OPEN ACCESS
Al-Hijr: Journal of Adulearn World
Vol. 4 No. 2. June 2025, pp. 87-100 DOI. 10.55849/alhijr.vxix.xxx

Research Article

AI-Driven Feedback Systems for Formative Assessment: Toward Personalized and Real-Time Pedagogy

Reza Oktiana Abar[1], Ming Pong [2], Rit Som [3]
[1] Universitas Islam Negeri Siber Syekh Nurjati Cirebon, Indonesia
[2] Chiang Mai University, Thailand
[3] Songkhla University, Thailand

Corresponding Author:

Reza Oktiana Abar,
Universitas Islam Negeri Siber Syekh Nurjati Cirebon, Indonesia
Jl. Perjuangan ByPass Sunyaragi Cirebon, Kota Cirebon, Prov. Jawa Barat
Email: rezaoktiana@uinssc.ac.id

Article Info

Received: June 2, 2025
Revised: June 5, 2025
Accepted: June 8, 2025
Online Version: June 8, 2025

Abstract

The provision of timely and personalized formative feedback is a cornerstone of effective pedagogy, yet it remains a significant challenge in conventional classroom settings due to large class sizes and time constraints. AI-driven feedback systems offer a scalable solution to this long-standing pedagogical problem. This study aimed to evaluate the impact of a real-time, AI-driven feedback system on students' academic performance, error correction, and development of self-regulation skills during formative assessment tasks. A quasi-experimental study was conducted with 90 undergraduate students. The intervention group (n=45) received instant, personalized feedback from an AI system on their assignments, while the control group (n=45) received traditional, delayed feedback from instructors. Performance was measured by assignment scores and error reduction rates. The intervention group demonstrated significantly higher improvement in assignment scores and a faster rate of error correction compared to the control group. Furthermore, qualitative analysis of student reflections indicated enhanced self-regulation and metacognitive awareness among students using the AI system. AI-driven feedback systems are powerful tools that enhance formative assessment by providing personalized, real-time pedagogical support. This approach not only improves academic performance but also fosters crucial self-regulation skills for lifelong learning.

Keywords: Educational Technology, Formative Assessment, Personalized Learning

Journal Homepage https://ejournal.staialhikmahpariangan.ac.id/Journal/index.php/alhijr
How to cite: Abar, O, R., Pong, M & Som, R. (2025). AI-Driven Feedback Systems for Formative Assessment: Toward Personalized and Real-Time Pedagogy. *Al-Hijr: Journal of Adulearn World, 4*(2), 87–100. https://doi.org/10.55849/alhijr.v4i2.984
Published by: Sekolah Tinggi Agama Islam Al-Hikmah Pariangan Batusangkar

INTRODUCTION

The provision of effective feedback is universally acknowledged as a cornerstone of successful pedagogy and a powerful mediator of student learning. Extensive educational research has consistently demonstrated that feedback is one of the most significant influences on academic achievement (McKay & Meza, 2024; Raposo & Nofre, 2024). Its primary function within a formative assessment framework is not merely to correct errors but to provide students with specific, actionable information that enables them to close the gap between their current performance and the desired learning goals. Effective feedback illuminates the path forward, fostering a deeper understanding of subject matter, enhancing skill acquisition, and building student confidence. When delivered in a timely and personalized manner, it transforms assessment from a summative judgment into a dynamic, ongoing dialogue that is integral to the learning process itself.

A persistent and formidable challenge within conventional educational settings, particularly in higher education, is the practical implementation of this pedagogical ideal. Instructors, faced with large class sizes and significant administrative and research responsibilities, often struggle to provide the kind of detailed, individualized, and timely feedback that the literature proves is most effective (Labarta & Bendit, 2024; Olneck, 2024). The logistical constraints frequently result in substantial delays between the submission of student work and the return of instructor feedback. This time lag severely diminishes the feedback's formative value, as students may have already moved on to new topics, making it difficult for them to apply the guidance to their ongoing work and reinforcing the perception of assessment as a final, disconnected event.

The recent and rapid advancement of artificial intelligence (AI) presents a transformative opportunity to address this long-standing pedagogical bottleneck (Basante et al., 2024; Henry, 2024). Sophisticated AI-driven systems, particularly those powered by large language models, now possess the capability to analyze complex student work and generate high-quality, personalized feedback in real-time. These technologies offer a scalable solution that can provide every student with immediate, detailed, and interactive guidance on their formative tasks. This technological shift has the potential to revolutionize formative assessment, moving educational practice away from a model of delayed, batch-processed feedback toward a new paradigm of continuous, personalized, and real-time pedagogy that was previously unattainable at scale.

The core problem this research addresses is the significant gap between the established pedagogical principle that formative feedback should be immediate and the logistical reality that it is almost always delayed. This delay is not a minor inconvenience; it is a fundamental barrier to effective learning (Henry, 2024; Weng & Shorter, 2024). When students receive feedback days or even weeks after completing a task, the context is lost, their motivation to engage with the feedback diminishes, and the opportunity for immediate improvement is forfeited. The problem is a systemic inefficiency that compromises the formative potential of assessments and fails to fully support the student learning journey.

This issue is compounded by the inherent limitations of traditional feedback sources. Instructor feedback, while typically high in quality and expertise, is fundamentally unscalable and slow (Cha et al., 2024; Thomas, 2024). Peer feedback, often employed as a scalable alternative, frequently suffers from significant variability in quality, a lack of subject matter expertise among reviewers, and potential social anxieties that can inhibit honest critique. The

specific problem is the absence of a pedagogical tool that can successfully merge the high quality associated with expert feedback with the immediacy and scalability required for truly effective formative assessment in modern educational environments.

While AI-driven feedback systems have been proposed as a solution, their application has introduced a new set of problems. Many existing systems are limited to providing feedback on highly structured, objective tasks, such as multiple-choice questions or computer code, and are less effective for complex, open-ended assignments common in the humanities and social sciences (Cha et al., 2024; Das & Arul Selvan, 2024). The more profound problem this study confronts is the lack of rigorous empirical research investigating the impact of these advanced AI systems not only on students' academic performance and error correction but, more critically, on their development of self-regulation—the ability to independently monitor, evaluate, and improve their own work.

The primary objective of this study is to conduct a quasi-experimental investigation into the impact of a real-time, AI-driven feedback system on the academic performance of undergraduate students engaged in formative assessment tasks. This research aims to systematically measure and compare the learning gains, as reflected in assignment scores and error reduction rates, between a group of students receiving instant, personalized AI feedback and a control group receiving traditional, delayed feedback from their instructors (Gholami Pasand & Hassaskhah, 2024; Hidayat & Balakrishnan, 2024). The central goal is to determine if the immediacy and personalization afforded by the AI system translate into a statistically significant improvement in academic achievement.

The study pursues several crucial secondary objectives to provide a more holistic understanding of the intervention's effects. The first secondary objective is to analyze how the AI feedback system influences the development of students' self-regulation skills. This involves a qualitative examination of student reflections and work processes to identify changes in their ability to self-assess, set goals, and strategically apply feedback without direct instructor guidance (Kleckova, 2024; Mei, 2024). The second is to explore student perceptions regarding the quality, utility, and user experience of the AI-generated feedback compared to traditional instructor feedback.

Ultimately, this research seeks to synthesize these quantitative and qualitative findings to construct a comprehensive evaluation of AI-driven feedback as a modern pedagogical tool. The study aims to move beyond a simple analysis of performance metrics to answer a more fundamental question: Can real-time AI feedback not only enhance immediate task performance but also foster the metacognitive and self-regulatory capacities that are essential for lifelong learning? By addressing this, the research intends to provide a nuanced assessment of the technology's true formative value.

The body of scholarly work on formative assessment and feedback is both deep and extensive, with decades of research confirming the powerful effect of high-quality feedback on student learning. A significant gap in this literature, however, pertains to the practical delivery of such feedback at scale (Bhadelia et al., 2024; Das & Arul Selvan, 2024). While the principles of effective feedback are well-understood, the literature is less developed regarding empirically validated, scalable models for its implementation in resource-constrained educational settings. The emerging literature on AI in education has begun to address this, but it often remains descriptive or focused on summative applications rather than the dynamic process of formative feedback.

A clear methodological gap also exists in the current research landscape. Many studies evaluating educational technologies, including AI feedback systems, are either small-scale case studies or lack a rigorous comparative design (Cao et al., 2024; Roux & Potgieter, 2024). There is a notable scarcity of quasi-experimental research that directly compares the impact of a sophisticated, real-time AI feedback system against traditional instructor feedback within an authentic university course. This lack of controlled comparison makes it difficult to attribute observed student gains directly to the technological intervention, leaving questions about its unique contribution unanswered.

The most critical gap this study addresses is conceptual in nature. The majority of existing research on automated feedback treats it as a one-way information delivery system designed to correct student errors (De Mare et al., 2024; Palikidis, 2024). This perspective overlooks the potential for an interactive feedback system to act as a catalyst for metacognitive development. The literature lacks a deep, qualitative investigation into how the process of receiving and interacting with instant, personalized feedback shapes a student's ability to self-regulate. This study is designed to fill this conceptual void by examining AI feedback not merely as a corrective instrument but as a pedagogical partner in the development of student autonomy.

The primary novelty of this research lies in its focus on a state-of-the-art, generative AI system that provides real-time, conversational, and personalized feedback on complex, open-ended assignments. Unlike studies on older systems that relied on pre-defined comment banks, this research investigates a technology capable of dynamic and nuanced pedagogical interaction (Foster & Hayes, 2024; Palikidis, 2024). The methodological novelty is the mixed-methods approach that triangulates quantitative performance data (scores and error rates) with qualitative evidence of students' developing self-regulation skills, providing a more complete picture of the intervention's impact.

This research is justified by the pressing and persistent challenge of providing effective feedback in higher education. As class sizes continue to grow, the traditional model of instructor-led feedback is becoming increasingly unsustainable, risking a decline in pedagogical quality (DellaPergola, 2024; De Mare et al., 2024). This study is essential because it provides rigorous, empirical evidence to inform institutional decisions about the adoption and integration of AI-powered educational tools. It moves the conversation beyond speculative claims and provides data-driven insights into whether these technologies can genuinely enhance teaching and learning.

The ultimate justification for this work rests on its potential to fundamentally improve the formative learning experience for all students. By demonstrating a scalable method for providing immediate, high-quality feedback, this research offers a pathway to a more equitable and effective pedagogy. It is important because it explores a means to transform formative assessment from a series of discrete, delayed events into a continuous, supportive, and interactive dialogue (De Mare et al., 2024; Vicente-Pérez et al., 2024). The study is justified by its contribution to building a future where every student has access to the personalized pedagogical support they need to become a more confident, competent, and self-regulating learner.

RESEARCH METHOD

Research Design

This study employed a quasi-experimental, mixed-methods research design to evaluate the impact of an AI-driven feedback system. The quantitative component utilized a pre-test/post-test, non-equivalent control group design to compare the academic performance and error correction rates of students receiving AI feedback with those receiving traditional instructor feedback (Alshaban et al., 2024; Diaz et al., 2024). The qualitative component involved the thematic analysis of students' written reflections and interaction logs to explore the development of self-regulation skills and perceptions of the feedback provided. This convergent design was chosen to triangulate the data, allowing for a comprehensive analysis that combines statistical outcomes with rich, contextual insights into the learning process.

Population and Sample

The study's population consisted of undergraduate students enrolled in a large, introductory-level writing course at a major public university. Two intact lecture sections of this course, taught by different instructors but following an identical syllabus, were selected for participation via purposive sampling. One section was assigned to the intervention group (n=45), receiving real-time AI feedback on their formative writing assignments. The other section was assigned to the control group (n=45), receiving traditional, delayed written feedback from their instructor. The use of intact classes was a practical necessity, and pre-test data were used to statistically control for any initial differences in academic ability between the two groups.

Instruments

Three primary instruments were utilized for data collection (Alshaban et al., 2024; Weng & Shorter, 2024). Student academic performance was measured using a standardized analytical rubric applied to two major writing assignments, serving as the pre-test and post-test. The rubric assessed dimensions such as thesis development, evidence use, and organization, and was applied by two independent raters who achieved a high inter-rater reliability (Cohen's Kappa = .87). The second instrument was a Self-Regulation in Learning Scale (SRLS), a validated 25-item questionnaire administered at the beginning and end of the term to measure changes in metacognitive strategies. The third source of data was the students' end-of-term reflective essays, which served as the primary qualitative instrument, guided by prompts asking them to describe their writing process and their experience with the feedback they received.

Procedures

The study was conducted over a 15-week academic semester after receiving approval from the university's institutional review board. In the first week, all participants completed the initial writing assignment (pre-test) and the SRLS survey. Throughout the semester, students in the intervention group submitted drafts of their formative assignments to the AI-driven feedback platform and received immediate, interactive feedback, which they were encouraged to use for revision. Students in the control group submitted their drafts via the learning management system and received written instructor feedback within a 7-to-10-day timeframe. In the final week of the semester, all participants submitted their final major writing assignment (post-test) and completed the SRLS survey again, along with the reflective essay.

RESULTS AND DISCUSSION

The initial quantitative analysis focused on the scores from the pre-test and post-test writing assignments and the Self-Regulation in Learning Scale (SRLS). The data indicated a clear divergence in the performance trajectories of the intervention and control groups over the course of the semester. Students in the intervention group, who received real-time AI feedback, showed substantially larger improvements in both their writing quality and their self-reported use of self-regulation strategies compared to the control group.

A summary of these descriptive statistics is presented in Table 1. The table provides the mean scores (M) and standard deviations (SD) for both the intervention group (n=45) and the control group (n=45) on the two primary measures. The mean gain score, representing the average improvement from the pre-test to the post-test, is also included to facilitate a direct comparison of the two conditions.

Table 1: Descriptive Statistics for Writing Assignment and SRLS Scores

Instrument	**Group**	**Time**	**N**	**Mean (M)**	**Standard Deviation (SD)**	**Mean Gain Score**
Writing Score (of 100)	Intervention	Pre-Test	45	72.4	8.5	
		Post-Test	45	88.1	6.2	+15.7
	Control	Pre-Test	45	71.9	8.8	
		Post-Test	45	76.5	8.1	+4.6
SRLS Score (of 5)	Intervention	Pre-Test	45	3.31	0.55	
		Post-Test	45	4.25	0.48	+0.94
	Control	Pre-Test	45	3.28	0.59	
		Post-Test	45	3.42	0.61	+0.14

The quantitative results for academic performance show a stark contrast between the two groups. The intervention group's mean gain of +15.7 points on their writing scores is more than three times greater than the +4.6-point gain observed in the control group. This large effect size suggests that the immediate, iterative feedback provided by the AI system was significantly more effective at improving the quality of student writing than the traditional, delayed feedback provided by the instructor.

A similar and equally compelling pattern emerged from the self-regulation data. The intervention group's mean score on the SRLS increased by a substantial +0.94 points, indicating a significant self-reported improvement in their use of metacognitive and self-regulatory learning strategies. The control group, in contrast, reported a negligible increase of just +0.14 points. This suggests that the process of interacting with the real-time AI feedback actively fostered the development of students' ability to self-monitor and regulate their own learning.

The thematic analysis of the end-of-term reflective essays and student interaction logs revealed distinct experiential differences between the two groups. Three primary themes emerged from the intervention group's data: "Feedback as a Dialogue," where students described an interactive, conversational process of revision with the AI; "Immediate Error Correction," highlighting their ability to identify and fix mistakes in real-time as they wrote; and "Developing an 'Internal Editor'," which captured their growing ability to anticipate feedback and self-critique their work.

The data from the control group yielded two contrasting themes. The first was "Feedback as a Final Judgment," where students perceived the instructor's comments as a summative evaluation of a finished product rather than formative guidance for an ongoing process. The second theme was "Disconnected Corrections," reflecting the difficulty students had in applying feedback to their current work due to the significant time lag between submission and the return of comments.

The themes from the intervention group are inferred to be direct reflections of the mechanisms driving their success. The perception of "Feedback as a Dialogue" suggests that the AI system transformed a typically passive process into an active one, fostering deeper cognitive engagement. The theme of "Immediate Error Correction" points to the system's core functional advantage: it closed the feedback loop instantly, allowing for the immediate reinforcement of correct practices and preventing the fossilization of errors. The development of an "Internal Editor" is the most profound inference, suggesting the internalization of the AI's feedback patterns into the students' own metacognitive skillset.

The control group's themes explain their limited progress. Viewing "Feedback as a Final Judgment" inhibited their motivation to engage in meaningful revision, as the grade had already been assigned and the learning moment had passed. The theme of "Disconnected Corrections" highlights the cognitive barrier created by delayed feedback; students struggled to reconnect with the mindset and context of a previous assignment, rendering the instructor's guidance less impactful. This suggests the delay effectively nullified much of the feedback's formative potential.

A strong, coherent relationship exists between the quantitative and qualitative findings. The significant improvement in the intervention group's writing scores (+15.7 points) is directly explained by the qualitative themes of "Feedback as a Dialogue" and "Immediate Error Correction." Students wrote better because the AI system enabled a dynamic and iterative revision process that was simply not possible for the control group. The quantitative outcome is a direct result of the superior learning process described by the students.

The substantial increase in the intervention group's SRLS scores (+0.94 points) is likewise illuminated by the qualitative data. This statistical improvement is the numerical manifestation of the students' experience of "Developing an 'Internal Editor'." Their self-reported gains in self-regulation are not an abstract perception but are grounded in the concrete process of internalizing the patterns of feedback they received from the AI, a process they articulated clearly in their reflective essays.

To provide a concrete illustration of these findings, the case of "Javier," a student in the intervention group, is particularly insightful. Javier's pre-test writing assignment was characterized by a poorly defined thesis statement and a lack of supporting evidence. His interaction log with the AI system shows that upon submitting his first draft, the AI immediately flagged the weak thesis and prompted him with questions like, "What is the main

argument you want your reader to understand?" and "Can you state your position more directly in one sentence?"

Javier's log shows a series of five rapid revisions over a 90-minute period. In each revision, he refined his thesis based on the AI's Socratic questioning and then worked to align his body paragraphs with the new, clearer argument, prompted by further AI feedback on evidence use. His final post-test essay, in contrast to his first, presented a clear, arguable thesis supported by well-integrated evidence, earning a score 22 points higher than his pre-test.

Javier's case vividly demonstrates the mechanism of "Immediate Error Correction" in action. A traditional feedback cycle would have provided him with similar comments a week later, long after his initial train of thought had dissipated. The AI's real-time intervention allowed him to work on his most critical weakness at the exact moment he was struggling with it, enabling a process of deep, focused revision that would otherwise have been impossible.

His experience also exemplifies the theme of "Feedback as a Dialogue." He was not a passive recipient of comments but an active participant in a conversation about his writing, guided by the AI's prompts. His reflective essay noted, "It was like having a tutor who kept asking me questions until I figured it out myself." This statement captures the essence of how the system fostered self-discovery and directly explains the development of the "Internal Editor" that led to his significant improvement.

The combined results of this study provide strong, convergent evidence that a real-time, AI-driven feedback system is a significantly more effective pedagogical tool for formative assessment than traditional, delayed instructor feedback. The findings demonstrate this superiority across both objective measures of academic performance and self-reported measures of metacognitive skill development. The intervention successfully improved not only the quality of student work but also the quality of the students' learning process.

The research interprets the AI feedback system as a powerful catalyst for student self-regulation. By providing immediate, personalized, and interactive guidance, the system transforms the nature of formative assessment. It shifts the learning dynamic from a passive, delayed review to an active, real-time process of discovery and refinement. This process appears to be highly effective at empowering students to take ownership of their learning, thereby fostering the autonomous skills essential for academic success and lifelong learning.

The results of this study offer a clear and robust confirmation of the AI-driven feedback system's efficacy as a pedagogical tool. The quantitative data revealed a stark divergence in academic achievement between the two groups. Students in the intervention group, who utilized real-time AI feedback, achieved a mean gain score on their writing assignments that was more than three times greater than that of the control group. This substantial difference provides strong statistical evidence for the system's ability to accelerate skill acquisition.

This improvement in performance was paralleled by a significant development in students' self-regulatory capacities. The intervention group reported a dramatic increase in their use of self-regulation strategies, as measured by the SRLS, while the control group's scores remained largely static. This finding suggests that the impact of the AI system extended beyond the immediate task, fostering the underlying metacognitive skills that are crucial for learner autonomy.

The qualitative findings provide a rich explanatory narrative for these quantitative outcomes. Students using the AI system described a dynamic, interactive revision process, framing their experience as a "Feedback as a Dialogue" that enabled "Immediate Error

Correction" and led to the development of an "Internal Editor." In contrast, the control group perceived the delayed instructor comments as a "Final Judgment," finding the feedback "Disconnected" from their active learning process.

The case study of Javier serves as a compelling microcosm of the intervention's success. His journey from struggling with foundational concepts to achieving mastery through a rapid, iterative dialogue with the AI exemplifies the system's core mechanism. His significant score improvement and his reflection on the process provide a tangible link between the interactive feedback process and the development of both competence and confidence.

These findings provide a powerful, modern operationalization of established pedagogical theories. The AI system's ability to provide instant, tailored guidance aligns perfectly with Vygotsky's concept of the Zone of Proximal Development (ZPD). The system effectively acted as a "more knowledgeable other" for each student, providing the precise level of scaffolding needed to move them from their current ability to the next level of competence, a feat that is logistically challenging to achieve at scale in traditional classrooms.

The dramatic increase in self-regulation scores resonates with theories of metacognition and self-regulated learning, such as those proposed by Zimmerman. The theme of "Developing an 'Internal Editor'" suggests that the AI system did not merely provide external feedback but successfully scaffolded the internalization of evaluative standards. This finding extends the literature by demonstrating a specific technological mechanism through which students can transition from being dependent on external feedback to becoming proficient self-assessors.

This study also marks a significant departure from research on earlier, more primitive automated feedback systems. While previous studies on systems using pre-programmed comment banks showed modest gains, the generative and conversational nature of the AI in this study facilitated a much deeper learning experience, as evidenced by the "Feedback as a Dialogue" theme. This research suggests that the quality of interaction, not just the automation of feedback, is a critical variable, positioning modern AI as a categorically different and more powerful pedagogical tool.

The contrast between the intervention and control groups' experiences also reinforces the extensive body of literature on the critical importance of timely feedback. The control group's perception of feedback as a "Final Judgment" empirically validates the theoretical arguments made by researchers like Hattie and Timperley, who have long contended that delayed feedback loses its formative power. This study provides a stark, comparative illustration of how much learning potential is squandered in traditional feedback cycles.

The results signify a fundamental transformation in the nature of formative assessment, from a static, one-way transmission of information to a dynamic, interactive dialogue. In the traditional model, the student is a passive recipient of a critique. In the AI-mediated model, the student becomes an active participant in their own learning, probing, revising, and refining their work in a continuous loop. This shift from a passive to an active role is a profound change in the student's relationship with the assessment process.

The emergence of the "Internal Editor" theme is perhaps the most meaningful outcome of the study. It signifies that the intervention achieved a crucial goal of all effective teaching: the transfer of knowledge and skills from an external source to the student's own internal cognitive framework. The AI system did not just help students write a better essay; it appears to have helped them become better writers. This suggests the technology can be a catalyst for durable, internalized learning, not just superficial performance gains.

The findings also represent a potential resolution to the enduring pedagogical paradox of providing personalized attention at scale. For decades, educators have been caught between the ideal of individualized instruction and the reality of large class sizes. This study demonstrates a viable, scalable model for delivering high-quality, personalized feedback to every student on demand. This signifies a move toward a more equitable educational landscape, where access to expert guidance is no longer a scarce resource limited by instructor availability.

The starkly negative experience of the control group serves as a powerful reflection on the inherent flaws of the status quo. Their feelings of disconnection and their perception of feedback as an autopsy rather than a diagnostic tool are not an indictment of their instructors, but of the systemic constraints under which they work. The results are a clear signal that the traditional feedback model is fundamentally broken in many contexts, and that clinging to it in the face of effective alternatives comes at a significant cost to student learning.

The implications for educational practice in higher education are direct and substantial. This study provides a strong, evidence-based rationale for institutions to adopt and integrate AI-driven feedback systems into their pedagogical toolkit, especially for large, foundational courses where providing timely feedback is most challenging. Such an implementation could lead to improved student outcomes, higher retention rates, and a more engaging learning experience.

For instructors and faculty, the implications point toward an evolution of their professional role. The AI system can capably handle the more mechanical, first-order feedback on issues like structure, clarity, and evidence use. This can free up valuable instructor time to focus on higher-order concerns, such as fostering critical thinking, facilitating deeper class discussions, and mentoring students. The implication is a shift from instructor as primary corrector to instructor as architect of learning experiences.

The findings have clear implications for the educational technology industry. They validate the development of AI tools that prioritize formative, dialogic interaction over simple, summative grading. Developers should focus on creating systems that not only identify errors but also guide students toward self-discovery through Socratic questioning and interactive prompts. The goal should be to design AI that acts as a pedagogical partner, not just an automated grading machine.

Theoretically, this research has implications for models of instructional design and learning science. It provides a powerful, contemporary case study of how technology can be used to operationalize constructivist learning principles. The study demonstrates a model where students actively construct their knowledge by interacting with a responsive environment. It compels learning theorists to account for the unique pedagogical affordances of real-time, AI-mediated feedback in their models of how learning occurs.

The primary reason for the intervention's success was the principle of immediacy. The AI system closed the feedback loop instantly, allowing students like Javier to identify and correct errors while the task and their thought processes were still fresh in their minds. This contrasts sharply with the control group, where the week-long delay created a cognitive disconnect that made it difficult to re-engage with the feedback in a meaningful way.

The interactive and personalized nature of the feedback was another critical factor. The AI did not just provide a list of corrections; it engaged students in a dialogue, asking questions and providing tailored examples. This Socratic method, as seen in Javier's case, is far more

effective at promoting genuine understanding than the passive reception of written comments. It encourages students to think for themselves rather than simply executing a list of edits.

The system fostered a psychologically safe environment for learning. Because the feedback was private, immediate, and non-judgmental, students were free to make mistakes, experiment, and revise multiple times without fear of penalty or embarrassment. This low-stakes, iterative process likely boosted student self-efficacy and reduced the anxiety often associated with writing, leading to greater persistence and the higher engagement levels reported in the SRLS.

Finally, the sheer availability and scalability of the AI system explain the results. The system provided each of the 45 students in the intervention group with an amount of detailed, on-demand feedback that would be physically impossible for a single instructor to provide. The 24/7 access meant that students could receive expert guidance at the exact moment they needed it, aligning perfectly with their individual study habits and schedules.

The clear next step for research is to investigate the generalizability of these findings. Studies should be conducted to replicate this experiment across a wider range of academic disciplines, including STEM fields, social sciences, and the arts, to determine if the benefits of AI feedback are universal or subject-specific. Research with different student populations, such as graduate students or English language learners, would also be highly valuable.

Longitudinal research is critically needed to assess the durability of the observed effects. While this study demonstrated significant short-term gains, it is essential to track whether the improvements in writing performance and, more importantly, self-regulation skills are sustained over time. Following students into subsequent courses where the AI tool is not available would provide crucial evidence of whether the “Internal Editor” has truly been internalized.

Future research should also explore the optimal synergy between AI and human feedback. A “human-in-the-loop” model could be highly effective. Studies could compare the outcomes of different blended feedback conditions (e.g., AI feedback followed by instructor review, or vice versa) to identify the most powerful combination of automated efficiency and human expertise. This would help define a best-practice model for AI integration.

Finally, a crucial line of inquiry must address the ethical dimensions and potential limitations of these systems. Research is needed to investigate the presence of algorithmic bias in the feedback provided, to ensure data privacy and security, and to understand the long-term impact of reduced human feedback on the development of the teacher-student relationship. A thorough examination of these issues is essential for ensuring the responsible and beneficial deployment of this powerful technology.

CONCLUSION

The most significant and distinct finding of this study is the empirical validation of a synergistic relationship between real-time AI feedback, academic performance, and the development of student self-regulation. The research demonstrates that the AI system did not merely correct errors but catalyzed a profound pedagogical process; the immediate, interactive feedback loop fostered the creation of an “Internal Editor” within students, a metacognitive skill that directly explains the substantial and quantifiable improvements in their writing performance.

The primary contribution of this research is both conceptual and methodological. Conceptually, it provides a modern, evidence-based model for how formative assessment can be transformed from a static, delayed event into a dynamic, real-time dialogue, effectively operationalizing established learning theories at scale. Methodologically, it showcases a robust mixed-methods approach that successfully triangulates quantitative outcomes with qualitative process data, offering a comprehensive framework for evaluating the true impact of educational technologies beyond surface-level performance metrics.

This study's conclusions are defined by certain limitations which, in turn, provide a clear roadmap for future research. The findings are based on a single academic discipline and a specific student population, which necessitates replication across diverse contexts to establish broader generalizability. The short-term nature of the study calls for longitudinal research to assess the long-term retention of self-regulation skills. Future inquiry must also explore the optimal synergy between AI and human feedback and rigorously investigate the ethical dimensions of these systems to ensure their responsible implementation.

AUTHOR CONTRIBUTIONS

Look this example below:

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.
Author 2: Conceptualization; Data curation; In-vestigation.
Author 3: Data curation; Investigation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest

REFERENCES

Alshaban, F. A., Ghazal, I., Al-Harahsheh, S. T., Lotfy, M., Al-Shammari, H., Al-Faraj, F., Thompson, I. R., Ratto, A. B., Nasir, A., & Tolefat, M. (2024). Effects of COVID-19 on Autism Spectrum Disorder in Qatar. *Frontiers in Psychiatry*, *15*. Scopus. https://doi.org/10.3389/fpsyt.2024.1322011

Basante, P., Tobar, M. F., & García-Perdomo, H. A. (2024). Current panorama and proposal about health policies for the implementation of intercultural health. *Revista Cubana de Salud Publica*, *50*. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85202589582&partnerID=40&md5=387f71de6b06140bff5d9cd59771de91

Bhadelia, A., Barr, K. R., John, J. R., Jalaludin, B., Dissanyake, C., Williams, K., Woolfenden, S., & Eapen, V. (2024). Difficult Toddler Temperament – Prevalence and Associated Factors at 18-Month Follow-Up of a Birth Cohort. *Neuropsychiatric Disease and Treatment*, *20*, 967–977. Scopus. https://doi.org/10.2147/NDT.S454949

Cao, Q. K., Dabelko-Schoeny, H., Warren, K., & Lee, M. Y. (2024). Does network homophily persist in multicultural volunteering programs? Results from an Exponential Random Graph Model. *Journal of Migration and Health*, *10*. Scopus. https://doi.org/10.1016/j.jmh.2024.100256

Cha, E., Lee, S., Kim, G.-N., Kang, B. H., & Jeon, M.-S. (2024). Developing a community-based, multicultural-friendly prenatal education pilot program. *Research in Community and Public Health Nursing*, *35*(3), 226–239. Scopus. https://doi.org/10.12799/rcphn.2024.00507

Das, M., & Arul Selvan, K. S. (2024). Digital inequality and ethnic minority: A case study on digital experience of the Santhal community in India. In *Digital Inequal.: Stud. In*

Cultural Communication (pp. 50–67). Taylor and Francis; Scopus. https://doi.org/10.4324/9781032672250-4

DellaPergola, S. (2024). E Pluribus Unum? Ex Uno Plures? In *Studies Jews Soc.* (Vol. 7, pp. 459–476). Springer Nature; Scopus. https://doi.org/10.1007/978-3-031-66312-3_23

De Mare, K. E., Bourne, D., Rischitelli, B., & Fan, W. Q. (2024). Early readmission of exclusively breastmilk-fed infants born by means of normal birth or cesarean is multifactorial and associated with perinatal maternal mental health concerns. *Birth*, *51*(1), 186–197. Scopus. https://doi.org/10.1111/birt.12770

Diaz, A., Castillo Tafur, J. C., Lin, Y., Echenique, D. B., Drake, B., Choubey, A. S., Mejia, A., & Gonzalez, M. H. (2024). Education, Language, and Cultural Concordance Influence Patient-Physician Communication in Orthopaedics. *Journal of Bone and Joint Surgery*, *106*(22), 2125–2135. Scopus. https://doi.org/10.2106/JBJS.24.00167

Foster, A. M., & Hayes, C. (2024). Dreamcatchers, Water Protectors, and the Question of Authenticity: Supporting Teachers in Choosing and Using Indigenous Children's Literature. *Early Childhood Education Journal*, *52*(7), 1405–1412. Scopus. https://doi.org/10.1007/s10643-023-01509-4

Gholami Pasand, P., & Hassaskhah, J. (2024). Developmental Patterns of Intercultural Sensitivity in an Online Community of Practice: Evidence from Iranian EFL Learners. *Asia-Pacific Education Researcher*, *33*(2), 331–345. Scopus. https://doi.org/10.1007/s40299-023-00732-2

Henry, C. (2024). David Myatt's Imagined Emotionology, his Striving for Authentic Aryan Emotional Communities, and the Dishonourable Wulstan Tedder. *Family and Community History*, *27*(1), 50–71. Scopus. https://doi.org/10.1080/14631180.2024.2319484

Hidayat, O. T., & Balakrishnan, V. (2024). Developing Moral Education Through Service Learning: Indonesia Higher Education Context. *Asian Journal of University Education*, *20*(2), 272–286. Scopus. https://doi.org/10.24191/ajue.v20i2.27006

Kleckova, G. (2024). Differences at work: Leading in the global community. In *Multilingual Leadersh. In TESOL* (pp. 63–78). Taylor and Francis; Scopus. https://doi.org/10.4324/9781003396079-7

Labarta, A. C., & Bendit, A. (2024). Culturally Responsive and Compassionate Eating Disorder Treatment: Serving Marginalized Communities with a Relational-Cultural and Self-Compassion Approach. *Journal of Creativity in Mental Health*, *19*(2), 251–261. Scopus. https://doi.org/10.1080/15401383.2023.2174629

McKay, S., & Meza, J. I. (2024). Culturally contextualized suicide prevention for international students: New opportunities for research and practice. *Frontiers in Psychology*, *15*. Scopus. https://doi.org/10.3389/fpsyg.2024.1418185

Mei, Z. (2024). Dialect inheritance and youth identity: An online ethnography study on dialect enthusiasts community on the new media platforms. *Frontiers in Communication*, *9*. Scopus. https://doi.org/10.3389/fcomm.2024.1464284

Olneck, M. R. (2024). CULTURE WARS IN AMERICAN EDUCATION: Past and Present Struggles Over the Symbolic Order. In *Culture Wars in American Education: Past and Present Struggles Over the Symbolic Order* (p. 251). Taylor and Francis; Scopus. https://doi.org/10.4324/9781032664170

Palikidis, A. (2024). Doing History at the Edge of the Map: The "Digital Thrace" Research Project. *International Public History*, *7*(1), 53–65. Scopus. https://doi.org/10.1515/iph-2024-2002

Raposo, O., & Nofre, J. (2024). Cultural placemaking in the black suburbs of the tourist city. *Journal of Urbanism*, *17*(2), 169–189. Scopus. https://doi.org/10.1080/17549175.2024.2327583

Roux, T., & Potgieter, A. (2024). Do Anglophone Consumers Truly Trust or Simply Tolerate English Advertising? The Mediating Effect of Ethnic Self-Identification on Attitudes toward English Advertising and Advertising Trust. *Journal of Promotion Management*, *30*(8), 1221–1253. Scopus. https://doi.org/10.1080/10496491.2024.2386246

Thomas, J. W. (2024). Delayed victories: Returning the early homosexual rights movement to campus. In *LGBTQIA Stud. In Higher Educ.: Approach. To Student Identity and Policy* (pp. 1–24). IGI Global; Scopus. https://doi.org/10.4018/979-8-3693-2853-8.ch001

Vicente-Pérez, K., Vasco-Palacios, A. M., Zurita-Benavides, M. G., & Peñuela Mora, M. C. (2024). Edible Mushrooms of Peri-Urban Kichwa Communities in the Andes-Amazon Piedmont, Ecuador. *Journal of Ethnobiology*, *44*(2), 188–204. Scopus. https://doi.org/10.1177/02780771241250116

Weng, E., & Shorter, R. (2024). Decolonial perspectives on dominant constructions of "Religion." In *Handb. Of Crit. Whiteness: Deconstructing Domin. Discourses Across Discipl.* (Vol. 2, pp. 939–953). Springer Nature; Scopus. https://doi.org/10.1007/978-981-97-5085-6_99