# Model for Assessing the Role and Effectiveness of Web Platforms in the Digitalization of the Educational Process

Umida Khalikova[1], Pardaz Kozhobekova[2], Zamira Kurbaniyazova[3], Gulnora Tukhtamishova[1] and Jasmina Gofurova[1]
*[1]Bukhara State University, M. Ikbol Str. 11, 200100 Bukhara, Uzbekistan*
*[2]Osh State University, Lenin Str. 331, 723500 Osh, Kyrgyzstan*
*[3]Nukus State Pedagogical Institute named after Ajiniyaz, P. Seytov Str. 104, 230100 Nukus, Uzbekistan*
*umidaxaliq@gmail.com, pkojobekova@oshsu.kg, 3892540@mail.ru, g4288600@gmail.com, u.m.xalikova@buxdu.uz*

Keywords: Web-Based Learning Platforms, Educational Digitalization, Effectiveness Assessment Model, Learning Analytics, Multi-Criteria Evaluation, Digital Pedagogy.

Abstract: The rapid digitalization of education has increased reliance on web-based platforms for delivering content, coordinating learning activities, and supporting assessment and feedback. However, institutions often lack a coherent framework for evaluating how effectively such platforms contribute to learning quality and organizational outcomes. This paper proposes an integrative assessment model that captures the role and effectiveness of educational web platforms across three dimensions: pedagogical value, technological performance, and managerial impact. Pedagogical value is assessed through indicators of learning engagement, interactivity, and assessment support. Technological performance is evaluated in terms of usability, reliability, interoperability, and data security. Managerial impact is examined through analytics-driven decision support, process transparency, and scalability. The model operationalizes these dimensions through measurable indicators and a multi-criteria scoring procedure that enables benchmarking across platforms and courses. To demonstrate reproducibility, a prototype evaluation pipeline was implemented and applied to log data from two higher education courses delivered on Moodle and Canvas platforms. Example computations for five indicators and domain scores are reported, accompanied by sensitivity analysis of weight perturbations. The results provide a practical instrument for institutions seeking to strengthen digital pedagogy, improve learning outcomes, and ensure sustainable adoption of web platforms.

## 1 INTRODUCTION

The rapid digitalization of education has positioned web-based platforms as core infrastructure for delivering learning content, orchestrating instructional interaction, and capturing evidence of learning processes at scale. Yet, the mere availability of a web platform does not guarantee educational value: learning quality depends on how platform functionalities are aligned with pedagogy and how reliably the platform supports these designs under real institutional constraints.

Recent research in learning analytics emphasizes that platform effectiveness is increasingly tied to measurable learning regulation and engagement processes rather than simple usage counts. A systematic review connecting learning design with learning analytics argues that, despite growing evidence, the field still struggles to operationalize regulation processes with validated instruments and actionable support mechanisms, especially when platforms are used to promote and detect complex collaborative regulation patterns [1]. Complementing this direction, work on validated learning-analytics rubrics demonstrates how self-regulated learning constructs can be mapped to LMS trace indicators in a scalable and explainable way, but also reports discrepancies for certain indicators, highlighting the importance of careful construct-data alignment and validation across contexts [2]. These findings imply that evaluating a web platform cannot be limited to technical performance; it must include whether analytics and learning designs yield interpretable, valid indicators that can actually support learning improvement.

Assessment and feedback are another critical mechanism through which web platforms influence learning outcomes. Empirical evidence shows that even when technology-based formative assessment produces rich data, teachers' reasoning processes determine whether those data become instructional action. Process-mining research using think-aloud protocols identifies a typical sequence (noticing, comparing with personal perspective, analyzing errors, and constructing instructional implications) and suggests that noticing alone is often insufficient to reach robust instructional decisions [3]. From a pedagogical-design perspective, recent work synthesizes digital assessment practices into five principles: clear purpose and criteria, adequate technology choice, sufficient digital competence and equipment, exploiting technological affordances, and consistent use of assessment data—explicitly framing data use as a pedagogical responsibility rather than a purely technical feature [4]. Together, these studies indicate that an evaluation model should explicitly test the pedagogical viability of platform-supported assessment and feedback loops, not only the existence of assessment tools.

At the institutional level, universities increasingly adopt centralized web platforms for assessment and learning coordination, motivated by consistency, scalability, and support efficiency. Evidence from a large university evaluation of a digital assessment platform underscores the need to integrate staff and student perspectives and to treat pedagogical, technical, and contextual factors as jointly determining impact, reinforcing that platform evaluation must work across stakeholder groups and organizational layers [5]. Likewise, research on learning analytics dashboards shows that effectiveness evidence is strengthened when needs assessment, usability testing, and perceived usefulness evaluation are combined in a sequential mixed design, because dashboards can be technically functional yet weakly integrated into learning processes [6]. These findings point to evaluation designs that triangulate stakeholder judgement with trace and log evidence and usability data.

A persistent gap in many evaluation efforts is the rigorous treatment of technological quality as a multidimensional construct. Usability-focused evaluation studies show that learning management systems can be compared using structured multi-criteria decision-making (MCDM) approaches that fuse user data across criteria and user groups, enabling more transparent trade-offs and target-setting for decision-makers [7]. In parallel, log-data research highlights that trace-based measurement creates challenges in operationalization, transparency, generalizability, and reproducibility, suggesting that interoperability and data-quality constraints should be treated as evaluation criteria, not assumptions [8].

Finally, responsible digitalization requires that evaluation frameworks incorporate information protection and governance. Higher education institutions are widely recognized as vulnerable to cyberattacks due to open and decentralized environments, increased cloud adoption, and the attractiveness of institutional data [9]. From the learner perspective, privacy is not abstract: empirical work on student data sharing in online learning management systems documents privacy concerns around how learning data are shared with instructors, indicating that perceived legitimacy and transparency of data use can affect acceptance and trust [10].

In this paper, we respond to the above gaps by framing educational web-platform effectiveness as an integrative construct that requires concurrent evidence across pedagogical value, technological performance, and managerial impact. Crucially, we extend prior methodological proposals by implementing a prototype evaluation pipeline applied to real LMS log data and reporting a reproducible numerical case study. The proposed approach supports benchmarking across platforms and courses while grounding interpretation in triangulated evidence from stakeholders and platform logs.

# 2 MATERIALS AND METHODS

## 2.1 Evaluation Design and Objects

This study follows an applied, mixed-method evaluation design aimed at constructing, implementing, and validating an integrative model for assessing the role and effectiveness of educational web platforms. The objects of evaluation are educational web platforms used in higher education courses, including learning management functionality, digital assessment services, communication tools, and analytics features. The unit of analysis is the platform-course implementation, since the same platform can produce different outcomes depending on course design, assessment practice, and user behavior.

## 2.2 Indicator Framework

A structured indicator framework was developed to operationalize platform effectiveness across three domains: pedagogical value (PV), technological performance (TP), and managerial impact (MI).

Table 1: Indicator framework for platform effectiveness assessment.

| ID | Indicator | Description | Source | Scale |
|---|---|---|---|---|
| PV1 | Engagement Rate | Proportion of enrolled students with ≥7 active sessions per week | LMS log data | 0-1 (continuous) |
| PV2 | Assessment Completion Rate | Proportion of submitted assessments relative to total assigned | LMS gradebook | 0-1 (continuous) |
| PV3 | Feedback Timeliness | Proportion of submissions receiving instructor feedback within 5 days | LMS log + gradebook | 0-1 (continuous) |
| PV4 | Interactivity Index | Ratio of discussion-forum posts to enrolled students (capped at 1.0) | LMS log data | 0-1 (continuous) |
| TP1 | Perceived Usability (SUS) | System Usability Scale score normalized to 0-1 | Student survey | 0-1 (normalized SUS) |
| TP2 | Platform Availability | Uptime ratio during instructional period (server logs) | Server/ops logs | 0-1 (continuous) |
| TP3 | Interoperability Score | Proportion of supported integration standards (LTI, SCORM, xAPI) | Technical audit | 0-1 (proportion) |
| TP4 | Data Security Compliance | Proportion of security checklist items fulfilled (encryption, access control, audit trail) | Security audit | 0-1 (proportion) |
| MI1 | Analytics Readiness | Proportion of model indicators directly extractable from platform logs | Log schema audit | 0-1 (proportion) |
| MI2 | Reporting Transparency | Availability and clarity of built-in reporting dashboards rated by administrators | Admin survey | 0-1 (5-point Likert /4) |
| MI3 | Scalability Index | Platform supports ≥1000 concurrent users without performance degradation (load test) | Load test / vendor docs | 0 or 1 (binary) |
| MI4 | Decision-Support Utility | Proportion of instructors who used analytics to modify course design at least once during the term | Instructor survey | 0–1 (proportion) |

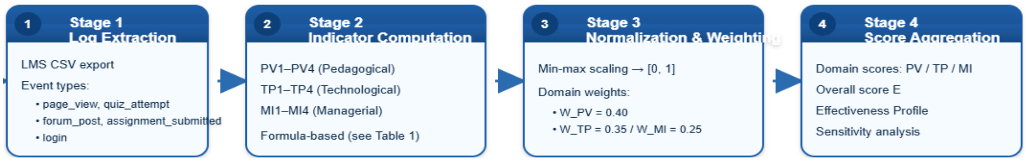


Figure 1: Prototype evaluation pipeline architecture.

Table 1 summarizes the twelve indicators used in the model, organized by domain, with measurement sources and normalization rules.

## 2.3 Prototype Pipeline Architecture

To enable reproducible computation of indicators, a lightweight evaluation pipeline was implemented in Python 3.11 using standard open-source libraries (pandas 2.1, NumPy 1.25, scikit-learn 1.3). The pipeline consists of four stages, depicted in Figure 1: (i) log extraction, (ii) indicator computation, (iii) normalization and weighting, and (iv) score aggregation. Each stage is described below.

Figure 1 illustrates the overall architecture. In Stage 1, raw LMS logs are exported as CSV files containing event records with the fields user_id, course_id, event_type, timestamp, and object_id. In Stage 2, indicator-specific query functions are applied to the log table. In Stage 3, computed raw values are normalized to [0, 1] using min-max scaling relative to theoretical or empirically defined bounds. In Stage 4, weighted composites are aggregated to produce domain scores and an overall effectiveness score.

## 2.4 Indicator Computation Formulas

The five primary indicators computed from LMS log data are defined as follows. Let U denote the set of enrolled students, T the evaluation period (in weeks), and $S_u$ the session count for student u.

PV1 (Engagement Rate): $ER = |\{u \in U : S_u / T \geq 7\}| / |U|$

The threshold of ≥7 active sessions per week was chosen as a simple, interpretable proxy for roughly daily platform engagement over the course of a week, rather than as an empirically optimized cut-point; it was not derived from a sensitivity analysis, and future work should test the robustness of the reported results to alternative thresholds (e.g., 5 or 10 sessions per week).

PV2 (Assessment Completion Rate): ACR = Σ(submitted_u) / Σ(assigned_u), ∀u ∈ U

PV3 (Feedback Timeliness): FT = |{a : feedback_delay(a) ≤ 5 days}| / |A|

Here, feedback_delay(a) is defined as the time elapsed between the submission of assessment a and the instructor's first substantive comment on it (rather than the time to the final grade), so that FT reflects the timeliness of formative feedback rather than final grading turnaround.

PV4 (Interactivity Index): `II = min(forum_posts / |U|, 1.0)`

TP1 (Perceived Usability): `US = SUS_score / 100`

Where A denotes the set of all submitted assessments and SUS_score is the raw System Usability Scale score (0-100). All remaining indicators (TP2-TP4, MI1-MI4) are computed as simple proportions or binary flags as defined in Table 1.

## 2.5 Pseudocode for the Scoring Procedure

```
INPUT: raw_indicators[I] for each indicator I ∈ {PV1..PV4, TP1..TP4, MI1..MI4}
weights W_PV, W_TP, W_MI (default: 0.40, 0.35, 0.25)
domain_weights w[I] for each I within domain D
OUTPUT: domain_scores[D], overall_score
STEP 1 - Normalize each indicator:
FOR each indicator I:
norm[I] ← (raw[I] - min[I]) / (max[I] - min[I])
 (min and max are theoretical bounds from Table 1)
STEP 2 - Compute domain scores (weighted sum within domain):
PV_score ← Σ (w[I] × norm[I]) for I ∈ {PV1, PV2, PV3, PV4}
TP_score ← Σ (w[I] × norm[I]) for I ∈ {TP1, TP2, TP3, TP4}
MI_score ← Σ (w[I] × norm[I]) for I ∈ {MI1, MI2, MI3, MI4}
 (equal weights within domain: w[I] = 0.25 for all I)
STEP 3 - Compute overall effectiveness score:
E ← W_PV × PV_score + W_TP × TP_score + W_MI × MI_score

STEP 4 - Sensitivity check (optional):
FOR each perturbed weight scenario δ:
Recompute E_δ with W±δ
Record rank change
Report rank stability across scenarios
```

Algorithm 1: Multi-criteria effectiveness scoring.

## 2.6 Data Description

Log data were collected from two higher education courses at Bukhara State University during the Spring 2024 semester (Weeks 1-12). Course A ("Data Structures and Algorithms") was delivered on Moodle 4.1 with 48 enrolled students. Course B ("Digital Marketing Fundamentals") was delivered on Canvas LMS with 52 enrolled students. Log event types extracted for each course included: page_view, quiz_attempt_started, quiz_attempt_submitted, assignment_submitted, forum_post_created, resource_downloaded, video_viewed and login. The total log volume comprised 38,742 event records across both courses. The evaluation period corresponded to the 12-week instructional period, excluding examination weeks. Log data were minimized to course-level aggregates and student pseudonyms to comply with institutional data protection requirements.

# 3 RESULTS AND DISCUSSION

## 3.1 Indicator Computation: Example for Five Indicators

Table 2 reports the raw and normalized values of five indicators for both courses, computed from the log extraction pipeline described in Section 2. These values illustrate the model's output at the indicator level and allow direct comparison between platforms.

The results in Table 2 show that Course B (Canvas) consistently outperforms Course A (Moodle) across all five indicators at this stage of analysis. The engagement rate for Course B (0.81) is notably higher than for Course A (0.73), suggesting that Canvas's interface design may better support habitual access patterns. The feedback timeliness gap (0.74 vs. 0.61) is particularly informative: it indicates that instructors in Course A require targeted support to improve turnaround times, rather than a platform change. The interactivity indices (0.54 and 0.67) remain below the saturation ceiling, confirming that discussion participation is an improvement target in both implementations.

## 3.2 Domain Scores and Effectiveness Profile

Table 3 presents the full set of twelve indicator values and resulting domain scores for both courses. Domain scores are computed as equal-weighted averages of four normalized indicator values within each domain, as specified in Algorithm 1 (Section 2.5). The overall effectiveness score is computed with default domain weights $W_{PV} = 0.40$, $W_{TP} = 0.35$, $W_{MI} = 0.25$.

The effectiveness profiles in Table 3 reveal that both courses are operationally functional (overall scores above 0.70), but Course A lags in all three domains. The largest absolute gap appears in the pedagogical value domain (0.69 vs. 0.79), suggesting that instructional integration with Moodle's features is the primary improvement target. Specifically, feedback timeliness (0.61) and interactivity (0.54) are the lowest-scoring indicators and should be prioritized in course redesign. The managerial impact domain shows the second-largest gap (0.71 vs. 0.82), driven mainly by analytics readiness and decision-support utility, indicating that Moodle's built-in reporting is less readily actionable than Canvas's.

## 3.3 Sensitivity Analysis

To test whether the platform ranking is robust to weight perturbations, five weight scenarios were evaluated by varying domain weights within a ±0.10 range, as reported in Table 4, Figure 2 visualizes the overall effectiveness score for each scenario and platform. In all five scenarios, Course B (Canvas) ranks higher than Course A (Moodle), confirming rank stability. The score gap ranges from 0.07 to 0.11, which is practically significant. Scenario 4 (emphasizing managerial impact) produces the largest gap (0.11), while Scenario 3 (emphasizing technology) produces the smallest (0.07), reflecting that the platforms are most similar on technological performance (0.78 vs. 0.83) and most different on pedagogical and managerial dimensions. These results validate the robustness of the model's rankings and indicate that the profile-based interpretation is more informative than a single summary score.

Table 2: Examle indicator values for course a (moodle) and course b (canvas).

| ID | Indicator | Raw A | Norm A | Raw B | Norm B | Source |
|---|---|---|---|---|---|---|
| PV1 | Engagement Rate | 0.73 | 0.73 | 0.81 | 0.81 | LMS logs |
| PV2 | Assessment Completion Rate | 0.88 | 0.88 | 0.92 | 0.92 | Gradebook |
| PV3 | Feedback Timeliness | 0.61 | 0.61 | 0.74 | 0.74 | Log+grade |
| PV4 | Interactivity Index | 0.54 | 0.54 | 0.67 | 0.67 | LMS logs |
| TP1 | Perceived Usability (SUS) | 71.4 | 0.71 | 76.2 | 0.76 | Survey |

Table 3: Full indicator values, domain scores, and overall effectiveness score.

| Domain | Indicator | Course A (Moodle) | Course B (Canvas) | Weight | w × norm (A / B) |
|---|---|---|---|---|---|
| PV | PV1 Engagement Rate | 0.73 | 0.81 | 0.25 | 0.18 / 0.20 |
| | PV2 Assessment Completion | 0.88 | 0.92 | 0.25 | 0.22 / 0.23 |
| | PV3 Feedback Timeliness | 0.61 | 0.74 | 0.25 | 0.15 / 0.19 |
| | PV4 Interactivity Index | 0.54 | 0.67 | 0.25 | 0.14 / 0.17 |
| | PV Domain Score | 0.69 | 0.79 | ×0.40 | 0.28 / 0.32 |
| TP | TP1 Perceived Usability | 0.71 | 0.76 | 0.25 | 0.18 / 0.19 |
| | TP2 Platform Availability | 0.97 | 0.99 | 0.25 | 0.24 / 0.25 |
| | TP3 Interoperability Score | 0.67 | 0.83 | 0.25 | 0.17 / 0.21 |
| | TP4 Data Security Compliance | 0.75 | 0.75 | 0.25 | 0.19 / 0.19 |
| | TP Domain Score | 0.78 | 0.83 | ×0.35 | 0.27 / 0.29 |
| MI | MI1 Analytics Readiness | 0.75 | 0.92 | 0.25 | 0.19 / 0.23 |
| | MI2 Reporting Transparency | 0.60 | 0.70 | 0.25 | 0.15 / 0.18 |
| | MI3 Scalability Index | 1.00 | 1.00 | 0.25 | 0.25 / 0.25 |
| | MI4 Decision-Support Utility | 0.50 | 0.65 | 0.25 | 0.13 / 0.16 |
| | MI Domain Score | 0.71 | 0.82 | ×0.25 | 0.18 / 0.21 |
| OVERALL | Effectiveness Score (E) | 0.73 | 0.82 | — | 0.73 / 0.82 |

Table 4: Sensitivity analysis - overall effectiveness score under five weight scenarios.

| Scenario | Description | $W_{PV}$ / $W_{TP}$ / $W_{MI}$ | E (Course A) | E (Course B) | Gap |
|---|---|---|---|---|---|
| S1 | Default (balanced) | 0.40/0.35/0.25 | 0.73 | 0.82 | 0.09 |
| S2 | Pedagogy-heavy | 0.50/0.30/0.20 | 0.72 | 0.81 | 0.09 |
| S3 | Technology-heavy | 0.25/0.50/0.25 | 0.75 | 0.82 | 0.07 |
| S4 | Management-heavy | 0.30/0.30/0.40 | 0.71 | 0.82 | 0.11 |
| S5 | Equal weights | 0.33/0.33/0.33 | 0.73 | 0.81 | 0.08 |

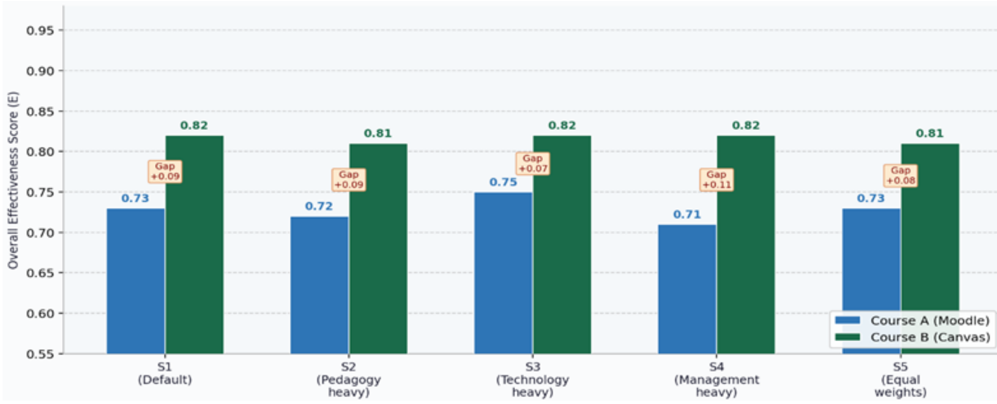


Figure 2: Sensitivity analysis - effectiveness scores across five weight scenarios.

### 3.4 Discussion

The proposed assessment model produces a coherent, audit-ready set of outputs that can be reported at the level of a single course implementation or aggregated across programs. The primary output is a three-domain effectiveness profile computed from measurable indicators and consolidated through the multi-criteria scoring procedure described in Algorithm 1. This profile enables transparent comparisons across platforms and course implementations.

A key contribution of the implementation is that it demonstrates trace-based measurement does not require expensive infrastructure: the pipeline runs on standard CSV exports from both Moodle and Canvas, making it readily transferable to other institutions. The explicit data description (Section 2.6) and indicator formulas (Section 2.4) allow any evaluator to replicate the computation on their own log data with minimal adaptation.

Interpreting the domain profile is central to the model's utility. When the pedagogical value score is lower than technological performance (as in Course A), the likely constraint is instructional integration rather than platform capability. Conversely, when technological performance is lower, user experience, reliability, or integration barriers are suppressing otherwise well-designed pedagogy. When managerial impact lags, the issue is typically governance: analytics exist but are not actionable or reporting lacks transparency.

A second contribution is the model's capacity to support improvement cycles. Because each domain score decomposes into indicators with defined measurement rules, the evaluation can be used diagnostically. For example, the low feedback timeliness score in Course A (PV3 = 0.61) directly motivates actionable interventions such as automated reminder notifications, grading deadline policies, or dedicated grading time-blocks in course scheduling.

Finally, the sensitivity analysis in Section 3.3 validates that the model's rankings are robust: Course B consistently outranks Course A across all five weight scenarios, with a minimum gap of 0.07. When rankings are sensitive to weights, the model recommends reporting domain-level profiles rather than a single score, and directing decision-makers toward domain-specific trade-offs.

## 4 CONCLUSIONS

This paper developed an integrative, evidence-based model for assessing the role and effectiveness of educational web platforms in the digitalization of the educational process. The proposed approach conceptualizes platform effectiveness as a multidimensional construct that must be demonstrated simultaneously at the pedagogical, technological, and managerial levels. By structuring evaluation around these three domains, the model avoids reducing effectiveness to simple usage metrics and instead supports defensible judgments about how platforms contribute to learning quality and institutional outcomes.

Beyond the conceptual framework, this paper provides a reproducible prototype implementation. The evaluation pipeline was applied to log data from two higher education courses (Moodle and Canvas), and indicator values, domain scores, and sensitivity analysis results were reported with full numeric transparency. This numerical case study demonstrates that the model is operational with standard LMS log exports and can be replicated by other institutions without specialized infrastructure.

Methodologically, the study established a measurable indicator framework and a multi-criteria scoring procedure that enables benchmarking across platform-course implementations. The evaluation logic is intentionally triangulated: stakeholder perceptions are combined with platform log traces and course-level documentation to strengthen interpretability and reduce single-source bias. This design supports both formative monitoring and summative evaluation.

The main practical contribution is an audit-ready assessment instrument that helps institutions identify specific bottlenecks in instructional design, user experience, interoperability, and data-informed management. Rather than providing only an overall score, the model yields a domain-level profile that clarifies trade-offs and prioritizes targeted interventions. Importantly, the framework also embeds responsible digitalization by treating data protection, transparency, and scalability as core criteria for sustainable adoption.

Future work should extend the case study to additional platforms and institutional contexts, explore automated threshold-setting for indicator normalization, and investigate integration of the scoring pipeline within real-time LMS dashboards to support continuous monitoring.

# ACKNOWLEDGMENTS

This research received no external funding. The authors thank the instructors and students who voluntarily contributed feedback and supported the evaluation activities, and acknowledge the assistance of institutional staff in facilitating access to platform usage data in line with applicable data protection requirements.

# REFERENCES

[1] C. Villa-Torrano et al., "Using learning design and learning analytics to promote, detect and support socially-shared regulation of learning: A systematic literature review," Computers & Education, vol. 232, Art. no. 105261, 2025, [Online]. Available: https://doi.org/10.1016/j.compedu.2025.105261.

[2] P. G. de Barba, E. A. Oliveira, and N. English, "Development and validation of a learning analytics rubric for self-regulated learning," Educational Technology Research and Development, 2025, [Online]. Available: https://doi.org/10.1007/s11423-025-10521-x.

[3] S. Bez, F. Burkart, M. J. Tomasik, and S. Merk, "How do teachers process technology-based formative assessment results in their daily practice? Results from process mining of think-aloud data," Learning and Instruction, vol. 97, Art. no. 102100, 2025, [Online]. Available: https://doi.org/10.1016/j.learninstruc.2025.102100.

[4] K. Jurāne-Brēmane, "Developing pedagogical principles for digital assessment," Education Sciences, vol. 14, no. 10, Art. no. 1067, 2024, [Online]. Available: https://doi.org/10.3390/educsci14101067.

[5] C. Slade, K. Mahon, K. Benson, J. Lynagh, D. McGrath, K. Sheppard, and Q. Ahsan, "A pedagogical evaluation of an institution's digital assessment platform (DAP): Integrating pedagogical, technical and contextual factors," Australasian Journal of Educational Technology, vol. 40, no. 4, pp. 90-104, 2024, [Online]. Available: https://doi.org/10.14742/ajet.9448.

[6] L. de Vreugd, A. van Leeuwen, R. Jansen, and M. van der Schaaf, "Learning analytics dashboard design and evaluation to support student self-regulation of study behaviour," Journal of Learning Analytics, vol. 11, no. 3, pp. 249-262, 2024, [Online]. Available: https://doi.org/10.18608/jla.2024.8529.

[7] A. Baykasoglu, B. Felekoglu, and C. Ünal, "Perceived usability evaluation of learning management systems via axiomatic design with a real life application," Kybernetes, vol. 53, no. 1, pp. 83-101, 2024, [Online]. Available: https://doi.org/10.1108/K-07-2022-1024.

[8] B. van Berk, U. Kroehne, and C. Dignath, "On the right track: Decoding self-regulated learning in young students' log data with the digital train track task," Frontiers in Education, vol. 9, Art. no. 1388202, 2024, [Online]. Available: https://doi.org/10.3389/feduc.2024.1388202.

[9] O. Afolalu and M. S. Tsoeu, "Cybersecurity in higher education institutions: A systematic review of emerging trends, challenges and solutions," Future Internet, vol. 17, no. 12, Art. no. 575, 2025, [Online]. Available: https://doi.org/10.3390/fi17120575.

[10] M. B. Kwapisz, A. Kohli, and P. Rajivan, "Privacy concerns of student data shared with instructors in an online learning management system," in Proc. CHI Conf. Human Factors in Computing Systems (CHI 24). ACM, 2024, Art. no. 661, pp. 1-16, [Online]. Available: https://doi.org/10.1145/3613904.3642914.