VNU Journal of Foreign Studies

Journal homepage: https://jfs.ulis.vnu.edu.vn/

RETHINKING ASSESSMENT: DIFFERENTIATED INSTRUCTION IN THE AI ERA

Trinh Hoai Thu*

English B Department, Hanoi University of Business and Technology,
No. 29A, Lane 124 Vinh Tuy Street, Vinh Hung Ward, Hanoi, Vietnam

Received 05 December 2025
Revised 26 February 2026; Accepted 24 July 2026

Abstract: This article re-examines the role of assessment within the rapidly evolving landscape of artificial intelligence (AI), focusing specifically on differentiated instruction. As AI tools become more integrated into education, traditional, one-size-fits-all assessment methods are proving to be inadequate for capturing the full scope of student learning. This paper argues for a paradigm shift, urging educators to rethink assessment by embracing a differentiated approach where students can demonstrate their mastery through diverse and creative products. We explore how AI can serve as a powerful ally in this process, offering new ways to automate feedback, analyze complex student outputs, and provide personalized support that goes beyond conventional grading. However, this shift also presents significant challenges, including concerns about academic integrity and the ethical use of AI. By reviewing current literature and proposing a forward-looking framework, this paper highlights the potential of AI to not only streamline the assessment process but also cultivate a more equitable and student-centered learning environment.

Keywords: artificial intelligence (AI), differentiated instruction, assessment products, literature, framework

* Corresponding author.
Email address: Hoaithu1985108@gmail.com
https://doi.org/10.63023/2525-2445/jfs.ulis.5681

TÁI CẤU TRÚC ĐÁNH GIÁ: TIẾP CẬN PHÂN HÓA TRONG KỶ NGUYÊN AI

Trịnh Hoài Thu

Khoa Tiếng Anh B, Trường Đại học Kinh doanh và Công nghệ Hà Nội,
Số 29A, Ngõ 124 Vĩnh Tuy, Phường Vĩnh Hưng, Hà Nội, Việt Nam

Nhận bài ngày 05 tháng 12 năm 2025
Chỉnh sửa ngày 26 tháng 02 năm 2026; Chấp nhận đăng ngày 24 tháng 7 năm 2026

Tóm tắt: Bài viết xem xét lại vai trò của đánh giá trong bối cảnh trí tuệ nhân tạo (AI) đang phát triển nhanh chóng, tập trung cụ thể vào giảng dạy phân hóa. Khi các công cụ AI ngày càng được tích hợp vào giáo dục thì các phương pháp đánh giá truyền thống đang tỏ ra không đủ để nắm bắt toàn bộ phạm vi học tập của người học. Bài viết ủng hộ một sự thay đổi mô hình, kêu gọi các nhà giáo dục cân nhắc một sự thay thế đánh giá truyền thống bằng cách áp dụng phương pháp tiếp cận phân hóa, nơi người học có thể thể hiện năng lực của mình thông qua các sản phẩm đa dạng và sáng tạo. Đồng thời, bài viết khám phá cách AI có thể đóng vai trò là một cộng sự đắc lực trong quá trình này, cung cấp những cách thức mới để tự động hóa phản hồi, phân tích các kết quả đầu ra phức tạp của người học và cung cấp hỗ trợ cá nhân hóa vượt ra ngoài cách đánh giá chấm điểm thông thường. Tuy nhiên, sự thay đổi này cũng đặt ra những thách thức đáng kể, bao gồm cả những lo ngại về liêm chính học thuật và việc sử dụng AI một cách có đạo đức. Bằng cách xem xét các tài liệu hiện tại và đề xuất một khuôn khổ hướng tới tương lai, bài viết này làm nổi bật tiềm năng của AI không chỉ trong việc hợp lý hóa quy trình đánh giá mà còn xây dựng một môi trường học tập công bằng hơn và lấy người học làm trung tâm.

Từ khóa: đánh giá, trí tuệ nhân tạo (AI), tiếp cận phân hóa, sản phẩm, khuôn khổ

1. Introduction

1.1. Contextual Background

The integration of Artificial Intelligence (AI) into the educational ecosystem marks a pivotal point in pedagogical history. AI tools, including large language models and intelligent tutoring systems, are increasingly reshaping how knowledge is accessed, synthesized, and produced in educational contexts (Alexandrowicz, 2024; Iqbal et al., 2024; Ning et al., 2024; Vieriu & Petrea, 2025). This technological disruption has simultaneously highlighted the urgent need for students to develop 21st century skills such as critical thinking, creativity, and complex problem-solving capacities that remain distinctly human and are vital for future success. The World Economic Forum's (WEF, 2023) Future of Jobs Report consistently emphasizes analytical thinking, creative thinking, and complex problem-solving ability as the fastest growing and most crucial skills for the global workforce in the era of automation. The UNESCO (2021) working paper, *AI and Education: Guidance for Policy Makers*, serves as a policy brief arguing that education must shift its focus toward developing unique human competencies (e.g., critical thinking, socio-emotional skills) so that students can effectively collaborate with AI. The U.S. Department of Education, Office of Educational Technology (2023) report, *Artificial Intelligence and the Future of Teaching and Learning*, strongly recommends that educators realign pedagogical priorities. Specifically, the document urges the focus to shift toward cultivating competencies that remain distinct from automated capabilities, notably creativity, critical thinking, and collaboration, as an imperative strategy for preparing students for an AI-integrated professional landscape.

1.2. Statement of the Problem

Against this backdrop, traditional, one-size-fits-all assessment methods such as standardized essay prompts or simple multiple choice tests have revealed their inherent vulnerabilities. Generative AI can easily produce highly polished, contextually relevant responses to these conventional tasks, leading to an assessment crisis where the evaluations risk measuring AI capabilities rather than the student's actual understanding. Consequently, these inadequate methods fail to capture the holistic, diverse, and often unique ways students learn and master material, resulting in a significant "Assessment Gap". This shift necessitates the adoption of new frameworks, such as the Process-Product Assessment Approach, which validates learning by emphasizing the student's interaction with AI tools and the demonstration of their unique critical analysis, rather than relying on easily simulated final outputs (Frontiers, 2024). The same technological forces driving this crisis are simultaneously presenting solutions, with research increasingly focused on employing Large Language Models (LLM) for automated and dynamic question generation to directly counter the static limitations of conventional assessments (Maity & Deroy, 2024). Manokore et al. (2025) assert that the rapid emergence of Generative AI renders conventional assessment methods, which primarily rely on standardized output, to be inadequate; they argue that this shift necessitates a fundamental rethinking of evaluative frameworks to uphold pedagogical soundness and ensure assessments authentically reflect genuine learner competencies.

1.3. Thesis Statement

This paper advances a reconceptualization of assessment logic in the AI era, proposing Differentiated Assessment Products (DAP) as a generative framework that synthesizes and extends existing models of differentiated instruction and authentic assessment. While existing scholarship has separately examined differentiated instruction, authentic assessment, and AI-assisted evaluation, this paper advances the field by conceptualizing "Differentiated Assessment Products" (DAP) as a distinct theoretical construct that extends differentiated instruction (DI) theory into the domain of assessment design under AI conditions. Unlike prior frameworks that primarily focus on task authenticity or AI-resistance, this model integrates product diversification, process documentation, and AI-supported evaluative analytics into a single coherent structure. The novelty of the framework lies not in introducing new technological tools, but in reconfiguring assessment logic: shifting from output validation to multi-modal, process-aware, AI-mediated evidence ecosystems. Furthermore, AI should be strategically leveraged not as a cheating threat, but as an ally in this process, helping educators analyze these complex outputs and provide personalized, sophisticated feedback. This fundamental requirement for institutional change is echoed in contemporary literature, which frames AI integration as a necessary "education paradigm shift" driven by the need for spatiotemporal analysis to capture the full spectrum of student learning (Zhong & Zhao, 2025).

Damyanov (2024) further emphasizes AI's role in delivering personalized feedback and data-driven instructional adaptation, which forms a foundational condition for differentiated assessment. Similarly, Ruslim and Khalid (2024) report that AI assists teachers in automating administrative functions, including reviewing and grading assignments, thereby supporting the operational feasibility of differentiated evaluation. This automation is crucial for supporting the shift to complex, time-consuming assessment products that require sophisticated analysis.

1.4. Roadmap of the Paper

This article will first establish the theoretical foundations of DI, followed by an exploration of how a new framework DAP with AI support can be implemented. Finally, it will address the critical challenges, particularly concerning academic integrity and ethical use, concluding with implications for policy and future research. This article adopts a conceptual and narrative literature review approach aimed at synthesizing recent theoretical and empirical discussions on assessment in the era of generative artificial intelligence. Rather than conducting a systematic review with statistical aggregation, the paper develops a theoretical synthesis that integrates scholarship from differentiated instruction, authentic assessment, and AI-supported pedagogy. The purpose is framework construction and conceptual advancement rather than effect-size comparison or meta-analytic generalization. Regarding data sources and evidentiary basis, this paper does not employ primary empirical data. All analytical claims are derived from secondary sources, including peer-reviewed research articles, institutional policy documents, assessment frameworks, and practitioner-oriented implementation reports. The study therefore operates at a conceptual and interpretive level, synthesizing existing evidence to construct a forward-looking framework rather than empirically testing intervention outcomes.

2. Theoretical Foundations and the Assessment Crisis

2.1. Literature Selection and Review Scope

The literature included in this review was selected through a purposive narrative search strategy focusing on recent scholarship at the intersection of artificial intelligence and educational assessment. Priority was given to publications from 2023 - 2025 to ensure conceptual relevance in the rapidly evolving generative AI landscape, while foundational works on differentiated instruction were retained for theoretical grounding. Sources were identified through academic databases (Scopus, Web of Science, ERIC, Google Scholar) using key search terms such as "generative AI assessment", "AI-resistant assessment", "differentiated assessment", and "authentic assessment in AI era".

Inclusion criteria consisted of:

(1) peer-reviewed journal articles, conference proceedings, and institutional policy reports.

(2) explicit discussion of AI's impact on assessment design, feedback, or academic integrity.

(3) conceptual or empirical relevance to differentiated instruction or authentic assessment.

Exclusion criteria included opinion pieces without theoretical grounding, non-educational AI applications, and studies focused solely on technical system development without pedagogical implications. In synthesizing the literature, greater analytical weight was assigned to peer-reviewed empirical studies and institutional policy frameworks, particularly those providing implementation evidence or evaluative data. Conceptual and practitioner-oriented publications were used to extend interpretive insights and practical design implications. This layered evidentiary approach supports the development of a theoretically grounded yet practice-responsive assessment framework. Given the secondary nature of the evidence base, conclusions regarding effectiveness, workload reduction, equity implications, and learning outcomes should be interpreted as analytically inferred patterns across the reviewed literature rather than direct empirical findings produced by this study. The framework proposed is therefore conceptual and heuristic, intended to guide future empirical validation.

2.2. Defining Differentiated Instruction (DI)

Differentiated instruction, championed by figures like Carol Ann Tomlinson, advocates for varying educational practices to suit the individual learning needs of students. The concept of differentiated instruction, advanced by Carol Ann Tomlinson, proposes a holistic educational approach that systematically adjusts content, process, and product to maximize the learning potential of every student by addressing their unique readiness levels, interests, and learning profiles. Silmi et al. (2025) conducted a systematic review solidifying the application of DI in modern education and concluded that differentiated learning was effective "through the adaptation of content, processes, products, and learning environments" (p. 61). Godor (2021) stated, "Differentiated instruction occurs when teachers use students' level of readiness, interests, and learning preferences to adjust the content, process, or products." (p. 43).

The core tenets of DI focus on adjusting the Content (what students learn), Process (how they learn it), and crucially, the Product (how they demonstrate learning). Importantly, if content, process, and product are systematically differentiated, assessment practices cannot remain standardized without compromising construct validity. When students are permitted to demonstrate learning through varied products and learning pathways, a single uniform assessment structure risks constructing underrepresentation or construct-irrelevant variance. Differentiated assessment is therefore not merely a complementary extension of differentiated instruction, but a logical requirement of it. Criteria, rubrics, and forms of acceptable evidence must also be differentiated in order to validly capture diverse demonstrations of learning. This theoretical interdependence becomes particularly salient in the AI era, where variability in process and product is further amplified through digital and generative tools. In the AI era, the product dimension gains paramount importance. Differentiated products move beyond simple essays to include multimedia presentations, digital portfolios, case study analyses, architectural models, or even code repositories, allowing students to express mastery using their unique strengths and interests. Spyropoulou et al. (2025) highlight the necessity for technological support so that teachers can apply DI "For DI to be implemented, there needs to be the necessary technological support, allowing secondary education teachers to apply it both in the production of materials and in teaching. In addition, students themselves must be able to use, apply, and create with these tools" (p. 74). Bouchrika (2026) explicitly links product differentiation to offering student choice in contemporary classroom contexts. The article suggests that teachers should "offer options such as writing a report, creating an organizer or a visual flow of the lesson, delivering an oral report, and putting up a show based on the lesson" (para. 12). It further identifies digital portfolios as a powerful alternative to traditional examinations, noting that these may be "part written, part visual, and part presentation" (para. 14). This supports the argument for incorporating varied digital formats within differentiated assessment design. Arqam and Asrifan (2024) conclude that integrating AI into Project-Based Learning (PBL) holds significant potential for transforming differentiated English language instruction. Since PBL inherently promotes complex outputs beyond simple essays, AI enhances this capability by providing tools such as personalized virtual assistants, smart grading systems, and an immersive learning environment.

2.3. The Limitations of Conventional Assessment in the AI Era

Traditional assessments are often designed for efficiency and ease of scoring, emphasizing recall and standardized application. The advent of LLMs, however, means that the cognitive load of producing standard text is dramatically reduced. When an LLM can generate a passing five-paragraph essay or solve a textbook problem in seconds, the assessment no longer reliably measures the student's unique knowledge or skill development. This erosion of validity necessitates a fundamental redesign. If the assessment can be easily completed by an AI, it is measuring the wrong skill or no skill at all. Gonsalves (2025) asserts that the emergence of Generative AI (GenAI) has critically undermined the validity and reliability of conventional assessment methodologies. This loss of validity stems from evaluative measures' inability to reliably differentiate between AI-generated content and authentic student production. The author therefore calls for deeper integration of contextualized assessment design to elicit critical thinking and real-world application skills. Similarly, Williams (2025) argues that the traditional academic essay is becoming increasingly outdated and inauthentic in light of advances in computational tools and artificial intelligence. Empirical grading studies suggest that AI-generated responses to open-ended university essay questions frequently achieve high classification scores under conventional assessment rubrics and, in some cases, receive first-class grades from independent examiners. Extending this critique, Matheis and John (2024) contend that conventional assessment paradigms fail to capture diverse forms of knowledge and are therefore vulnerable to exploitation through student misuse of Generative AI tools. Collectively, these arguments underscore the growing difficulty of distinguishing between student- and AI-generated work, raising serious concerns about the validity and authenticity of traditional evaluation processes. Consequently, a strategic shift toward more authentic, multimodal, and interactive assessment designs appears increasingly necessary, particularly as static and unimodal tasks are more easily replicated by contemporary AI systems.

3. The Framework: Differentiated Assessment and AI as an Ally

3.1. Principle of Differentiated Assessment Products

To thrive in the AI era, assessment tasks must evolve to demand uniquely human attributes. This requires a focus on forgery-resistant tasks that are anchored in personal experience, local context, or complex, multi-modal synthesis. Recent scholarship converges on the need to redesign assessment tasks to foreground higher-order cognitive engagement and contextual specificity. These studies further emphasize the importance of contextualized and process-oriented evaluation in AI-mediated learning environments (Farag et al., 2024; Khlaif et al., 2025; Lewis et al., 2024). Collectively, these frameworks emphasize authenticity, creativity, and applied reasoning as dimensions less susceptible to AI replication. Rather than treating these contributions as isolated strategies, they can be interpreted as part of a broader shift toward competency-centered, performance-based evaluation structures. Specifically, they identify creativity, critical thinking, and advanced problem-solving abilities as core, uniquely human attributes essential for contemporary academic success and future workforce readiness. This shift requires educators to develop assessment instruments rooted in authenticity that specifically target these higher-order cognitive skills.

3.1.1. Designing AI-Resistant Tasks

Tasks must require students to synthesize information or apply knowledge in ways an AI cannot easily replicate, such as proposing a policy solution for a specific local school district, or integrating self-reflection on their personal learning journey. Gonsalves (2025) proposes an expanded theoretical framework for contextual learning, encompassing five dimensions: practical, situational, experiential, interdisciplinary, and ethical. The framework mandates the integration of diverse and specific contexts (e.g., social, personal) into assessment design to promote higher-order thinking and real-world application, directly challenging AI's reliance on generic data. Tucker (2024) offers practical strategies centered on personalization. Key recommendations include: Mandating personal reflection or lived experiences (capitalizing on

AI's lack of a personal narrative), requiring the use of specific, obscure local, school, or community resources (information inaccessible to generic LLMs) and challenging students to justify their process (reinforcing metacognition). The University of Chicago (2024) focuses on enhancing metacognition and authenticity in assessment design. Their resources recommend implementing learning journals or reflective diaries that prompt students to self-evaluate their progress against Intended Learning Outcomes (ILOs) and formulate improvement plans. Furthermore, the university advocates for context-specific assessments, such as analyzing case studies relevant to obscure local features. Tomisu et al. (2025) propose a paradigm shift to the "Cognitive Mirror", where AI is reframed as a "teachable novice" (p. 1). This model compels the learner to teach the AI a specific concept, with the constrained quality of the AI's response reflecting the quality and completeness of the learner's explanation. This process inherently forces deep reflection and refinement of the knowledge base.

3.1.2. Diversification of Products

Educators must offer a menu of assessment formats. The strategic shift in evaluative focus, moving from the final product to process documentation, is critical. This approach simultaneously enhances the authenticity of tasks, fosters opportunities for differentiated assessment, and acts as an effective barrier against the misuse of generative artificial intelligence tools.

Digital Portfolios characterized by the documentation of the complete learning trajectory rather than the exclusive focus on the summative final outcome, constitute a core pedagogical strategy. This modality is particularly critical when confronting the challenges posed by Generative Artificial Intelligence, and furthermore, it functions as a powerful mechanism for implementing differentiated assessment. E-portfolios are widely positioned as vehicles for authentic, process-oriented assessment that foreground iterative development and reflective documentation (Gonsalves, 2025; Todeschini & Sollberger, 2023). Rather than merely serving as repositories of artifacts, they function as structured evidence ecosystems that capture both cognitive progression and contextual application. E-Portfolios effectively provide strong evidence for the validity, reliability, and authenticity of demonstrated student skills, as they necessitate continuous, sustained engagement with course materials and learning objectives. The requirement to document iterative development processes, reflective decision-making, and contextualized learning trajectories significantly reduces the likelihood that generative AI alone can produce fully authentic portfolio evidence, particularly when assessment includes process verification and dialogic review, thereby enhancing the trustworthiness of the assessment process.

Simulations/Prototype requires tangible application of knowledge (e.g., building a small scale model or developing a user interface). Khlaif et al. (2025) advocate for AI-resistant assessment methodologies that center on tasks mandating the demonstration of personal application of knowledge, original ideation, and non-replicable human competencies. The authors specifically emphasize performance-based assessments, highlighting those that require students to construct or create a tangible output, such as a physical model or a digital product. RMIT University (2025) underscores six characteristics integral to authentic assessment, specifically highlighting the requirement for an artefact or performance as the submitted evidence, and the necessity of demonstrating transferable knowledge and skills with real-world application.

Oral Defence/Presentation can test extemporaneous knowledge and critical thinking through real-time dialogue. The Oral Defence/Presentation is a crucial strategy within both differentiated assessment and AI-resistant assessment frameworks, as it tests extemporaneous knowledge and critical thinking through mandatory real-time dialogue, thus requiring the spontaneous articulation of understanding rather than the submission of pre-generated content. Khlaif et al. (2025) explicitly propose oral evaluations as an integral component of a novel assessment framework designed to mitigate the risks associated with generative AI. This strategy is classified as highlighting the demonstration of personal application and higher-order thinking skills, which remain challenging for AI systems to replicate effectively within a live, dialogic environment. Crucially, the paper emphasizes that this real-time dialogue grants the assessor the ability to conduct deeper probes into student comprehension. Interactive Oral Assessment is characterized as a "robust, authentic assessment design approach" (O'Riordan et al., 2026, p. 1). They assert that real-time dialogue enables the assessor to conduct deeper probes, posing peripheral questions or demanding immediate clarification, thereby presenting a significant difficulty for students reliant on AI-generated content.

3.1.3. Distinction From Existing AI-Resistant and Authentic Assessment Models

While recent frameworks (e.g., AI-resistant assessment models, contextual assessment designs, and authentic performance-based approaches) emphasize task redesign to mitigate AI misuse, the DAP framework proposed here extends beyond resistance. First, it systematizes product diversification as a structural principle rather than a supplementary strategy. Second, it integrates AI not merely as a constraint to be managed but as an analytic partner in evaluating multimodal and process-based evidence. Third, it explicitly aligns differentiated product design with scalable AI-supported feedback mechanisms, addressing the longstanding critique that differentiated assessment is pedagogically sound but operationally unsustainable.

Thus, the conceptual advancement lies in repositioning AI from an external threat requiring defensive redesign toward an embedded evaluative infrastructure that enables differentiated assessment at scale.

3.2. AI's Role in Streamlining Differentiated Assessment

The biggest challenge for differentiated assessment is the increased teacher workload involved in evaluating diverse, complex, and creative outputs. This is where AI moves from adversary to indispensable ally. Thomas (2025) directly explores how several implementation studies and institutional pilot reports indicate the potential of AI-driven assessment systems to reduce elements of teacher workload, particularly in relation to formative feedback and routine evaluation tasks. The Tulsyan (2025) report from the Observer Research Foundation (ORF), titled "Every Teacher's Ally: Harnessing AI for Equity and Excellence", designates AI as a crucial partner for educators. The report specifically underscores AI's capacity to deliver tailored student feedback and facilitate the management of multi-grade teaching, which serves as a vital component of differentiated instruction. Several emerging AI-supported platforms illustrate the operational feasibility of differentiated assessment by assisting educators in generating customized instructional materials and automating formative feedback processes. These implementations demonstrate how AI can support scalability without displacing pedagogical judgment. These tools are explicitly marketed as making "differentiation doable" by specifically addressing the challenges inherent in differentiated assessment and resource creation. They are purpose-built to assist teachers in generating and customizing instructional materials and assessments for diverse classroom populations.

3.2.1. Automated Formative Feedback

AI can automate the assessment of fundamental components across diverse products (e.g., checking code syntax, analyzing clarity and organization in a presentation script, or checking citation consistency). This frees the teacher to focus on higher-order cognitive analysis the depth of the student's insight, the novelty of their solution, and their critical application of the concepts. Thomas (2025) argues that AI-based systems streamline the evaluation of fundamental components, thereby allowing educators to reallocate their time toward mentoring and higher-order tasks, such as nurturing creativity and critical thinking. Florida International University (FIU) Center for the Advancement of Teaching (2025) designed strategies for assessing learning with AI. This resource advocates for using AI to automate repetitive tasks (like spotting common errors), which in turn, allows faculty time for deeper engagement mentoring, dialogue, and individualized guidance. It stresses that authentic assignments requiring critical thinking, reflection, and application to real-world contexts are essential in the age of AI. Fajardo-Ramos et al. (2025) highlight a significant shift in the teacher's role, stressing that authentic assignments requiring critical thinking, reflection, and application to real-world contexts are essential in the age of AI. It notes that AI transfers routine assistance and parts of integrity checking to the platform. The teacher's new central task evolves toward setting transparent criteria, explaining automated checks, and documenting decisions, ensuring the human-in-the-loop maintains focus on validating the construct of the assessment (i.e., whether the assignment genuinely tests critical thinking).

3.2.2. Analysis of Complex Outputs

Advanced AI and machine learning tools offer new capacities for assessment. Sentiment and tone analysis can evaluate the persuasiveness and coherence of argumentative essays or speeches, while pattern recognition can identify common learning trajectories or misconceptions across a large, diverse dataset of student work, including video submissions and design concepts (Hafdi & El Kafhali, 2025; Naik et al., 2025). Demonstrating this capability, Alawad (2025) utilized AI-powered tools to evaluate the argumentative structure of essays across sequential. It moves beyond simple grading to analyze high-level features like the strength of the thesis, evidence quality, and integration of counterarguments. By tracing performance across submissions, the AI identifies learning trajectories (progression of skills) and patterns of argumentation strength, which is crucial for delivering differentiated feedback to students on their persuasive writing coherence. According to Naik et al. (2025), students' interaction with and evaluation of generative AI tools in HCI design courses often involve treating design concepts as complex outputs. It identifies new forms of judgment students make, such as agency-distribution judgment (deciding how much creative responsibility to assign to the AI) and reliability judgment. This pattern recognition of student interaction and evaluation provides data on their metacognitive learning trajectories within a complex, non-textual domain (Naik et al., 2025). Hafdi and El Kafhali (2025) apply various machine learning (ML) algorithms, such as long short-term memory (LSTM) and support vector machine (SVM), to predict student success in coding courses by combining academic history with weekly behavioral performance data. The ML models recognize patterns in early behavior that correlate with later success or failure. This predictive capacity is the foundation of differentiated assessment, allowing instructors to implement early, tailored interventions and personalized scaffolding for students predicted to struggle, ensuring instruction meets their individual readiness levels.

3.2.3. Personalized learning support

By analyzing continuous, diverse assessment data, Adaptive AI systems have demonstrated the capacity, within controlled implementations, to identify student learning gaps and recommend targeted resources, thereby supporting differentiated instructional and assessment practices. Ye (2025) develops an AI model for foreign language learning that constructs tailor-made learning paths and resource recommendations by continuously analyzing multi-dimensional student behavior. The system dynamically adjusts the difficulty of learning content and suggests specific resources using deep collaborative filtering and attention mechanisms. The goal is to optimize the learning path and enhance efficiency by ensuring students receive the exact support needed to close gaps before they face summative assessments. Dagunduro et al. (2024) explore how adaptive learning systems leverage advanced algorithms to provide personalized instruction tailored to individual student needs. By processing data from student interactions and continuous formative assessment, these algorithms identify each student's Zone of Proximal Development (ZPD) and learning gaps. This allows the system to instantly provide targeted interventions (like instructional videos or practice exercises) or present more advanced content, effectively ensuring that the evaluation measures true mastery achieved through differentiated support. Tensen et al. (2026) investigate how Generative AI can serve as a supplementary source of expedient formative feedback on complex academic outputs (thesis drafts). By using AI to evaluate drafts against specific criteria, students receive instant, structured feedback on areas needing improvement (like theoretical foundation or clarity). This immediate, non-judgmental scaffolding encourages students to iterate and address learning gaps before official human review or final submission (summative assessment), making the final evaluation a true measure of refined work. Taken together, these developments suggest that assessment transformation in the AI era requires not merely task modification, but structural integration of differentiated product design, process transparency, and AI-supported analytics.

4. Challenges and Ethical Considerations

4.1. Academic Integrity Redefined

The traditional focus on detecting cheating is becoming obsolete. The new challenge is redefining academic integrity as the transparent and ethical use of AI. The solution lies in pedagogical design: moving from an "AI detection arms race" to designing tasks where the value is placed on the contribution and reflection of the student, and the use of AI is explicitly acknowledged and referenced as a co-pilot tool. The redefinition of academic integrity in the era of generative AI necessitates a shift in pedagogical design, prioritizing transparency and student accountability. As articulated by the University of Alberta (2025), contemporary academic policy must establish explicit definitions regarding the permissible and impermissible uses of AI across diverse assessment types to ensure clarity on expectations. Furthermore, to uphold the ethical use of digital tools, institutions must mandate transparency and reflection, requiring students to fully document AI usage, including the specific purpose for its employment and the precise prompts utilized, typically submitted in a supplemental reflective statement. Crucially, this framework emphasizes AI as a Tool, asserting that while AI necessitates explicit acknowledgment and proper citation, students retain ultimate responsibility for the accuracy, ethical implications, and scholarly integrity of all incorporated outputs. This approach moves beyond an 'AI detection arms race' toward a model where AI is treated as a co-pilot, and the value is placed on the student's contribution, critical judgment, and honest methodological disclosure. The Office of Academic Integrity and Student Conduct at Harvard University advocates for a proactive pedagogical strategy that encourages instructors to "Design with AI in Mind". Their recommendations center on maintaining scholarly integrity by fostering critical engagement rather than prohibition. Specifically, this approach calls for ethical use evaluations, where students are required to submit a component of the assignment that outlines precisely how AI supported their work, with a focused assessment on the ethical implications and judgment applied in its use. Furthermore, instructors are advised to design for critical thinking, developing assignments that strategically leverage AI for idea generation or problem-solving assistance while simultaneously demanding substantial individual contribution and personal evaluation of the AI-produced content. This framework reframes the challenge of AI integration into an opportunity to cultivate advanced student literacy and ethical reasoning.

4.2. Equity and Algorithmic Bias

The deployment of AI in assessment is fraught with ethical risks. Algorithmic bias where AI systems inadvertently penalize non-native speakers, students with unique learning styles, or minority groups due to biased training data is a grave concern. To maintain an equitable learning environment, any AI-powered assessment tool must undergo rigorous bias audits, and human oversight must remain the final arbiter of student grades. Furthermore, institutions must address the emerging digital divide, ensuring equitable access to the very AI tools required to complete the new differentiated tasks. The increased deployment of AI in educational assessment introduces significant ethical risks, most notably the threat of algorithmic bias. As Boateng and Boateng (2025) demonstrated in their examination of AI-driven decision-making, such bias, often rooted in unrepresentative historical training data, can inadvertently perpetuate or exacerbate existing educational disparities, disproportionately penalizing marginalized communities, including non-native speakers or students with non-traditional learning styles. Consequently, to safeguard an equitable learning environment, any AI-powered assessment tool must undergo rigorous bias audits to ensure fairness across demographic groups. Crucially, the final decision-making authority must remain with human oversight to act as the ultimate arbiter of student grades, thereby mitigating the systemic failures of biased algorithms. Furthermore, institutions bear the responsibility of addressing the emerging digital divide, ensuring that all students have equitable access to the very AI tools required to successfully complete the new, differentiated learning tasks, preventing technology from creating new systemic barriers to opportunity.

4.3. Implementation and Training

The transition to a differentiated, AI-supported assessment framework is a massive logistical challenge. It requires substantial investment in teacher professional development, equipping educators not just with new AI tools, but with the pedagogical skills to design truly creative, differentiated tasks and to interpret the AI-generated feedback accurately. A growing number of professional development (PD) models and curricula are being proposed, notably the cohort-based PD approach. This model mandates that educators engage in sustained, semester-long experimentation with specific AI applications within their subject matter. Crucially, this framework also addresses the necessity of designing AI-incorporated assessments. These innovative assessments serve a dual function: to evaluate student proficiency and simultaneously act as a powerful professional learning instrument for the educators themselves. Borović et al. (2025) noted that dedicated AI platforms, such as Brisk Teaching, are being engineered to support instructional processes, streamline feedback generation, and enhance student motivation, thereby making personalized learning more achievable. Simultaneously, structured learning modules, including the Generative AI for Educators with Gemini course, provide the essential foundation, focusing on practical skills to enhance efficiency, personalize instruction, and generate comprehensive student feedback. This dual focus on specialized tooling and practical PD is crucial to ensure educators move beyond mere tool usage to acquire the sophisticated pedagogical skills needed to design truly creative, differentiated tasks and accurately interpret the complex data derived from AI-generated feedback.

4.4. Theoretical Contributions

This paper contributes to assessment theory in three primary ways:

First, it reconceptualizes differentiated instruction's "product" dimension as the central axis of assessment redesign in AI-mediated environments. While DI theory traditionally frames product as one of three adjustable components (content, process, product), this study elevates product diversification into a foundational assessment principle.

Second, it proposes an integrated model that connects differentiated assessment products with AI-supported analytic capacity. Existing authentic assessment frameworks emphasize realism and performance but often neglect scalability; AI-resistant frameworks focus on integrity but under-theorize differentiation. The present model bridges these domains by linking multimodal assessment design with AI-assisted evaluative analytics.

Third, it advances a normative shift in academic integrity discourse, reframing AI use from prohibition toward transparent co-agency. This conceptual repositioning contributes to emerging scholarship that views AI literacy and ethical disclosure as core assessment competencies rather than compliance mechanisms.

Together, these contributions move the discussion from reactive adaptation toward proactive assessment redesign.

5. Conclusion

5.1. Summary of Findings

The reviewed literature suggests that traditional assessment methods are increasingly challenged in their capacity to validly capture student learning within AI-mediated environments. While not entirely obsolete, their effectiveness appears constrained when applied to contexts where generative AI can replicate conventional outputs. This paper establishes that the future of evaluation lies in adopting differentiated assessment, which leverages diverse student products to capture holistic learning. AI, when deployed ethically and strategically, is the essential tool for making this individualized evaluation scalable, efficient, and data-informed.

5.2. Policy and Pedagogical Implications

Educational policy bodies must prioritize the development of clear guidelines on the ethical use and required transparency of AI in assessment. Educators, in turn, must embrace their new role as curators of AI-resistant, high-challenge learning experiences, focusing their energy on evaluating creativity, synthesis, and critical reflection, rather than rote output.

5.3. Limitations of the Evidence Base

As this study is grounded in secondary literature synthesis, its conclusions are constrained by the scope, methodological diversity, and contextual variability of the reviewed sources. Many cited studies report pilot implementations, localized institutional trials, or discipline-specific interventions, which may limit generalizability across broader educational systems. Consequently, the proposed framework should be interpreted as a conceptual model requiring empirical validation through longitudinal and comparative assessment research.

5.4. Future Research Directions

Future research should focus on empirical studies comparing student outcomes and equity metrics under traditional versus AI-supported differentiated assessment models. Furthermore, work is needed to develop universally accepted standards and best practices for auditing AI-driven grading systems to mitigate bias and ensure fairness across all student populations. The ultimate goal is not just to measure knowledge more accurately, but to actively cultivate a system that fosters genuine student growth in a world co-created with intelligent machines. This paper proposes a conceptual framework suggesting that differentiated assessment, when structurally integrated with differentiated instruction and supported by AI-enabled analytics, may enhance construct validity, equity, and process transparency in contemporary learning environments.

Rather than presenting definitive empirical claims, this framework advances a set of theoretical propositions:

(1) that differentiated instruction logically necessitates differentiated assessment criteria.

(2) that AI can function as a process-tracing and feedback mechanism rather than merely a content generator.

(3) that academic integrity in the AI era may be more sustainably grounded in transparency than prohibition.

Future empirical research is required to test, refine, and contextualize these propositions across disciplinary and institutional settings.

References

Alawad, E. A. (2025). AI-driven analysis: A case study on assessing argumentation proficiency and skill development through sequential submissions. *Linguistic and Philosophical Investigations, 24*(1), 108–117. https://www.researchgate.net/publication/389503034_AI-Driven_Analysis_A_Case_Study_on_Assessing_Argumentation_Proficiency_and_Skill_Development_Through_Sequential_Submissions

Alexandrowicz, V. (2024). Artificial intelligence integration in teacher education: Navigating benefits, challenges, and transformative pedagogy. *Journal of Education and Learning, 13*(6), 346–364. https://doi.org/10.5539/jel.v13n6p346

Arqam, & Asrifan, A. (2024). Integrating AI in project-based learning for differentiated English language instruction: A scoping review. *Journal of English Education and Teaching, 8*(3), 586–608. https://doi.org/10.33369/jeet.8.3.586-608

Awadallah Alkouk, W., & Khlaif, Z. N. (2024). AI-resistant assessments in higher education: Practical insights from faculty training workshops. *Frontiers in Education, 9*, Article 1499495. https://doi.org/10.3389/feduc.2024.1499495

Bagida Urazaliyevna, Y. (2025). Enhancing differentiated instruction: how ai tools support personalized lesson planning for diverse learners. *Hamkor Konferensiyalar*, *1*(20), 108–114. https://academicsbook.com/index.php/konferensiya/article/view/2700

Boateng, O., & Boateng, B. (2025). Algorithmic bias in educational systems: Examining the impact of AI-driven decision making in modern education. *World Journal of Advanced Research and Reviews*, *25*(1), 2012–2017. https://doi.org/10.30574/wjarr.2025.25.1.0253

Borović, F., Kovačić, M., & Aleksić-Maslać, K. (2025, July). *The use of the Gen AI tool Brisk Teaching in the educational process and its impact on student motivation* [Paper presentation]. 48th International Convention on Information, Communication and Electronic Technology (MIPRO 2025), Opatija, Croatia. https://doi.org/10.1109/MIPRO65660.2025.11131911

Bouchrika, I. (2026, March 14). *Differentiated instruction: Definition, examples & strategies for the classroom for 2026.* Research.com. https://research.com/education/differentiated-instruction

Dagunduro, A. O., Chikwe, C. F., Ajuwon, O. A., & Ediae, A. A. (2024). Adaptive learning models for diverse classrooms: Enhancing educational equity. *International Journal of Applied Research in Social Sciences, 6*(9), 2228–2240. https://doi.org/10.51594/ijarss.v6i9.1588

Damyanov, K. (2024). Differentiation of educational content through artificial intelligence systems in inclusive education. *International Journal of Education, 12*(3), 13-20. https://doi.org/10.5121/ije2024.12302

Fajardo-Ramos, D. C., Andrés, C., & Mella-Norembuena, J. (2025). Human-in-the-loop assessment with AI: Implications for teacher education in Ibero-American universities. *Frontiers in Education, 10*, Article 1710992. https://doi.org/10.3389/feduc.2025.1710992

Farag, W. A., Nadeem, M., & Helal, M. (2024). Assessment transformation in the age of AI: Moving beyond the influence of generative tools. In *Proceedings of the 2024 Mediterranean Smart Cities Conference (MSCC)* (pp. 1–6). IEEE. https://doi.org/10.1109/MSCC62288.2024.10697011

Florida International University - Center for the Advancement of Teaching. (2025). *Design strategies for assessing learning with AI*. https://cat.fiu.edu/resources/teaching-with-ai/design-strategies-for-assessing-learning-with-ai/

Godor, B. P. (2021). The many faces of teacher differentiation: Using Q methodology to explore teachers' preferences for differentiated instruction. *The Teacher Educator, 56*(1), 43–60. https://doi.org/10.1080/08878730.2020.1785068

Gonsalves, C. (2025). Contextual assessment design in the age of generative AI. *Journal of Learning Development in Higher Education, 34,* 1-14. https://files.eric.ed.gov/fulltext/EJ1464858.pdf

Hafdi, Z. S., & El Kafhali, S. (2025). A comparative evaluation of machine learning methods for predicting student outcomes in coding courses. *AppliedMath, 5*(2), 75. https://doi.org/10.3390/appliedmath5020075

Iqbal, M., Khan, N. U., & Imran, M. (2024). The role of artificial intelligence in transforming educational practices: Opportunities, challenges, and implications. *Qlantic Journal of Social Sciences, 5*(2), 348–359. https://doi.org/10.55737/qjss.349319430

Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences, 15*(2), 174. https://doi.org/10.3390/educsci15020174

Lewis, H., Johnson, S. M., & Phillips, E. (2024). Redefining student achievement: Authentic assessment in the era of artificial intelligence. In L. Marron (Ed.), *Cases on authentic assessment in higher education* (pp. 17–36). IGI Global. https://doi.org/10.4018/979-8-3693-1001-4.ch002

Maity, S., & Deroy, A. (2024). The future of learning in the age of generative AI: Automated question generation and assessment with large language models. *arXiv*. https://arxiv.org/abs/2410.09576

Manokore, K., Manokore, A. S., & Chigora, T. B. (2025). Educational assessments in the age of generative AI: Transforming Educational Assessment in Africa - Opportunities and challenges of generative AI. In P. Wachira, X. Liu, & S. Koc (Eds.), *Educational assessments in the age of generative AI* (pp. 59–92). IGI Global. https://doi.org/10.4018/979-8-3693-6351-5.ch003

Matheis, P., & John, J. J. (2024). Reframing assessments: Designing authentic assessments in the age of generative AI. In S. Mahmud (Ed.), *Academic integrity in the age of artificial intelligence* (pp. 139–161). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-0240-8.ch008

Naik, S., Shukla, P., Obi, I., Backus, J., Rasche, N., & Parsons, P. C. (2025). Tracing the invisible: Understanding students' judgment in AI-supported design work. In *Proceedings of the 2025 Conference on Creativity and Cognition* (pp. 438–442). Association for Computing Machinery. https://doi.org/10.1145/3698061.3734399

Ning, Y., Zhang, C., Xu, B., Zhou, Y., & Wijaya, T. T. (2024). Teachers' AI-TPACK: Exploring the relationship between knowledge elements. *Sustainability, 16*(3), 978. https://doi.org/10.3390/su16030978

O'Riordan, F., Thangaraj, J., Girme, P., & Ward, M. (2025). Interactive oral assessment: Staff perceptions, challenges and benefits of this robust, authentic assessment design approach. *Innovations in Education and Teaching International*, 63(2), 494-507. https://doi.org/10.1080/14703297.2025.2477160

Ruslim, M. I., & Khalid, F. (2024). The use of artificial intelligence in differentiated instruction classrooms. *International Journal of Academic Research in Business and Social Sciences, 14*(8), 680–695. https://dx.doi.org/10.6007/IJARBSS/v14-i8/22435

Silmi, T. A., Lubis, A., & Maulidi, A. (2025). Model differentiated learning in 21st century education: A systematic review of strategies, results, and challenges. *Al Ulya: Jurnal Pendidikan Islam, 10*(1), 60–80. https://doi.org/10.32665/alulya.v10i1.3572

Spyropoulou, E., Wallace, M., & Poulopoulos, V. (2025). Differentiated education using technology in junior and high school classrooms. *Encyclopedia, 5*(2), 71. https://doi.org/10.3390/encyclopedia5020071

Tensen, D., Grainger, P., & Graham, W. (2026). Using AI to generate formative feedback in doctoral education. *Assessment & Evaluation in Higher Education, 51*(3), 476-492. https://doi.org/10.1080/02602938.2025.2536558

Thomas, P. (2025). AI-driven assessment systems: Reducing teacher workload in grading and feedback. https://www.researchgate.net/publication/396245710_AI-Driven_Assessment_Systems_Reducing_Teacher_Workload_in_Grading_and_Feedback

Todeschini, B., Sollberger, E., & Hugo, A. (2023). E-portfolios for student assessment in the age of artificial intelligence. In *Proceedings of 16th Annual International Conference of Education, Research and Innovation* (p. 76). Academic Conferences International. https://doi.org/10.21125/iceri.2023.2144

Tomisu, H., Ueda, J., & Yamanaka, T. (2025). The cognitive mirror: A framework for AI-powered metacognition and self-regulated learning. *Frontiers in Education, 10*, Article 1697554. https://doi.org/10.3389/feduc.2025.1697554

Tucker, C. (2024). *5 tips for designing AI-resistant tasks.* https://catlintucker.com/2024/10/ai-resistant-tasks/

Tulsyan, A. (2025). *Every teacher's ally: Harnessing AI for equity and excellence*. Observer Research Foundation. https://www.orfonline.org/expert-speak/every-teacher-s-ally-harnessing-ai-for-equity-and-excellence

UNESCO. (2021). *AI and education: Guidance for policy-makers*. UNESCO Publishing. https://doi.org/10.54675/PCSP7350

University of Chicago. (2024). *Academic technology solutions*. https://academictech.uchicago.edu/

U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf

Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence on students' academic development. *Education Sciences, 15*(3), 343. https://doi.org/10.3390/educsci15030343

Williams, A. (2025). Integrating artificial intelligence into higher education assessment. *Intersection: A Journal at the Intersection of Assessment and Learning, 6*(1), 128-154. https://doi.org/10.61669/001c.131915

World Economic Forum. (2023). *The future of jobs report 2023*. https://www.weforum.org/publications/the-future-of-jobs-report-2023/

Ye, L. (2025). AI-driven personalized learning path optimization and adaptive resource recommendation for foreign language education. In R. Su (Ed.), *Proceedings of the Seventh International Conference on Image, Video Processing, and Artificial Intelligence* (Vol. 13731, Article 137311B). SPIE. https://doi.org/10.1117/12.3076209

Zhong, R., & Zhao, Y. (2025). Paradigm shifts in education: An ecological analysis. *ECNU Review of Education, 8*(1), 21-40. https://doi.org/10.1177/20965311241296162