Extracting Institutional Analytics features from LMS data: Towards bridging Learning Design Analytics and Learning Analytics

Ariel Ortiz-Beltrán[1,*], Davinia Hernández-Leo[1,*]

[1]Universitat Pompeu Fabra, Plaça de la Mercè 10-12, 08002 Barcelona, Spain

Abstract

Learning Management Systems (LMS) generate extensive interaction data that, when strategically combined with Learning Analytics (LA) and Learning Design Analytics (LDA), hold significant potential for institutional analytics. However, conventional LA analyses often focus exclusively on isolated course-level behaviors, thereby neglecting important pedagogical design contexts. Addressing this gap, our study systematically aligns fundamental LMS-derived engagement indicators with three theoretically grounded constructs: (1) Massive vs. Distributed Learning, (2) Workload, and (3) Active Learning. We analyzed anonymized LMS interaction logs from courses of a Spanish brick-and-mortar university. Employing exploratory data analysis, theory-informed and expert-guided refinement, as well as rigorous feature engineering, we extracted a concise set of easily replicable indicators. Our primary contribution lies in proposing a structured, theoretically aligned approach for deriving meaningful, pedagogically contextualized insights from minimal LMS data, using only standard activity-level log fields commonly available at most institutions. This lightweight analytic approach not only facilitates broader institutional adoption but also supports targeted instructional interventions and informed course design improvements. Future research directions include validating the robustness of these indicators across multiple institutional contexts and exploring their predictive capabilities regarding key educational outcomes such as student satisfaction, engagement quality, and academic performance.

Keywords

Institutional analytics, learning analytics, learning design, feature extraction

1. Introduction

Learning Analytics (LA) and Learning Design (LD) have grown as complementary approaches to improving teaching and learning in higher education. Learning design is typically defined as the planned sequence of learning activities, resources, and assessments that reflect an instructor's pedagogical intent for a course. Learning analytics, on the other hand, entails collecting and analyzing data about learners' interactions, typically from Learning Management System (LMS) logs, in order to gain insight and enhance learning processes. Early on, researchers recognized the significant potential synergy between these fields. In fact, [1] argued that learning analytics should "take up where learning design leaves off", using data to test whether student behavior matches the intended design and recommend interventions when it does not. They hoped that combining LD and LA would provide a critical context for interpreting student data and evidence that well-designed learning experiences actually improve outcomes. Over the last decade, this vision has fueled both theoretical frameworks and empirical studies that connect learning design decisions to analytics on student engagement and performance [2]. Our research is motivated by this promising but underexplored link. We want to investigate common and widely available LMS student usage data to determine what insightful features can be extracted easily using the most common log data and tools available to any institution.

LMS platforms generate large volumes of interaction data, typically analyzed through Learning Analytics (LA) to extract behavioral patterns. However, analyses commonly focus on single courses and often overlook their interpretation in the frame of their Learning Design (LD), limiting their relevance to

support institutional decision-making, for example, related to the pedagogical model of the institution or the needs for teacher training [3, 4]. Cross-course datasets offer broader insights, but a key challenge remains: extracting comparable features that reflect both student behavior and pedagogical context.

This work examines student interaction logs from a university's Moodle platform encompassing 85 courses and over 68,000 activity records. Such a cross-course dataset provides a rich basis for institutional analytics, revealing patterns that single-course analyses might overlook [4]. To enable meaningful comparisons across diverse courses, we extract a set of high-level features from the raw logs that characterize each course's engagement profile. For example, we compute the total number of activity visits per course (as an indicator of overall student engagement), the time between each student's first and last interactions in a course (capturing how distributed or concentrated their participation is), and the diversity of task types accessed (reflecting the breadth of learning activities experienced). These features could be used to transform raw, unused data into interpretable metrics that carry implications for both student behavior and the underlying learning design.

This type of data may supply two complementary types of insights: Learning Design (LD) indicators and Learning Analytics (LA) indicators. LD indicators reflect instructors' intended student engagement, particularly the nature and diversity of learning resources and tasks assigned to students. Such indicators provide information about teachers' design choices and intended teaching methods. LA indicators, on the other hand, provide empirical insights into actual student engagement, such as LMS usage frequency, interaction patterns, inferred workload, and engagement patterns over time (e.g., distributed vs. massive). Analyzing these indicators together can help institutions better understand how course design choices align with, or differ from, actual student engagement behaviors and outcomes.

In this context, despite significant advances in Learning Analytics and its growing institutional importance, the practical challenge remains: how can institutions systematically utilize readily accessible LMS data to generate meaningful insights linked explicitly to educational designs and strategies? More specifically, there is a critical gap in translating basic, available activity-level LMS data into pedagogically relevant engagement metrics. Addressing this gap is essential for enabling institutional stakeholders to better understand and enhance student learning experiences. Thus, the research question guiding this paper is: **How can basic activity-level LMS variables be systematically transformed into engagement indicators that map onto the constructs of Massive vs Distributed learning, Workload, and Active Learning?**

2. Data & Methods

To systematically address our research question—How can basic activity-level LMS variables be systematically transformed into engagement indicators that map onto the constructs of Massive vs Distributed learning, Workload, and Active Learning?—we adopted a methodological approach grounded in clearly defined extraction rules and modern analytical techniques. This section describes the LMS-derived dataset used in this study, outlines the theoretical constructs guiding our analysis, and specifies formal procedures and mathematical definitions developed to transform raw LMS log data into pedagogically meaningful indicators.

Interaction data collected by Learning Management Systems (LMS) is usually standardized and available to system administrators across educational institutions. Such datasets are ideal for institutional analytics due to their structured nature and the ease with which meaningful metrics can be extracted and interpreted. The dataset used in this study is representative of standard LMS data of this type, consisting of student interaction logs from Moodle aggregated at the Activity or Resource (A/R) level. It contains six key variables described in Table 1. Although the original dataset contained information regarding the time spent on each task, for our proposal, we decided not to consider it. Time spent on tasks in an LMS might seem like a straightforward measure, but it often doesn't tell the whole story. Students might leave a page open while doing other activities, or spend more time on a task due to difficulty without being engaged. This can lead to misleading conclusions about their involvement in learning [5, 6].

Table 1

Summary of raw fields defining the students' interaction with the LMS activities.

Field	Description	Format
Anonymized ID (u_i)	Pseudonymised learner identifier	**VARCHAR** (opaque hash)
Course (c_i)	Pseudonymised Course Identificator	**VARCHAR (opaque hash)**
Activity (a_i)	Unique activity- or resource-identifier	**VARCHAR** (Or integer)
Visits (v_i)	Total number of times the learner accessed the activity	**INTEGER** (only positive values)
First Access (f_i)	First calendar date on which the learner accessed the activity	**DATE** (YYYY-MM-DD)
Last Access (l_i)	The last calendar date on which the learner accessed the activity	**DATE** (YYYY-MM-DD)

3. Proposed Indicators

The selected indicators, as well as their classification into clear theoretical constructs, present a methodological reflection that connects Learning Design Analytics (LD) and Learning Analytics (LA). LD focuses mostly on instructors' intentions and pedagogical choices, such as the variety and type of activities or resources offered. In contrast, LA reflects actual student behaviors and interactions as captured by LMS data, such as engagement patterns, frequency of interaction, and distribution of work. By systematically aligning these empirically measured behaviors (LA indicators) with their corresponding pedagogical intentions (LD indicators), institutions can better assess the effectiveness and alignment of their instructional designs with actual student engagement outcomes. Analyzing discrepancies or congruencies between designed engagement (intended LD) and observed interaction patterns (empirical LA) allows for targeted instructional interventions and data-driven course design.

3.1. Theoretical Constructs

Following a basic statistical exploration of the dataset and theory-oriented alignment with ultimate learning design intentions, a set of key analytical indicators was identified and defined to help characterize student engagement [2, 7]. These indicators were organized into three analytical constructs based on their hypothetical relevance, massive vs. distributed learning, workload, and active learning, because of their established relevance for institutional analytics in higher education. The Massive vs. Distributed Learning construct has significance because it distinguishes student engagement behaviors associated with deeper, sustained learning from superficial, short-term engagement, allowing institutions to identify at-risk courses or students prone to last-minute study habits [8, 9, 10]. Workload, based on cognitive load theory and student experience research, enables institutions to identify when high cognitive demands from course activities exceed learners' capacity, potentially triggering superficial learning approaches or disengagement, and thus informing interventions for workload calibration [11, 12]. Finally, the Active Learning concept, rooted in constructivist educational theory, emphasizes meaningful student interactions and engagement in learning tasks rather than passive consumption, highlighting how institutional LMS analytics can reveal the effectiveness of pedagogical designs in fostering deeper learning and active student participation. [13, 14].

Massive versus Distributed Learning Indicators aim to distinguish between highly concentrated patterns of interaction (massive learning) versus more distributed, sustained engagement patterns over time. Indicators in this category include the temporal spread of interactions with Activities/Resources (**A/R Interaction Span**), the frequency of visits per activity/resource (**A/R Interaction Frequency**), intervals between students' first-time access to activities/resources (**A/R First-time Access Intervals**), and overall course-level frequency of interactions (**Course Interaction Frequency**).

Workload Indicators were identified based on their potential link to student-reported satisfaction with course workload. These indicators include the breadth or variety of resources accessed, measured by the **Diversity Activity/Resource Index (Diversity A/R Index)**—higher values suggesting greater cognitive demand—and two frequency-based metrics: the total frequency of interactions per course

(**Course Interaction Frequency**) and the average daily frequency of student access (**Daily Access Frequency**).

Active Learning Indicators are defined as measures reflective of student engagement behaviors that involve deliberate, frequent, and diverse interactions. Indicators considered here include the total number of unique activities/resources accessed (**Diversity A/R Index**), the total number of days students remained actively engaged with course materials (**Course Interaction Span** in days), regularity of intervals between initial student accesses to new activities/resources (First-time Access Interval), and the daily frequency of LMS access (**Daily Access Frequency**).

3.2. Extraction rules

This section provides a standardized specification of the extraction rules required for computing the proposed engagement indicators from standard LMS logs. To ensure reproducibility, we formalize the raw data present in the event log for each activity, denoted as:

$$L_a = \langle u_i, a_i, c_i, v_i, f_i, l_i \rangle \ : \ i = 1, ..., N_c$$

Where L_a is the whole log for the activity a, u_i represents an anonymized (or pseudoanonymized) student identifier, a_i denotes the Activity/Resource identifier (A/R), v_i is the number of visits for an event, and f_i and l_i are the dates (in days, YYYY-MM-DD format) of the first and the last student access to the activity a_i , respectively. At the course level, the extraction rules for the indicators are defined as follows:

The Course Interaction Span (**CIS**) measures the total span in days of student engagement across the entire course and is calculated by subtracting the earliest first-access date $F_{min}^c = min(f_i)$ from the latest last-access date. Formally,

$$CIS_c = 1 + (L_{max}^c - F_{min}^c)$$

With dates computed as calendar-day differences, inclusive of both endpoints. In parallel, Course Interaction Frequency (**CIF**) reflects overall engagement intensity as total visits v_i of a student u_i to any A/R a_i during the course, divided by the course-interaction span, expressed as:

$$CIF_c = \frac{\sum_i v_i}{CIS_c}$$

The Daily Access Frequency (**DAF**) quantifies the regularity of student engagement, calculated by counting the number of distinct calendar days on which at least one activity was accessed within the course. To compute this, we create a set of distinct dates of access per student per activity (from f_i to l_i) and then count unique days across all activities in the course. To define this indicator formally, we need to define the set of days where the engagement happened:

$$\Gamma_i = \{d \in \mathbb{Z} \ : \ f_i \leqslant d \leqslant l_i\}$$

That is, every calendar day on which a student u_i registered an interaction with any activity a_i. Then, we need to aggregate across all records in course c.

$$\Gamma_c = \bigcup_{i=1}^{N_c} \Gamma_i$$

So, finally, we can define the daily access frequency for course c as the cardinality of this union:

$$DAF_c = |\Gamma_c|$$

On the other hand, the Diversity A/R index (**DAI**) is a Learning Design indicator, specially designed to measure, to some extent, the intended breadth of usage of A/R from the perspective of the teacher. It could be calculated simply by the cardinality of the distinct A/R accessed within the course. Formally,

$$DAI_c = |\{a_i \,:\, i = 1, ..., N_c\}|$$

We can also aggregate some A/R indicators to the course level to allow a better alignment with other institutional insights. For instance the A/R-interaction span (**AIS**) quantifies the span of student engagement per A/R, but the average of all the accessed activities $a \in A_c$ reports, for course c, the average number of calendar days over which students engaged with each A/R, let's denote by F_a^c and L_a^c the earliest first-access date and the latest last-access date, respectively, First compute the span of each activity a as:

$$S_a^c = 1 + (L_a^c - F_a^c)$$

defined in calendar days (inclusive). The course-level A/R-interaction span (AIS) is then the mean of these activity spans.

$$AIS_c = \frac{1}{|A_c|} \sum_{a \in A_c} S_a^c$$

Similarly, the A/R interaction frequency (AIF) follows the same aggregation process. Specifically, for every course c, let A_c denote the set of distinct activities/resources that registered at least one access, and let,

$$V_a^c = \sum_{\{i \,:\, a_i = a\}}$$

Be the total number of visits to the activity a (Summing up the visit count v_i over all rows i whose Activity ID equals a). With F_a^c and L_a^c as the earliest first-access and latest last-access dates for the activity a, the visit rate for that activity is:

$$\phi_a^c = \frac{V_a^c}{(L_a^c - F_a^c)}$$

The course-level A/R-interaction frequency (**AIF**) is then the mean of these per-activity rates:

$$AIF_c = \frac{1}{|A_c|} \sum_{a \in A_c} \phi_a^c$$

Therefore, AIF_c reports the average of the students' visits to each activity/resource aggregated at the course level. Lastly, the First-time access interval (**FAI**) measures the regularity with which students engage for the first time with different activities/resources. First, we sequence all activities/resources in chronological order according to their earliest first-access date. Then we calculate the average interval (in days) between successive first-access dates $F_{a_j}^c$, $F_{a_{j+1}}^c$, $F_{a_{j+2}}^c$, etc.

$$FAI_c = \frac{1}{|A_c|-1} \sum_{j=1}^{|A_c|-1} (F_{a_{j+1}}^c - F_{a_j}^c)$$

Where activities a_j are ordered in increasing $F_{a_j}^c$ order. Note that this indicator requires at least two distinct activities to be interpretable.

These extraction rules rely exclusively on standardized activity-level LMS data commonly available to institutions. They offer straightforward implementation using standard computational tools (e.g., Python, R, SQL). Importantly, while the indicators proposed here are robust and coherent across typical LMS datasets, their interpretative clarity may vary slightly depending on the course structure (e.g., fully asynchronous or self-paced courses). Nonetheless, the described approach offers a comprehensive and reproducible method for systematically assessing student engagement, facilitating integration between Learning Analytics and Learning Design Analytics at the institutional level.

3.3. Implementation Notes

All the proposed indicators rely exclusively on fields commonly provided by standard LMS exports (e.g., Moodle or Blackboard), namely student ID, course ID, activity ID, first and last access dates, and total visit counts. No advanced data streams, such as xAPI or detailed server-session logs, are necessary, making this approach particularly viable for institutions with limited LMS data availability. The computational extraction of indicators can be efficiently carried out using standard tools such as SQL queries (leveraging common functions like **GROUP BY, MIN, MAX**, and date differences), as well as Python's *Pandas* library (**groupby().agg()**), or R's *dplyr* package (**summarise()**); our 68,000-record dataset required less than 10 seconds of processing time using each of these methods using a single *Apple Silicon M1* SoC. To ensure consistency, timestamps should always be converted to the institutional time zone before computing spans, thus preventing inflated intervals due to mixed time zones. If explicit visit counts are missing from LMS logs, assigning a default visit value of one per access event maintains indicator stability; our sensitivity tests showed this simplification had minimal impact on indicator rankings. Although the LMS collects a Time Spent *"Time on Task"* field (HH:MM:SS per activity), we deliberately exclude it from our indicator because raw dwell time recorded by LMSs is widely acknowledged as an often unreliable standalone indicator of genuine, multidimensional student engagement. [5, 6]. Courses with fewer than two activities, which prevent meaningful computation of the First-time Access Interval (**FAI**), or those with zero-day interaction spans should be flagged and excluded from comparative analyses.

4. Discussion

The central aim of this study was to address the research question: **How can basic activity-level LMS variables be systematically transformed into engagement indicators that map onto the constructs of Massive-versus-Distributed Learning, Workload, and Active Learning** The approach presented in Section 2 describes a practical, rigorously specified workflow that transforms a minimal set of LMS log fields—first and last access dates, visit counts, and activity identifiers—into interpretable engagement indicators. By expressly tying each metric to a pedagogical construct, the paper illustrates, in a standardized and replicable manner, how design-aware analytics can be derived without relying on complex or sensitive data. Temporal dispersion measures (Course-Interaction Span, A/R-Interaction Span, First-time Access Interval) characterize massed versus distributed learning behaviours, while frequency-based metrics (Course-Interaction Frequency, Daily-Access Frequency) and breadth metrics (Diversity A/R Index) offer proxies for workload demands and active-learning opportunities. Together, these examples demonstrate the analytical power latent in even the most rudimentary LMS exports and provide a concrete template that institutions can adopt or adapt to their contexts [3]. Moreover, prior research has illustrated that this kind of tada could be associated with student outcomes such as satisfaction and academic success [7]. Integrating such outcome variables with our indicators would therefore provide a richer institutional picture

However, despite the methodological feasibility, the broader adoption of these indicators faces significant practical challenges. Higher Education Institutions (HEIs) are often cautious about sharing fine-grained LMS data with external or even internal analytics providers due to legitimate privacy concerns, regulatory compliance issues such as GDPR, administrative overhead, and institutional risk perceptions [3], [4]. Institutional skepticism is reinforced by reviews indicating limited demonstrable impact of expensive dashboard initiatives [15]; low-risk, anonymised, course-level indicators such as ours provide a pragmatic first step. Our proposed indicator set mitigates these concerns by emphasizing minimal and anonymized data requirements, aiming at course-level indicators rather than student-level fields typically already available to LMS administrators. Nonetheless, successful implementation will depend heavily on continued dissemination, clear communication, and demonstration of tangible value to stakeholders [16]. To this end, it is essential to emphasize through workshops, policy briefs, and pilot studies that the benefits derived from this minimal-data analytical approach (e.g., early-warning systems, workload diagnostics, and pedagogical design evaluations) outweigh the perceived risks associated

with sharing aggregated and de-identified LMS activity data.

Additionally, our analysis underscores a fundamental limitation: LMS behavioral data alone are insufficient to capture the complete complexity of the theoretical constructs we aim to measure. Indicators derived solely from LMS logs provide valuable but inherently partial views of student engagement, as they lack insights into underlying motivations, cognitive processes, and affective responses. For instance, indicators measuring the breadth and frequency of resource access (e.g., Diversity A/R Index) cannot independently distinguish between meaningful exploratory behavior and superficial resource navigation driven by confusion or lack of clarity. To address acknowledged limitations in current analytics practice, we propose expanding our minimal-data indicators into a multidimensional evidence set that combines (i) additional LA traces—such as assessment scores, formative-quiz attempts, and forum interaction networks, (ii) explicit LDA descriptors drawn from design documentation—planned weekly workload, task complexity, collaboration mode, and (iii) outcome variables—including course grades and validated satisfaction. This richer data integration directly addresses the shortcomings identified in recent reviews: Kaliisa et al. highlight how small samples, single-source logs, and non-standardised outcomes reduce the evidential value of many dashboard studies [15], while Topali et al. show that instructor-facing dashboards frequently raise awareness but rarely translate into actionable pedagogical feedback because the displayed metrics lack a clear theoretical link to learning design [17]. Our approach answers this missing connection by explicitly mapping each indicator to the constructs of Massive-versus-Distributed Learning, Workload, and Active Learning, resulting in a stronger foundation for actionable, design-aware analysis.

In summary, the paper offers three key contributions: (1) a rigorously defined, minimal-data workflow that converts standard LMS logs into theoretically grounded engagement indicators; (2) a discussion of practical considerations for institutional adoption, particularly around privacy and administrative feasibility; and (3) a call for multidimensional analytics that integrate additional Learning Analytics and Learning Design data to overcome the explanatory limits of single-source logs. Addressing these challenges will enhance higher-education institutions' capacity to generate actionable, design-aware insights that ultimately improve teaching and learning outcomes.

5. Conclusions & Future Work

This paper addressed the practical research question: **How can a minimal set of activity-level LMS variables be systematically transformed into engagement indicators that map onto the constructs of Massive-versus-Distributed Learning, Workload, and Active Learning?** To answer this, we developed a clearly defined and rigorously specified extraction workflow that converts only four standard LMS fields—first-access date, last-access date, visit counts, and activity identifiers—into three coherent indicator categories: temporal-dispersion, frequency, and breadth. By explicitly aligning each metric with established learning-design theory, this approach provides institutions with a standardized, replicable, and privacy-conscious analytic approach suitable for immediate implementation. Demonstrated feasibility across 85 university courses illustrates the practical potential of these indicators to yield meaningful insights, including identification of engagement patterns (massed versus distributed learning), workload peaks, and active learning opportunities, without requiring sensitive student data or advanced tracking capabilities.

Nevertheless, this study also highlighted important areas for future research. First, expanding towards a multidimensional analytics strategy is crucial: subsequent studies should integrate our minimal LMS indicators with complementary learning analytics signals (e.g., quiz results, forum interactions) and explicit learning-design metadata (planned weekly workload, collaborative task tags), providing richer interpretations and deeper insights. Second, validating the predictive power of these indicators by linking them to critical student-level outcomes such as performance (grades), satisfaction, and retention will further solidify their practical utility. Third, replicating the analysis across multiple institutions and diverse LMS platforms will test indicator robustness and generalizability. Fourth, although we excluded dwell-time measures due to validity concerns, future efforts could explore calibrated measures

of engagement duration as supplementary indicators. Lastly, developing and trialing low-overhead dashboards that embed these indicators directly into instructor workflows will illuminate how design-aware analytics influence pedagogical decisions and ultimately student learning outcomes. Taken together, these future directions will transform this initial proof-of-concept into a robust, scalable, and institutionally actionable analytics toolkit for enhancing higher education teaching and learning.

Declaration on Generative AI

This manuscript includes content that was revised with the assistance of ChatGPT, a large language model developed by OpenAI. The model was used to assist with improving the clarity, conciseness, and style of the manuscript. All content was critically reviewed and validated by the authors to ensure accuracy and academic integrity.

References

[1] L. Lockyer, E. Heathcote, S. Dawson, Informing pedagogical action: Aligning learning analytics with learning design, American Behavioral Scientist 57 (2013) 1439–1459. doi:10.1177/0002764213479367.

[2] G. Ahmadi, A. Mohammadi, S. Asadzandi, M. Shah, R. Mojtahedzadeh, What are the indicators of student engagement in learning management systems? a systematized review of the literature, The International Review of Research in Open and Distributed Learning 24 (2023) 117–136. doi:10.19173/irrodl.v24i1.6453.

[3] O. Viberg, Å. Grönlund, Desperately seeking the impact of learning analytics in education at scale: Marrying data analysis with teaching and learning, in: Online Learning Analytics, Auerbach Publications, 2021.

[4] T. Hildebrandt, L. Mehnen, Cross-course process mining of student clickstream data- aggregation and group comparison, in: 2024 26th International Conference on Business Informatics (CBI), 2024, pp. 90–98. doi:10.1109/CBI62504.2024.00020.

[5] C. R. Henrie, R. Bodily, R. Larsen, C. R. Graham, Exploring the potential of lms log data as a proxy measure of student engagement, J Comput High Educ 30 (2018) 344–362. doi:10.1007/s12528-017-9161-1.

[6] V. Kovanović, D. Gašević, S. Dawson, S. Joksimović, R. S. Baker, M. Hatala, Penetrating the black box of time-on-task estimation, in: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, LAK '15, Association for Computing Machinery, New York, NY, USA, 2015, pp. 184–193. doi:10.1145/2723576.2723623.

[7] N. Korobova, S. S. Starobin, A comparative study of student engagement, satisfaction, and academic success among international and american students, Journal of International Students 5 (2015). doi:10.32674/jis.v5i1.444.

[8] N. J. Cepeda, H. Pashler, E. Vul, J. T. Wixted, D. Rohrer, Distributed practice in verbal recall tasks: A review and quantitative synthesis, Psychol Bull 132 (2006) 354–380. doi:10.1037/0033-2909.132.3.354.

[9] N. Kornell, Optimising learning using flashcards: Spacing is more effective than cramming, Applied Cognitive Psychology 23 (2009) 1297–1317. doi:10.1002/acp.1537.

[10] F. Rodriguez, M. J. Rivas, L. H. Matsumura, M. Warschauer, B. K. Sato, How do students study in stem courses? findings from a light-touch intervention and its relevance for underrepresented students, PLOS ONE 13 (2018). doi:10.1371/journal.pone.0200767.

[11] J. Sweller, Cognitive load during problem solving: Effects on learning, Cognitive Science 12 (1988) 257–285. doi:10.1207/s15516709cog120_4.

[12] C. C. Bonwell, J. A. Eison, Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports, Technical Report, ERIC Clearinghouse on Higher Education, The

George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183, 1991. Accessed: May 31, 2025. [Online]. Available: https://eric.ed.gov/?id=ED336049.

[13] S. Freeman, et al., Active learning increases student performance in science, engineering, and mathematics, Proceedings of the National Academy of Sciences 111 (2014) 8410–8415. doi:10.1073/pnas.1319030111.

[14] M. Prince, Does active learning work? a review of the research, Journal of Engineering Education 93 (2004) 223–231. doi:10.1002/j.2168-9830.2004.tb00809.x.

[15] R. Kaliisa, K. Misiejuk, S. López-Pernas, M. Khalil, M. Saqr, Have learning analytics dashboards lived up to the hype? a systematic review of impact on students' achievement, motivation, participation and attitude, in: Proceedings of the 14th Learning Analytics and Knowledge Conference, ACM, Kyoto, Japan, 2024, pp. 295–304. doi:10.1145/3636555.3636884.

[16] L.-A. Lim, A. Atif, K. Heggart, N. Sutton, In search of alignment between learning analytics and learning design: A multiple case study in a higher education institution, Education Sciences 13 (2023). doi:10.3390/educsci13111114.

[17] P. Topali, I.-A. Chounta, A. Martínez-Monés, Y. Dimitriadis, Delving into instructor-led feedback interventions informed by learning analytics in massive open online courses, Journal of Computer Assisted Learning 39 (2023) 1039–1060. doi:10.1111/jcal.12799.