Associate of Digital Learning Behaviours and Academic Performance in Blended Learning in Higher Education environment: A Systematic Review from a Learning Analytics Perspective

Sun Wenrui, Tajularipin Sulaiman, Mohd Mokhtar Muhamad

Abstract - Blended learning has been widely used in the field of higher education, and learning analysis technology can quantify students' digital learning behaviour through learning management systems. However, there is still a lack of a comprehensive and systematic review of the correlation between such quantitative behaviour and academic performance. This systematic review was conducted following the PRISMA 2020 statement to : 1) identify and sort out the most commonly studied digital learning behaviours in the hybrid learning environment of higher education that can be reliably quantified through learning analysis technology; 2) integrate these quantitative digital behaviours with the results of students' academic performance. Empirical evidence of the relevant; 3) Classify and evaluate the learning analysis techniques used in existing research, including its reporting advantages, practical limitations and typical application scenarios; 4) Critically evaluate the methodological limitations in the current literature, and put forward the priority direction of future research based on evidence. Studies were retrieved from Scopus and Web of Science. A total of 383 records were initially identified, and 33 studies were finally included after screening. This review synthesizes high-quality evidence to clarify the relationships between digital learning behaviour s and academic performance. The findings will provide evidence-based insights for optimizing blended instructional design, implementing targeted academic interventions, and guiding future rigorous learning analytics research in higher education.

Sun Wenrui, University Putra Malaysia (Email address: GS71642@student.upm.edu.my).

Tajularipin Sulaiman, University Putra Malaysia (Email address: tajulas@upm.edu.my).

Mohd Mokhtar Muhamad, University Putra Malaysia (Email address: mk_mokhtar@upm.edu.my).

Keywords - Blended Learning; Higher Education; Digital Learning Behaviour s; Academic Performance; Learning Analytics; Systematic Review

I. INTRODUCTION

The digital transformation of higher education is in progress, and blended learning is already a mainstream instructional model that combines online and offline teaching components (Müller & Mildenberger, 2021). Blended learning generates a variety of digital learning behaviour s reflecting students' engagement and self-regulatory skills, which greatly influence their academic performance, and is supported by learning management systems (LMS) and adaptive learning platforms (Ali & Hanna, 2021). Learning analytics allows practitioners to quantify, analyze, and predict this behaviour al patterns to support the delivery of personalized learning interventions (González-Nucamendi et al., 2022). While numerous empirical studies have explored digital learning behaviour s and academic performance via learning analytics, the literature remains fragmented, with inconsistent indicators, methods, and findings. This systematic review fills these gaps by synthesizing a decade of research from a learning analytics perspective, to clarifycore behaviour al dimensions, behaviour –performance relationships, applicable analytics techniques, and future research directions.

This study addresses four research questions:

RQ1. What are the most frequently examined learning analytics–quantifiable digital learning behaviour s in higher education blended learning contexts?

RQ2. How are digital learning behaviour s associated with academic performance in higher education blended learning environments?

RQ3. What learning analytics techniques are employed to examine digital learning behaviour s, and what are their reported strengths and limitations?

RQ4. What methodological limitations and priority directions for future research are identified in the existing literature?

II. MATERIALS AND METHODS

This systematic review followed the PRISMA 2020 statement to ensure rigor and reproducibility (Page et al., 2021).

Search Strategy

Literature searches were performed in Web of Science (SSCI/ESCI) and Scopus. The search string was:("digital learning behaviour *" OR "online learning behaviour *" OR "learning log*" OR "log data" OR "LMS data" OR "learning analytics" OR "educational data mining" OR "clickstream data" OR "trace data" OR "learning management system*" OR Moodle OR Blackboard) AND ("academic performance" OR "student performance" OR "academic achievement*" OR grade* OR GPA OR "learning outcome*") AND ("blended learning" OR "hybrid learning") AND ("higher education" OR universit* OR college OR "tertiary education"). The publication period was limited to 2017–2026.

Inclusion Criteria

Studies were included if they focused on digital learning behaviour s, learning analytics, and their links with academic performance in higher education blended learning; were empirical quantitative, qualitative, or mixed-methods studies; were peer-reviewed journal articles indexed in Web of Science or Scopus; were written in English; had full-text available; and were published between 2017 and 2026.

Exclusion Criteria

Studies were excluded if they were book chapters, conference papers, dissertations, or gray literature; were review articles such as systematic reviews, meta-analyses, or scoping reviews; did not focus on digital learning behaviour s, learning analytics, or academic performance; were conducted in non-blended learning environments; focused on K-12 or non-higher education populations; were not written in English; were duplicated across databases; were preprints or published outside the 2017–2026 time frame; and did not apply learning analytics techniques.

Study Selection

The initial search retrieved 383 records. After removing duplicates, three researchers independently screened titles, abstracts, and full texts. Discrepancies were resolved via discussion. Finally, 33 studies were included. The screening process is shown in the PRISMA 2020 flow diagram (Figure 1).

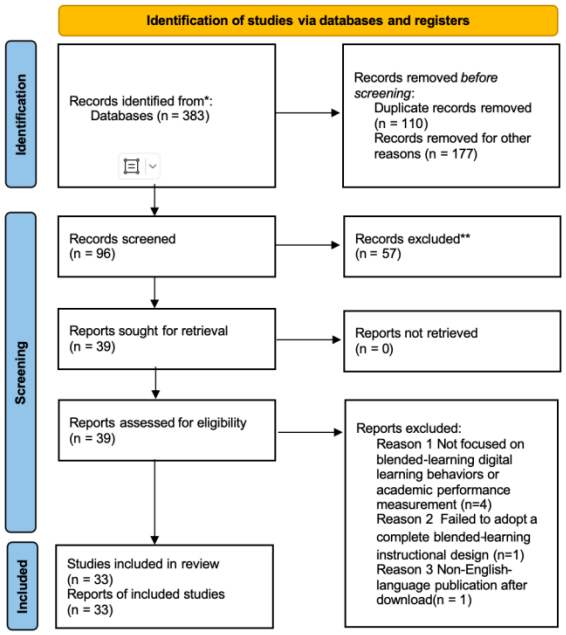


Figure 1. PRISMA 2020 flow diagram

Data Extraction

A standardized form was used to extract study details, quantifiable digital learning behaviour s, behaviour – performance associations, learning analytics techniques, limitations, and future directions.

III. RESULTS AND DISCUSSION

Frequently quantified digital learning behaviour s (RQ1)

Across the 33 included studies, four categories of learning analytics–quantifiable digital learning behaviour s emerged as the most frequently examined constructs in blended higher education. Task completion behaviour s were reported in 28 out of 33 studies, with common indicators including assignment submission rate, timely submission, quiz completion, task completion ratio, and learning objective attainment. Interactive engagement behaviour s were documented in 22 studies, covering forum participation, peer collaboration, instructor–student feedback exchange, resource access, and active tool interaction. In 19 studies, temporal usage behaviour s were studied, which were measured by login frequency, access duration, time distribution, learning pace, and time of procrastination. In 15 studies, self-regulated or problematic learning behaviour s were studied, with indicators such as academic procrastination, self-monitoring via dashboard, cramming, and long-term inactivity as a proxy for dropout risk

Associations Between Behaviour s and Academic Performance (RQ2)

All 33 studies found significant associations between digital learning behaviour s and academic outcomes. The direction of the associations was consistent, but the effect size differed for different types of behaviour s. Task completion was the strongest and most stable positive correlate. Higher and more timely task completion was correlated with better grades and lower likelihood of failure or dropout. Interactive use demonstrated moderate positive associations, especially for meaningful interactions such as forum discussions and feedback seeking. Passive browsing of resources showed weak or non-significant associations. Temporal use indicators demonstrated weak and inconsistent correlations. The total number of logins or total access time rarely predicted performance. Only consistent timing added small explanatory value. Procrastination and low engagement were strong predictors of lower grades. Irregular engagement also predicted higher risk of dropping out. Across studies, quality of engagement (e.g. task completion, purposeful interaction) was more consistently predictive of performance than quantity of engagement (e.g. time on task, click counts). Most studies were cross-sectional designs, so causal conclusions cannot be drawn. It is unclear whether positive behaviour s lead to better performance or whether higher-ability students simply engage in more effective behaviour s. Most analyses also did not control for confounding factors such as prior grades, self-efficacy and course difficulty. This may have led to overestimates of the independent effects of digital learning behaviour s.

Strengths and Limitations of Learning Analytics Techniques (RQ3)

These 33 studies widely adopted four major learning analytics techniques. 19 studies have used the supervised machine learning for predicting the student grades and identifying the learners at risk of underperforming. This technique has high prediction accuracy but is not interpretable and requires a large amount of labeled data. 11 studies have used clustering and other unsupervised techniques for analyzing the student behaviour al patterns. These techniques enable researchers to discover hidden learning trends without labeled data, but the results may be inconsistent across different courses. 9 studies have employed traditional statistical approaches including correlation, regression and structural equation modeling. These methods are easy to interpret and align well with educational theories, but they have limited ability to model non-linear learning relationships. Only six papers in this review integrated multiple sources of data. Merging different datasets involves complicated processing steps. This complexity stops the method from being widely used in real teaching settings. Still, past research supports that combining LMS activity logs with exam marks or survey responses can make study findings more valid.

Most research in this field prioritises predictive accuracy at the cost of model interpretability. Only a small portion of these models fit the daily demands of classroom teachers. Many projects do not run validation tests across separate courses or schools, so it is uncertain if these models work reliably in varied educational contexts. The most prominent issue is the large disconnect between academic research and

classroom practice. Data-driven findings hardly lead to tangible adjustments in routine teaching.

Methodological Limitations and Future Directions (RQ4)

This review examined all thirty-three included papers and found flaws in research methods. Twenty-seven of these studies only used log data from learning management systems and left out offline learning activities, students' personal traits and different learning conditions, which weakens the explanatory power of their research results. Another 24 studies only collected data from a single course or a single school, so it was found that it was difficult to be widely used in disciplines and institutions with different cultural backgrounds. The analytical models constructed by 21 studies are only applicable to students with learning difficulties. These models cannot generate reliable predictions for intermediate and superior learners. The lack of real classroom intervention tests in 30 studies means that all predictive data have not been converted into actual teaching plans or student support solutions. At the same time, each paper does not use unified rules to define and measure students' learning behaviour, which makes it difficult to compare different research results horizontally, and the credibility of the overall comprehensive research conclusions is reduced.

What this points to is a clear research agenda. Future work must combine multi-modal, multi-source data to improve measurement validity and completeness. There is an urgent need for interpretable learning analytics tools that teachers can genuinely use in real classrooms, not just in controlled studies. Generalizability will only improve if validation is carried out across diverse courses and institutions. And without intervention-focused studies, learning analytics will remain disconnected from teaching practice.

The deeper problem is one of orientation: the field is still more interested in describing and predicting outcomes than in developing theoretical understanding. Very few studies explore the causal or theoretical mechanisms through which digital learning behaviours influence academic achievement. Moving forward, blended learning research must go beyond cataloguing behavioural patterns to build frameworks that are both theoretically grounded and practically deployable.

IV. CONCLUSION

To unpack digital learning behaviours quantifiable through learning analytics and their correlations with academic performance in blended postsecondary education, we synthesised data drawn from 33 relevant empirical studies in this systematic review. Our collective results yielded four primary classifications of these online learning behaviours. This study first found steady connections between students' repeated online actions and their grades. It sorted out common learning analytics tools used in past papers. The work also listed each tool's advantages and real-world drawbacks. Lastly, it pointed out repeated method problems seen in most published research.

This review organises a complete set of empirical evidence. The organised results help education researchers measure students' digital learning participation and figure out how such participation affects learners' academic performance.

This work also has two primary limitations. It only selects English journal articles published from 2017 to 2026. High-quality studies written in other languages are not included in the analysis. Besides, the paper fails to conduct full comparisons of contextual factors across different studies. These factors can alter the connection between students' learning behaviours and their academic performance.

When researchers carry out their work, they can first integrate multiple data sources. The analytical framework constructed by this method is clearer and easier to understand. Cross-situational tests should also be conducted for the same research direction, which can ensure stable tracking of behavioural trends in different learning environments. Future research needs to transform abstract data indicators into actionable classroom teaching practises. Such continuous improvement can gradually provide stronger learning support for students and improve academic performance in a hybrid learning environment.

REFERENCES

Pecuchova, J., & Drlik, M. (2024). Enhancing the early student dropout prediction model through clustering analysis of students' digital traces. *IEEE Access*, 12,

159336–159367. https://doi.org/10.1109/ACCESS.2024.3486762

Akram, A., Fu, C., Li, Y., Javed, M. Y., Lin, R., Jiang, Y., & Tang, Y. (2019). Predicting students' academic procrastination in blended learning course using homework submission data. *IEEE Access*, 7, 102487–102498. https://doi.org/10.1109/ACCESS.2019.2930867

Ali, A., & Hanna, W. (2021). Predicting Students' Achievement in a Hybrid Environment Through Self-Regulated Learning, Log Data, and Course Engagement: A Data Mining Approach. *Journal of Educational Computing Research*, 60, 960 - 985. https://doi.org/10.1177/07356331211056178.

Alwadei, F. H., Brown, B. P., Alwadei, S. H., Harris, I. B., & Alwadei, A. H. (2023). The utility of adaptive eLearning data in predicting dental students' learning performance in a blended learning course. *International Journal of Medical Education*, 14, 137–144. https://doi.org/10.5116/ijme.64f6.e3db

Al-Zawqari, A., Peumans, D., & Vandersteen, G. (2022). A flexible feature selection approach for predicting students' academic performance in online courses. *Computers and Education: Artificial Intelligence,* 3, Article 100103. https://doi.org/10.1016/j.caeai.2022.100103

Cagliero, L., Canale, L., Farinetti, L., Baralis, E., & Venuto, E. (2021). Predicting student academic performance by means of associative classification. *Applied Sciences*, 11(4), Article 1420. https://doi.org/10.3390/app11041420

Chango, W., Cerezo, R., & Romero, C. (2021). Multi-source and multimodal data fusion for predicting academic performance in blended learning university courses. *Computers and Electrical Engineering*, 89, Article 106908. https://doi.org/10.1016/j.compeleceng.2020.106908

Chen, L.-Q., Wu, M.-T., Pan, L.-F., & Zheng, R.-B. (2021). Grade prediction in blended learning using multisource data. *Scientific Programming*, 2021, Article 4513610. https://doi.org/10.1155/2021/4513610

Chiu, M.-S., & Hsiao, Y. P. A. (2023). Process indicators for grading group essays: Learning analytics of assessment data and online behaviour. *Athens Journal of Education,* 10(4), 575–592. https://doi.org/10.30958/aje.10-4-2

Deeva, G., De Smedt, J., Saint-Pierre, C., Weber, R., & De Weerdt, J. (2022). Predicting student performance using sequence classification with time-based windows. *Expert Systems with Applications*, 209, Article 118182. https://doi.org/10.1016/j.eswa.2022.118182

Fahd, K., Miah, S. J., & Ahmed, K. (2025). Predicting student performance in a blended learning environment using learning management system interaction data. *Applied Computing and Informatics*, 21(3/4), 220–231. https://doi.org/10.1108/ACI-06-2021-0150

González-Nucamendi, A., Noguez, J., Neri, L., Robledo-Rella, V., García-Castelán, R., & Escobar-Castillejos, D. (2022). Learning Analytics to Determine Profile Dimensions of Students Associated with their Academic Performance. *Applied Sciences*. https://doi.org/10.3390/app122010560.

Green, R. A., Whitburn, L. Y., Zacharias, A., Byrne, G., & Hughes, D. L. (2018). The relationship between student engagement with online content and achievement in a blended learning anatomy course. *Anatomical Sciences Education*, 11(5), 471–477. https://doi.org/10.1002/ase.1761

Hamadneh, N. N., Atawneh, S., Khan, W. A., Almejalli, K. A., & Alhomoud, A. (2022). Using artificial intelligence to predict students' academic performance in blended learning. Sustainability, 14(18), Article 11642. https://doi.org/10.3390/su141811642

Han, F., & Ellis, R. A. (2023). The relations between self-reported perceptions of learning environment, observational learning strategies, and academic outcome. Journal of Computing in Higher Education, 35(1), 111–125. https://doi.org/10.1007/s12528-022-09333-2

Han, F., Pardo, A., & Ellis, R. A. (2020). Students' self-report and observed learning orientations in blended university course design: How are they related to each other and to academic performance? *Journal of Computer Assisted Learning*, 36(6), 969–980. https://doi.org/10.1111/jcal.12453

Hsu, W.-S. (2025). Analyzing student engagement and learning outcomes in a gamified blended cybersecurity course using LMS-based behaviour al data.

International Journal of Engineering Pedagogy, 15(6), 111–122. https://doi.org/10.3991/ijep.v15i6.57015

Jayashanka, R., Hettiarachchi, E., & Hewagamage, K. P. (2022). Technology enhanced learning analytics dashboard in higher education. Electronic Journal of e-Learning, 20(2), 151–170. https://doi.org/10.34190/ejel.20.2.2189

Kohnke, L., Foung, D., & Chen, J. (2022). Using learner analytics to explore the potential contribution of multimodal formative assessment to academic success in higher education. SAGE Open, 12(2), Article 21582440221089957. https://doi.org/10.1177/21582440221089957

Le, B., Lawrie, G. A., & Wang, J. T. H. (2022). Student self-perception on digital literacy in STEM blended learning environments. *Journal of Science Education and Technology*, 31(3), 303–321. https://doi.org/10.1007/s10956-022-09956-1

Manzanares, M. C. S., Llamazares, M. C. E., & González, Á. A. (2020). Effectiveness of blended learning in nursing education. *International Journal of Environmental Research and Public Health*, 17(5), Article 1589. https://doi.org/10.3390/ijerph17051589

Masibigiri, P., Ntsime, B. P., & Mampana, M. (2026). The impact of learning management system on academic performance in a statistics module within a hybrid learning environment in Gauteng, South Africa. Discover Education, 5(1), Article 876. https://doi.org/10.1007/s44217-025-00876-x

Matcha, W., Gasevic, D., Uzir, N. A., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Pérez-Sanagustín, M., & Tsai, Y. S. (2020). Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics*, 7(2), 45–71. https://doi.org/10.18608/jla.2020.72.3

Müller, C., & Mildenberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: a systematic review of blended learning in higher education. *Educational Research Review*, 34, 100394. https://doi.org/10.1016/j.edurev.2021.100394.

Mwalumbwe, I., & Mtebe, J. S. (2017). Using learning analytics to predict students' performance in moodle learning management system: A case of Mbeya University of science and technology. Electronic *Journal of Information Systems in Developing Countries*, 79(1), 1–13.

Nuankaew, P. (2022). Self-regulated learning model in educational data mining. *International Journal of Emerging Technologies in Learning*, 17(17), 4–27. https://doi.org/10.3991/ijet.v17i17.23623

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS medicine,* 18(3), e1003583. https://doi.org/10.1371/journal.pmed.1003583

Pardo, A., Han, F., & Ellis, R. A. (2017). Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance. *IEEE Transactions on Learning Technologies*, 10(1), 82–92. https://doi.org/10.1109/TLT.2016.2639508

Phong, P. D., Lan, P. T., & Thanh, T. X. (2025). Linguistic summarization and outlier detection of blended learning data. *Applied Sciences*, 15(12), Article 6644. https://doi.org/10.3390/app15126644

Regueras, L. M., Verdú, M. J., de Castro, J. P., & Álvarez-Álvarez, S. (2025). Techno-pedagogical approaches and academic performance: A quantitative study based on LMS log data. *Education Sciences*, 15(11), Article 1533. https://doi.org/10.3390/educsci15111533

Regueras, L. M., Verdu, M. J., De Castro, J.-P., & Verdu, E. (2019). Clustering analysis for automatic certification of LMS strategies in a university virtual campus. *IEEE Access*, 7, 137680–137690. https://doi.org/10.1109/ACCESS.2019.2943212

Saqr, M., & Alamro, A. (2019). The role of social network analysis as a learning analytics tool in online problem based learning. *BMC Medical Education*, 19(1), Article 1599. https://doi.org/10.1186/s12909-019-1599-6

Saqr, M., Fors, U., & Tedre, M. (2018). How the study of online collaborative learning can guide teachers and predict students' performance in a medical course.

BMC Medical Education, 18(1), Article 1126. https://doi.org/10.1186/s12909-018-1126-1

Setiadi, H., Josi, F. A. I., Doewes, A., & Damayanti, R. W. (2025). Optimizing student performance prediction: A comparative analysis of regression algorithms and feature selection techniques on LMS log data. *Vietnam Journal of Computer Science*, 12(02), Article 25500265. https://doi.org/10.1142/S2196888825500265

Smirani, L. K., Yamani, H. A., Menzli, L. J., & Boulahia, J. A. (2022). Using ensemble learning algorithms to predict student failure and enabling customized educational paths. Scientific Programming, 2022, Article 3805235. https://doi.org/10.1155/2022/3805235

Sukhbaatar, O., Usagawa, T., & Choimaa, L. (2019). An artificial neural network based early prediction of failure-prone students in blended learning course. *International Journal of Emerging Technologies in Learning*, 14(19), 77–92. https://doi.org/10.3991/ijet.v14i19.10366

Xue, H., & Niu, Y. (2023). Multi-output-based hybrid integrated models for student performance prediction. *Applied Sciences*, 13(9), Article 5384. https://doi.org/10.3390/app13095384