance as a continuous variable, instead recognizing the complex, multidimensional nature of academic achievement in digital learning environments.

2.4.1 Information components and academic excellence

We hypothesize a positive relationship between information components and excellence based on several theoretical foundations. Cognitive load theory [32] suggests that well-structured information presentation reduces extraneous cognitive load, allowing learners to allocate more cognitive resources to germane load associated with schema construction. Information components in Moodle provide organizational scaffolding through clear course descriptions, structured syllabi, and teacher presence indicators. Self-regulated learning theory [33] indicates that high-achieving students benefit from clear expectations and transparent course structures. The presence of multiple teachers may provide diverse expertise and increased availability for guidance, supporting the autonomous learning strategies characteristic of excellent students. Research on deep learning approaches shows positive correlations with academic achievement [34–36], and students who engage in deep learning require comprehensive understanding and integration of knowledge, which is facilitated by well-organized information architecture [37]. Therefore, we hypothesize that information components will show their strongest relationship with excellence, moderate relationships with good performance, and weaker relationships with lower grade categories. Supporting this, studies in Moodle environments reveal that comprehensive information resources have stronger impacts on performance for graduate students – who often possess greater self- regulation skills – compared to undergraduates [8], suggesting that higher-achieving learners leverage structured information more effectively, while lower performers may exhibit weaker relationships due to differential engagement.

2.4.2 Resource components and performance levels

Resource components serve different functions across performance levels, leading to hypothesized differential relationships. For excellence-oriented students, diverse resource formats support the constructionist principle of learning through multiple representations [18]. These students typically demonstrate effective self-regulation and can navigate varied resources to construct deep understanding, leading us to hypothesize a moderate positive relationship between resource components and excellence. For students achieving satisfactory performance, the relationship may be weaker, as research on surface learning approaches suggests that students employing memorization strategies may become overwhelmed by resource diversity without proper scaffolding. However, resource diversity shows promise for failure prevention, as Universal Design for Learning principles emphasize multiple means of representation to accommodate different learning preferences [38].

2.4.3 Activity components and scaffolded learning

Activity components operationalize the Zone of Proximal Development through structured practice opportunities. Our hypothesized relationships are grounded in formative assessment theory [39]. For satisfactory performance, we hypothesize a strong positive relationship, as students operating within their ZPD for most course objectives benefit significantly from scaffolded activities that provide regular practice with feedback. Assignment components offer the structured support these students need to achieve basic competencies. For excellence, the relationship may be weaker, as high-achieving students often demonstrate effective self-regulation and may rely less on structured activities, preferring to engage with complex, open-ended challenges. However, Quiz components may maintain a moderate relationship with excellence by providing self-assessment opportunities that support metacognitive development.

2.4.4 Communication components and social learning

Communication components embody social constructivist principles, but their relationships with different grade categories reflect varying needs for peer support. We hypothesize moderate relationships with both satisfactory performance and failure prevention. Social learning theory [40] suggests that peer interaction supports knowledge construction, particularly for students who benefit from collaborative scaffolding within their zone of proximal development [41]. Forums enable asynchronous reflection and peer support, while synchronous chat may help identify and address struggling students in real-time. The relationship with excellence may be weaker, as high-achieving students might engage selectively with communication tools, using them for deeper discussions rather than basic support. However, research suggests mixed findings regarding high achievers' use of communication tools. While Webb [42] found that high-ability students often provide explanations rather than receive them in peer interactions, potentially limiting their learning gains from forums, other studies [43] indicate that structured collaborative activities can benefit all achievement levels, including excellent students engaged in complex problem-solving tasks.

2.4.5 Interaction effects

Beyond individual component effects, we hypothesize specific interaction effects based on complementary learning mechanisms. The Activity × Communication interaction is hypothesized to support satisfactory performance particularly well, as the combination of structured practice with peer discussion creates a supportive environment addressing both individual practice needs and social learning opportunities. The Resource × Activity interaction is hypothesized for failure prevention, as diverse resource formats combined with structured practice opportunities may create multiple pathways to understanding, reducing the likelihood of complete disengagement or failure.

2.4.6 Limitations of the conceptual model

While we ground our hypotheses in established learning theories, several relationships remain exploratory due to limited empirical evidence specific to LMS contexts. The strength of relationships represents theoretical predictions requiring empirical validation. The model assumes relatively homogeneous student populations within courses, though individual differences in learning approaches may moderate these relationships. Additionally, the quality of component implementation is not captured in our model, though it likely influences the strength of relationships.

Figure 1 presents our conceptual model with hypothesized relationships. These hypotheses serve as a starting point for empirical investigation rather than definitive claims, acknowledging that the actual relationships may vary based on context, student characteristics, and implementation quality.

2.4.7 Research model summary

Our conceptual model proposes that different Moodle components will show varying relationships with different grade categories based on the cognitive demands and

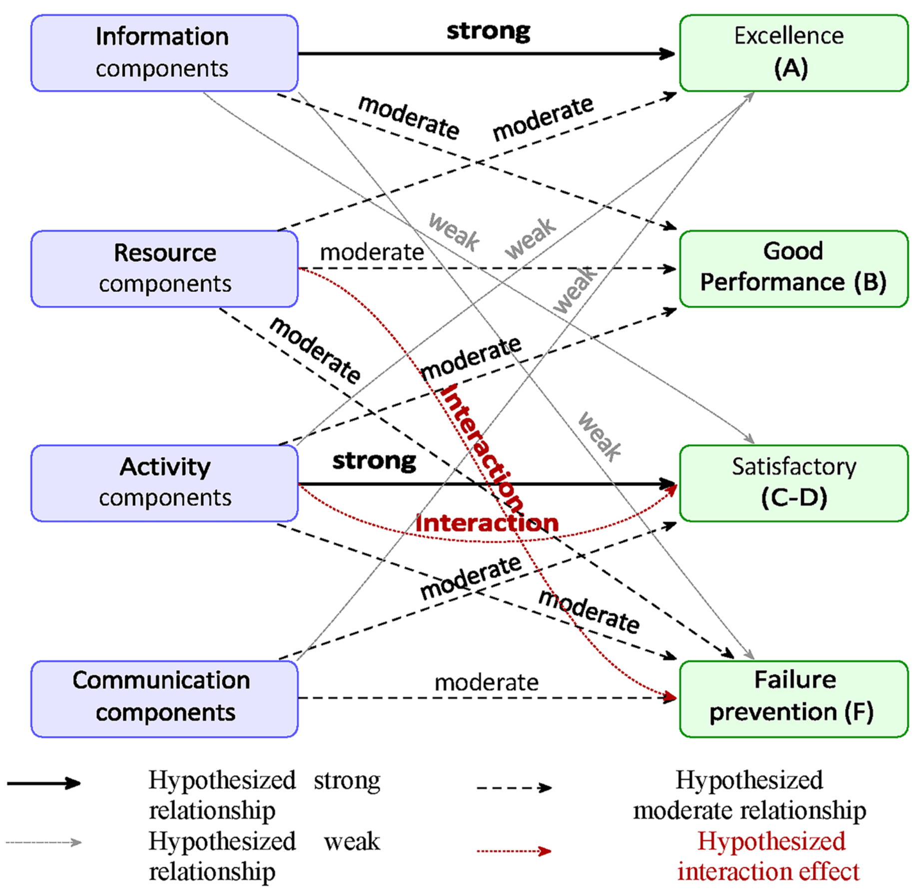

Fig. 1 Conceptual model of hypothesized relationships between Moodle components and grade categories based on learning theories. Relationship strengths are theoretical predictions requiring empirical validation. The model integrates cognitive load theory, self-regulated learning, formative assessment theory, and social constructivism to predict differential effects of LMS components on various achievement levels

support needs associated with each achievement level. These relationships are moderated by the alignment between component affordances and student learning approaches, whether deep, strategic, or surface. Component interactions may produce effects beyond individual components, particularly for supporting middle-achieving students and preventing failure. Course characteristics such as faculty, educational level, and semester may further moderate these relationships. This theoretically-grounded framework guides our empirical investigation while acknowledging the exploratory nature of specific relationship strengths. The model extends traditional learning analytics approaches by connecting established learning theories to LMS design elements, providing a foundation for evidence-based adaptive learning strategies.

3 Research questions

Building on the conceptual framework and addressing gaps identified in the literature review, this study investigates the following research questions:

1. RQ1 – Excellence predictors: How do specific Moodle components differentially associate with excellence in academic performance (A grades)? Which components are most strongly associated with excellence, and what combinations of components best predict high proportions of A grades?
2. RQ2 – Good performance factors: What relationships exist between Moodle components and good performance (B grades)? How do the predictors of good performance differ from those of excellence, and are there components that specifically support the transition from satisfactory to good performance?
3. RQ3 – Satisfactory performance support: What relationships exist between Moodle components and satisfactory performance (C-D grades)? Are there specific components that primarily support basic competency achievement, and how do the predictors of satisfactory performance differ from those of good and excellent performance?
4. RQ4 – Failure prevention mechanisms: Which Moodle design elements contribute to failure prevention (reducing F grades)? What components are associated with lower proportions of failing grades, and are there interaction effects between components that enhance support for struggling students?
5. RQ5 – Course design patterns: What course design patterns can be identified through multivariate analysis of component combinations and grade distributions? Can we identify distinct course design archetypes with characteristic grade distribution profiles, and which design patterns best support diverse learning needs across the performance spectrum?

These research questions address a critical gap in current educational technology research by investigating the complex, multivariate relationships between course design elements and differentiated learning outcomes. The answers will contribute to the development of more sophisticated context-aware adaptive learning environments that can dynamically respond to diverse educational objectives and learner needs.

4 Methodology

4.1 Research design

This study employs a quantitative, non-experimental design using existing data from completed courses. The approach combines correlational analysis with predictive modeling to examine relationships between Moodle components and grade distributions. This methodology aligns with current approaches in learning analytics research, which increasingly emphasize multivariate analysis and pattern recognition [44].

The study extends beyond traditional learning analytics approaches in several ways. First, we treat grade categories as separate dependent variables rather than using aggregated performance measures, allowing us to identify differential effects across achievement levels. Second, we employ canonical correlation analysis to examine multivariate relationships between components and outcomes, revealing how combinations of components relate to overall grade patterns. Third, we use cluster analysis to identify course design archetypes with distinctive grade distribution profiles, pro- viding practical templates for course design. Finally, we connect analytical insights to adaptive design strategies, bridging the gap between analysis and application.

This multi-faceted methodological approach enables us to move from descriptive analytics (what happened) through diagnostic analytics (why it happened) to predictive analytics (what might happen) and ultimately toward prescriptive analytics (what should be done), following the maturity model proposed by Tepge¸c and Ifenthaler [45].

4.2 Dataset and context

This study utilizes data from Kryvyi Rih State Pedagogical University (KSPU) in Ukraine, collected and digitized across all university specialties for the 2020–2021 (winter and summer sessions) and 2021–2022 (winter session) academic years. The dataset includes 985 Moodle courses with complete information on course components and student grade distributions.

The context of data collection encompasses a period affected by the COVID-19 pandemic, during which higher education institutions worldwide rapidly transitioned to online and blended learning formats. This contextual factor is important for interpreting the findings, as it represents a period of accelerated digital transformation in educational practices. While this context may limit generalizability to normal operating conditions, it provides valuable insights into LMS usage patterns when digital learning becomes central rather than supplementary to instruction.

The data collection and anonymization process was approved by the KSPU rector's decision and the university's ethical committee on January 26, 2024, ensuring compliance with ethical standards for educational data analysis. The dataset is publicly available at Zenodo [46], promoting transparency and reproducibility. The public availability of this dataset enables other researchers to validate our findings and extend the analysis in new directions.

4.3 Variables and operationalization

4.3.1 Dependent variables: grade category proportions

The dependent variables are proportions of each grade category within each course, calculated as the percentage of enrolled students receiving each grade. These proportions were arcsin-square-root transformed to address the bounded nature of percentage data and improve normality for statistical analysis. The transformation formula applied was $Y' = \arcsin\left(\sqrt{Y/100}\right)$, which is standard for proportion data in educational research.

The grade categories analyzed include A% (90–100%, excellence level), B% (80–89%, good performance), C% (70–79%, satisfactory performance), D% (60–69%, satisfactory performance), E% (50–59%, lowest passing grade), and F% (below 50%, failing grades including both F and FX designations). For certain analyses, we combine C% and D% into a single "satisfactory performance" category based on findings from Fadieieva and Semerikov [7] that showed these categories function similarly in the KSPU assessment system. The E grade category was not modeled separately due to its concentration at exactly 50 points (the minimum passing threshold), which limited meaningful variation for regression modeling.

4.3.2 Independent variables: moodle components

All Moodle component variables were operationalized as raw counts representing the number of each component type present in a course. This count-based approach was chosen over binary presence/absence coding to capture the intensity of component usage, as recommended by Rienties et al. [22] for learning analytics research. The independent variables were organized into four theoretically-grounded constructs.

Information construct includes variables capturing course context and structure: Form of education (1 = full-time, 2 = part-time, 3 = both), Educational level (1 = undergraduate, 2 = graduate), Semester (1–8 continuous), Status (1 = required, 2 = elective), and Number of teachers (1–9 continuous). These variables establish the learning context and institutional parameters within which courses operate.

Resources construct encompasses content delivery components: Label (text/-multimedia displayed directly), Page (web-based content), URL (external links), Book (multipage resources), Folder (organized collections), and File (downloadable documents). Each is measured as a count of instances within the course.

Activities construct includes interactive learning components: Assignment (submitted work), Quiz (assessments), SCORM (interactive packages), Glossary (collaborative definitions), Lesson (branching content), Feedback (surveys), H5P (HTML5 activities), HotPot (interactive exercises), Survey (questionnaires), Database (structured collections), Choice (polls), Visiting (attendance tracking), Wiki (collaborative pages), LTI External tool (integrated applications), and Workshop (peer assessment). All are measured as counts of instances.

Communication construct comprises interaction tools: Forum (asynchronous discussions) and Chat (synchronous conversations), both measured as counts. Additionally, we calculated a "Resource diversity" variable representing the number of different resource types used in each course, as recent research by Asriadi A. M. et al. [47] suggests that differentiated instruction approaches can positively impact learning outcomes.

4.3.3 Control variables

To manage inter-course differences, we included several control variables. Faculty (coded 1–9) was dummy-coded to control for disciplinary differences in grading practices and pedagogical traditions. Educational level controls for student maturity and preparation differences between undergraduate and graduate courses. Semester (1–8 continuous) controls for curriculum progression effects. Course status (required/elective) controls for potential differences in student motivation and engagement. Form of education controls for delivery mode differences between full-time and part-time programs. Variance inflation factors (VIF) were calculated for all variables to ensure multicollinearity was not problematic, with all values below 3.0 indicating acceptable levels.

4.3.4 Contextual factors and study limitations

We acknowledge that multiple contextual factors influence grade distributions beyond LMS component usage. Table 1 summarizes these factors and our approach to addressing them.

These limitations mean our findings should be interpreted as associations within a specific institutional context rather than universal relationships. The heterogeneity in unmeasured factors likely attenuates our observed relationships, suggesting that true effects within homogeneous course groups may be stronger than reported here.

4.4 Analytical approaches

4.4.1 Alignment of research questions and analytical methods

Our analytical strategy employs three complementary methods that address different aspects of the research questions. Table 2 presents the alignment between research questions and analytical methods.

This multi-method approach provides both detailed component-level insights through regression and holistic system-level understanding through canonical correlation and clustering. The analytical progression moves from examining individual grade categories to understanding multivariate patterns and finally to identifying design archetypes, providing multiple levels of insight for adaptive learning system design.

4.4.2 Data preparation and exploratory analysis

Initial data preparation involved comprehensive quality checks and transformations. Missing value analysis revealed complete data for all variables in the final dataset. Outlier

Table 1 Contextual factors affecting grade distributions and study approaches

Factor	Potential impact	Our approach	Limitations
Course discipline	Different fields have different grading norms	Included faculty (9 categories) as control variable; conducted separate analyses by faculty	Cannot capture within-faculty variation
Course type	Required courses may have different grade distributions than electives	Included as control variable (required/elective)	Binary classification oversimplifies
Difficulty level	Advanced courses typically have different distributions	Controlled viaeducational level (UG/grad) and semester (1–8)	No direct difficulty measure
Instructor quality	Teaching effectiveness strongly impacts grades	Number of teachers included; instructor effects partially captured	Cannot measure teaching quality directly
Grading standards	Instructors have different grading philosophies	Faculty-level controls capture some departmental norms	Individual instructor variance not captured
Assessment methods	Exam-based vs. project-based affects distributions	Moodle components partially proxy for this (Quiz vs. Assignment)	Cannot distinguish assessment quality
Class size	Large classes often have different distributions	Not directly measured	Important omission
Course content	Technical vs. theoretical content affects grades	Not measured	Significant limitation

Table 2 Mapping of research questions to analytical methods

RQ	Focus	Primary method	Justification
RQ1	Excellence (A grades)	Multiple regression	Identifies specific components and their relative contributions to high achievement; allows testing of interaction effects
RQ2	Good performance (B grades)	Multiple regression	Enables comparison with excellence predictors; identifies unique support needs for good performers
RQ3	Satisfactory performance (C-D grades)	Multiple regression	Reveals components supporting basic competency; allows direct comparison across performance levels
RQ4	Failure prevention (F grades)	Multiple regression	Identifies protective factors; interaction terms reveal synergistic effects for at-risk students
RQ1-4	Multivariate relationships across all grades	Canonical correlation analysis	Examines holistic patterns between component sets and grade distributions; validates regression findings through multivariate lens
RQ5	Course design patterns	Cluster analysis	Identifies naturally occurring design archetypes; reveals how component combinations create distinct grade profiles

detection using Mahalanobis distance identified five extreme outliers (0.5% of cases) that were removed as they represented data entry errors. Normality testing through Shapiro-Wilk tests indicated that grade proportion variables required transformation, leading to the arcsin-square-root transformation described earlier. Multicollinearity diagnostics showed acceptable VIF values for all predictors.

Exploratory analysis included examination of grade distributions across courses revealing substantial variation in patterns, correlation analysis between Moodle components and grade categories identifying preliminary relationships, and visual exploration through scatterplots and boxplots revealing non-linear relationships for some variables. These exploratory findings informed subsequent model specifications and variable transformations.

4.4.3 Multiple regression models

To address RQ1-RQ4, we developed separate multiple regression models for each grade category following the general form:

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_{ij} + \sum_{j=1}^{m} \sum_{l=j+1}^{m} \beta_{jl}(X_{ij} \times X_{il}) + \epsilon_i,$$

where Y_i represents the transformed grade category percentage for course *i*, X_{ij} represents the value of Moodle component *j* for course *i*, β_j represents the main effect coefficient, β_{jl} represents the interaction effect coefficient, and ϵ_i represents the error term.

Model development followed a systematic process. We began with all main effects and control variables, then used stepwise regression with AIC criterion to identify significant predictors. Interaction terms were added based on theoretical predictions and preliminary analyses. Model diagnostics included residual analysis, influence diagnostics, and cross-validation. Models were evaluated based on R^2, adjusted R^2, significance of predictors, and effect sizes following procedures recommended by Josephine et al. [48] for educational data analysis.

4.4.4 Canonical correlation analysis

To provide a multivariate perspective complementing the separate regression models, canonical correlation analysis (CCA) examined relationships between two variable sets:

Set 1 containing Moodle components (Resources, Activities, Communication) and Set 2 containing Grade categories (A%, B%, C-D%, F%). CCA identifies linear combinations of variables in each set (canonical variates) that maximize correlation between sets, revealing patterns of association between component combinations and grade distribution profiles.

The CCA procedure involved several steps. First, we standardized all variables to ensure comparability. Then we computed canonical correlations and tested significance using Wilks' lambda. We interpreted canonical functions through examination of canonical loadings (correlations between variables and their canonical variates) and cross-loadings (correlations between variables and opposite canonical variates). Finally, we validated findings through split-sample analysis. This approach follows recommendations by López-Pernas and Saqr [23] for analyzing educational data.

4.4.5 Cluster analysis

To address RQ5, we employed k-means cluster analysis to identify distinct course design patterns. This method reveals naturally occurring groupings based on component profiles, connects component combinations to grade distributions, and provides actionable design archetypes.

The clustering procedure included standardizing all Moodle component variables to prevent scale effects, determining optimal cluster number using elbow method and silhouette analysis, running k-means algorithm with 25 random starts to ensure stability, and validating clusters through split-sample reproducibility testing. For each cluster, we characterized component usage patterns through mean comparisons, examined average grade distribution profiles, and identified distinctive features through discriminant analysis. This approach aligns with recent research by Wang et al. [28] on identifying learning design patterns.

4.4.6 Statistical assumptions and diagnostic tests

Prior to main analyses, we performed comprehensive diagnostic tests. For multiple regression, we verified linearity through partial regression plots, independence through Durbin-Watson statistics (range 1.87–2.13), homoscedasticity through Breusch-Pagan tests (all $p > 0.05$), normality of residuals through Q-Q plots and Shapiro-Wilk tests, and absence of multicollinearity through VIF values (all < 3.0).

For canonical correlation, we confirmed sample size adequacy (985 cases, 27 variables exceeding 20:1 ratio), linearity through bivariate scatterplots, and acceptable multivariate normality through Mardia's test (skewness $p = 0.082$, kurtosis $p = 0.067$). For cluster analysis, we standardized variables to prevent range domination, examined multivariate outliers through Mahalanobis distance, and validated stability through 85% agreement in split-sample assessment.

4.4.7 Analytical software and implementation

Analyses were conducted using R (version 4.2.0) with packages including tidyverse for data manipulation, stats for regression models, CCA for canonical correlation, cluster for k-means clustering, and ggplot2 for visualization. All analysis scripts are available upon request to ensure reproducibility.

4.5 Ethical considerations

This study adheres to ethical standards for learning analytics research as outlined by Willis et al. [49]. Data privacy was ensured through complete anonymization with no personally identifiable information. Institutional approval was obtained from KSPU administration and ethics committee. Transparency is promoted through public dataset availability. The research aims to improve educational practices demonstrating beneficence. Care was taken to avoid potential misinterpretation ensuring non-maleficence. These considerations are particularly important in learning analytics where data-driven approaches must balance innovation with ethical responsibility.

5 Results

Before presenting our findings, we emphasize that the relationships reported below represent average effects across a heterogeneous sample of courses varying in discipline, difficulty, instructor quality, and pedagogical approach. Our results should therefore be interpreted as general patterns that may vary substantially in specific contexts. To partially address this heterogeneity, we conducted robustness checks including separate analyses by faculty, subgroup analyses for required versus elective courses, and sensitivity analyses excluding potential outlier disciplines.

5.1 RQ1: Components associated with excellence (A grades)

Our analysis of excellence predictors reveals distinct patterns differentiating high achievement from other performance levels. The regression model for excellence (A%) explains 32.4% of the variance (adjusted $R^2 = 0.324$, F (5, 979) = 95.14, $p < 0.001$) (see Table 3).

The prominence of Number of teachers as the strongest predictor ($\beta = 0.41$, $p < 0.001$) suggests that courses with multiple instructors provide more diverse expertise and support, enabling more students to achieve excellence. This finding aligns with research on team teaching approaches highlighting benefits of multiple perspectives for student learning [29].

Content-focused components (URL and Page resources) show significant positive relationships with excellence, suggesting that well-structured, accessible content supports higher achievement. These components typically deliver core course materials and facilitate self-directed learning for high-performing students, aligning with research by Wang et al. [50] on the importance of well-organized content presentation.

Quiz components emerge as significant predictors, possibly serving dual functions of providing formative assessment opportunities and motivating regular engagement. This supports research on the importance of regular formative assessment in promoting deeper learning [51].

Table 3 Multiple regression results for excellence model (A%)

Predictor	B	SE	β	*p*-value
(Intercept)	14.56	2.37	–	< 0.001
Number of teachers	5.27	0.63	0.41	< 0.001
Educational level	-7.83	1.69	-0.22	< 0.001
URL resources	0.21	0.08	0.14	0.007
Quiz components	0.18	0.07	0.12	0.009
Page resources	0.24	0.12	0.10	0.042

Table 4 Multiple regression results for good performance model (B%)

Predictor	B	SE	β	*p*-value
(Intercept)	16.28	2.05	–	< 0.001
Number of teachers	3.54	0.55	0.33	< 0.001
Quiz components	0.19	0.06	0.15	0.002
Page resources	0.25	0.10	0.12	0.014
Forum components	0.48	0.22	0.11	0.029
Educational level	-3.02	1.47	-0.10	0.040

Table 5 Multiple regression results for satisfactory performance model (C-D%)

Predictor	B	SE	β	*p*-value
(Intercept)	25.37	2.19	–	< 0.001
Assignment components	0.27	0.05	0.25	< 0.001
Forum components	1.07	0.23	0.22	< 0.001
Choice components	2.85	0.76	0.18	< 0.001
SCORM components	0.89	0.29	0.15	0.002
Assignment × Forum interaction	0.06	0.02	0.14	0.003

The negative coefficient for Educational level indicates that undergraduate courses have higher proportions of A grades than graduate courses when controlling for other factors, possibly reflecting different grading standards or student preparation levels.

5.2 RQ2: Components supporting good performance (B grades)

The model for good performance reveals both similarities and differences compared to excellence predictors, explaining 26.8% of the variance (adjusted $R^2 = 0.268$, F (5,979) = 72.82, $p < 0.001$) (see Table 4).

While Number of teachers remains the strongest predictor, its coefficient is lower than for excellence (β = 0.33 vs. 0.41), suggesting diminishing returns for good versus excellent performance. The addition of Forum components as a significant predictor distinguishes good performance from excellence, indicating that communication elements play a more important role in supporting good but not excellent performance. This aligns with social constructivist principles emphasizing peer interaction for knowledge construction [27].

5.3 RQ3: Components supporting satisfactory performance (C-D grades)

Satisfactory performance shows a distinct pattern of predictors, with the model explaining 29.5% of the variance (adjusted $R^2 = 0.295$, F (5, 979) = 82.94, $p < 0.001$) (see Table 5).

Interactive components dominate the predictors for satisfactory performance. Assignment components emerge as the strongest predictor ($\beta = 0.25$), suggesting that structured practice opportunities help students develop basic competencies. Communication components (Forums) play a significant role ($\beta = 0.22$), supporting social constructivist principles.

The significant interaction between Assignment and Forum components ($\beta = 0.14$, $p = 0.003$) indicates that combining structured practice with discussion is particularly effective for supporting satisfactory performance. This interaction effect highlights the potential of intelligently combining different Moodle components to enhance learning outcomes.

Table 6 Multiple regression results for failure prevention model (F%)

Predictor	B	SE	β	*p*-value
(Intercept)	21.86	1.93	–	< 0.001
Assignment components	-0.19	0.05	-0.20	< 0.001
Resource diversity	-2.95	0.81	-0.18	< 0.001
Chat components	-3.76	1.25	-0.15	0.003
Glossary components	-1.22	0.50	-0.12	0.015
Resource diversity × Assignment	-0.04	0.02	-0.11	0.034

Table 7 Summary of canonical correlation analysis results

Function	Canonical correlation	Wilks' lambda	*p*-value	Primary pattern
1	0.58	0.43	< 0.001	Excellence orientation
2	0.46	0.64	< 0.001	Satisfactory performance
3	0.39	0.81	< 0.001	Failure prevention

5.4 RQ4: Components preventing failure (F grades)

The failure prevention model reveals protective factors, explaining 24.7% of the variance (adjusted $R^2 = 0.247$, $F\,(5, 979) = 65.35$, $p < 0.001$) (see Table 6).

Negative coefficients indicate factors that reduce failure rates. Assignment components show the strongest protective effect (β = −0.20), suggesting that structured practice opportunities help prevent failure. Resource diversity (β = −0.18) indicates that providing content in various formats helps reach diverse learners, supporting Universal Design for Learning principles [38, 52].

Communication components (Chat) are associated with lower failure rates (β = −0.15), suggesting that synchronous communication may help identify and address struggling students. The significant interaction between Resource diversity and Assignment (β = −0.11) indicates that combining diverse resources with structured practice is particularly effective at reducing failure rates.

5.5 RQ1-4: Multivariate relationships through canonical correlation analysis

Canonical correlation analysis revealed three significant canonical functions that explain relationships between Moodle components and grade distributions, providing a multivariate perspective that complements the regression findings.

Function 1 (Excellence Orientation, r_c = 0.58) is characterized by high loadings on Number of teachers (0.72), URL resources (0.45), and Page resources (0.42) in the component set, corresponding to high A% (0.87) and moderate B% (0.51) in the grade set. This pattern confirms the regression findings that information and content components primarily support excellence (see Table 7).

Function 2 (Satisfactory Performance, r_c = 0.46) shows high loadings on Assignment (0.64), Forum (0.58), and Choice components (0.47), corresponding to high C% (0.78) and D% (0.73). This validates the distinct support mechanisms for satisfactory performance identified in the regression analysis.

Function 3 (Failure Prevention, r_c = 0.39) exhibits high loadings on Resource diversity (0.65) and Assignment components (0.47), corresponding to low F% (-0.85). This confirms that diverse resources combined with structured activities effectively prevent failure.

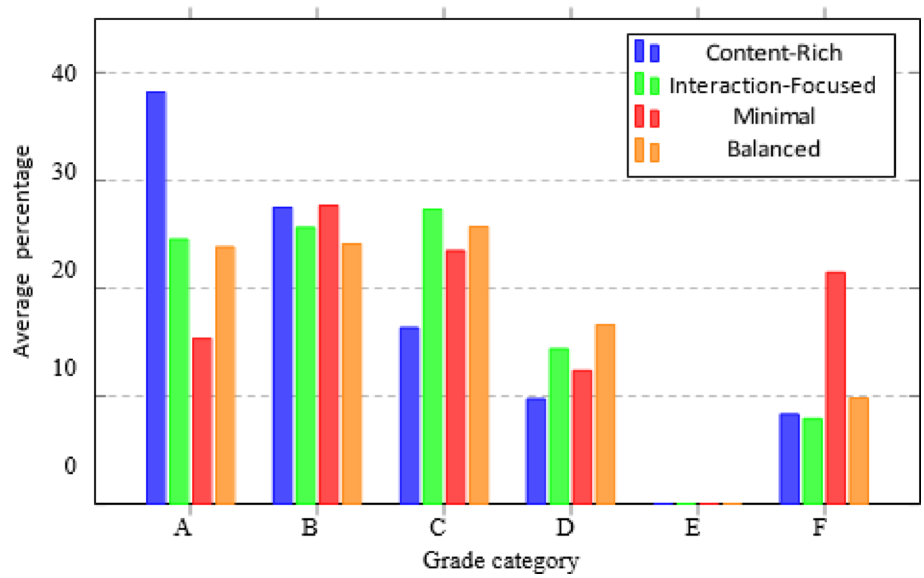

Fig. 2 Average grade distributions for each course design cluster identified through k-means analysis. Content-Rich designs promote excellence, Interaction-Focused designs support satisfactory performance, Minimal designs show higher failure rates, and Balanced designs produce even grade distributions

Table 8 Characteristics of course design clusters

Cluster	Key characteristics	% of courses
Content-Rich Design	High in Resources (especially URL, Page) Moderate in Activities (especially Quiz) Low in Communication Grade pattern: High A% (38.2%), moderate B% (27.5%), low F% (8.3%)	31%
Interaction-Focused Design	Moderate in Resources High in Activities (especially Assignment, Choice) High in Communication (especially Forum) Grade pattern: Moderate A% (24.6%), high C-D% (41.8%), low F% (7.9%)	25%
Minimal Design	Low in all components Grade pattern: Low A% (15.3%), moderate C-D% (35.7%), high F% (21.4%)	29%
Balanced Design	Moderate in Resources Moderate in Activities (diverse types) Moderate in Communication Grade pattern: Balanced distribution with slightly higher B-C% (42.5%), low F% (9.8%)	15%

5.6 RQ5: Course design patterns through cluster analysis

Cluster analysis identified four distinct course design patterns based on Moodle component usage, each associated with characteristic grade distributions.

Figure 2 visualizes the grade distributions associated with each cluster.

The Content-Rich Design cluster (31% of courses) emphasizes resource provision with moderate assessment, producing the highest proportion of excellent grades. The Interaction-Focused Design cluster (25%) prioritizes active engagement and communication, effectively supporting satisfactory performance with low failure rates. The Minimal Design cluster (29%) shows insufficient utilization of LMS components, resulting in the highest failure rate (21.4%). The Balanced Design cluster (15%) employs moderate levels of diverse components, producing relatively even grade distributions suitable for heterogeneous student populations (see Table 8).

6 Discussion

6.1 Theoretical implications

This study makes several important theoretical contributions to the learning analytics and adaptive learning literature. Our findings demonstrate that treating academic performance as a unidimensional construct obscures important relationships between course design and learning outcomes. By decomposing grade distributions and analyzing them separately, we reveal that different LMS components serve distinct pedagogical functions at different achievement levels. Excellence requires information-rich environments with multiple perspectives and formative assessment, good performance benefits from a balance of content and communication, satisfactory performance depends on structured practice with peer support, and failure prevention requires resource diversity and multiple engagement pathways. This differentiated view challenges the implicit assumption in much learning analytics research that predictors of "good performance" operate uniformly across the achievement spectrum.

Our findings provide empirical validation for several established learning theories within LMS contexts. Cognitive load theory receives support through the strong association between well-structured information components and excellence, confirming that reducing extraneous load enables deeper learning among high-achieving students. The Zone of Proximal Development framework is validated by the differential effects of activity components, with structured activities strongly predicting satisfactory performance but showing weaker relationships with excellence. Social constructivism gains empirical support through the significant interaction effects between communication and activity components, particularly for middle-achieving students.

These empirical validations strengthen the theoretical foundations of educational technology design by demonstrating how abstract principles manifest in actual LMS usage patterns. The findings suggest that effective LMS design must consider not just what components to include, but how different components support different types of learning and achievement levels.

6.2 Practical implications for course design

Our findings provide evidence-based guidance for educators and instructional designers seeking to achieve specific grade distribution objectives. For promoting excellence, courses should prioritize multiple instructor involvement to provide diverse expertise and perspectives, well-structured content resources (URLs, Pages) that support self-directed learning, and regular formative assessment through quizzes that promote metacognitive development. The Content-Rich Design cluster exemplifies this approach, achieving 38.2% A grades while maintaining low failure rates.

For supporting satisfactory performance and ensuring basic competency achievement, the emphasis shifts to interactive assignments combined with discussion forums that provide scaffolded practice with peer support, choice activities that engage students in decision-making, and SCORM packages that deliver structured interactive content. The Interaction-Focused Design cluster demonstrates this approach, effectively supporting 41.8% of students at satisfactory levels with minimal failure.

For preventing failure and supporting at-risk students, critical elements include resource diversity to accommodate different learning preferences and abilities, structured assignments that provide regular engagement checkpoints, synchronous communication tools that enable timely intervention, and glossary activities that help students master fundamental terminology. The combination of these elements, as shown in the Resource diversity × Assignment interaction, creates multiple pathways to understanding that reduce the likelihood of complete disengagement.

6.3 Implications for adaptive learning systems

Our findings enable a shift from generic adaptation to targeted strategies based on desired outcomes. Adaptive systems can now adjust course configurations based on real-time grade distribution monitoring. If excellence rates are below targets, the system can increase information component visibility and add formative assessment opportunities. If failure rates are concerning, it can diversify resource formats and increase assignment scaffolding. For balanced performance across all levels, the system can adopt the Balanced Design pattern with moderate use of all component types.

This differentiated approach to adaptation aligns with emerging research on context-aware learning environments. Rather than treating all students uniformly, adaptive systems can recognize that students at different achievement levels benefit from different types of support. The identification of specific component combinations that produce interaction effects provides additional sophistication for adaptive algorithms, suggesting that components should not be adjusted in isolation but as coordinated sets.

6.4 The challenge of heterogeneity and contextual factors

Our study reveals a fundamental tension in learning analytics research between identifying general principles and acknowledging contextual specificity. The heterogeneity in our results across faculties and course types illustrates that "best practices" in LMS design are likely context-dependent. Our aggregate findings may underestimate true effects within homogeneous contexts, as averaging across diverse courses introduces noise that attenuates relationships. Practitioners working within specific disciplines may find stronger effects than we report.

The optimal combination of Moodle components likely depends on factors we could not measure directly, including instructor pedagogical philosophy, student preparation levels, course learning objectives, and assessment approaches. Our design patterns should be considered starting points requiring local adaptation rather than rigid templates. The unmeasured heterogeneity also limits causal interpretation, as the associations between components and grade distributions may partly reflect selection effects where instructors who use certain components may also differ in grading philosophy or teaching quality.

6.5 Limitations and future research directions

Several limitations should be considered when interpreting our findings. The study was conducted at a single Ukrainian pedagogical university during 2020–2022, a period marked by pandemic-related disruptions and emergency remote teaching. Generalizability to other contexts, particularly post-pandemic normal operations, requires further investigation. While our analysis identifies significant relationships, the observational nature of our data precludes definitive causal conclusions. Our reliance on component counts rather than quality measures means that a course with many poorly-designed assignments may be less effective than one with fewer, high-quality assignments.

The absence of student-level variables such as prior achievement, motivation, and learning preferences represents a significant limitation. These individual differences likely moderate the relationships we observed and should be incorporated in future multilevel analyses. Similarly, we could not directly measure instructor quality, teaching experience, or pedagogical approaches, all of which undoubtedly influence both LMS usage patterns and grade distributions.

Future research should address these limitations through several approaches. Longitudinal studies examining how component-grade relationships evolve over time would provide insights into dynamic adaptation needs. Experimental interventions testing whether adopting identified design patterns causally impacts grade distributions would strengthen causal claims. Cross-cultural validation examining whether patterns hold across different educational systems would enhance generalizability. Quality-focused analyses incorporating rubrics or engagement metrics would provide more nuanced understanding of effective implementation. Multilevel modeling examining how student characteristics interact with course design would reveal important moderating effects.

7 Conclusion

This study has examined the relationships between Moodle components and specific grade distributions across 985 courses, providing valuable insights for developing context-aware adaptive learning environments. Through multivariate analysis techniques, we identified distinct patterns of association between course design elements and different grade categories, moving beyond traditional approaches that treat assessment as a unidimensional construct.

Our findings reveal that excellence is strongly associated with Information components and Content-focused elements, satisfactory performance is primarily predicted by interactive components and their synergistic combinations, and failure prevention is supported by resource diversity combined with structured activities. The identification of four distinct course design patterns with characteristic grade distributions provides practical templates for achieving specific pedagogical objectives.

These differentiated relationships provide a foundation for more nuanced adaptive learning approaches that can target specific educational objectives rather than applying generic adaptations. By demonstrating that different achievement levels require different types of support, this research contributes to both theoretical understanding and practical application of learning analytics in higher education.

The study advances learning analytics theory by empirically validating learning theories in LMS contexts, revealing the importance of component interactions, and demonstrating the value of analyzing grade distributions rather than aggregate performance. Practically, it provides evidence-based guidance for course design, templates for different pedagogical objectives, and a framework for implementing context-aware adaptations.

As educational technology continues to evolve, this nuanced understanding of how course design elements differentially impact various aspects of student performance will be essential for creating truly adaptive, inclusive, and effective learning environments. The shift from analyzing aggregate performance to understanding grade distributions represents more than a methodological choice – it reflects a philosophical commitment to recognizing and supporting the full diversity of student achievement.

Acknowledgements
The authors would like to express their sincere gratitude to Kryvyi Rih State Pedagogical University for providing access to the educational data used in this research. We acknowledge the support of the KSPU rector and ethics committee who approved this research. We are grateful to the educational technology specialists and administrators who helped with data collection and anonymization procedures. Special thanks to all faculty members whose courses were included in this analysis for their contribution to digital education development. We also appreciate the technical support provided by the University's IT department in extracting and processing the Moodle data. Finally, we thank the anonymous reviewers for their valuable feedback that helped improve the quality of this manuscript.

Author contributions
SOS conceptualized the study, developed the theoretical framework, and wrote the main manuscript text. PPN contributed to the methodological approaches and data collection process. TAV contributed to the discussion section and interpretation of findings. ISM performed the statistical analysis and co-developed the theoretical framework. LOF processed the data and prepared the visualizations. All authors reviewed, edited, and approved the final manuscript.

Funding
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability
The dataset analysed during the current study is available in the Zenodo repository, https:/doi.org/10.5281/zenodo.10938019 .

Declarations

Ethics approval and consent to participate
This research was conducted in accordance with ethical standards for educational data analysis. The study utilized anonymized course data from Kryvyi Rih State Pedagogical University's Learning Management System. The need for ethical approval and informed consent was waived by the ethics committee of Kryvyi Rih State Pedagogical University on January 26, 2024, due to the retrospective nature of the study involving secondary analysis of anonymized institutional data. No personally identifiable information was included in the dataset, ensuring privacy protection for all students and instructors whose courses were analyzed. All data was processed and stored in compliance with applicable data protection regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 13 May 2025 / Accepted: 28 November 2025
Published online: 06 December 2025

References

1. Wang Y, Lai Y, Huang X. Innovations in online learning analytics: A review of recent research and emerging trends. IEEE Access. 2024;12:166761–66775. https://doi.org/10.1109/ACCESS.2024.3493621.
2. Alnasyan B, Basheri M, Alassafi M. The power of deep learning techniques for predicting student performance in virtual learning environments: A systematic literature review. Computers Education: Artif Intell. 2024;6:100231. https://doi.org/10.1016/j.caeai.2024.100231.
3. Hasanov A, Laine TH, Chung TS. A survey of adaptive context-aware learning environments. J Ambient Intell Smart Environ. 2019;11(5):403–28. https://doi.org/10.3233/AIS-190534.
4. Yu R, Yao T, Bai F. Educational big data analytics for futuristic smart learning using deep learning techniques. Scalable Comput. 2024;25(4):2728–35. https://doi.org/10.12694/scpe.v25i4.2988.
5. Sergeev KA, Mironenko OI, Krivich OY, Petrov AA, Kozlov MV. Using the module analytics and machine learning in LMS moodle at training students of specialty rolling stock. AIP Conference Proceedings. 2023 12;2624(1):050056. https://doi.org/10.1063/5.0133923
6. Abuzinadah N, Umer M, Ishaq A, Hejaili AA, Alsubai S, Eshmawi AA, et al. Role of convolutional features and machine learning for predicting student academic performance from MOODLE data. PLoS ONE. 2023;18:e0293061. https://doi.org/10.1371/journal.pone.0293061.
7. Fadieieva L, Semerikov S et al. Exploring the Interplay of Moodle Tools and Student Learning Outcomes: A Composite-Based Structural Equation Modelling Approach. In: Faure E, Tryus Y, Vartiainen T, Danchenko O, Bondarenko M, Bazilo C, editors. Information Technology for Education, Science, and Technics. vol. 222 of Lecture Notes on Data Engineering and Communications Technologies. Cham: Springer Nature Switzerland; 2024. pp. 418–435. Available from: https://doi.org/10.1007/978-3-031-71804-5_28
8. Semerikov SO, Nechypurenko PP, Vakaliuk TA, Mintii IS, Fadieieva LO. Differential effects of moodle course design on student subpopulations: advancing personalized learning in higher education. Smart Learn Environ. 2025;12(1):46. https://doi.org/10.1186/s40561-025-00400-6.
9. Aziz A, Khoukhi F. An Ontological Approach for Service Adaptation in Context Aware Environment. In: Ezziyyani M, editor. Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). vol. 1102 of Advances in Intelligent Systems and Computing. Cham: Springer International Publishing; 2020. pp. 91–100. Available from: https://doi.org/10.1007/978-3-030-36653-7_9

10. Ramaswami G, Susnjak T, Mathrani A. Supporting students' academic performance using explainable machine learning with automated prescriptive analytics. Big Data Cogn Comput. 2022;6(4):105. https://doi.org/10.3390/bdcc6040105.
11. Moodle. Philosophy. Available from: https://web.archive.org/web/20060901040155/http://docs.moodle.org/en/Philosophy
12. Pascual-Leone J. Piaget as a Pioneer of Dialectical Constructivism: Seeking Dynamic Processes for Human Science. In: Martí E, Rodríguez C, editors. After Piaget. New York: Routledge; 2017. Available from: https://doi.org/10.4324/9781315082899
13. Solomon C, Harvey B, Kahn K, Lieberman H, Miller ML, Minsky M, et al. History of logo. Proc ACM Program Lang. 2020;4(HOPL):79. https://doi.org/10.1145/3386329.
14. Prawat RS, Mind. Cult Activity. 1999;6(4):255–73. https://doi.org/10.1080/10749039909524731.
15. Harrer A, Moskaliuk J, Kimmerle J, Cress U. Visualizing wiki-supported knowledge building: co-evolution of individual and collective knowledge. In: Proceedings of the 4th International Symposium on Wikis. WikiSym '08. New York, NY, USA: Association for Computing Machinery; 2008. p. 19. Available from: https://doi.org/10.1145/1822258.1822284
16. Ramos E, Ramirez M, Hern'Ndez J, Morales L, Garcia M, Rodriguez J. Toward supporting constructivist instruction through CSCL activities embedded in LMSs. In: Proceedings – 8th International Conference on Intelligent Environments, IE 2012; 2012. pp. 259–266. Available from: https://doi.org/10.1109/IE.2012.36
17. Gnadlinger F, Selmanagić A, Simbeck K, Kriglstein S. Adapting Is Difficult! Introducing a Generic Adaptive Learning Framework for Learner Modeling and Task Recommendation Based on Dynamic Bayesian Networks. In: Proceedings of the 15th International Conference on Computer Supported Education - Volume 1: CSEDU. INSTICC. SciTePress; 2023. pp. 272–280. Available from: https://doi.org/10.5220/0011964700003470
18. Jordan C. Comparison of international baccalaureate (IB) chemistry students' preferred vs actual experience with a constructivist style of learning in a moodle e-learning environment. Int J Lesson Learn Stud. 2013 jun;2(2):155–67. https://doi.org/10.1108/20468251311323397.
19. Liu R, Shi C. Investigating Collaborative Learning Effect in Blended Learning Environment by Utilizing Moodle and WeChat. In: Cheung SKS, Kwok Lf, Shang J, Wang A, Kwan R, editors. Blended Learning: Aligning Theory with Practices. Cham: Springer International Publishing; 2016. pp. 3–13. Available from: https://doi.org/10.1007/978-3-319-41165-1_1
20. Mirata V, Bergamin P. Developing an implementation framework for adaptive learning: A case study approach. In: Proceedings of the European Conference on e-Learning, ECEL. vol. 2019-November; 2019. pp. 668–673. Available from: https://doi.org/10.34190/EEL.19.024
21. Hilali K, Chergui M, Ammoumou A. Adaptive Learning Systems: A Comprehen- sive Overview and Identification of Challenges. In: 2023 IEEE International Conference on Technology Management, Operations and Decisions, ICTMOD 2023; 2023. Available from: https://doi.org/10.1109/ICTMOD59086.2023.10438128
22. Rienties B, Toetenel L, Bryan A. Scaling up learning design: impact of learning design activities on LMS behavior and performance. In: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. LAK'15. New York, NY, USA: Association for Computing Machinery; 2015. pp. 315–319. Available from: https://doi.org/10.1145/2723576.2723600
23. López-Pernas S, Saqr M. From Variables to States to Trajectories (VaSSTra): A Method for Modelling the Longitudinal Dynamics of Learning and Behaviour. In: García-Pen˜alvo FJ, García-Holgado A, editors. Proceedings TEEM 2022: Tenth International Conference on Technological Ecosystems for Enhancing Multiculturality. Lecture Notes in Educational Technology. Singapore: Springer Nature Singapore; 2023. pp. 1169–1178. Available from: https://doi.org/10.1007/978-981-99-0942-1_123
24. Uwimpuhwe G, Singh A, Higgins S, Coux M, Xiao Z, Shkedy Z, et al. Latent class evaluation in educational trials: what percentage of children benefits from an intervention? J Exp Educ. 2022;90(2):404–18. https://doi.org/10.1080/00220973.2020.1767021.
25. Van der Velde M, Sense F, Borst JP, Rijn H. Large-scale evaluation of cold-start mitigation in adaptive fact learning: knowing what matters more than knowing who. User Model User-Adapt Interact. 2024;34(4):1467–91. https://doi.org/10.1007/s11257-024-09401-5.
26. Lawrence R, Foss S, Urazova T. Evaluation of submission limits and regression penalties to improve student behavior with automatic assessment systems. ACM Trans Comput Educ. 2023;23(3):31. https://doi.org/10.1145/3591210.
27. Rienties B, Nguyen Q, Holmes W, Reedy K. (2017) A review of ten years of implementation and research in aligning learning design with learning analytics at the open university UK. Interact Des Archit. 33:134–54.
28. Wang H, Huang T, Zhao Y, Hu S. The impact of dashboard feedback type on learning Effectiveness, focusing on learner differences. Sustainability. 2023;15(5):4474. https://doi.org/10.3390/su15054474.
29. Muñoz JLR, Ojeda FM, Jurado DLA, Peña PFP, Carranza CPM, Berríos HQ, et al. Systematic review of adaptive learning technology for learning in higher education. Eurasian J Educational Res. 2022;2022(98):221–33. https://doi.org/10.14689/ejer.2022.98.014.
30. Sihabudin Sahid DS, Nugroho LE, Santosa PI. A model of personalized context aware E-learning based on psychological experience. J Telecommunication Electron Comput Eng. 2018;10(1–5):137–43. https://jtec.utem.edu.my/jtec/article/view/3645.
31. Salazar OM, Ovalle DA, Duque ND et al. Incorporating Context-Awareness Services in Adaptive U-MAS Learning Environments. In: Bajo J, Hallenborg K, Pawlewski P, Botti V, Sánchez-Pi N, Duque Méndez ND, editors. Highlights of Practical Applications of Agents, Multi-Agent Systems, and Sustainability - The PAAMS Collection. vol. 524 of Communications in Computer and Information Science. Cham: Springer International Publishing; 2015. pp. 331–339. Available from: https://doi.org/10.1007/978-3-319-19033-4_28
32. Sweller J. CHAPTER TWO - Cognitive Load Theory. In: Mestre JP, Ross BH, editors. Psychology of Learning and Motivation. vol. 55. Academic Press; 2011. pp. 37–76. Available from: https://doi.org/10.1016/B978-0-12-387691-1.00002-8
33. Zimmerman BJ. Chapter 2 - Attaining Self-Regulation: A Social Cognitive Perspective. In: Boekaerts M, Pintrich PR, Zeidner M, editors. Handbook of Self-Regulation. San Diego: Academic Press; 2000. pp. 13–39. Available from: https://doi.org/10.1016/B978-012109890-2/50031-7
34. Marton F, Säljö R. On qualitative differences in learning: I—Outcome and process. Br J Educ Psychol. 1976;46(1):4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x.
35. Richardson M, Abraham C, Bond R. Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychol Bull. 2012;138(2):353–87. https://doi.org/10.1037/a0026838.

36. Takase M, Yoshida I. The relationships between the types of learning approaches used by undergraduate nursing students and their academic achievement: A systematic review and meta-analysis. J Prof Nurs. 2021;37(5):836–45. https://doi.org/10.1016/j.profnurs.2021.06.005.
37. Biggs J, Tang C. Teaching for quality learning at university: what the student does. 4th ed. Maidenhead: Society for Research into Higher Education & Open University; 2011.
38. Alajlani N, Crabb M, Murray I. A systematic review in Understanding stakeholders' role in developing adaptive learning systems. J Computers Educ. 2024;11(3):901–20. https://doi.org/10.1007/s40692-023-00283-x.
39. Black P, Wiliam D. Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan. 2010;92(1):81–90. https://doi.org/10.1177/003172171009200119.
40. Bandura A. Social Learning Theory. Englewood Cliffs, NJ: Prentice-Hall; 1977. Available from: https://www.asecib.ase.ro/mps/Bandura_SocialLearningTheory.pdf
41. Vygotsky LS. Mind in Society: Development of Higher Psychological Processes. Harvard University Press; 1978. Available from:https://w.pauldowling.me/rtf/2021.1/readings/LSVygotsky_1978_MindinSocietyDevelopmentofHigherPsycholo.pdf
42. Webb NM. Peer interaction and learning in small groups. Int J Educational Res. 1989;13(1):21–39. https://doi.org/10.1016/0883-0355.
43. Johnson DW, Johnson RT. Making cooperative learning work. Theory into Pract. 1999;38(2):67–73. https://doi.org/10.1080/00405849909543834.
44. Wang Y, Huang X, Yu Q, Lai Y. Emerging Techniques for Online Learning Analytics. In: Cheung SKS, Wang FL, Paoprasert N, Charnsethikul P, Li KC, Phusavat K, editors. Technology in Education. Innovative Practices for the New Normal. vol. 1974 of Communications in Computer and Information Science. Singapore: Springer Nature Singapore; 2024. pp. 109–118. Available from: https://doi.org/10.1007/978-981-99-8255-4_10
45. Tepgec¸ M, Ifenthaler D. learning analytics based interventions: a systematic review of experimental studies. In: Sampson DG, Ifenthaler D, Isaías P, editors. Proceedings of the 19th International Conference on Cognition and Exploratory Learning in the Digital Age, CELDA 2022. IADIS Press; 2022. pp. 327–330. Available from: https://www.iadisportal.org/digital-library/learning-analytics-based-interventions-a-systematic-review-of-experimental-studies
46. Fadieieva L, Semerikov S. KSPU Moodle activities and marks 2020–2022. Zenodo. Available from: https://doi.org/10.5281/zenodo.10938019
47. Asriadi AMM, Hadi S, Istiyono E, Retnawati H. Does differentiated instruction affect learning outcome? Systematic review and meta-analysis. J Peda- Gogical Res. 2023;7(5):18–33. https://doi.org/10.33902/JPR.202322021.
48. Josephine N, Cedric A, McDaniel ST, Kanagwa B, Joseph T, Marvin G. Interpretable Machine Learning Techniques for Students Grade Prediction. In: Proceedings of the 2024 Sixteenth International Conference on Contemporary Computing. IC3-2024. New York, NY, USA: Association for Computing Machinery; 2024. pp. 213–225. Available from: https://doi.org/10.1145/3675888.3676055
49. Willis JE, Slade S, Prinsloo P. Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. Education Tech Research Dev. 2016;64(5):881–901. https://doi.org/10.1007/s11423-016-9463-4.
50. Wang P, Zhao P, Li Y, Design of Education Information Platform on Education Big Data Visualization. Wirel Commun Mob Comput. 2022;2022:6779105. https://doi.org/10.1155/2022/6779105.
51. Fernández Alemán JL. Automated assessment in a programming tools course. IEEE Trans Educ. 2011;54(4):576–81. https://doi.org/10.1109/TE.2010.2098442.
52. Barna OV, Kuzminska OH, Semerikov SO. Enhancing digital competence through STEM-integrated universal design for learning: a pedagogical framework for computer science education in Ukrainian secondary schools. Discover Educ. 2025;4(1):357. https://doi.org/10.1007/s44217-025-00821-y.

Publisher's note