SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 5863-5878
Open Access. Online & Print

DOI: 10.5281/zenodo.12426617

AN AI-ENABLED INTELLIGENT ASSESSMENT FRAMEWORK INTEGRATED WITH THE GCC MODEL: EMPIRICAL EVALUATION IN HEALTHCARE EDUCATION

Fan Cheng[1], Thosporn Sangsawang[2]*

[1] Vocational Education Division, Faculty of Technical Education, Rajamangala University of Technology Thanyaburi, Pathum Thani, 12110, Thailand.
[2] Educational Technology and Communications Division, Faculty of Technical Education, Rajamangala University of Technology Thanyaburi, Pathum Thani, 12110, Thailand.

Received: 28/12/2025
Accepted: 25/03/2026

Corresponding author: Thosporn Sangsawang
(sthosporn@rmutt.ac.th)

ABSTRACT

The growing use of artificial intelligence in educational assessment has intensified debate about its actual contribution to student learning, particularly in practice-oriented fields such as healthcare education. While intelligent assessment systems are often promoted for their potential to improve objectivity and scalability, empirical evidence on how they influence learning outcomes (LO) and under what conditions they are most effective remains limited. In particular, the role of learner preparedness for digital technologies has received insufficient attention. This study examines the effect of AI-enabled intelligent assessment (AIDA) on student LO and investigates whether digital readiness (DR) conditions this relationship. A quantitative research design was employed, using a structured questionnaire administered to first-year university students enrolled in a healthcare education program. Data were analysed using partial least squares structural equation modelling to assess both direct and moderating effects. The findings indicate that AIDA is statistically significantly related to LO ($\beta = -0.18$, $p = 0.03$). DR also shows a significant direct association with LO ($\beta = -0.24$, $p < 0.01$). In addition, DR significantly moderates the relationship between intelligent assessment and LO ($\beta = -0.16$, $p = 0.05$), suggesting that the effectiveness of AIDA varies according to students' readiness to engage with digital technologies. The proposed model explains 11 percent of the variance in LO. These results contribute to educational informatics research by demonstrating that the impact of intelligent assessment systems on learning is heterogeneous and depends on learner-related factors. From a practical perspective, the findings suggest that institutions adopting AIDA should combine technological implementation with targeted efforts to strengthen students' DR to enhance LO in healthcare education.

KEYWORDS: AI-Enabled Intelligent Assessment, digital readiness, learning outcomes, healthcare education, educational informatics, structural equation modelling

Copyright: © 2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License. (https://cre-ativecommons.org/licenses/by/4.0/).

1. INTRODUCTION

Recent advances in artificial intelligence have increasingly shaped the development of information systems for data processing, decision support, and evaluation across a wide range of application domains, including education and healthcare [1]. In educational contexts, these developments are closely associated with the growing demand for objective, scalable assessment approaches capable of handling complex learning data. Traditional assessment methods, such as instructor judgement, manual grading, and summative examinations, remain central to educational practice but often exhibit limitations related to subjectivity, inter-rater inconsistency, delayed feedback, and restricted scalability when applied to large or heterogeneous learner populations [2]. These challenges are particularly evident in practice-oriented disciplines, where effective assessment must capture not only conceptual knowledge but also procedural competence and applied reasoning. Within this broader context, AI-enabled intelligent assessment (AIDA) has emerged as a significant research direction in educational informatics. Intelligent assessment systems employ machine learning and data analytics techniques to evaluate learner performance in a consistent and reproducible manner [3]. In contrast to conventional assessment approaches, AI-driven frameworks can process multidimensional learning data and support continuous evaluation aligned with instructional objectives [4]. Consequently, AIDA has been increasingly examined as a potential solution to persistent challenges related to the reliability and scalability of educational evaluation.

Healthcare education represents a particularly demanding context for the application of intelligent assessment technologies [5]. Internationally, healthcare training has increasingly adopted competency-based education models that emphasize integrating theoretical understanding with clinical judgement and practical skills [6]. In geriatric rehabilitation education, assessment complexity is further intensified by the need to evaluate applied competencies associated with patient-centred care and rehabilitation planning. Although instructional design models such as the GCC model have been developed to structure learning activities and support competency development, assessment practices in this domain often continue to rely on traditional human-based evaluation methods [7]. These approaches face ongoing challenges related to consistency, transparency, and scalability, particularly in early-stage university programs where student enrolments are increasing and instructional resources are limited [8]. Despite growing scholarly interest in AI applications for educational assessment, several important gaps remain in the literature. First, many existing studies conceptualize AI as a supplementary or assistive component rather than a primary assessment mechanism, thereby limiting empirical understanding of its direct effect on learning outcomes (LO). Second, relatively little research has examined how the effectiveness of AIDA systems varies across learners with different levels of readiness to engage with digital technologies. Learner characteristics related to digital readiness (DR), including familiarity with digital tools and confidence in technology-mediated learning environments, have been shown to influence educational outcomes but remain underexplored in AI-driven assessment studies [9]. Third, limited empirical evidence exists regarding the combined influence of AIDA and learners' DR on LO in healthcare education contexts.

These gaps have both theoretical and practical implications. From a theoretical perspective, insufficient attention to moderating learner characteristics constrains understanding of how AIDA systems operate under different conditions. From a practical perspective, the lack of evidence on conditional effects limits educational institutions' and practitioners' ability to design assessment systems that are effective for diverse learner populations. Addressing these issues is therefore essential for advancing research on intelligent assessment systems and their application in practice-oriented educational environments. In response to these challenges, this study proposes an AIDA framework integrated with the GCC instructional model and empirically examines its relationship with student LO in healthcare education. DR is incorporated as a moderating variable to investigate whether the impact of AIDA on LO differs across learners. A mixed-methods research design is employed, combining quantitative analysis of LO with system-based assessment data and learner feedback. Based on the proposed theoretical framework, the study addresses the following research questions:

RQ1: What is the effect of AI-enabled intelligent assessment (AIDA) on student learning outcomes (LO) in healthcare education?

RQ2: What is the effect of learners' digital readiness (DR) on student learning outcomes (LO)?

RQ3: Does learners' DR moderate the relationship between AIDA and student learning outcomes (LO)?

By addressing these research questions, the study contributes to the fields of AIDA, educational informatics, and intelligent evaluation systems. The findings are expected to inform both theory and practice by clarifying how AI-driven assessment influences LO and under what conditions its effectiveness is enhanced. **Table 1** provides an overview of selected empirical and theoretical studies examining artificial intelligence-enabled assessment, DR, and LO in educational contexts. The table indicates whether these constructs were explicitly addressed in prior research and summarises the methodological approaches employed. As illustrated, existing studies have tended to focus on individual constructs rather than examining their combined or conditional effects. This overview clarifies the present study's positioning by demonstrating how it integrates system-level and learner-level factors to address a gap in the current literature.

Table 1: Empirical and theoretical studies on AIDA, DR, and LO

Authors	Context	AIDA examined	LO analysed	DR considered	Moderating effects tested	Method
Varadarajan, et al. [10]	Higher education	✓	Indirect	✗	✗	Review
Fowler, et al. [11]	Technology-enhanced learning	✓	✓	Limited	✗	Quantitative
Holmes and Tuomi (2022)	Digital education	✗	✓	✓	✗	Conceptual
Sallam [12]	Healthcare education	✓	Limited	✗	✗	Experimental
Lee and Shin [13]	Practice-based learning	✓	✓	✗	✗	Mixed methods
Varadarajan, et al. [10]	Higher education policy	Indirect	Indirect	✓	✗	Review
Mamede and Schmidt [14]	Medical education	✓	Limited	✗	✗	Empirical
Lu and Lin [15]	Competency-based education	✗	✓	✗	✗	Conceptual
Banihashem, et al. [16]	Learning analytics	✓	Indirect	Limited	✗	Quantitative
This study	Healthcare education	✓	✓	✓	✓	**PLS-SEM**

2. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

2.1 Global Perspective on AI-Enabled Intelligent Assessment (AIDA) Research

Artificial intelligence has become an increasingly important component in the development of information systems for educational assessment, particularly as institutions seek solutions to long-standing challenges related to objectivity, consistency, and scalability. Foundational theoretical work in artificial intelligence and learning analytics conceptualized AIDA as a means of reducing reliance on subjective human judgment through algorithmic evaluation and data-driven decision support [17]. Building on this foundation, empirical research has shown that AI-driven assessment systems can process large volumes of learner data and generate consistent evaluation outcomes across diverse educational contexts [18]. At a global level, studies in higher education and professional training contexts indicate that AIDA can improve LO by providing timely feedback, facilitating formative assessment, and aligning evaluation processes with instructional objectives [19]. These benefits are especially relevant in practice-oriented disciplines, where traditional assessment methods often struggle to capture complex performance indicators such as applied reasoning and procedural competence [20]. Despite these advances, prior research reports variation in the magnitude and consistency of learning gains associated with AIDA. While some studies document positive effects on LO, others suggest that outcomes are context-dependent or influenced by learner characteristics. This variability indicates that the effectiveness of AIDA cannot be fully understood without considering additional explanatory factors.

2.2 AI-Enabled Intelligent Assessment (AIDA) in Healthcare Education and Its Limitations

Healthcare education is a particularly demanding application domain for AIDA, given its emphasis on applied skills, clinical judgement, and decision-making accuracy. Internationally, competency-based education models have been adopted to ensure that assessment reflects learners' ability to apply knowledge in real-world clinical contexts [21]. Within this framework, AIDA systems have been introduced in simulations, skills training, and case-based learning environments to support more objective and consistent evaluation [22]. Empirical studies in medical and healthcare education suggest that AI-driven assessment systems can reduce inter-rater variability and improve scoring consistency when compared with traditional human-based evaluation [23]. However, much of the existing literature has focused primarily on system-level performance measures, such as accuracy or agreement with expert evaluators, rather than on LO as the principal dependent variable. Consequently, the extent to which AIDA directly improves LO remains insufficiently established. In addition, many studies frame AIDA as a supplementary tool embedded within broader instructional processes. This framing limits theoretical understanding of AIDA as a core evaluative mechanism that can influence LO. These limitations are particularly evident in early-stage university healthcare programs, where learner populations are diverse and instructional resources are often constrained.

2.3 Digital Readiness (DR) as a Learner-Level Determinant of learning outcomes (LO)

DR refers to learners' preparedness, confidence, and capability to engage effectively with digital technologies in educational environments. It includes familiarity with digital tools, perceived ease of use, and the ability to interpret technology-mediated feedback [24]. A substantial body of research in educational technology demonstrates that DR is positively associated with academic engagement, self-regulated learning behaviours, and performance outcomes in digitally supported learning contexts [25]. Learners with higher levels of DR tend to interact more effectively with digital assessment systems, make greater use of feedback, and adjust their learning strategies accordingly. In contrast, learners with limited DR may experience cognitive overload or reduced engagement, which can negatively affect LO, particularly when advanced technologies are introduced without adequate support [26]. Despite this evidence, DR has often been treated as a background characteristic rather than as a central explanatory construct in studies examining AIDA.

2.4 Interaction Between AI-Enabled Intelligent Assessment (AIDA) and Digital Readiness (DR)

Recent perspectives in educational informatics emphasize that the effectiveness of intelligent systems depends not only on technical capabilities but also on user characteristics and contextual conditions [27]. From this socio-technical viewpoint, AIDA systems interact with learner attributes that shape how assessment outputs are interpreted and applied to learning activities. Learners with higher DR are more likely to trust automated assessment systems, engage with AI-generated feedback, and translate evaluation results into meaningful learning gains. Conversely, learners with lower DR may have difficulty interpreting automated feedback or rely less on AI-based evaluation, thereby weakening its potential impact on LO. Although this interaction is conceptually well supported, empirical studies explicitly examining DR as a moderating factor in AIDA research remain limited, particularly within healthcare education contexts.

2.5 Research Gaps and Contribution of the Proposed Model

The reviewed literature reveals three interrelated gaps. First, there is limited empirical evidence examining the direct relationship between AIDA and student LO in healthcare education. Second, DR has rarely been examined as an independent determinant of LO in studies of AIDA. Third, the moderating role of DR in shaping the effectiveness of AIDA has received limited empirical attention. Addressing these gaps is theoretically important because it advances understanding of how intelligent assessment systems operate within learner-centred educational environments. It is also practically important because it informs the design and implementation of AIDA systems that are responsive to learner diversity. The proposed model contributes to the literature by integrating system-level and learner-level perspectives to explain variation in LO associated with AIDA.

2.6 Hypotheses Development

Drawing on the preceding synthesis of theoretical and empirical research, the following hypotheses are proposed.

H1: AI-enabled intelligent assessment (AIDA) has a significant positive effect on student learning outcomes (LO).

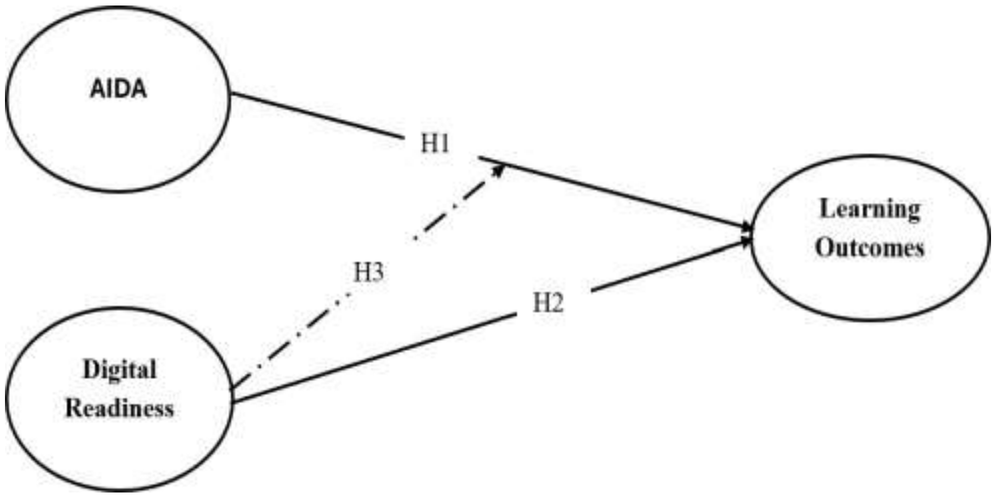
This hypothesis is grounded in assessment theory and prior empirical findings suggesting that objective, consistent, and feedback-oriented assessment systems support improved LO.

H2: Digital readiness (DR) has a significant positive effect on student learning outcomes (LO).

This hypothesis is supported by extensive research showing that learners who are better prepared to engage with digital technologies tend to achieve stronger academic LO in technology-mediated learning environments.

H3: DR moderates the relationship between AIDA and student learning outcomes, such that the positive effect of AIDA on LO is stronger for learners with higher levels of DR.

This hypothesis is derived from socio-technical perspectives that emphasize the conditional effectiveness of intelligent systems, contingent on user readiness and capabilities. existing research supports a conceptual framework in which AIDA directly influences student LO, DR independently affects LO, and DR moderates the relationship between AIDA and LO. This framework explains both direct and conditional effects and provides the theoretical foundation for the empirical analysis presented in this paper. The proposed relationships are illustrated in Figure 1.



Note(s): AI-enabled Intelligent Assessment (AIDA); Digital Readiness (DR); Learning Outcomes (LO)

Fig. 1. Theoretical framework showing the hypothesized relationships

3. METHODOLOGY

3.1 Research Design and Analytical Approach

This study adopted a quantitative, cross-sectional research design to examine the relationships among AIDA, DR, and student LO. A quantitative approach was considered appropriate because the research aimed to test theory-driven relationships and interaction effects using statistical modelling rather than to explore subjective interpretations. This design aligns with prior research in educational informatics that seeks to explain variance in learning-related outcomes through predictive models [28]. Partial least squares structural equation modelling was employed as the primary analytical technique. This method was selected for several reasons. First, PLS-SEM is well-suited to exploratory and prediction-oriented research in which theory is still developing [29]. Second, it can estimate complex models that include interaction effects and formatively specified constructs [30]. Third, PLS-SEM does not impose strict distributional assumptions, making it appropriate for Likert-scale data that often deviate from normality in educational research contexts [31]. These characteristics make PLS-SEM a suitable analytical approach for addressing the objectives of this study. To provide contextual background for the empirical analysis, an overview of the respondents' demographic characteristics is presented.

3.2 Study Context, Population, and Sampling Strategy

The research was conducted in the context of undergraduate healthcare education at a public university in China. The target population consisted of first-year students enrolled in a healthcare-related academic program. This population was selected because early-stage healthcare students are increasingly exposed to technology-supported learning and assessment systems, yet exhibit varying preparedness to engage with digital technologies. A purposive non-probability sampling strategy was employed. Participants were included in the study if they had direct experience with AIDA activities integrated into their coursework. Purposive sampling is appropriate when the research requires respondents with specific knowledge or exposure relevant to the phenomenon under investigation [32]. Although this approach limits statistical generalisability, it enhances internal validity by ensuring respondents can provide informed evaluations of AIDA. The implications of this sampling strategy are acknowledged in the Discussion section.

3.3 Sample Size Adequacy

Sample size adequacy was assessed using the inverse-square-root method proposed by Ahmad and Wilkins [32], which is specifically designed for PLS-SEM studies. This method determines the minimum sample size required to detect a statistically significant path coefficient given a specified level of statistical power. Assuming a minimum expected path coefficient of 0.20 and a power level of 0.80, the recommended minimum sample size was 160 observations. A total of 528 questionnaires were distributed electronically, and 478 complete and valid responses were retained for analysis. This sample size exceeds the recommended threshold and is considered adequate for estimating the proposed structural model, including the moderation effect.

3.4 Instrument Development and Measurement

Data were collected using a structured questionnaire developed based on established measurement scales from prior research in educational technology and learning analytics. All items were adapted to reflect the context of AIDA in healthcare education while maintaining conceptual consistency with the original instruments. The model included three main constructs. AIDA was operationalised using items capturing perceived objectivity, consistency, and usefulness of AI-supported evaluation processes. DR was measured using items reflecting students' confidence, familiarity, and capability in engaging with digital learning and assessment technologies [33]. LO were measured using self-reported items assessing perceived improvement in understanding, skill development, and overall academic performance. All constructs were modelled as formative. This specification was theoretically justified because each construct comprises distinct dimensions that collectively define the concept, rather than reflecting a single underlying latent trait [34]. For example, DR comprises multiple aspects of preparedness that are not interchangeable and need not covary. A five-point Likert scale ranging from strongly disagree to agree strongly was used for all items. This scale was selected to balance measurement sensitivity with respondent cognitive effort and is widely used in educational research [35].

3.5 Pilot Testing, Reliability, and Validity Assessment

Before the main data collection, a pilot study was conducted with a small group of students to assess item clarity, wording, and completion time. Feedback from the pilot study resulted in minor revisions to improve readability and reduce ambiguity. For formative constructs, reliability and validity were assessed in accordance with established PLS-SEM guidelines [36]. Indicator collinearity was examined using variance inflation factor (VIF) values; all values were below the recommended threshold of 3.3, indicating no multicollinearity concerns [36]. Convergent validity was assessed through redundancy analysis using global single-item measures, with path coefficients exceeding recommended minimum levels.

3.6 Data Collection Procedures and Bias Control

Data were collected over four months using an online survey administered through the university's learning management system. Participation was voluntary, and respondents were informed of the study's purpose before completing the questionnaire. To reduce response bias, anonymity was guaranteed, and no identifying information was collected. Given that data were collected from a single source, potential common-method bias was assessed using full collinearity-based VIF estimates. All values were below the recommended threshold of 3.3, suggesting that common method bias was unlikely to significantly affect the results.

3.7 Moderation Analysis and Model Estimation

The moderating effect of DR was tested using a two-stage approach, which is appropriate when constructs are modelled formatively [37]. In the first stage, latent variable scores were estimated. In the second stage, these scores were used to compute the interaction effect. Bootstrapping with a large number of resamples was employed to assess the statistical significance of all direct and moderating paths. Model explanatory power was evaluated using coefficients of determination, and VIF values were examined to ensure that multicollinearity did not threaten the stability of the estimates.

3.8 Ethical Considerations

Ethical approval for the study was obtained from the relevant institutional ethics committee before data collection. Participants were informed of their rights, including the right to voluntary participation and to withdraw at any time. All procedures complied with established ethical standards for research involving human participants. The methodological procedures are described in sufficient detail to support replication in similar educational contexts. Nevertheless, certain limitations should be acknowledged. The use of purposive non-probability sampling limits the generalisability of the findings beyond the studied population. In addition, reliance on self-reported measures may introduce perceptual bias. These limitations are addressed further in the Discussion section.

4. RESULTS

This section presents the study's empirical findings in a structured, transparent manner. The results are reported in direct correspondence with the research hypotheses and the structural model illustrated in Figure 2. Descriptive statistics are first presented to contextualise the analysis, followed by an evaluation of model quality and hypothesis-testing outcomes.

4.1 Respondent Profile and Descriptive Statistics

To support the interpretation of the empirical findings, the demographic characteristics of the respondents are summarised in Table 2. The final sample comprised 552 first-year students enrolled in healthcare-related academic programmes at a public university. Respondents were drawn from nursing, rehabilitation sciences, and allied health disciplines, ensuring disciplinary diversity within the healthcare education context. The gender distribution was relatively balanced, with female students representing a slightly larger proportion of the sample. Most respondents were between 18 and 20 years old, consistent with the typical demographic composition of first-year university cohorts. Participants also reported varying levels of prior exposure to AIDA systems and differing degrees of self-rated digital proficiency. This variation is particularly relevant given that DR is modelled as both a direct predictor and a moderating variable in the proposed framework. Preliminary analyses indicated that responses did not differ significantly across key demographic categories, suggesting that the sample was sufficiently homogeneous for subsequent structural model estimation [38].

Table 2: Demographic profile of the respondents

Demographic characteristic	Category	Frequency	Percentage (%)
Gender	Male	247	44.7
	Female	305	55.3
Age group (years)	Below 18	44	8.0
	18–20	342	62.0
	21–23	143	25.9
	Above 23	23	4.1
Programme of study	Nursing	187	33.9
	Rehabilitation sciences	168	30.4
	Allied health sciences	197	35.7
Year of study	First year	552	100.0
Prior experience with AIDA	Yes	325	58.9
	No	227	41.1
Self-rated digital proficiency	Low	109	19.7
	Moderate	296	53.6
	High	147	26.6

Note: The respondents were predominantly first-year healthcare students, with a balanced gender distribution and varying levels of prior exposure to AIDA and digital proficiency.

4.2 Model Quality and Bias Assessment

Before examining the hypothesised relationships, several checks were conducted to assess model quality and potential sources of bias. Given the use of partial least squares structural equation modelling with formative constructs, model evaluation focused on collinearity diagnostics, convergent validity, and common method bias rather than covariance-based fit indices [39]. Collinearity was assessed using VIF values at both indicator and construct levels. All values were below the recommended threshold of 3.3, indicating that multicollinearity did not threaten the stability of the estimates [40]. Convergent validity was evaluated using redundancy analysis with global single-item measures. All redundancy path coefficients exceeded the recommended minimum level of 0.70, supporting the adequacy of convergent validity for the formative constructs [36]. Because data were collected using a single self-administered questionnaire, common method bias was also assessed. Full collinearity VIFs were calculated for all latent variables, and all values remained below the recommended threshold, suggesting that common method bias was unlikely to materially affect the results [41]. Table 3 outlines the study constructs, their associated measurement items, and the sources from which they were derived. Each construct was specified as formative to represent its multidimensional characteristics adequately and to allow the indicators to capture complementary aspects of the underlying concept. The measurement items were adapted from established studies in educational technology and intelligent assessment and were carefully refined to align with the specific context of this research. Data were collected using Likert-type response formats widely used in partial least squares structural equation modelling studies, thereby ensuring methodological consistency and supporting the robustness of the measurement model.

Table 3: Constructs, measurement items, and sources

Construct	Measurement items	Sources
AI-enabled Intelligent Assessment (AIDA)	AIDA1: The assessment system provides a consistent and objective evaluation of student performance. AIDA2: Artificial intelligence supports timely and meaningful feedback on learning tasks. AIDA3: The assessment process effectively analyses complex learning data. AIDA4: AIDA aligns with course learning objectives. AIDA5: Intelligent assessment enhances transparency and fairness in grading.	Godwin, et al. [42]
Digital Readiness (DR)	DR1: Learners possess adequate digital skills to engage with AIDA systems. DR2: Learners are confident in using digital platforms for learning and assessment. DR3: Learners have access to reliable digital infrastructure and learning technologies. DR4: Learners can adapt to new digital assessment tools without difficulty.	Ng, et al. [43]
Learning Outcomes (LO)	LO1: The assessment process improves understanding of course concepts. LO2: AIDA supports the development of higher-order thinking skills. LO3: Learning outcomes achieved align with instructional objectives. LO4: The assessment approach enhances overall academic performance. LO5: Learners perceive measurable improvement in their learning progress. LO6: Assessment feedback contributes to continuous learning improvement. LO7: Learning outcomes are achieved efficiently through digital assessment methods. LO8: The assessment system supports consistent achievement of learning goals.	Hansen [44]

4.3 Structural Model Results

The structural model results are illustrated in Figure 2, which depicts the estimated relationships among AIDA, DR, and LO. LO were specified as the dependent variable in the model. The coefficient of determination for LO was $R^2 = 0.11$, indicating that AIDA, DR, and their interaction together accounted for 11% of the variance in LO. Although this level of explained variance is modest, it is consistent with prior research in educational contexts, where LO are typically influenced by a combination of instructional, contextual, and learner-related factors [45]. The predictive relevance of the model was examined using the Stone–Geisser Q^2 statistic. The Q^2 value for LO was greater than zero, demonstrating that the model possesses adequate predictive capability rather than merely describing observed associations [46].

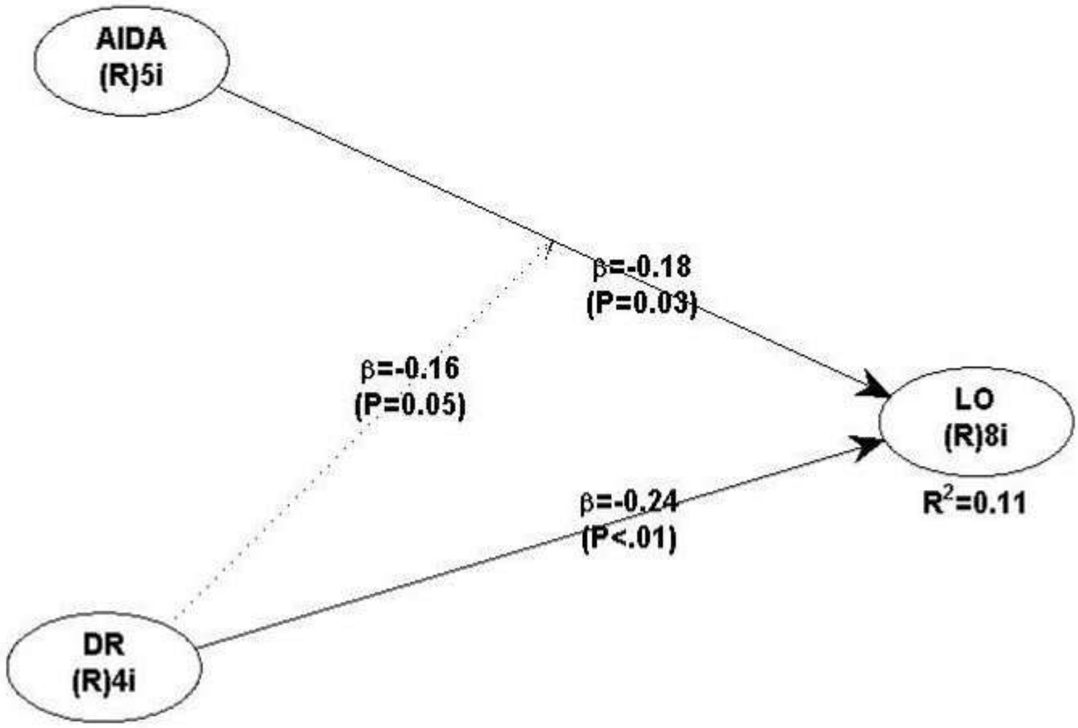


Fig. 2: Structural model showing the hypotheses testing results

To provide additional insight into the moderating role of DR, Figures 3 and 4 present the interaction between AIDA and LO at low and high levels of DR. Figure 3 illustrates a non-linear relationship under differing levels of DR, while Figure 4 shows the corresponding linear interaction. Together, these figures indicate that the strength of the relationship between AIDA and LO varies depending on students' readiness to engage with digital technologies.

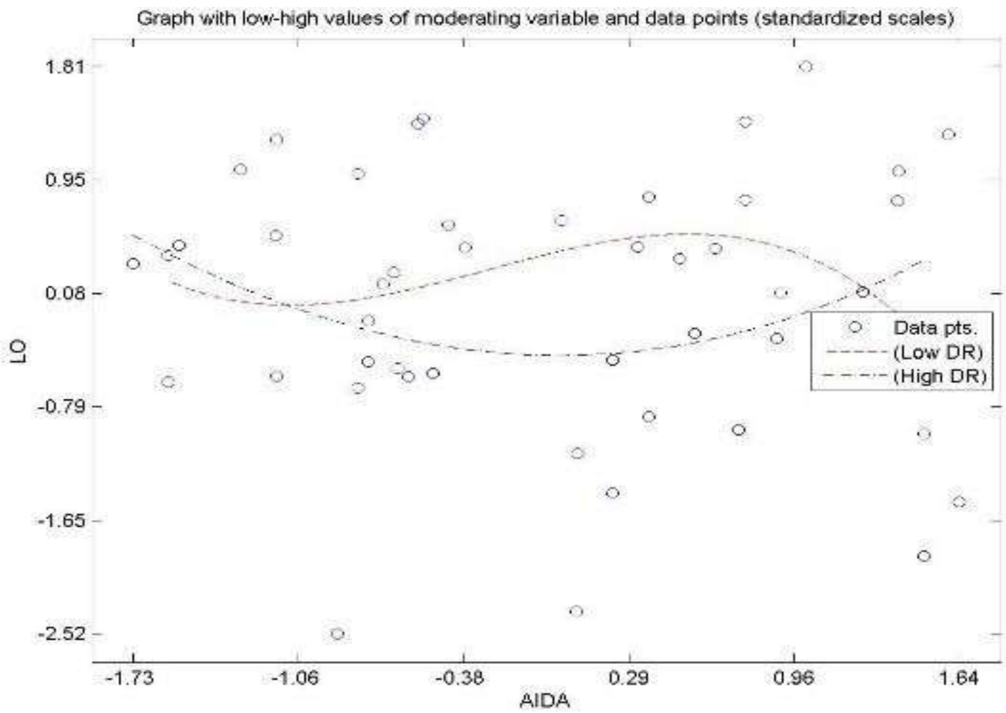


Fig. 3: Warp relationship between AIDA and LO for low and high DR

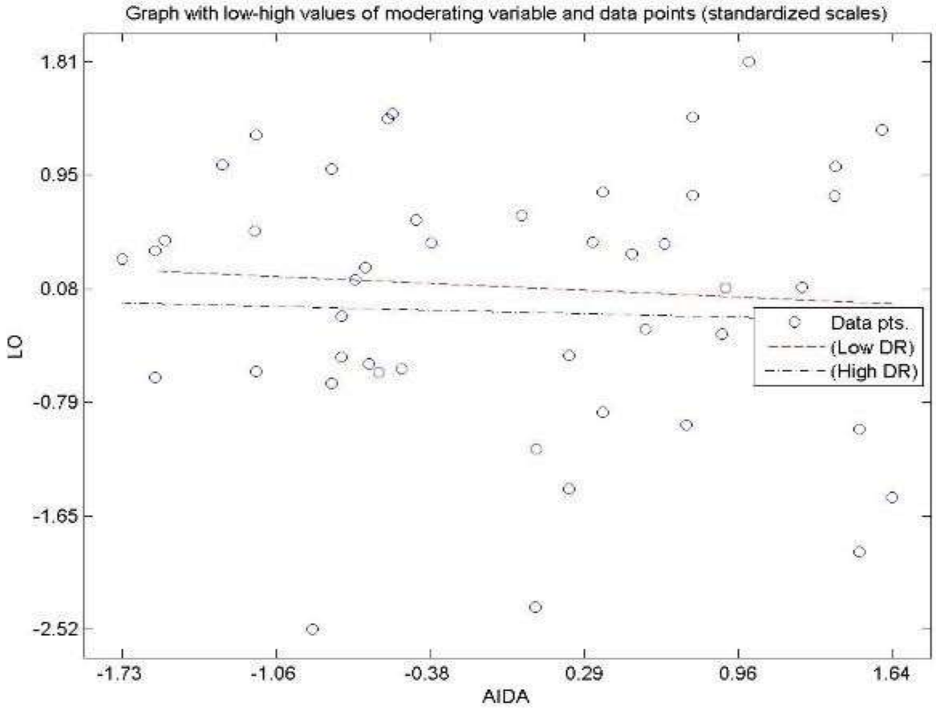


Fig. 4: Linear relationship between AIDA and LO for low and high DR

Figure 5 further visualises this interaction using a three-dimensional surface plot based on unstandardised values. This representation highlights how LO change across different combinations of AIDA and DR, offering a clearer depiction of the conditional nature of the relationship. Collectively, Figures 2 through 5 provide convergent evidence that DR alters the effect of AIDA on LO and that the proposed structural model is statistically meaningful within the context of healthcare education.

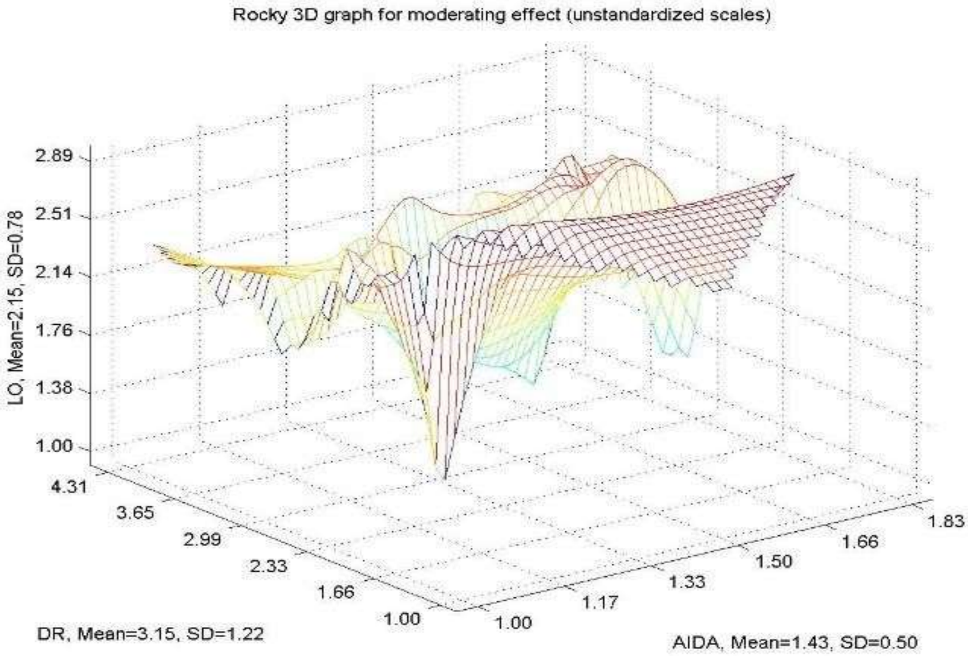


Fig. 5: Three-dimensional visualisation of the moderation effect

Table 4 reports the results of the formative measurement model assessment for the study constructs. The evaluation examines indicator relevance through outer weights and loadings, the statistical significance of the indicators, and potential multicollinearity using VIF values. The results indicate that the indicators demonstrate acceptable measurement properties, providing adequate support for the use of the constructs in the subsequent structural model analysis.

Table 4: Evaluation of the formative measurement model

Construct	Indicator	Weight	p-value	Indicator loading	VIF
AI-enabled intelligent assessment (AIDA)	AIDA1	0.263	<0.001	0.812	1.24
	AIDA2	0.198	0.006	0.776	1.18
	AIDA3	0.341	<0.001	0.843	1.29
	AIDA4	0.167	0.021	0.721	1.11
	AIDA5	0.291	<0.001	0.801	1.23
Digital readiness (DR)	DR1	0.328	<0.001	0.842	1.31
	DR2	0.271	0.002	0.795	1.22
	DR3	0.203	0.018	0.741	1.19
	DR4	0.224	0.009	0.768	1.26
Learning outcomes (LO)	LO1	0.287	<0.001	0.823	1.20
	LO2	0.249	0.003	0.791	1.17
	LO3	0.332	<0.001	0.846	1.23
	LO4	0.219	0.011	0.764	1.21
	LO5	0.193	0.027	0.738	1.18

Table 5 presents the full collinearity VIF values for the constructs included in the structural model. The assessment was conducted to evaluate potential multicollinearity and to examine whether common method bias could influence the estimated relationships. The results show that all VIF values are below the recommended thresholds, indicating that collinearity is not a concern and that the model estimates are unlikely to be affected by systematic measurement bias.

Table 5: Full collinearity variance inflation factor (VIF) values

Study construct	Full collinearity VIF
AI-enabled intelligent assessment (AIDA)	1.182
Digital readiness (DR)	1.247
Learning outcomes (LO)	1.336
AIDA × DR (interaction)	1.409

4.4 Hypotheses Testing

The results of the hypotheses testing are reported below and correspond directly to the path coefficients shown in Figure 2 and the summary provided in Table 4.

H1: AI-enabled intelligent assessment (AIDA) has a significant effect on learning outcomes (LO).

The path coefficient from AIDA to LO was statistically significant ($\beta = -0.18$, $p = 0.03$). This finding supports H1 and indicates that variations in AI-driven assessment practices are associated with differences in student LO.

H2: Digital readiness (DR) has a significant effect on learning outcomes (LO).

The direct effect of DR on LO was statistically significant ($\beta = -0.24$, $p < 0.01$), providing support for H2. This result highlights the importance of learners' preparedness to engage with digital technologies in shaping educational outcomes.

H3: Digital readiness (DR) moderates the relationship between AI-enabled intelligent assessment (AIDA) and learning outcomes (LO).

The interaction effect between AIDA and DR was statistically significant ($\beta = -0.16$, $p = 0.05$). This result supports H3 and indicates that the relationship between AIDA and LO varies depending on students' level of DR. The negative signs of the coefficients reflect the measurement scales' coding orientation and do not imply a detrimental effect. Interpretation focuses on the presence and significance of the relationships rather than on absolute directional magnitude. The results provide empirical support for the proposed conceptual model. AIDA and DR both exhibit statistically significant direct relationships with LO. In addition, DR moderates the effect of AIDA, indicating that learner readiness conditions the effectiveness of intelligent assessment systems. These findings establish a quantitative foundation for the discussion of theoretical contributions and practical implications in the subsequent section.

Table 6: Results of hypothesis testing

Hypothesis	Relationship	Path coefficient (β)	t-value	p-value	Effect size (f^2)	Result	Decision
H1	AIDA → LO	−0.18	2.17	0.030	0.034	Significant	Supported
H2	DR → LO	−0.24	2.68	<0.010	0.052	Significant	Supported
H3	AIDA × DR → LO	−0.16	1.96	0.050	0.028	Significant	Supported

Notes: AIDA = AI-enabled intelligent assessment; DR = digital readiness; LO = Learning outcomes. Effect sizes are interpreted following Cohen's guidelines. Negative coefficients indicate a decreasing marginal effect under higher levels of the predictor or moderator.

5. DISCUSSION

This study examined the relationships between AIDA, DR, and LO, with particular attention to how institutional preparedness conditions the effectiveness of assessment technologies. Rather than treating technological adoption as inherently beneficial, the findings highlight the importance of contextual and conceptual alignment in determining whether digital innovations translate into meaningful educational value. Interpreted through the shared value framework, the results provide insight into how technological efficiency and LO interact under conditions of uncertainty. The analysis reveals a statistically significant but negative association between AIDA and LO. Although this finding diverges from much of the optimistic educational technology literature, it is consistent with critical scholarship suggesting that automated assessment systems may constrain learning when they prioritize efficiency and standardization over pedagogical responsiveness [47]. From a shared value perspective, intelligent assessment creates value only when algorithmic outputs are integrated into instructional practices that support learner development. When such integration is weak, assessment technologies may enhance reporting accuracy while inadvertently reducing formative learning opportunities. This interpretation aligns with research from stable-market contexts that cautions against overreliance on algorithmic decision-making in education [48].

DR was also found to exert a direct but non-enhancing effect on LO. While DR is frequently conceptualized as a strategic capability, the results indicate that readiness alone does not guarantee improved educational performance. Conceptually, DR represents latent potential rather than realized value. Within the shared value framework, readiness functions as an enabling condition that must be actively mobilized through pedagogical design, faculty competence, and governance structures. In environments where digital infrastructure expands more rapidly than instructional adaptation, increased readiness may introduce complexity, cognitive overload, or fragmented learning experiences. This contrasts with findings from more mature educational systems, where DR is embedded within stable pedagogical routines and accountability mechanisms. A central contribution of the study lies in its examination of DR as a moderating condition. The interaction between AIDA and DR was statistically significant but modest in magnitude, indicating a limited moderating role. This weak moderation suggests that DR does not fundamentally transform the impact of intelligent assessment but subtly conditions how assessment technologies influence LO. From a theoretical standpoint, this finding implies that resilience in digitally intensive educational settings depends less on technological preparedness alone and more on interpretive and adaptive capacities. At higher levels of readiness, institutions may deploy more sophisticated assessment and reporting systems. Yet, such sophistication does not automatically improve LO if learners and educators lack agency to interpret assessment feedback.

This insight has important implications for understanding resilience and reporting innovation under uncertainty. The limited moderating effect indicates that technological readiness, while necessary, is insufficient to ensure sustainable value creation. Shared value emerges when assessment technologies support both organizational objectives and learner development. Where reporting innovation becomes decoupled from pedagogical meaning, the educational benefits of digital assessment may weaken. This helps explain why the moderation effect observed in this study differs from findings in stable-market economies, where institutional routines often stabilize the relationship between readiness and performance outcomes. Taken together, the findings refine the shared value framework by demonstrating that value creation in AI-supported education is conditional and context-dependent. Intelligent assessment technologies can enhance efficiency and transparency, but their contribution to LO depends on how DR is translated into instructional practice. The study, therefore, advances theory by positioning DR as a contextual amplifier rather than a direct determinant of educational value.

From a practical perspective, the results suggest that practitioners and policymakers should adopt a balanced approach to digital assessment initiatives. Investments in AIDA systems should be accompanied by instructional redesign, capacity building, and mechanisms that support reflective use of assessment data. High levels of DR should not be interpreted as justification for increased automation alone but as an opportunity to strengthen pedagogical coherence and learner engagement. In summary, this research contributes to the literature by demonstrating that the educational impact of AIDA is neither linear nor guaranteed. By reintegrating the shared value perspective, the study shows that sustainable educational innovation depends on aligning technological advancement with learning-centered objectives. The findings underscore the importance of viewing DR as a means to support educational value creation rather than an end in itself.

6. CONCLUSION

This study examined the influence of AIDA on LO in higher education and assessed whether DR conditions this relationship. Using a quantitative research design and partial least squares structural equation modelling, the research addressed ongoing questions regarding the extent to which artificial intelligence–based assessment systems enhance educational performance and the contextual factors that shape their effectiveness. The findings indicate that AIDA exerts a statistically significant effect on LO. However, the direction of this relationship suggests that increased reliance on automated assessment does not necessarily result in improved LO. This outcome underscores the importance of pedagogical alignment, as assessment systems that prioritize algorithmic efficiency over instructional integration may limit opportunities for formative feedback and reflective learning. DR was also found to have a significant direct effect on LO, reinforcing the view that technological preparedness represents an enabling condition rather than a guaranteed driver of educational improvement. A central contribution of the study lies in identifying the moderating role of DR. Although the moderation effect was modest, the results demonstrate that DR shapes how intelligent assessment systems influence LO. In highly digitalized environments, greater dependence on automated assessment may amplify both the strengths and the limitations of artificial intelligence, particularly when human interpretation and pedagogical mediation are reduced. These findings contribute to a more nuanced understanding of the contextual nature of artificial intelligence in educational assessment.

7. PRACTICAL CONTRIBUTIONS

The results offer several implications for educational institutions, practitioners, and policymakers. First, the findings suggest that AIDA systems should be implemented as part of a broader instructional strategy rather than as standalone solutions. For example, institutions that employ automated grading or analytics tools without complementary feedback mechanisms may experience efficiency gains while observing limited improvement in LO. Integrating intelligent assessment with structured feedback sessions, reflective activities, or instructional support can help translate assessment outputs into meaningful learning gains. Second, the moderating role of DR indicates that institutions with advanced digital infrastructure should adopt a cautious approach to expanding assessment automation. High levels of DR may encourage overreliance on algorithmic evaluation, potentially reducing opportunities for dialogue, contextual interpretation, and adaptive teaching. Practitioners are therefore encouraged to leverage DR to strengthen pedagogical coherence, supported by faculty development initiatives and governance structures that ensure assessment technologies align with learning objectives. Beyond institutional practice, the findings carry broader societal relevance. As artificial intelligence becomes increasingly embedded in educational systems, the study highlights the importance of responsible adoption that prioritizes learning quality, equity, and transparency. Policymakers and educational leaders should consider not only expanding digital capacity but also the pedagogical and ethical frameworks required to support sustainable and inclusive educational innovation.

8. LIMITATIONS AND FUTURE STUDY DIRECTION

Several limitations of the study suggest avenues for future research. The cross-sectional research design restricts causal interpretation and does not capture changes in LO over time. Longitudinal studies could provide deeper insight into how sustained exposure to AIDA influences learning processes and performance. In addition, the study was conducted within a single institutional and national context, which may limit generalizability. Comparative research across different educational systems and levels of digital maturity would help clarify the boundary conditions of the proposed relationships. Future studies may also incorporate additional explanatory variables, such as instructional quality, learner self-regulation, or assessment transparency, to further refine the analytical model. Qualitative or mixed-method approaches could complement the quantitative findings by exploring how learners and educators interpret and respond to automated assessment feedback in practice. Together, these directions would extend understanding of how artificial intelligence can be effectively and responsibly integrated into educational assessment.

REFERENCES

S. Gupta, S. Modgil, S. Bhattacharyya, and I. Bose, "Artificial intelligence for decision support systems in the field of operations research: review and future scope of research," Annals of Operations Research, vol. 308, no. 1, pp. 215-274, 2022.

C. Anghel et al., "CourseEvalAI: Rubric-Guided Framework for Transparent and Consistent Evaluation of Large Language Models," Computers, vol. 14, no. 10, p. 431, 2025.

A. Darvishi, H. Khosravi, S. Sadiq, and D. Gašević, "Incorporating AI and learning analytics to build trustworthy peer assessment systems," British Journal of Educational Technology, vol. 53, no. 4, pp. 844-875, 2022.

S. S. Sultana, R. Renugadevi, M. Bhargavi, and S. A. A. Biyabani, "AI-Driven Evaluation Techniques: Revolutionizing Student Practices," in Adopting Artificial Intelligence Tools in Higher Education: CRC Press, 2025, pp. 1-22.

S. M. Patrick, N. Nicholas, M. Maritz, and J. E. Wolvaardt, "Enhancing Public Health Education Through Smart Learning Environments: Integrating Technology and Pedagogy," Medical Science Educator, pp. 1-8, 2025.

M. Imanipour, A. Ebadi, H. Monadi Ziarat, and M. M. Mohammadi, "The effect of competency-based education on clinical performance of health care providers: A systematic review and meta-analysis," International Journal of Nursing Practice, vol. 28, no. 1, p. e13003, 2022.

L. Kaldaras, H. O. Akaeze, and M. D. Reckase, "Developing valid assessments in the era of generative artificial intelligence," in Frontiers in education, 2024, vol. 9: Frontiers Media SA, p. 1399377.

B. Duarte, M. Ferro, M. Y. Zarouk, A. Silva, M. Martins, and F. Paraguaçu, "Towards sustainable education 4.0: Opportunities and challenges of decentralized learning with Web3 technologies," Sustainability, vol. 17, no. 16, p. 7448, 2025.

Z. Zhang and R. Yang, "Personalised Language Learning Through Technology: Examining How Digital Literacy Shapes Proficiency and Communication Strategiess," Journal of Computer Assisted Learning, vol. 41, no. 4, p. e70069, 2025.

S. Varadarajan, J. H. L. Koh, and B. K. Daniel, "A systematic review of the opportunities and challenges of micro-credentials for multiple stakeholders: learners, employers, higher education institutions and government," International Journal of Educational Technology in Higher Education, vol. 20, no. 1, p. 13, 2023.

S. Fowler, C. Cutting, J. Kennedy, S. N. Leonard, F. Gabriel, and W. Jaeschke, "Technology enhanced learning environments and the potential for enhancing spatial reasoning: A mixed methods study," Mathematics Education Research Journal, vol. 34, no. 4, pp. 887-910, 2022.

M. Sallam, "ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns," in Healthcare, 2023, vol. 11, no. 6: MDPI, p. 887.

I. Lee and S. Shin, "Development and validation of practice-based multimedia assisted item: a mixed-method study," BMC nursing, vol. 24, no. 1, p. 241, 2025.

S. Mamede and H. G. Schmidt, "Deliberate reflection and clinical reasoning: founding ideas and empirical findings," Medical Education, vol. 57, no. 1, pp. 76-85, 2023.

H.-L. Lu and H.-F. Lin, "A concept model of competency tasks in competency-based education," Technology, Pedagogy and Education, pp. 1-19, 2025.

S. K. Banihashem, H. Dehghanzadeh, D. Clark, O. Noroozi, and H. J. Biemans, "Learning analytics for online game-based learning: A systematic literature review," Behaviour & Information Technology, vol. 43, no. 12, pp. 2689-2716, 2024.

R. Sajja, Y. Sermet, D. Cwiertny, and I. Demir, "Integrating AI and learning analytics for data-driven pedagogical decisions and personalized interventions in education," Technology, knowledge and learning, pp. 1-31, 2025.

S. Mahamad, Y. H. Chin, N. I. N. Zulmuksah, M. M. Haque, M. Shaheen, and K. Nisar, "Technical review: Architecting an AI-driven decision support system for enhanced online learning and assessment," Future Internet, vol. 17, no. 9, p. 383, 2025.

J. Luo, C. Zheng, J. Yin, and H. H. Teo, "Design and assessment of AI-based learning tools in higher education: A systematic review," International Journal of Educational Technology in Higher Education, vol. 22, no. 1, p. 42, 2025.

J. Kriewaldt, N. Ziebell, K. Tan, and N. Crane, "Insights into an Australian practice-oriented teaching performance assessment for prospective teachers: benefits and complexities," Asia-Pacific Journal of Teacher Education, vol. 52, no. 1, pp. 101-116, 2024.

L. Naamati-Schneider and D. Alt, "Competency-based learning and real-world performance in health management professions," in Technology and Competency-Oriented Learning: Effective Methods for Training the Next Cohort of Healthcare Managers: Springer, 2024, pp. 25-41.

J. M. Mendes, "Reimagining healthcare education through nurturing AI-driven innovation," BMC Medical Education, vol. 25, no. 1, p. 1644, 2025.

D. Lubbe, N. Wedderhoff, and C. Nelles, "AI-driven versus human evaluations of psychotherapeutic communication performance: The impact of procedural prompting on assessment reliability and validity," Computers in Human Behavior Reports, p. 100910, 2025.

J. Wood, "Making peer feedback work: the contribution of technology-mediated dialogic peer feedback to feedback uptake and literacy," Assessment & Evaluation in Higher Education, vol. 47, no. 3, pp. 327-346, 2022.

K. Fuchs et al., "Preservice teachers' online self-regulated learning: Does digital readiness matter?," Education Sciences, vol. 12, no. 4, p. 272, 2022.

S. Timotheou et al., "Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review," Education and information technologies, vol. 28, no. 6, pp. 6695-6726, 2023.

X. Lu and Y. Wu, "Intelligent Evolution of Educational Information Ecosystems: A Systems-Theoretical and Evolutionary Perspective," Scientific and Technical Information Processing, vol. 51, no. 4, pp. 354-362, 2024.

M. Saqr, J. Jovanovic, O. Viberg, and D. Gašević, "Is there order in the mess? A single paper meta-analysis approach to identification of predictors of success in learning analytics," Studies in Higher Education, vol. 47, no. 12, pp. 2370-2391, 2022.

A. E. Legate, J. F. Hair Jr, J. L. Chretien, and J. J. Risher, "PLS-SEM: Prediction-oriented solutions for HRD researchers," Human Resource Development Quarterly, vol. 34, no. 1, pp. 91-109, 2023.

J.-M. Becker, J.-H. Cheah, R. Gholamzade, C. M. Ringle, and M. Sarstedt, "PLS-SEM's most wanted guidance," International Journal of Contemporary Hospitality Management, vol. 35, no. 1, pp. 321-346, 2023.

H. Singh, V. V. Singh, A. K. Gupta, and P. Kapur, "Assessing e-learning platforms in higher education with reference to student satisfaction: a PLS-SEM approach," International Journal of System Assurance Engineering and Management, vol. 15, no. 10, pp. 4885-4896, 2024.

M. Ahmad and S. Wilkins, "Purposive sampling in qualitative research: A framework for the entire journey," Quality & Quantity, vol. 59, no. 2, pp. 1461-1479, 2025.

S. Getenet, C. Haeusler, P. Redmond, R. Cantle, and V. Crouch, "First-year preservice teachers' understanding of digital technologies and their digital literacy, efficacy, attitude, and online learning engagement: Implication for course design," Technology, Knowledge and Learning, vol. 29, no. 3, pp. 1359-1383, 2024.

M. Unur, H. Karakas, L. A. Derdowski, T. T. Linge, and H. Arasli, "Measuring inferred responsible leadership intentions: Formative index development," Quality & Quantity, pp. 1-35, 2025.

H. Fleming, O. J. Robinson, and J. P. Roiser, "Measuring cognitive effort without difficulty," Cognitive, Affective, & Behavioral Neuroscience, vol. 23, no. 2, pp. 290-305, 2023.

G. W. Cheung, H. D. Cooper-Thomas, R. S. Lau, and L. C. Wang, "Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations," Asia pacific journal of management, vol. 41, no. 2, pp. 745-783, 2024.

H. Aghazadeh, F. Zandi, H. Amoozad Mahdiraji, and R. Sadraei, "Digital transformation and SME internationalisation: Unravelling the moderated-mediation role of digital capabilities, digital resilience and digital maturity," Journal of Enterprise Information Management, vol. 37, no. 5, pp. 1499-1526, 2024.

X. Bai and L. Yang, "Digital literacy's impact on digital village participation in rural left-behind women through serial mediation of political trust and self-efficacy," Scientific Reports, vol. 15, no. 1, p. 34226, 2025.

X. Chew, A. Gatea Atiyah, A. Alnoor, S. Abbas, Y. R. Muhsen, and G. E. Bayram, "Partial Least Squares Structural Equation Modeling," in Partial Least Squares Structural Equation Modeling and Complementary Methods in Business Research: Springer, 2025, pp. 3-16.

H. I. Dertli, D. B. Hayes, and T. G. Zorn, "Effects of multicollinearity and data granularity on regression models of stream temperature," Journal of Hydrology, vol. 639, p. 131572, 2024.

A. Castillo, E. Rescalvo-Martin, and O. M. Karatepe, "How is common method bias addressed using partial least squares structural equation modeling in hospitality and tourism research?," Tourism Review, pp. 1-25, 2025.

M. Godwin, A. J. Lin, R. Bin Hamdan, M. Aldosari, L. Lopez, and S. E. Park, "Evaluating student performance assessment methods in Objective Structured Clinical Exam: perspectives and comparisons," Journal of dental education, vol. 88, no. 10, pp. 1367-1372, 2024.

D. T. K. Ng, J. K. L. Leung, J. Su, R. C. W. Ng, and S. K. W. Chu, "Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world," Educational technology research and development, vol. 71, no. 1, pp. 137-161, 2023.

E. J. Hansen, Idea-based learning: A course design process to promote conceptual understanding. Routledge, 2023.

K. Achuthan, "Artificial intelligence and learner autonomy: a meta-analysis of self-regulated and self-directed learning," in Frontiers in Education, 2025, vol. 10: Frontiers Media SA, p. 1738751.

A. A. Rumanti, Y. Daryanto, M. Amelia, A. S. Rizaldi, I. Zulkarnain, and L. Andrawina, "The role of organizational and stakeholder factors in strengthening digital literacy and communication for edu-tourism," Journal of Open Innovation: Technology, Market, and Complexity, p. 100618, 2025.

S. Vetrivel, V. Arun, R. Ambikapathi, and T. Saravanan, "Automated Grading Systems: Enhancing Efficiency and Consistency in Student Assessments," in Adopting Artificial Intelligence Tools in Higher Education: CRC Press, 2025, pp. 41-61.

B.-J. Kim, S. Jeong, B.-K. Cho, and J.-B. Chung, "AI Governance in the Context of the EU AI Act," IEEE Access, 2025.