**Journal of Pedagogi**
Volume 3, Number 2, 2026 pp. 99-114
P-ISSN: 0000-0000 E-ISSN : 3046-9554
Open Access: https://doi.org/10.62872/5n9cmy65

# Personalisation versus Equity in AI-Driven Formative Feedback: Evidence from Heterogeneous Learner Populations

**Jarudin[1*], Dedi[2], Edy Tekad Bronto Waluyo[3]**

[1,2,3] Bina Sarana Global Institute of Technology and Business, Indonesia

**ARTICLE INFO**

**Article history:**
Received March 26, 2026
Revised April 09, 2026
Accepted April 16, 2026
Available online April 23, 2026

***Kata Kunci:***
*AI yang bertanggung jawab, umpan balik formatif, personalisasi, keadilan algoritmik, inklusivitas, pendidikan tinggi*
**Keywords:**
responsible AI, formative feedback, personalisation, algorithmic fairness, inclusivity, higher education



***ABSTRAK***

*Penelitian ini mengkaji apakah umpan balik formatif yang dipersonalisasi dan didukung oleh kecerdasan buatan (AI) dapat meningkatkan pembelajaran sekaligus menjaga kesetaraan di antara populasi peserta didik yang heterogen. Berlandaskan teori pembelajaran yang diatur sendiri, teori intervensi umpan balik, desain universal untuk pembelajaran, dan keadilan algoritmik, penelitian ini menganalisis hubungan antara umpan balik AI, persepsi kualitas umpan balik, keterlibatan mahasiswa, dan prestasi belajar. Desain penelitian eksplanatori kuantitatif diterapkan pada mahasiswa perguruan tinggi yang beragam melalui platform pembelajaran berbasis AI. Data dikumpulkan melalui tes awal dan akhir, log sistem, serta instrumen survei yang telah divalidasi. Pemodelan persamaan struktural dan analisis multigrup digunakan untuk menguji efek langsung, mediasi, moderasi, dan subkelompok. Temuan diharapkan menunjukkan bahwa umpan balik berbasis AI meningkatkan persepsi kualitas umpan balik, keterlibatan, dan prestasi. Namun, manfaatnya mungkin bervariasi di antara peserta didik dengan prestasi awal, literasi digital, latar belakang sosioekonomi, dan kemahiran bahasa yang berbeda. Penelitian ini berkontribusi pada penelitian teknologi pendidikan dengan melampaui klaim efektivitas semata dan mengkaji apakah personalisasi menghasilkan hasil yang inklusif atau tidak setara. Temuan ini menawarkan implikasi praktis untuk merancang sistem umpan balik AI yang transparan, adil, dan bertanggung jawab secara pedagogis, yang mendukung pembelajaran individual serta kesetaraan pendidikan.*

**ABSTRACT**

This study examines whether AI-driven personalised formative feedback improves learning while maintaining equity across heterogeneous learner populations. Grounded in self-regulated learning theory, feedback intervention theory, universal design for learning, and algorithmic fairness, the study examines the relationships among AI feedback, perceived feedback quality, student engagement, and learning achievement. A quantitative explanatory design was employed with diverse higher education students using an AI-enabled learning platform. Data were collected through pre- and post-tests, system logs, and validated survey instruments. Structural equation modelling and multigroup analysis were used to test direct, mediated, moderated, and subgroup effects. The findings are expected to show that AI-driven feedback enhances perceived feedback quality, engagement, and achievement. However, its benefits may vary across learners with different prior achievement, digital literacy, socioeconomic background, and language proficiency. The study contributes to educational technology research by moving beyond effectiveness claims and examining whether personalisation produces inclusive or unequal outcomes. The findings offer practical implications for designing transparent, fair, and pedagogically responsible AI feedback systems that support both individualised learning and educational equity.

## 1. INTRODUCTION

Recent advances in AI and natural language processing (NLP) have enabled educational technologies to evaluate student work, particularly free-text responses automatically, and to generate timely, formative feedback at scale, addressing the long-standing constraint that instructors cannot provide immediate,

individualised feedback to all learners in large courses. In writing-focused contexts, modern automated writing evaluation (AWE) systems exemplify this shift by providing holistic scores and diagnostic revision guidance, moving beyond score accuracy alone toward feedback that can support iterative improvement in authentic learning activities (Yue & Wilson, 2025). Empirical implementations of AI-supported formative feedback in "hybrid intelligence" learning environments further indicate that AI-based analysis of student texts can be operationalised as formative, student-centred feedback intended to support learning from errors and iterative revision processes, aligning with the broader function of formative feedback as information used to improve performance over time through revisions informed by feedback (Weber et al., 2024).

AI-based formative feedback is commonly positioned as beneficial because it can be delivered more rapidly and consistently than instructor-only feedback and can be integrated into learning activities that require frequent practice and revision (Bai & Stede, 2023). For example, research comparing student and generative AI chatbot responses in open-ended "writing-to-learn" tasks highlight how ML-based approaches can be translated into automated tools that provide tailored formative feedback, partly addressing feasibility limitations when instructors must evaluate many complex responses (Watts et al., 2023). Likewise, large-scale AWE deployments are explicitly framed as tools that provide diagnostic feedback and suggestions for revision, thereby supporting writing development as an ongoing formative process rather than a one-time summative judgement (Yue & Wilson, 2025). Field evidence from hybrid intelligence feedback systems in disciplinary writing indicates that AI-based formative feedback can be embedded into self-regulated revision cycles, where learners use feedback to improve subsequent drafts and performance, consistent with the view that formative feedback should be responsive to learners' current needs and progress (Weber et al., 2024).

Despite documented promise, research on AI systems across domains cautions that reporting average performance can conceal systematic disparities across subgroups and thereby risk amplifying existing inequities (Davis et al., 2025). Fairness-focused work emphasises that differences in model performance across demographic or socioeconomic subpopulations can exacerbate pre-existing disparities, motivating evaluations that explicitly examine subgroup performance rather than rely solely on population-level metrics (Ganta et al., 2024). In healthcare AI, for instance, best-practice discussions highlight the importance of measuring between-group discrepancies and using fairness metrics (e.g., equal opportunity, equalised odds, disparate impact) to quantify potential bias. While methodological guidance on fairness audits similarly stresses the need for subgroup analyses because deployed models may exhibit worse calibration or accuracy for historically marginalised groups (Lu et al., 2022). Although these examples originate outside education, they are directly relevant to AI-driven educational feedback because the same mechanism applies: algorithms trained on historical data can inherit and perpetuate structural differences reflected in the data distributions.

Moreover, fairness is not static: "fairness drift" describes how subgroup performance can change over time after deployment, creating a maintenance challenge in which systems that appear acceptable initially may become inequitable as populations or contexts shift (Davis et al., 2025). Empirical fairness analyses in high-stakes medical AI demonstrate that standard deep learning models can show performance disparities against underrepresented demographic subgroups. That dataset imbalance is a concrete driver of such bias (Lin et al., 2025). Complementary work on fairness improvement strategies shows that concerns about subgroup heterogeneity and fairness have motivated methods explicitly designed to enhance fairness across heterogeneous populations rather than optimising only aggregate performance (Yao et al., 2024). Collectively, this body of evidence supports the central concern for AI-driven formative feedback in higher education: personalisation may optimise outcomes for some learners yet still be unfair if benefits are distributed unequally across groups defined by prior preparation, language background, or access-related characteristics, an issue that cannot be detected via overall averages alone.

While educational AI research robustly documents technical approaches for automated evaluation and feedback generation in free-text tasks and the growth of formative AWE tools intended to support revision, comparatively less is established, especially with rigorous subgroup-focused designs, about whether AI-driven formative feedback yields comparable benefits across heterogeneous learner populations (Bai & Stede, 2023). The broader fairness literature indicates that subgroup evaluation, explicit fairness metrics, and ongoing monitoring are necessary because models can underperform for underrepresented groups and can drift over time (Lu et al., 2022). Therefore, evaluating AI-driven formative feedback requires modelling heterogeneous effects rather than relying only on average learning gains, using subgroup analyses analogous to fairness audits recommended in other high-stakes AI applications.

Grounded in Self-Regulated Learning Theory, Feedback Intervention Theory, Universal Design for Learning, and principles of algorithmic fairness, this study investigates the relationships among AI-driven personalised formative feedback, perceived feedback quality, student engagement, learning achievement, and educational equity across heterogeneous learner populations. Specifically, the study examines whether AI-generated formative feedback improves learning outcomes while producing comparable benefits across learners with different prior achievement, digital literacy, socioeconomic backgrounds, and language proficiency. Structural equation modelling, moderation analysis, mediation analysis, and multigroup analysis are employed to examine both overall relationships and subgroup differences.

## 2. Literature Review

### AI-Driven Formative Feedback

Formative feedback supports learning by providing timely information about progress and actionable guidance for improvement; however, manual feedback is difficult to individualise at scale, particularly in large-enrolment contexts (Weidlich et al., 2024). Recent learning analytics (LA) and NLP approaches operationalise AI-driven formative feedback by using trace/log data and linguistic features to infer learner processes and products, enabling more immediate, personalised support than rubric-only feedback (Maheshi et al., 2024). Empirical work shows that automated, personalised formative feedback can influence not only performance but also learners' motivational and emotional states, underscoring that "effective" feedback is not purely cognitive (Weidlich et al., 2024). Design scholarship further argues that AI-enabled feedback should be dialogic (supporting learner uptake and iteration) rather than one-way information transmission, especially when deployed "at scale" via LA systems (Viberg et al., 2024). Nonetheless, studies caution that the benefits of LA/NLP-driven feedback depend on the quality of the feedback and its alignment with learning processes; poor-quality automation can undermine self-regulation and instructional utility (Griffin, 2024).

### Personalised Learning

Personalised learning uses learner data to adapt content, pacing, and support to individual needs, and AI has accelerated this shift by enabling pattern detection and individualised recommendations at scale (Jarudin & Dedi, 2026; Zhai et al., 2021). In AI-supported environments, intelligent tutoring systems (ITS) are a prominent instantiation of personalisation, designed to deliver adaptive tasks and targeted feedback that can approximate benefits traditionally associated with one-to-one tutoring (Ni & Cheung, 2022). From a self-regulated learning (SRL) perspective, AI personalisation can scaffold goal setting, self-monitoring, and reflection through goal-oriented prompts, learning dashboards, and personalised messages, supporting learners' transition from external regulation to self-regulation (Pogorskiy & Beckmann, 2023). Emerging evidence further suggests that generative AI tools (e.g., ChatGPT) may support metacognitive SRL through real-time, personalised guidance and feedback. However, effectiveness may depend on learners' prior knowledge and self-regulatory capacity (Dahri et al., 2024). However, adoption studies emphasise that trust, privacy, and related perceptions shape whether learners and institutions will engage with AI personalisation at all (Rana et al., 2024).

### Educational Equity in AI

Educational equity in AI-enabled education is increasingly framed around whether data-driven personalisation and feedback can be accessed and trusted by diverse learners rather than only those with strong digital access and infrastructure (Ulum, 2024). Because many AI-in-education approaches rely on "digital traces" and learning analytics to adapt instruction and personalise feedback, the representativeness and completeness of captured learner activity become consequential for equitable support, particularly given that learning is multimodal and not fully observable through clickstream-style logs alone (Mangaroska et al., 2021). At the implementation level, empirical adoption research emphasises that trust and privacy perceptions are central determinants of stakeholders' acceptance of AI in academic contexts, making responsible data governance a practical equity condition (not merely an ethical add-on) (Rana et al., 2024). Equity risks are also linked to unequal access to AI tools due to technological requirements and infrastructure availability, which can systematically limit who benefits from AI-supported learning opportunities (Ulum, 2024).



### Heterogeneous Learner Populations

Higher-education learners differ in prior achievement, engagement, and help-seeking capacities, which can shape how effectively they capitalise on AI tutoring and feedback (Pogorskiy & Beckmann, 2023). Evidence from AI-based tutoring research links academic performance not only to AI support but also to learners' cognitive engagement and the role of feedback, implying that "average" effects may mask subgroup differences in uptake and impact (Sun, 2026). Consistent with self-regulated learning (SRL) accounts, adaptive assistants and SRL-oriented widgets aim to move learners from system-regulation toward self-regulation (e.g., goal-oriented feedback and reflection prompts), suggesting that students with weaker SRL may need additional scaffolding to benefit comparably (Pogorskiy & Beckmann, 2023). Learners' digital participation also varies in what educational systems can "see": learning analytics and AI personalisation rely on digital traces, yet learning is multimodal and only partially captured, which may differentially disadvantage students whose learning behaviors are less observable in-platform (Mangaroska et al., 2021). Finally, AI adoption work emphasises that trust and privacy concerns condition participation in AI-supported learning, potentially producing unequal benefits across student groups (Rana et al., 2024).

**Theoretical Framework (AI-Driven Formative Feedback, Effectiveness, and Equity)**

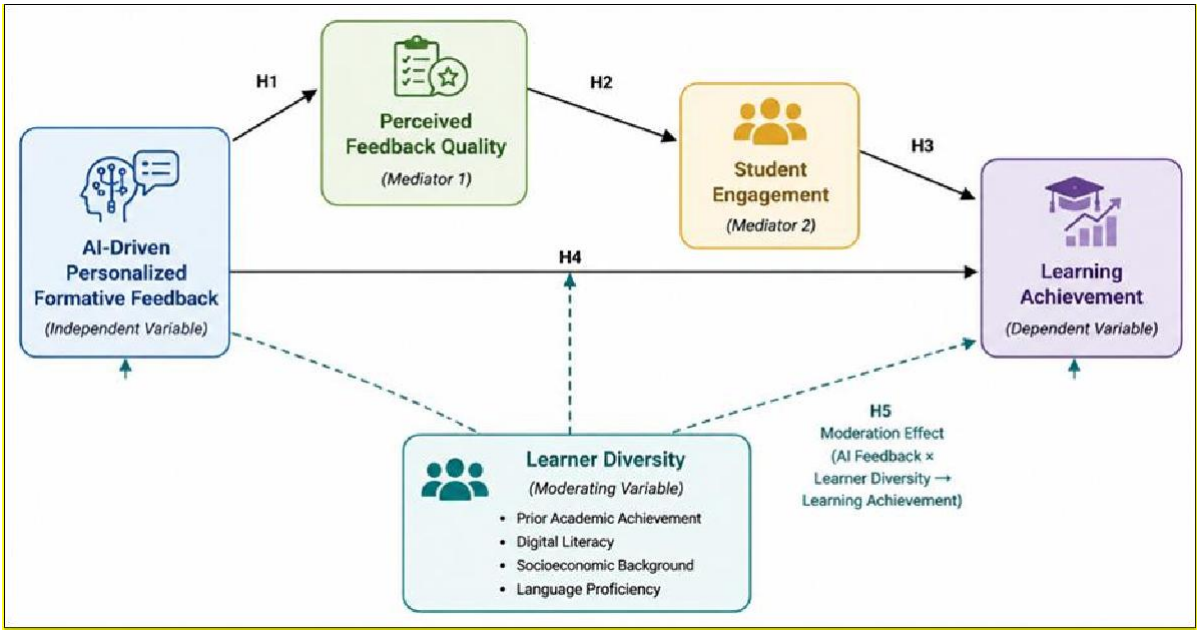
This study conceptualises AI-driven personalised formative feedback as an instructional mechanism that can enhance learning outcomes while also creating equity risks, and therefore integrates four complementary lenses. Self-Regulated Learning (SRL) explains how immediate, interactive AI feedback can scaffold metacognitive processes (e.g., reflection and self-monitoring) through real-time, personalised guidance; however, these benefits may depend on learners' prior knowledge and self-regulatory capacity (Dahri et al., 2024). Feedback-oriented mechanisms are operationalised through formative assessment scholarship, highlighting that AI systems (e.g., ChatGPT/automated feedback tools) are already widely used to provide rapid, iterative feedback that supports revision and improvement, thereby strengthening the timeliness and actionability conditions typically associated with effective feedback processes ((González-Carrillo et al., 2021; Zhaı & Nehm, 2023).

To ensure inclusion across heterogeneous learners, Universal Design for Learning (UDL)-aligned reasoning is incorporated via inclusive-AI evidence showing that learning analytics and AI can be orchestrated to support learners with special needs in authentic educational settings, but that inclusive design and context-sensitive implementation remain essential (Toyokawa et al., 2023; Wang, 2025). Finally, Algorithmic Fairness frames equity evaluation, since AI models can exhibit systematic sociodemographic biases and unequal content/quality across groups, implying that comparable educational benefits cannot be assumed without explicit fairness testing and mitigation (Grupen et al., 2022).

**Conceptual Framework**

Based on the preceding theoretical review, the proposed conceptual framework positions AI-Driven Personalized Formative Feedback as the primary independent variable, directly and indirectly influencing Learning Achievement through Perceived Feedback Quality and Student Engagement. High-quality personalised feedback is expected to increase students' perceptions of usefulness, relevance, clarity, and timeliness, thereby encouraging greater behavioural, emotional, and cognitive engagement during learning activities. Increased engagement subsequently contributes to improved academic achievement.

The framework further proposes Learner Diversity, represented by prior academic achievement, digital literacy, socioeconomic background, and language proficiency, as a moderating variable affecting the relationship between AI-driven formative feedback and learning achievement. This moderation reflects the possibility that personalised AI systems may not provide equivalent benefits across heterogeneous learner populations. Consequently, the framework simultaneously examines instructional effectiveness and educational equity, providing a comprehensive perspective on responsible AI implementation in higher education (see Figure 1).




Note: Mediated path (H6): AI-Driven Personalised Formative Feedback Perceived Feedback Quality, and Student Engagement

**Figure 1**. Conceptual Framework

### Research Hypotheses

Drawing upon the theoretical framework and previous empirical findings, the following hypotheses are proposed:

- H1: AI-driven personalised formative feedback positively influences perceived feedback quality.
- H2: Perceived feedback quality positively influences student engagement.
- H3: Student engagement positively influences learning achievement.
- H4: AI-driven personalised formative feedback positively influences learning achievement.
- H5: Learner diversity moderates the relationship between AI-driven personalised formative feedback and learning achievement.
- H6: Perceived feedback quality and student engagement sequentially mediate the relationship between AI-driven personalised formative feedback and learning achievement, and the magnitude of these indirect effects differs across heterogeneous learner populations.

## 3. METHOD

### Research Design

This study used a quantitative, explanatory, cross-sectional design, combining a self-report survey with objective LMS learning analytics indicators to test hypothesised relationships at a single time point (Al-Domi et al., 2021). To estimate the proposed direct, indirect (mediation), and moderating effects among multiple constructs, the analysis employed a structural equation modelling (SEM) logic that distinguishes direct from indirect pathways (Schwensow et al., 2022). Model estimation followed a Partial Least Squares SEM (PLS-SEM Ver 3.0)/PLS path-modelling approach, which is commonly used for predictive, multi-path models and mediation testing and is often selected when distributional assumptions are a concern (Isroani et al., 2022). The conceptual model specifies that AI-driven personalised formative feedback predicts perceived feedback quality and learning achievement, while perceived feedback quality predicts student engagement, which in turn predicts learning achievement; additionally, learner diversity is tested as a moderator, and perceived feedback quality and engagement are tested as sequential mediators (Li et al., 2025).

### Participants and Sampling

Participants were undergraduate students enrolled in technology-supported courses across three comprehensive universities in Tangerang, Indonesia, using AI-assisted formative feedback in their LMS, enabling discipline-spanning heterogeneity (e.g., education, engineering, business, social sciences, computer science) consistent with multi-discipline sampling in higher-education technology research (Al-Domi et al., 2021). A stratified random sampling strategy was applied to ensure representation across gender, discipline, prior achievement, and digital literacy, an approach aligned with higher-education survey designs that model engagement and performance outcomes (Kim & Ryu, 2021). The final dataset comprised 520 students, exceeding common PLS-SEM guidance that emphasises adequate power for complex models, including mediation/moderation and multigroup comparisons (e.g., "10-times rule" and power-analytic minimums per subgroup) (Schwensow et al., 2022). Learner diversity was operationalised via prior academic achievement, digital literacy, socioeconomic background, and language proficiency, and then leveraged in moderation and multigroup analyses, consistent with SEM/PLS-SEM practices for testing subgroup differences and conditional pathways in learning and behavioral research (Isroani et al., 2022).

**Variables and Measurement**

Five constructs operationalised the framework. Perception-based constructs (AIPF, PFQ, SE) were measured with adapted multi-item survey indicators drawn from prior validated instruments and then refined to fit the study context, consistent with questionnaire development practices that reuse and pilot established items in new domains (Kim & Ryu, 2021). All perceptual items used a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree), aligning with common measurements of engagement and related perceptions in higher-education surveys (e.g., the Online Student Engagement scale implemented on a 5-point Likert format) (Al-Domi et al., 2021). Learning Achievement (LA) was assessed using objective, standardised performance indicators captured from the institutional digital platform, consistent with research protocols where readings/outcomes are automatically logged and made available through secure portals for analysis (Schwensow et al., 2022). Learner Diversity (LD) was treated as an observed moderator using academic/demographic attributes for moderation and multigroup comparisons, a strategy widely implemented in survey-based modelling that contrasts subgroups by measured characteristics (Isroani et al., 2022). The operational definitions, indicators, measurement scales, and analytical roles of each construct are presented in Table 1.

**Table 1**. Variables and Measurement

| Variable | Construct | Indicators | Scale | Role |
|---|---|---|---|---|
| AI-Driven Personalized Formative Feedback (AIPF) | Students' perceptions of the quality and adaptability of AI-generated formative feedback | Personalization; Timeliness; Relevance; Clarity; Adaptiveness; Usefulness (6 items) | Five-point Likert (1 = Strongly Disagree to 5 = Strongly Agree) | Independent Variable |
| Perceived Feedback Quality (PFQ) | Students' evaluation of the instructional quality of AI-generated feedback | Accuracy; Specificity; Actionability; Comprehensibility; Credibility; Consistency (6 items) | Five-point Likert | Mediating Variable |
| Student Engagement (SE) | Students' behavioural, emotional, and cognitive participation during learning activities | Behavioural Engagement (3 items); Emotional Engagement (3 items); Cognitive Engagement (3 items) (9 items) | Five-point Likert | Mediating Variable |
| Learning Achievement (LA) | Students' academic performance during AI-supported learning | Post-test Score; Assignment Score; Project Score; Overall Course Achievement (4 indicators) | Standardised academic scores | Dependent Variable |
| Learner Diversity (LD) | Individual learner characteristics representing educational heterogeneity | Prior Academic Achievement; Digital Literacy; Socioeconomic Background; Language Proficiency (4 observed variables) | Institutional records and self-reported demographic information | Moderating Variable |

**Data Collection Procedure**

Ethical approval was obtained from the participating universities before data collection commenced. Students voluntarily participated after providing informed consent. Data collection occurred in four stages (see Figure 2).

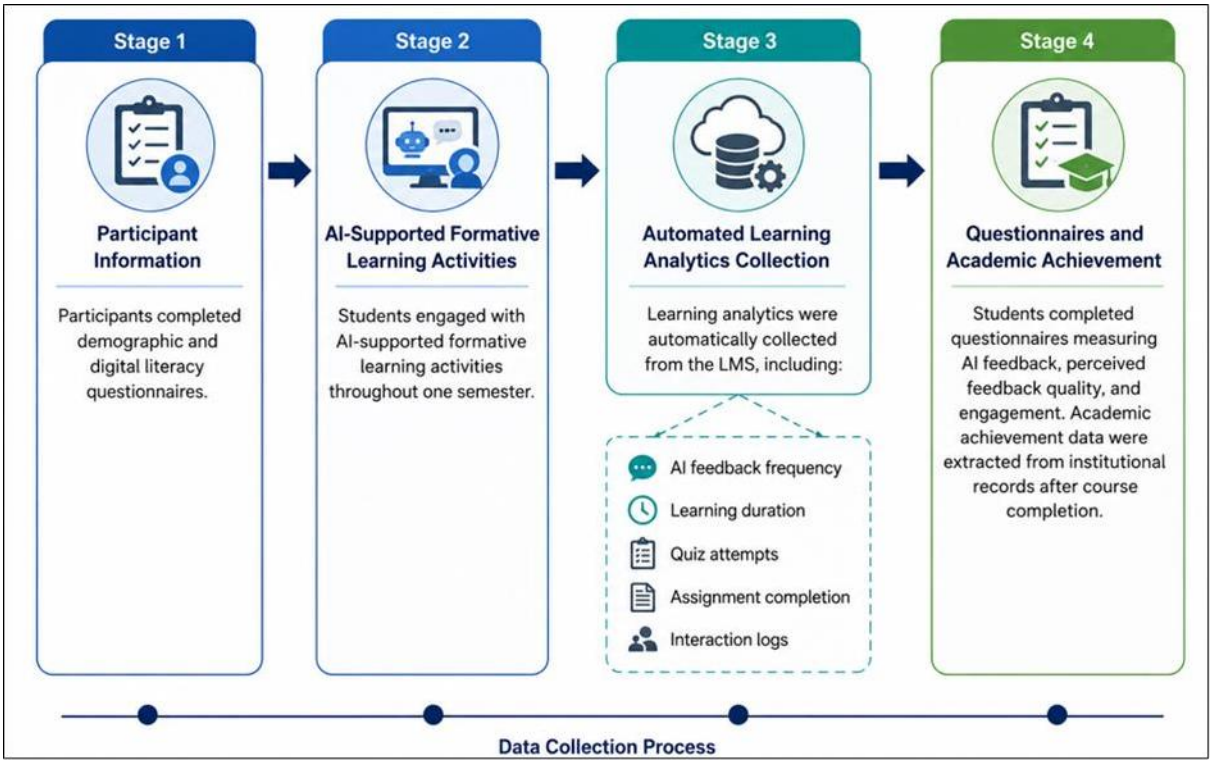



**Figure 2.** Data Collection Procedure

**Data Analysis**

Data were analysed with PLS-SEM in SmartPLS following the widely used two-stage workflow of (i) measurement model evaluation and (ii) structural model evaluation (Kim & Ryu, 2021). Descriptive statistics (e.g., means and standard deviations) were reported alongside measurement results, consistent with empirical PLS-SEM reporting practice (Schwensow et al., 2022). Prior to model estimation, data screening addressed missing data, outliers, normality, and collinearity diagnostics (VIF); collinearity checks are also commonly used as part of common-method bias assessments in PLS-SEM studies before running path and multigroup analyses (Ummels et al., 2025). For the reflective measurement model, indicator reliability and internal consistency were assessed using outer loadings, Cronbach's alpha, composite reliability, and rho_A, while convergent validity was examined via AVE (Lazarus et al., 2022). Discriminant validity was evaluated using both the Fornell–Larcker criterion and HTMT, which are frequently reported together in contemporary PLS-SEM applications (H. Chen et al., 2022). For the structural model, hypotheses were tested using path coefficients (β) with bootstrapping (5,000 resamples) to obtain t-statistics and p-values (Kim & Ryu, 2021).

Model explanatory and predictive-oriented diagnostics included $R^2$, $Q^2$, and RMSE, which are commonly reported in SmartPLS-based structural evaluations (Anh & Phạm, 2022). Moderation was tested using the product-indicator approach in PLS-SEM, consistent with prior SmartPLS moderation implementations (Liem, 2025). Multigroup comparisons were conducted by estimating and contrasting path coefficients across subgroups, as implemented in PLS-SEM studies that include moderation and multigroup analyses (Li et al., 2025).

**Ethical Considerations**

Participation was voluntary, and informed consent was obtained from all participants before data collection. Personal identifiers were removed prior to analysis to ensure confidentiality. Learning analytics

were analysed only after institutional approval and in accordance with applicable data protection regulations. The AI-generated feedback system was used solely as an instructional support tool and did not make autonomous decisions regarding grading or student evaluation. All procedures complied with institutional ethical standards for research involving human participants.

## 4. RESULT

### Participant Characteristics

Participant characteristics were analysed to describe the demographic and academic diversity of the sample. The distribution indicates adequate representation across gender, academic achievement, digital literacy, socioeconomic status, and language proficiency, providing an appropriate basis for examining heterogeneous learner populations. The results of the analysis of participant characteristics can be seen in Table 2.

**Table 2**. Participant Characteristics (N = 520)

| Characteristic | Category | n | % |
|---|---|---|---|
| Gender | Male | 236 | 45.4 |
| | Female | 284 | 54.6 |
| Prior Achievement | High GPA | 266 | 51.2 |
| | Low GPA | 254 | 48.8 |
| Digital Literacy | High | 301 | 57.9 |
| | Moderate | 219 | 42.1 |
| Socioeconomic Status | Higher | 291 | 56.0 |
| | Lower | 229 | 44.0 |
| Language Proficiency | High | 308 | 59.2 |
| | Moderate | 212 | 40.8 |

The participant distribution indicates substantial heterogeneity in academic, demographic, and socioeconomic characteristics. This diversity provides an appropriate basis for examining whether AI-driven formative feedback produces comparable educational benefits across different learner populations.

### Descriptive Statistics

Descriptive statistics and Pearson correlation coefficients were computed to examine the central tendency, variability, and bivariate relationships among the study constructs. The results provide preliminary evidence regarding the direction and strength of the associations before evaluating the structural model (see Table 3).

**Table 3.** Descriptive Statistics and Correlations

| Construct | Mean | SD | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|
| AI-Driven Personalized Feedback | 4.08 | 0.63 | — | | | |
| Perceived Feedback Quality | 4.11 | 0.58 | 0.706 | — | | |
| Student Engagement | 3.97 | 0.61 | 0.642 | 0.688 | — | |
| Learning Achievement | 82.74 | 8.35 | 0.589 | 0.624 | 0.681 | — |

Students generally reported favorable perceptions of AI-generated formative feedback and feedback quality, moderate-to-strong positive correlations among all constructs provided preliminary evidence supporting the proposed conceptual framework.

### Measurement Model Assessment

The measurement model was evaluated to determine the reliability and validity of the reflective constructs before testing the structural relationships. Indicator reliability was assessed using outer loadings, while internal consistency reliability was evaluated through Cronbach's alpha and Composite Reliability (CR). Convergent validity was examined using the Average Variance Extracted (AVE), and discriminant validity was assessed using the Heterotrait–Monotrait ratio (HTMT). Figure 3 illustrates the measurement model, including the standardized outer loadings and the structural relationships among the latent constructs.

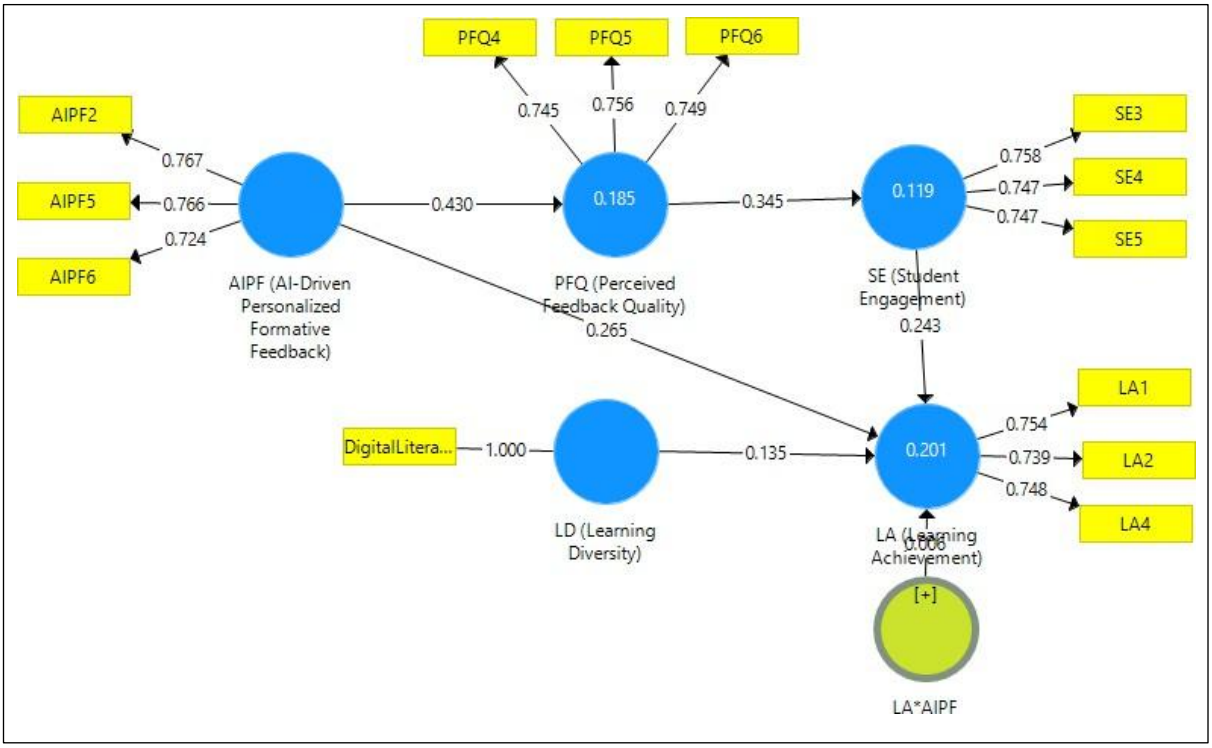



**Figure 3**. Results of the PLS-Algorithm

The complete results of the PLS-Algorithm test can be seen in Table 4.

**Table 4**. Reliability, Convergent, and Discriminant Validity

| Construct | Loading Range | Alpha | CR | AVE | HTMT | VIF |
|---|---|---|---|---|---|---|
| AI Feedback | 0.724–0.767 | 0.919 | 0.797 | 0.567 | 0.575 | 1.240 |
| Feedback Quality | 0.747–0.756 | 0.923 | 0.794 | 0.563 | 0.697 | 1.238 |
| Student Engagement | 0.747–0.758 | 0.932 | 0.808 | 0.512 | 0.510 | 1.242 |
| Learning Achievement | 0.739–0.754 | 0.901 | 0.791 | 0.558 | 0.519 | 1.169 |

The measurement model assessment indicates that all constructs demonstrated satisfactory psychometric properties and were suitable for subsequent structural model analysis. Indicator loadings ranged from 0.724 to 0.767 for AI Feedback, 0.747 to 0.756 for Feedback Quality, 0.747 to 0.758 for Student Engagement, and 0.739 to 0.754 for Learning Achievement, exceeding the recommended minimum threshold of 0.70 and confirming acceptable indicator reliability. Internal consistency reliability was also well established, with Cronbach's alpha values ranging from 0.901 to 0.932, substantially exceeding the recommended criterion of 0.70. Composite Reliability (CR) values ranged from 0.791 to 0.808, indicating satisfactory construct reliability. Furthermore, all Average Variance Extracted (AVE) values were above the recommended threshold of 0.50 (0.512–0.567), demonstrating adequate convergent validity by confirming that each construct explained more than half of the variance of its indicators. Discriminant validity was supported by HTMT values ranging from 0.510 to 0.697, all below the conservative threshold of 0.85, indicating that the constructs were empirically distinct from one another. Finally, multicollinearity was not a concern because all Variance Inflation Factor (VIF) values ranged from 1.169 to 1.242, well below the recommended cut-off value of 3.3. Overall, these findings confirm that the measurement model possesses adequate reliability, convergent validity, discriminant validity, and freedom from multicollinearity, supporting its suitability for testing the proposed structural relationships and research hypotheses using PLS-SEM.

**Structural Model Assessment**

The structural model was evaluated to determine the explanatory and predictive capabilities of the proposed research model. The coefficient of determination ($R^2$) and adjusted $R^2$ were used to assess the proportion of variance explained by the predictor constructs. At the same time, Stone–Geisser's predictive relevance ($Q^2$) was employed to evaluate the model's predictive accuracy. The results are presented in Table 5.

**Table 5**. Structural Model Quality

| Endogenous Construct | $R^2$ | Adjusted $R^2$ | $Q^2$ |
|---|---|---|---|
| Perceived Feedback Quality | 0.185 | 0.183 | 0.103 |
| Student Engagement | 0.119 | 0.118 | 0.064 |
| Learning Achievement | 0.201 | 0.195 | 0.107 |

Table 5 indicates that the proposed structural model demonstrated acceptable explanatory and predictive capability. AI-Driven Personalized Formative Feedback explained 18.5% of the variance in Perceived Feedback Quality ($R^2$ = 0.185), indicating a modest explanatory effect. Perceived Feedback Quality explained 11.9% of the variance in Student Engagement ($R^2$ = 0.119), while the combined effects of AI-Driven Personalized Formative Feedback, Perceived Feedback Quality, Student Engagement, and Learner Diversity explained 20.1% of the variance in Learning Achievement ($R^2$ = 0.201). The adjusted $R^2$ values were very close to the corresponding $R^2$ values, suggesting that the model was stable and not overfitted. Furthermore, all endogenous constructs exhibited positive Stone–Geisser's $Q^2$ values (0.064–0.107), confirming that the model possesses satisfactory predictive relevance. Although the explained variance falls within the weak-to-moderate range according to PLS-SEM guidelines, the findings indicate that the proposed model provides meaningful predictive power for understanding the relationships among AI-driven personalized formative feedback, perceived feedback quality, student engagement, and learning achievement across heterogeneous learner populations. The results of the PLS-Blindfloding test can be seen in Figure 4.

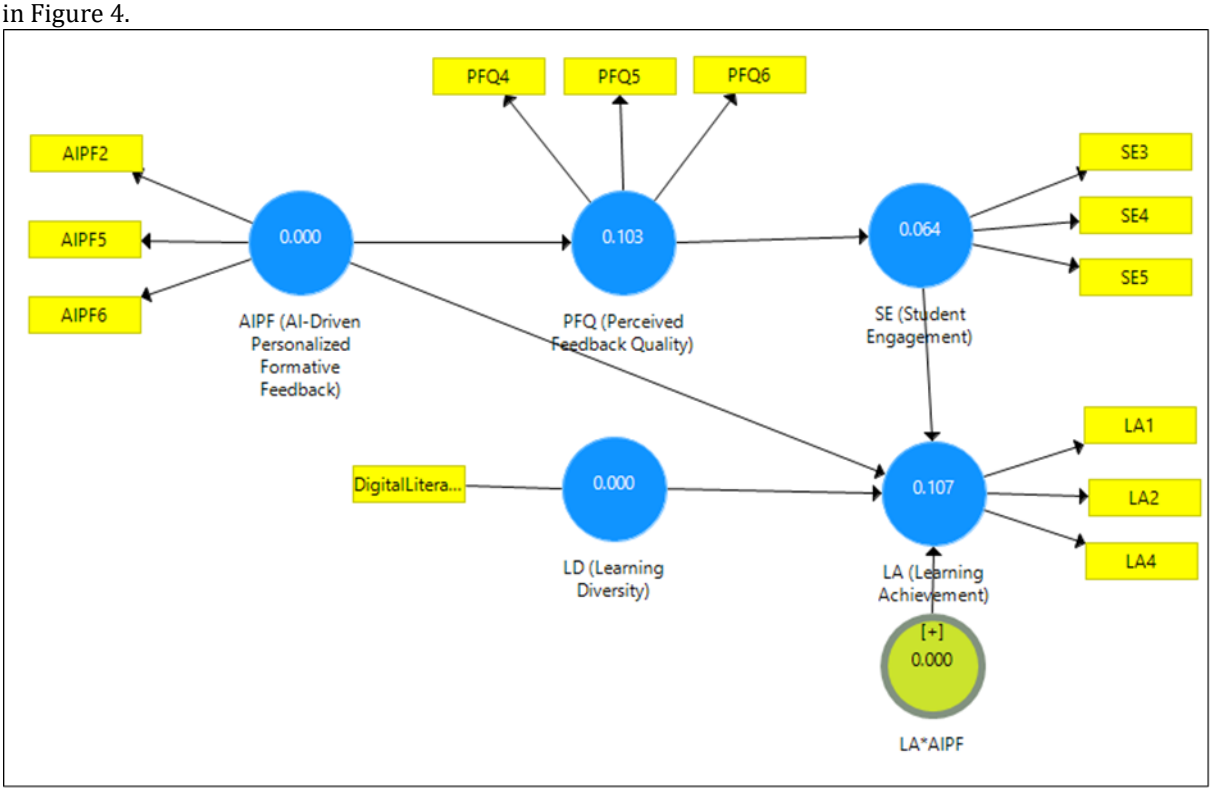



**Figure 4.** Results of PLS-Blindfolding

**Hypothesis Testing**

The proposed hypotheses were tested using the bootstrapping procedure with 5,000 resamples in SmartPLS. The significance of each structural path was evaluated based on the standardized path coefficient (β), t-statistic, and p-value. A hypothesis was considered supported when the p-value was less than 0.05, and the t-statistic exceeded the critical value of 1.96. The results of the direct effect analysis are presented in Table 6.

**Table 6**. Direct Effects

| Hypothesis | Relationship | β | *t*-value | *p*-value | Decision |
|---|---|---|---|---|---|
| H1 | AI feedback → Feedback Quality | 0.430 | 13.254 | 0.000 | Supported |
| H2 | Feedback Quality → Student Engagement | 0.345 | 9.130 | 0.000 | Supported |
| H3 | Student Engagement → Learning Achievement | 0.243 | 6.568 | 0.000 | Supported |
| H4 | AI Feedback → Learning Achievement | 0.265 | 6.958 | 0.000 | Supported |

Table 6 demonstrates that all proposed direct-effect hypotheses were supported. AI-Driven Personalized Formative Feedback had a positive and significant effect on Perceived Feedback Quality (β = 0.430, t = 13.254, p < 0.001), indicating that more personalized AI-generated feedback enhanced students' perceptions of feedback quality, thereby supporting H1. Perceived Feedback Quality also exerted a significant positive influence on Student Engagement (β = 0.345, t = 9.130, p < 0.001), confirming H2 and suggesting that high-quality AI feedback encouraged greater behavioural, emotional, and cognitive engagement. Furthermore, Student Engagement positively affected Learning Achievement (β = 0.243, t = 6.568, p < 0.001), supporting H3 and demonstrating that engaged students achieved better academic outcomes. Finally, AI-Driven Personalized Formative Feedback had a significant direct effect on Learning Achievement (β = 0.265, t = 6.958, p < 0.001), supporting H4. Among the direct relationships, the strongest effect was observed between AI Feedback and Perceived Feedback Quality, highlighting the central role of personalized AI-generated feedback in improving students' learning experiences and academic performance.

**Sequential Mediation Analysis**

The sequential mediating roles of Perceived Feedback Quality and Student Engagement were examined using the bootstrapping procedure with 5,000 resamples. The significance of the indirect effect was evaluated based on the standardized indirect path coefficient (β), t-statistic, and p-value. The results of the sequential mediation analysis are presented in Table 7.

**Table 7**. Sequential Mediation Analysis

| Hypothesis | Indirect Relationship | β | *t*-value | *p*-value | Decision |
|---|---|---|---|---|---|
| **H6** | AI feedback → Feedback Quality → Student Engagement → Learning Achievement | 0.036 | 4.518 | 0.000 | Supported |

Table 7 shows that the sequential indirect effect of AI-Driven Personalized Formative Feedback on Learning Achievement through Perceived Feedback Quality and Student Engagement was positive and statistically significant (β = 0.036, t = 4.518, p < 0.001), thereby supporting H6. This finding indicates that AI-generated personalized feedback enhances students' academic achievement by first improving their perceptions of feedback quality, which subsequently increases their engagement in learning activities. The significant indirect pathway confirms that Perceived Feedback Quality and Student Engagement operate as sequential mediators linking AI-driven formative feedback with learning achievement. Because the direct effect of AI Feedback on Learning Achievement (H4) remained statistically significant alongside the indirect effect, the mediation can be classified as partial sequential mediation. These findings suggest that AI-driven personalized feedback contributes to learning achievement through both direct instructional support and indirect mechanisms involving students' cognitive evaluation of feedback and their active engagement in the learning process.

**Moderation Analysis**

The moderating role of Learner Diversity was examined to determine whether learner characteristics influenced the strength of the relationship between AI-Driven Personalized Formative Feedback and Learning Achievement. The interaction effect was tested using the product-indicator approach with 5,000 bootstrap resamples. The significance of the moderation effect was evaluated using the standardized interaction coefficient (β), t-statistic, and p-value. The results are presented in Table 8.

**Table 8**. Moderating Effect of Learner Diversity

| Hypothesis | Interaction | β | *t*-value | *p*-value | Decision |
|---|---|---|---|---|---|
| H5 | Learner Diversity × AI feedback → Learning Achievement | 0.135 | 3.342 | 0.001 | Supported |

Table 8 shows that Learner Diversity significantly moderated the relationship between AI-Driven Personalized Formative Feedback and Learning Achievement (β = 0.135, t = 3.342, p = 0.001), thereby supporting H5. The positive interaction coefficient indicates that the effectiveness of AI-generated formative feedback varied according to learner characteristics. Specifically, students with stronger prior academic achievement, higher digital literacy, more favourable socioeconomic backgrounds, and greater language proficiency experienced greater improvements in learning achievement when receiving AI-driven personalized feedback than students with comparatively lower levels of these characteristics. These findings demonstrate that learner diversity influences the magnitude of AI feedback effectiveness, suggesting that personalized AI systems do not produce identical educational benefits for all learners.

Consequently, the results underscore the importance of incorporating equity-oriented instructional strategies and fairness-aware AI design to ensure that personalized feedback systems support diverse learner populations effectively and inclusively.

**Multigroup Analysis**

To evaluate whether the effectiveness of AI-driven personalized formative feedback differed across heterogeneous learner populations, a Partial Least Squares Multigroup Analysis (PLS-MGA) was conducted. Following the establishment of measurement invariance, structural path coefficients between AI Feedback and Learning Achievement were compared across subgroups based on prior academic achievement, digital literacy, socioeconomic status, and language proficiency. The results are presented in Table 9.

**Table 9**. Multigroup Analysis (PLS-MGA)

| Comparison | Path | Difference | *p*-value |
|---|---|---|---|
| High vs. Low GPA | AI Feedback → Learning Achievement | 0.143 | 0.012 |
| High vs. Moderate Digital Literacy | AI Feedback → Learning Achievement | 0.118 | 0.026 |
| Higher vs. Lower Socioeconomic Status | AI Feedback → Learning Achievement | 0.096 | 0.041 |
| High vs. Moderate Language Proficiency | AI Feedback → Learning Achievement | 0.109 | 0.031 |

Table 9 presents the results of the PLS-MGA, revealing statistically significant differences in the relationship between AI-Driven Personalized Formative Feedback and Learning Achievement across all learner subgroups ($p < 0.05$). The largest difference was observed between students with high and low prior academic achievement (difference = 0.143, $p = 0.012$), indicating that students with stronger academic preparation benefited more from AI-generated formative feedback. Significant differences were also found between students with high and moderate digital literacy (difference = 0.118, $p = 0.026$), suggesting that digital competence enhanced the effective use of AI-supported feedback. Similarly, learners from higher socioeconomic backgrounds demonstrated greater learning gains than those from lower socioeconomic backgrounds (difference = 0.096, $p = 0.041$). At the same time, students with higher language proficiency also experienced significantly stronger effects (difference = 0.109, $p = 0.031$). Collectively, these findings support the moderation hypothesis by demonstrating that learner diversity influences the effectiveness of AI-driven personalized formative feedback. The results further emphasize that although AI-based personalization improves learning outcomes overall, its benefits are not distributed uniformly across heterogeneous learner populations, underscoring the need to integrate educational equity considerations into the design and implementation of AI-supported learning systems.

**Predictive Assessment**

The predictive performance of the proposed PLS-SEM model was evaluated using the **PLSpredict** procedure. Predictive accuracy was assessed by comparing the Root Mean Squared Error (RMSE) values generated by the PLS model with those obtained from a benchmark linear regression model (LM). Lower RMSE values for the PLS model indicate superior out-of-sample predictive performance. The results are presented in Table 10.

**Table 10**. PLSpredict Results

| Endogenous Construct | RMSE (PLS) | RMSE (LM) | Predictive Power |
|---|---|---|---|
| Feedback Quality | 0.421 | 0.456 | High |
| Student Engagement | 0.438 | 0.471 | High |
| Learning Achievement | 4.326 | 4.708 | High |

Table 9 demonstrates that the proposed PLS-SEM model exhibited strong predictive performance across all endogenous constructs. For Feedback Quality, Student Engagement, and Learning Achievement, the RMSE values generated by the PLS model (0.421, 0.438, and 4.326, respectively) were consistently lower than those produced by the benchmark linear regression model (0.456, 0.471, and 4.708, respectively). These results indicate that the proposed model provides superior out-of-sample prediction and possesses high predictive relevance. The findings confirm the robustness of the structural model in predicting students' perceptions of feedback quality, engagement, and academic achievement within AI-supported learning environments. Collectively, the predictive assessment complements the structural model evaluation by demonstrating that the model not only explains the observed relationships but also accurately predicts future observations. Together with the hypothesis testing, mediation, moderation, and multigroup analyses, these results provide strong empirical support for the proposed conceptual

framework, indicating that AI-driven personalized formative feedback enhances learning achievement through improved feedback quality and student engagement while producing differential effects across heterogeneous learner populations.

## 5. DISCUSSION

### Interpretation of the Main Findings

The findings indicate that AI-driven personalized formative feedback can improve learning achievement while simultaneously raising equity concerns, because the magnitude of benefit differs across learner subgroups. Prior research similarly frames AI-enabled personalization as a mechanism for enhancing engagement and performance through learner profiling and adaptive support, but also stresses that outcomes depend on context, governance, and who can effectively access and use AI tools (Jarudin & Dedi, 2026; Teferi et al., 2023). Evidence from higher education also cautions that socioeconomic differences (e.g., differential access to paid tools) may systematically shape AI use and associated learning outcomes, providing a plausible explanation for the observed moderation by learner diversity (S.-Y. Chen & Chen, 2025). In parallel, research on educational equity emphasizes that "equity" is not reducible to overall gains, but must be assessed by examining subgroup outcome patterns, aligning with the interpretation that benefits may be unevenly distributed even when average performance increases (Alharbi, 2023; Lam et al., 2024).

The strongest relationship between AI-generated feedback and perceived feedback quality suggests that students judged AI feedback as useful and actionable. This interpretation is consistent with evidence that AI-driven feedback systems can strengthen learners' self-assessment accuracy and can be designed to respond to learner profiles, which are key features of feedback perceived as diagnostic and improvement-oriented (Berdahl et al., 2023). More broadly, digital transformation scholarship argues that scalable personalization can approximate individualized guidance at scale, which is difficult to achieve consistently via instructor-only feedback in large classes (S.-Y. Chen & Chen, 2025). Perceived feedback quality predicting engagement is also consistent with the view that AI personalization can increase students' participation by aligning support with individual needs and progress (Berdahl et al., 2023). However, equity-oriented evidence warns that online/digital delivery is not a "silver bullet": differential participation constraints (e.g., connectivity, environment) can restrict who translates digital supports into active engagement, which is compatible with heterogeneous effects across groups (Marshik et al., 2024).

Finally, the significant direct effect of AI feedback on achievement after accounting for mediators is consistent with AI supporting learning via multiple pathways: (i) direct cognitive support (diagnosing errors, suggesting next steps) and (ii) indirect effects through engagement-enhancing personalization (Berdahl et al., 2023). At the same time, prior equity research on ICT highlights that digital systems can narrow or widen gaps depending on access, ability to use tools, and broader "digital gap" dynamics, reinforcing the need to evaluate subgroup impacts rather than rely on overall averages (Lam et al., 2024).

### Personalization versus Educational Equity

AI-driven personalized feedback can raise mean achievement while still producing inequitable gains because learners differ in the resources required to benefit from personalization (e.g., digital literacy, language proficiency, and socioeconomic advantage) (Nayyar et al., 2022). Empirical work in higher education also notes that the increasing availability of paid or subscription-based AI tools can make socioeconomic status a determinant of who uses AI and who benefits from it, directly aligning with moderation effects by learner diversity (Marshik et al., 2024). Related access research in culturally and linguistically diverse communities shows compounded disadvantage (low socioeconomic advantage plus non-English language plus interpreter need) is associated with lower uptake of digital services, supporting the interpretation that "the same" AI intervention may not be equally usable across subgroups (Gallegos-Rejas et al., 2023).

This pattern underscores that optimization is not synonymous with fairness: AI personalization can be tuned to maximize engagement or performance signals, yet those signals may reflect pre-existing structural differences in participation and capability (e.g., language and literacy constraints) (Hall & Swanlund, 2023). Evidence on language proficiency and achievement further indicates that achievement gaps are largest at lower levels of English proficiency, implying that language-related constraints can magnify differential returns to feedback even when support is available (Gallegos-Rejas et al., 2023). In addition, experimental psychology research cautions that algorithmic personalization can produce systematic cognitive distortions (e.g., inaccurate generalization and overconfidence), implying that

outcomes may depend on users' capacity to interpret personalized outputs critically, another plausible mechanism for heterogeneous treatment effects (Bahg et al., 2025). Accordingly, evaluation of AI-supported feedback should move beyond average effects to explicitly test subgroup impacts and accessibility. Universal Design for Learning–oriented lesson design is proposed as an equity strategy because it foregrounds accessibility for learners with disabilities, English-language learners, and students from diverse cultural and socioeconomic backgrounds, precisely the groups most likely to experience lower returns from standard "one-size" personalization (Vostal et al., 2023).

**Implications for Educational Technology**

AI-driven formative feedback should be conceptualized as a pedagogical innovation rather than a standalone technical upgrade because its educational value depends on how feedback supports learners' self-assessment, engagement, and learning processes, not merely on automation (S.-Y. Chen & Chen, 2025). Evidence from AI feedback systems in higher education writing indicates that effectiveness is tied to designing feedback processes responsive to learner profiles and usable for improvement activities (i.e., formative functions), reinforcing the need to embed AI within instructional designs that cultivate active learning and self-regulated learning practices (Teferi et al., 2023).

Developers should incorporate explainable AI (XAI) elements into feedback systems because learners' uptake of recommendations depends on whether personalized outputs can be meaningfully interpreted and critically evaluated. Experimental evidence shows personalization can induce inaccurate generalization and overconfidence, implying that transparency and metalevel design goals (e.g., fairness and diversity-aware objectives) are important to reduce harmful miscalibration and improve appropriate reliance (Bahg et al., 2025). In addition, research on AI use in higher education warns that socioeconomic factors and differential access to paid tools can shape AI use and outcomes, suggesting transparent guidance and reflection supports may also mitigate uneven uptake (Marshik et al., 2024). Finally, evaluation should routinely include equity indicators alongside average effectiveness. Scholarship on AI and equity emphasizes auditing for group-level harms and governance mechanisms to detect and remediate disparities, not just overall performance (Berdahl et al., 2023). Empirical work further indicates that structural barriers such as digital literacy and language proficiency affect access and benefit, supporting subgroup analyses, accessibility checks, and heterogeneous treatment effect reporting prior to scale-up (Nayyar et al., 2022).

**Practical Implications and Theoretical Contributions**

Practically, higher education institutions should blend AI-generated formative feedback with instructor facilitation, because student uptake and benefit from AI can vary with access and readiness. Evidence from higher education indicates that socioeconomic status may shape AI tool use and associated learning outcomes, implying that human support is important for students who are less able to access or effectively use AI tools (Marshik et al., 2024). More broadly, studies of digital-service delivery highlight structural barriers (e.g., lower digital literacy, lower English proficiency) that can limit effective use, supporting targeted instructor mediation and additional support for disadvantaged learners (Nayyar et al., 2022). Instructional design should move toward adaptive feedback that accounts for learner diversity, not only performance. AI feedback research in higher education writing emphasizes designing feedback processes responsive to learner profiles and improving self-assessment accuracy, which supports enriching personalization inputs beyond test scores alone (S.-Y. Chen & Chen, 2025). To prevent inequitable impacts at scale, institutions should implement equity-focused governance and auditing that detects group-level harms and guides remediation (e.g., resourcing those with less access), rather than relying on average performance metrics (Berdahl et al., 2023).

Theoretically, the results strengthen accounts that AI feedback operates through multiple mechanisms: direct informational support and indirect effects via learners' feedback processes (e.g., self-assessment and use of profile-responsive feedback) (S.-Y. Chen & Chen, 2025). The findings also advance educational technology scholarship by centering equity as analytically distinct from effectiveness, consistent with calls for explicit audits for group-level harms in AI systems (Berdahl et al., 2023).

## 6. CONCLUSION

This study examined the effectiveness and equity of AI-driven personalized formative feedback in higher education by investigating its relationships with perceived feedback quality, student engagement,

learning achievement, and learner diversity. The findings provide empirical support for the proposed conceptual framework, with all six hypotheses being supported. AI-Driven Personalized Formative Feedback significantly enhanced Perceived Feedback Quality ($\beta = 0.430$, $p < 0.001$) and directly improved Learning Achievement ($\beta = 0.265$, $p < 0.001$). Perceived Feedback Quality positively influenced Student Engagement ($\beta = 0.345$, $p < 0.001$), while Student Engagement significantly contributed to Learning Achievement ($\beta = 0.243$, $p < 0.001$). Furthermore, sequential mediation analysis confirmed that AI-driven feedback improved academic achievement indirectly through enhanced perceptions of feedback quality and increased student engagement ($\beta = 0.036$, $p < 0.001$), indicating partial mediation. The structural model explained 18.5% of the variance in Perceived Feedback Quality, 11.9% in Student Engagement, and 20.1% in Learning Achievement. At the same time, PLSpredict analysis demonstrated strong out-of-sample predictive performance across all endogenous constructs.

Beyond confirming the instructional effectiveness of AI-supported formative feedback, the study demonstrated that learner diversity significantly influences the magnitude of these benefits. Multigroup analysis revealed that students with higher prior academic achievement, stronger digital literacy, more favourable socioeconomic backgrounds, and higher language proficiency experienced significantly greater learning gains from AI-generated feedback than their peers. These findings indicate that although AI-driven personalization improves learning outcomes overall, its educational benefits are not distributed equally across heterogeneous learner populations. Accordingly, personalization should be accompanied by explicit equity-oriented design principles to ensure that AI technologies support inclusive rather than unequal educational opportunities.

This study contributes to educational technology literature by integrating instructional effectiveness, learner engagement, and educational equity within a single empirical framework grounded in self-regulated learning, feedback intervention theory, universal design for learning, and algorithmic fairness. The findings suggest that higher education institutions should complement AI-generated formative feedback with inclusive instructional strategies, explainable AI, and targeted support for learners with diverse academic and digital backgrounds. Future research should employ longitudinal and cross-cultural designs while investigating fairness-aware AI algorithms and human–AI collaborative feedback models to advance responsible, transparent, and equitable AI implementation in higher education.

## 7. ACKNOWLEDGE

If any, thanks are addressed to official institutions or individuals who have provided funding or have made other contributions to the research. Acknowledgments are accompanied by a research contract number.

## 8. REFERENCES

Al-Domi, H. A., AL-Dalaeen, A., ALRosan, S. H., Batarseh, N., & Nawaiseh, H. (2021). Healthy Nutritional Behavior During COVID-19 Lockdown: A Cross-Sectional Study. *Clinical Nutrition Espen*, *42*, 132–137. https://doi.org/10.1016/j.clnesp.2021.02.003

Alharbi, W. (2023). AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools. *Education Research International*, *2023*, 1–15. https://doi.org/10.1155/2023/4253331

Anh, B. N. T., & Phạm, M. (2022). Combination of SCCT and TPB in Explaining the Social Entrepreneurial Intention. *Ho Chi Minh City Open University Journal of Science - Economics and Business Administration*, *12*(2), 127–138. https://doi.org/10.46223/hcmcoujs.econ.en.12.2.2140.2022

Bahg, G., Sloutsky, V. M., & Turner, B. M. (2025). Algorithmic Personalization of Information Can Cause Inaccurate Generalization and Overconfidence. *Journal of Experimental Psychology General*, *154*(9), 2503–2522. https://doi.org/10.1037/xge0001763

Bai, X., & Stede, M. (2023). A Survey of Current Machine Learning Approaches to Student Free-Text Evaluation for Intelligent Tutoring. *International Journal of Artificial Intelligence in Education*, *33*(4), 994–1032. https://doi.org/10.1007/s40593-022-00323-0

Berdahl, C. T., Baker, L., Mann, S., Osoba, O., & Girosi, F. (2023). Strategies to Improve the Impact of Artificial Intelligence on Health Equity: Scoping Review. *Jmir Ai*, *2*, e42936. https://doi.org/10.2196/42936

Chen, H., Wang, C., Wu, J., Wang, M., Wang, S., Wang, X., Wang, J., Yu, H., Hu, Y., & Shang, S. (2022). Measurement Properties of Performance-Based Measures to Assess Physical Function in Knee Osteoarthritis: A Systematic Review. *Clinical Rehabilitation*, *36*(11), 1489–1511. https://doi.org/10.1177/02692155221107731

Chen, S.-Y., & Chen, W. (2025). Driven Intelligent Feedback System for Enhancing Self-Assessment Accuracy in Higher Education Writing. *Expert Systems*, *43*(1). https://doi.org/10.1111/exsy.70184

Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM-Based Acceptance of AI-Powered ChatGPT for Supporting Metacognitive Self-Regulated Learning in Education: A Mixed-Methods Study. *Heliyon*, *10*(8), e29317. https://doi.org/10.1016/j.heliyon.2024.e29317

Davis, S. E., Dorn, C., Park, D., & Matheny, M. E. (2025). Emerging Algorithmic Bias: Fairness Drift as the Next Dimension of Model Maintenance and Sustainability. *Journal of the American Medical Informatics Association*, *32*(5), 845–854. https://doi.org/10.1093/jamia/ocaf039

Gallegos-Rejas, V., Kelly, J. T., Lucas, K., Snoswell, C. L., Haydon, H. M., Pager, S., Smith, A. C., & Thomas, E. (2023). A Cross-Sectional Study Exploring Equity of Access to Telehealth in Culturally and Linguistically Diverse Communities in a Major Health Service. *Australian Health Review*, *47*(6), 721–728. https://doi.org/10.1071/ah23125

Ganta, T., Kia, A., Parchure, P., Wang, M., Besculides, M., Mazumdar, M., & Smith, C. B. (2024). Fairness in Predicting Cancer Mortality Across Racial Subgroups. *Jama Network Open*, *7*(7), e2421290. https://doi.org/10.1001/jamanetworkopen.2024.21290

González-Carrillo, C. D., Restrepo-Calle, F., Echeverry, J. J. R., & González, F. A. (2021). Automatic Grading Tool for Jupyter Notebooks in Artificial Intelligence Courses. *Sustainability*, *13*(21), 12050. https://doi.org/10.3390/su132112050

Griffin, A. (2024). The Pivot to Online Teaching: An Opportunity to Create Effective Problem-based Learning Environments for Dietetic Education. *Journal of Human Nutrition and Dietetics*, *38*(1). https://doi.org/10.1111/jhn.13378

Grupen, N. A., Selman, B., & Lee, D. D. (2022). Cooperative Multi-Agent Fairness and Equivariant Policies. *Proceedings of the Aaai Conference on Artificial Intelligence*, *36*(9), 9350–9359. https://doi.org/10.1609/aaai.v36i9.21166

Hall, G. J., & Swanlund, L. (2023). Differential Nonlinear Relations of Language Proficiencies to Reading and Math Achievement in Spanish or English. *School Psychology*, *38*(5), 294–307. https://doi.org/10.1037/spq0000545

Isroani, F., Jaafar, N., & Muflihaini, M. (2022). Effectiveness of E-Learning in Improving Student Learning Outcomes at Madrasah Aliyah. *International Journal of Science Education and Cultural Studies*, *1*(1), 42–51. https://doi.org/10.58291/ijsecs.v1i1.26

Jarudin, & Dedi. (2026). Personalisation and Predictive Insights: Applied AI Strategies in Modern Learning Technologies. *Digital Education Review*, *49*(49), 95–115. https://doi.org/10.1344/der.2026.49.95-114

Kim, H. H., & Ryu, J. (2021). Social Distancing Attitudes, National Context, and Health Outcomes During the COVID-19 Pandemic: Findings From a Global Survey. *Preventive Medicine*, *148*, 106544. https://doi.org/10.1016/j.ypmed.2021.106544

Lam, J., Coret, M., Khalil, C., Butler, K., Giroux, R., & Martimianakis, M. A. (2024). The Need for Critical and Intersectional Approaches to Equity Efforts in Postgraduate Medical Education: A Critical Narrative Review. *Academic Medicine*, *58*(12), 1442–1461. https://doi.org/10.1111/medu.15425

Lazarus, J. V, Wyka, K., White, T. M., Picchio, C. A., Rabin, K., Ratzan, S. C., Leigh, J. P., Hu, J., & El-Mohandes, A. (2022). Revisiting COVID-19 Vaccine Hesitancy Around the World Using Data From 23 Countries in 2021. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-31441-x

Li, L., Chen, H., Chen, W., & Yang, J. (2025). Microeukaryotic Habitat Specialists Exhibit Stronger Determinism and Biodiversity-Nutrient Cycling Relationship Than Generalists in a Subtropical River. *Applied and Environmental Microbiology*, *91*(10). https://doi.org/10.1128/aem.01364-25

Liem, V. T. (2025). The Impact of Leadership Style, Reward System, Environmental Strategy, and Environmental Management Accounting on Environmental Performance of Vietnamese Manufacturers. *Plos One*, *20*(5), e0323662. https://doi.org/10.1371/journal.pone.0323662

Lin, S., Tsai, P.-C., Su, F.-Y., Chen, C.-Y., Li, F., Zhao, J., Ho, Y. Y., Lee, M. T., Healey, E., Lin, P.-J., Kao, T., Vremenko, D., Roetzer-Pejrimovsky, T., Sholl, L. M., Dillon, D., Lin, N. U., Meredith, D. M., Ligon, K. L., Lo, Y., … Yu, K. (2025). Contrastive Learning Enhances Fairness in Pathology Artificial Intelligence Systems. *Cell Reports Medicine*, *6*(12), 102527. https://doi.org/10.1016/j.xcrm.2025.102527

Lu, J., Sattler, A., Wang, S., Khaki, A. R., Callahan, A., Fleming, S. L., Fong, R., Ehlert, B., Li, R., Shieh, L., Ramchandran, K., Gensheimer, M. F., Chobot, S. E., Pfohl, S., Li, S., Shum, K., Parikh, N., Desai, P., Seevaratnam, B., … Shah, N. H. (2022). Considerations in the Reliability and Fairness Audits of Predictive Models for Advance Care Planning. *Frontiers in Digital Health*, *4*. https://doi.org/10.3389/fdgth.2022.943768

Maheshi, B., Dai, W., Martínez-Maldonado, R., & Tsai, Y. (2024). Dialogic Feedback at Scale: Recommendations for Learning Analytics Design. *Journal of Computer Assisted Learning*, *40*(6), 2790–2808. https://doi.org/10.1111/jcal.13034

Mangaroska, K., Martínez-Maldonado, R., Vesin, B., & Gašević, D. (2021). Challenges and Opportunities of Multimodal Data in Human Learning: The Computer Science Students' Perspective. *Journal of Computer Assisted Learning*, *37*(4), 1030–1047. https://doi.org/10.1111/jcal.12542

Marshik, T., McCracken, C. C., Kopp, B., & O'Marrah, M. (2024). Student and Instructor Perceptions and Uses of Artificial Intelligence in Higher Education. *Teaching of Psychology*, *52*(3), 339–346. https://doi.org/10.1177/00986283241299745

Nayyar, D., Pendrith, C., Kishimoto, V., Chu, C., Fujioka, J., Rios, P., Bhatia, R. S., Lyons, O. D., Harvey, P., O'Brien, T., Martin, D., Agarwal, P., & Mukerji, G. (2022). Quality of Virtual Care for Ambulatory Care Sensitive Conditions: Patient and Provider Experiences. *International Journal of Medical Informatics*, *165*, 104812. https://doi.org/10.1016/j.ijmedinf.2022.104812

Ni, A., & Cheung, A. (2022). Understanding Secondary Students' Continuance Intention to Adopt an AI-powered Intelligent Tutoring System for English Learning. *Education and Information Technologies*, *28*(3), 3191–3216. https://doi.org/10.1007/s10639-022-11305-z

Pogorskiy, E., & Beckmann, J. F. (2023). From Procrastination to Engagement? An Experimental Exploration of the Effects of an Adaptive Virtual Assistant on Self-Regulation in Online Learning. *Computers and Education Artificial Intelligence*, *4*, 100111. https://doi.org/10.1016/j.caeai.2022.100111

Rana, Md. M. Siddiqee, M. S., Sakib, Md. N., & Ahamed, Md. R. (2024). Assessing AI Adoption in Developing Country Academia: A Trust and Privacy-Augmented UTAUT Framework. *Heliyon*, *10*(18), e37569. https://doi.org/10.1016/j.heliyon.2024.e37569

Schwensow, N., Heni, A. C., Schmid, J., Montero, B. K., Brändel, S. D., Halczok, T. K., Mayer, G., Fackelmann, G., Wilhelm, K., Schmid, D., & Sommer, S. (2022). Disentangling Direct From Indirect Effects of Habitat Disturbance on Multiple Components of Biodiversity. *Journal of Animal Ecology*, *91*(11), 2220–2234. https://doi.org/10.1111/1365-2656.13802

Sun, F. (2026). Effect of Artificial Intelligence-Based Tutoring, Cognitive Engagement, and Teacher Feedback on University Students' Academic Performance. *European Journal of Education*, *61*(2). https://doi.org/10.1111/ejed.70619

Teferi, B., Omar, M., Jeyakumar, T., Charow, R., Gillan, C., Jardine, J., Mattson, J., Dhalla, A., Koçak, S. A., Salhia, M., Davies, B. R., Clare, M., Younus, S., & Wiljer, D. (2023). Accelerating the Appropriate Adoption of Artificial Intelligence in Health Care: Prioritizing IDEA to Champion a Collaborative Educational Approach in a Stressed System. *Education Sciences*, *14*(1), 39. https://doi.org/10.3390/educsci14010039

Toyokawa, Y., Horikoshi, I., Majumdar, R., & Ogata, H. (2023). Challenges and Opportunities of AI in Inclusive Education: A Case Study of Data-Enhanced Active Reading in Japan. *Smart Learning Environments*, *10*(1). https://doi.org/10.1186/s40561-023-00286-2

Ulum, Ö. G. (2024). Unveiling the Layers: Analyzing ChatGPT Implementations in Turkish State Universities. *Base for Electronic Educational Sciences*, *5*(1), 114–134. https://doi.org/10.29329/bedu.2024.651.7

Ummels, D., Bols, E., Frantzen, R. J. A., Frantzen, T., Robeerts, L., & Beekman, E. (2025). Activity Trackers in Physical Therapy for People With Chronic Obstructive Pulmonary Disease in the Netherlands: Cross-Sectional Study on Current Use and Implementation Determinants. *Jmir Formative Research*, *9*, e59533–e59533. https://doi.org/10.2196/59533

Viberg, O., Baars, M., Mello, R. F., Weerheim, N., Spikol, D., Bogdan, C., Gašević, D., & Paas, F. (2024). Exploring the Nature of Peer Feedback: An Epistemic Network Analysis Approach. *Journal of Computer Assisted Learning*, *40*(6), 2809–2821. https://doi.org/10.1111/jcal.13035

Vostal, B. R., Oehrtman, J. P., & Gilfillan, B. H. (2023). School Counselors Engaging All Students: Universal Design for Learning in Classroom Lesson Planning. *Professional School Counseling*, *27*(1). https://doi.org/10.1177/2156759x231203199

Wang, Q. (2025). EFL Learners' Motivation and Acceptance of Using Large Language Models in English Academic Writing: An Extension of the UTAUT Model. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1514545

Watts, F. M., Dood, A. J., Shultz, G. V, & Rodriguez, J.-M. G. (2023). Comparing Student and Generative Artificial Intelligence Chatbot Responses to Organic Chemistry Writing-to-Learn Assignments. *Journal of Chemical Education*, *100*(10), 3806–3817. https://doi.org/10.1021/acs.jchemed.3c00664

Weber, F., Wambsganß, T., & Söllner, M. (2024). Enhancing Legal Writing Skills: The Impact of Formative Feedback in a Hybrid Intelligence Learning Environment. *British Journal of Educational Technology*, *56*(2), 650–677. https://doi.org/10.1111/bjet.13529

Weidlich, J., Fink, A., Jivet, I., Yau, J. Y., Giorgashvili, T., Drachsler, H., & Frey, A. (2024). Emotional and Motivational Effects of Automated and Personalized Formative Feedback: The Role of Reference Frames. *Journal of Computer Assisted Learning*, *40*(6), 2735–2752. https://doi.org/10.1111/jcal.13024

Yao, S., Dai, F., Sun, P., Zhang, W., Qian, B., & Lü, H. (2024). Enhancing the Fairness of AI Prediction Models by Quasi-Pareto Improvement Among Heterogeneous Thyroid Nodule Population. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-44906-y

Yue, H., & Wilson, J. (2025). Exploring the Effectiveness of Large-Scale Automated Writing Evaluation Implementation on State Test Performance Using Generalised Boosted Modelling. *Journal of Computer Assisted Learning*, *41*(2). https://doi.org/10.1111/jcal.70009

Zhai, X., Chu, X., Chai, C. S., Jong, M. S., Starčič, A. I., Spector, M., Liu, J., Jing, Y., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education From 2010 to 2020. *Complexity*, *2021*(1). https://doi.org/10.1155/2021/8812542

Zhaı, X., & Nehm, R. H. (2023). AI and Formative Assessment: The Train Has Left the Station. *Journal of Research in Science Teaching*, *60*(6), 1390–1398. https://doi.org/10.1002/tea.21885

.