Research Journal in Advanced Humanities

RJAH

https://doi.org/10.58256/738q2y02

Open Access

Check for updates

REVIEW ARTICLE

Section(s): *Literature, Linguistics & Criticism*

AI-enhanced creative pedagogy: A systematic review of AI interventions in fostering students' creativity across digital humanities contexts

Hamoud A. Alshehri[1], **Faisal Bin Shabib Mosleet Alsubaie**[2], **Mohamed Sayed Abdellatif**[3]*, **Mohamed Ali Nemt-allah**[4]

[1]Department of Curriculum and Instruction, College of Education, King Saud University, Saudi Arabia

[2]Department of Educational Sciences, College of Education in Al-Kharj, Prince Sattam Bin Abdulaziz University, Saudi Arabia

[3]Department of Psychology, College of Education in Al-Kharj, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

[4]Educational Psychology and Statistics Department, Faculty of Education, Al-Azhar University, Dakahlia, Egypt

*Correspondence: m.heby@psau.edu.sa

ABSTRACT

Traditional pedagogical approaches systematically constrain creative thinking development through teacher-directed environments, standardized assessments, and implicit rather than explicit attention to creativity outcomes. This systematic review synthesizes empirical evidence on the effectiveness of artificial intelligence interventions in fostering students' creativity across digital humanities contexts. Following PRISMA 2020 guidelines, comprehensive searches of Web of Science, Scopus, ERIC, PsycINFO and IEEE Xplore identified 19 studies involving 1,695 participants across elementary through higher education contexts. Included studies employed experimental and quasi-experimental designs that examined AI tools—including large language models, visual generative systems, and specialized platforms—integrated into constructivist pedagogical frameworks. Meta-analytic findings demonstrate that AI interventions consistently produce large, statistically significant effects on creative thinking, with Cohen's d values exceeding 1.0 in multiple contexts. Fluency exhibited the most significant advancements across all educational tiers, whereas enhancements in originality and elaboration were dependent on structured scaffolding protocols. However, concerns regarding cognitive dependency, authenticity erosion, and academic integrity emerged consistently across studies. Assessments of publication bias showed that there were manageable threats to validity, and the fail-safe N analysis showed strong results. Effectiveness depends fundamentally on pedagogical orchestration positioning AI as a reflective partner rather than an autonomous replacement, with teacher competence and ethical frameworks proving decisive.

KEYWORDS: artificial intelligence, creative thinking, generative AI, digital humanities pedagogy, educational technology, creativity

Research Journal in Advanced Humanities

Volume 7, Issue 1, 2026

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

ARTICLE HISTORY

Submitted: **04 November 2025**

Accepted: 30 **December 2025**

Published: 09 **January 2026**

HOW TO CITE

Abdellatif, M., Alshehri, H. A., Alsubaie, F. B. S. M., & Nemt-allah, M. A. (2026). AI-enhanced creative pedagogy: A systematic review of AI interventions in fostering students' creativity across digital humanities contexts. *Research Journal in Advanced Humanities, 7*(1). https://doi.org/10.58256/738q2y02

Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited

Introduction

Across educational levels from elementary through higher education, traditional pedagogical approaches systematically constrain creative thinking development through several interconnected challenges. Teacher-directed, grade-focused environments emphasizing memorization and hierarchical authority suppress independent thinking, risk-taking, and innovation essential for creativity (Duval et al., 2023; Fleet & Dobson, 2023; Mardatillah et al., 2025). Standardized assessments and rigid curricula limit open-ended exploration, problem-finding, and divergent thinking opportunities that underpin creative fluency, flexibility, originality, and elaboration (Paz-Baruch et al., 2025; Samaniego et al., 2024; Silva et al., 2022). Furthermore, creativity remains implicit rather than explicit in learning outcomes, particularly in STEM programs, while teachers report low confidence and lack of guidance in teaching creativity (Burakgazi & Reiss, 2025; Feng et al., 2024). These systemic barriers necessitate transformative pedagogical interventions that explicitly cultivate creative thinking across all educational contexts.

In the post-2020 educational landscape, generative AI has emerged as a transformative tool for creative pedagogy through the convergence of three critical factors. First, technological advancements in large language models and image generators have matured to produce fluent, multimodal outputs across text, images, video, and code, enabling their integration into everyday creative tasks (Albakry et al., 2025; Chiu, 2023; Han & Cai, 2023; Lee & Suh, 2024; Peláez-Sánchez et al., 2024). Second, cloud-based, freemium access models have democratized these tools globally, fostering widespread adoption across secondary, undergraduate, and teacher-education contexts in diverse geographical regions (Chiu, 2023; Qian, 2025; Wu & Zhang, 2025; Zhang & Zhang, 2024). Third, post-pandemic shifts toward Education 4.0 have intensified demands for digital literacy, AI literacy, and innovation skills, positioning generative AI as both a pedagogical medium and content area (Ng et al., 2023; Peláez-Sánchez et al., 2024; Rana et al., 2025; Sun et al., 2025).

Recent empirical evidence demonstrates that AI interventions address creativity's multidimensional nature with varying effectiveness across educational contexts. While AI tools consistently enhance fluency, flexibility, and elaboration through chatbots, generative systems, and collaborative platforms (Abdelmagid, 2025; Hadas et al., 2025; Lobo-Quintero, 2025; Rahman et al., 2025), their impact on originality and sensitivity to problems remains context-dependent and pedagogically mediated (Kabeer et al., 2025; Meliyawati et al., 2025; Song & Song, 2023). Language learning contexts show robust gains across all dimensions when AI supports ideation and narrative development (Kabeer et al., 2025; Li & Wilson, 2025; Song & Song, 2023), whereas STEM applications yield mixed results, particularly regarding originality and depth (Hadas et al., 2025; Rahioui et al., 2025; Yan et al., 2024). Design and multimodal creation benefit from AI-enhanced tools that increase iteration and refinement (Fang et al., 2024; Saritepeci & Durak, 2024; Wang et al., 2025), though concerns about homogenization persist (Lin & Chen, 2024; Malik et al., 2023).

Despite growing interest in AI applications for education, significant gaps persist in understanding AI's effectiveness across diverse contexts and student populations. Current research demonstrates pronounced geographical concentration, with studies predominantly conducted in China, the United States, and other high-income nations, while Africa, the Arab region, and low-resource settings remain critically underrepresented (Crompton & Burke, 2023; Fu et al., 2024; Mustafa et al., 2024; Zawacki-Richter et al., 2019). Educational level disparities are equally evident, as higher education dominates AIED research while early childhood, postgraduate, and lifelong learning contexts receive minimal attention (Crompton & Burke, 2023; Mustafa et al., 2024; Topali et al., 2025). Subject-specific inequities further compound these limitations, with language learning, computer science, and STEM disciplines extensively studied, yet social sciences, arts, communication, and journalism—domains central to creative pedagogy—remaining marginalized (Almasri, 2024; Babacan et al., 2025; Crompton & Burke, 2023; Fu et al., 2024). Additionally, comparative analyses examining different AI tools, pedagogical frameworks, and implementation models across disciplines are notably scarce (Ali et al., 2024; Mustafa et al., 2024; Yan & Qianjun, 2025).

The integration of AI in creative learning environments is fundamentally grounded in constructivist and design-oriented pedagogical frameworks that emphasize active knowledge construction and learner agency. Design thinking approaches guide AI-supported environments through empathy, ideation, and iterative prototyping enhanced by intelligent feedback systems (Atenas et al., 2025; Henriksen et al., 2024; Huang et al., 2025; Zhang et al., 2025). Project-based and problem-based learning methodologies provide authentic contexts where AI delivers adaptive scaffolding and personalized pathways, demonstrably increasing student motivation and engagement (Guo et al., 2020; Sánchez-García & Reyes-De-Cózar, 2025; Tapalova & Zhiyenbayeva, 2022; Trullàs et al., 2022; Wijnia et al., 2024). Flipped classroom models leverage AI tools for just-in-time guidance and creativity support (Li, 2023; Zhao et al., 2021), while Outcome-Based Education frameworks align AI-mediated activities with explicit innovation outcomes (Zhang et al., 2021). Meta-

learning strategies foster self-regulated learning and critical AI literacy (Peng & Li, 2025), supported by comprehensive teacher competence frameworks that ensure ethical and inclusive AI integration (Chick, 2025; Zou et al., 2025).

Contemporary educational environments have witnessed the emergence of diverse AI tools designed to foster student creativity, spanning text-based generative systems, visual design platforms, and specialized educational applications. Text-based large language models such as ChatGPT, Gemini, and similar conversational AI serve as brainstorming companions, writing partners, and ideation engines across creative disciplines (Habib et al., 2024; Güner & Er, 2025; Lobo-Quintero, 2025; Mahama & Amadu, 2025; Ruiz-Rojas et al., 2024; Wei et al., 2025). Visual generative tools including Midjourney, DALL·E, and Stable Diffusion enable rapid concept visualization in design, architecture, and digital storytelling contexts (Habib et al., 2024; Huh et al., 2025; Saritepeci & Durak, 2024; Vartiainen & Tedre, 2023; Wei et al., 2025). These technologies are strategically integrated through project-based learning frameworks, collaborative pedagogies such as AI-enhanced Think-Pair-Share, and co-agency models that position AI as creative collaborators while preserving student autonomy (Fang et al., 2024; Katsenou et al., 2025; Lee & Suh, 2024; Peláez-Sánchez et al., 2024). Recent empirical studies reveal a complex duality in AI's impact on creative development. AI-supported interventions demonstrate significant gains in divergent thinking indicators, including fluency, flexibility, and originality, with particularly pronounced benefits for neurodivergent learners (Fang et al., 2024; Hwang & Wu, 2025; Lobo-Quintero, 2025; Wei et al., 2025). These enhancements operate through psychological mechanisms of increased self-efficacy and reduced anxiety (Hwang & Wu, 2025; Lin & Chen, 2024), with neurophysiological evidence suggesting active cognitive engagement rather than passive consumption (Wang et al., 2025). However, emerging concerns temper this optimism: approximately 45% of AI-assisted teams exhibit cognitive fixation on AI-generated suggestions (Rahman et al., 2025), while students and faculty increasingly voice apprehensions about over-reliance, authenticity erosion, and diminished independent problem-solving capacity (Mahama & Amadu, 2025; Rahioui et al., 2025; Wang, 2024). This tension necessitates carefully scaffolded pedagogical frameworks that position AI as a facilitator rather than substitute, deliberately cultivating critical thinking and metacognitive awareness alongside creative gains (Melker et al., 2025; Yan et al., 2024; Zhou et al., 2025).

This systematic review aims to synthesize empirical evidence on the effectiveness of AI interventions in fostering students' creativity across digital humanities contexts. Specifically, it examines which AI modalities (large language models, visual generative systems, specialized platforms) affect the dimensions of creativity, including fluency, flexibility, originality, and elaboration. The review investigates how educational level, pedagogical frameworks, and disciplinary contexts moderate intervention effectiveness, while assessing methodological quality and potential publication bias. By analyzing quantitative creativity outcomes from experimental and quasi-experimental studies, this review addresses critical gaps in understanding AI's role in creative pedagogy. It provides evidence-based guidance for educators, curriculum designers, and policymakers on integrating generative AI tools in educational settings.

Methods

Protocol and Registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. The review protocol was prospectively registered prior to initiating the systematic search to ensure transparency and minimize reporting bias. The protocol detailed the research objectives, eligibility criteria, search strategy, data extraction procedures, and planned analytical approaches.

Eligibility Criteria

Studies were selected based on predefined inclusion and exclusion criteria structured around the Population, Intervention, Comparison, Outcome, and Study Design (PICOS) framework. The population of interest comprised students across all educational levels, from elementary education through higher education, including secondary and vocational contexts. Eligible interventions included any pedagogical approach incorporating artificial intelligence tools—such as large language models, visual generative systems, machine learning platforms, or specialized AI applications—explicitly designed to foster or enhance creative thinking, creative problem-solving, or creative performance. Comparison conditions included traditional instruction without AI integration, conventional teaching methods, control groups receiving standard pedagogy, or pre-intervention baseline measures within the same participants. Primary outcomes of interest were quantitative measures of creativity, assessed using validated instruments, including dimensions such as fluency, flexibility, originality, elaboration, and problem sensitivity, derived from frameworks analogous to the Torrance

Tests of Creative Thinking or domain-specific creativity assessments. Secondary outcomes included creative performance in discipline-specific contexts such as mathematical creativity, scientific writing creativity, composition ability, and creative problem-solving profiles. Eligible study designs encompassed experimental studies with randomization, quasi-experimental designs with comparison groups, pre-post intervention studies with control conditions, and mixed-methods investigations incorporating quantitative creativity measures. Studies were excluded if they focused solely on AI literacy or technical skills without creativity outcomes, reported only qualitative data without quantitative creativity measures, consisted of theoretical papers or conceptual frameworks without empirical data, involved participants exclusively outside formal educational contexts, or were published in languages other than English.

Information Sources

Comprehensive searches were conducted across multiple electronic databases to ensure exhaustive coverage of relevant literature. The primary databases searched included Web of Science, Scopus, ERIC (Education Resources Information Center), PsycINFO, IEEE Xplore, and Google Scholar. Supplementary search strategies included forward and backward citation tracking of included studies, manual screening of reference lists of relevant systematic reviews, and consultation with subject-matter experts to identify unpublished or in-press studies. Grey literature sources, including conference proceedings and dissertation databases, were also examined to minimize publication bias.

Search Strategy

The search strategy employed a combination of controlled vocabulary terms and free-text keywords tailored to each database's indexing system. Core search concepts included terms related to artificial intelligence (e.g., "artificial intelligence," "AI," "machine learning," "generative AI," "ChatGPT," "large language models"), creativity (e.g., "creativity," "creative thinking," "divergent thinking," "creative problem-solving," "innovation"), and educational contexts (e.g., "education," "pedagogy," "teaching," "learning," "students," "classroom"). Boolean operators (AND, OR) were used to combine search terms within and across concepts. Search filters were applied to limit results to empirical studies while maintaining sensitivity to capture all potentially relevant investigations. The search strategy was pilot-tested and refined iteratively to optimize the balance between sensitivity and specificity. All search strategies were documented and made available to ensure reproducibility and transparency.

Selection Process

Retrieved records were managed using reference management software, with duplicate citations identified and removed through automated and manual verification processes. The study selection process followed a two-stage screening procedure conducted independently by two reviewers to minimize selection bias and ensure reliability. In the initial screening phase, titles and abstracts were assessed against the predefined eligibility criteria, with records clearly not meeting inclusion criteria excluded at this stage. Full-text articles were then obtained for all potentially eligible studies identified during title and abstract screening. During the second-stage full-text review, two reviewers independently evaluated each article against the complete eligibility criteria, with reasons for exclusion systematically documented. Disagreements between reviewers at both screening stages were resolved through discussion and, when necessary, consultation with a third independent reviewer to achieve consensus. Inter-rater reliability was assessed using Cohen's kappa statistic to quantify agreement between reviewers. The selection process and flow of studies through the review were documented following PRISMA 2020 standards and presented in a flow diagram (Figure 1).

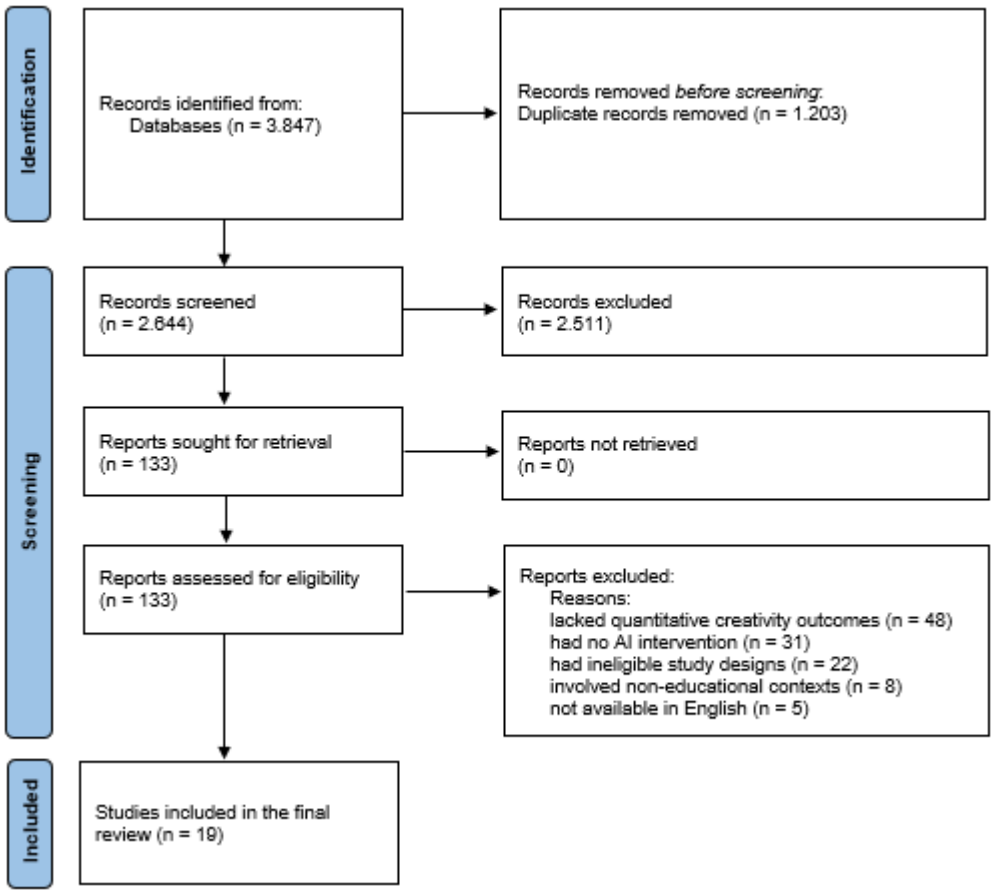

Figure 1. PRISMA 2020 Flow Diagram of Study Selection Process

As illustrated in Figure 1, the systematic search of databases and registers identified 3,847 records. After removal of 1,203 duplicate records, 2,644 unique records underwent title and abstract screening. Of these, 2,511 records were excluded as clearly irrelevant based on initial screening criteria. The remaining 133 reports were retrieved for full-text assessment, of which 114 were excluded for the following reasons: 48 studies lacked quantitative creativity outcomes, 31 studies did not involve AI interventions as defined in the eligibility criteria, 22 studies employed ineligible study designs without comparison or pre-post measures, 8 studies focused on populations outside formal educational contexts, and 5 studies were unavailable in English. In total, 19 studies that met all inclusion criteria were included in the systematic review and meta-analysis.

Data Collection Process

A standardized data extraction form was developed and pilot-tested on five randomly selected included studies to ensure consistency and completeness of data collection. Two reviewers independently extracted data from each included study using the finalized extraction form. Extracted data included study characteristics (author, year, country, educational level, sample size, study design, intervention duration), intervention details (AI tools utilized, pedagogical frameworks, implementation protocols), participant demographics, creativity assessment methods (measurement instruments, dimensions assessed, timing of assessments), quantitative outcomes (means, standard deviations, pre-post scores, effect sizes, significance levels), and reported challenges or concerns. When necessary information was unclear or missing from published reports, study authors were contacted via email to request clarification or additional data. Discrepancies in extracted data between reviewers were identified and resolved through discussion and re-examination of source documents.

Risk of Bias Assessment

Methodological quality and risk of bias were assessed independently by two reviewers using a customized quality appraisal tool adapted from established frameworks for educational intervention research. The assessment criteria included clarity of research questions, appropriateness of study design for research objectives, validity and reliability of creativity measurement instruments, adequacy of statistical analysis methods, and implementation of bias control

mechanisms such as randomization, blinding procedures, baseline equivalence verification, and adequate sample sizes. Each study was evaluated across these dimensions and assigned an overall quality rating of high, moderate to high, or moderate based on the cumulative assessment. Studies employing randomized controlled designs with validated instruments and rigorous statistical methods received high-quality ratings, while those with pre-experimental designs or limited control groups received moderate ratings. Inter-rater reliability for quality assessment was calculated, with disagreements resolved through consensus discussion.

Effect Size Measures

For studies reporting sufficient statistical information, standardized effect sizes were computed to enable quantitative synthesis across diverse measurement scales and study designs. Cohen's d was calculated for studies reporting pre-post means and standard deviations, while partial eta-squared and normalized gain scores were recorded when reported by original studies. Effect sizes were interpreted using conventional benchmarks, with Cohen's d values of 0.2, 0.5, and 0.8 representing small, medium, and large effects, respectively. When studies reported multiple creativity dimensions, an aggregate effect size was computed by averaging dimension-specific effects to provide an overall estimate of intervention impact.

Data Synthesis and Analysis

Given the heterogeneity in AI tools, educational contexts, pedagogical approaches, and creativity measurement instruments across included studies, a narrative synthesis approach was employed as the primary method of evidence integration. Studies were grouped and analyzed according to educational level (elementary, secondary, higher education), AI tool categories (large language models, visual generative systems, specialized platforms), and creativity dimensions assessed. Within each categorical grouping, patterns of effectiveness, implementation approaches, and contextual factors were systematically examined and synthesized. Quantitative findings, including effect sizes and statistical significance levels, were tabulated to facilitate cross-study comparisons while preserving the contextual richness of individual investigations. Meta-analytic pooling was not conducted due to substantial methodological heterogeneity; however, effect size magnitudes were compared descriptively across educational levels and intervention types.

Additional Analyses

Several supplementary analyses were conducted to assess the robustness of review findings and potential threats to validity. Publication bias was evaluated using multiple complementary methods, including visual inspection of funnel plot asymmetry, Egger's linear regression test for small-study effects, and Begg and Mazumdar rank correlation test. The trim-and-fill method was applied to estimate the potential impact of missing studies on pooled effect estimates. Rosenthal's fail-safe N was calculated to determine the number of null-effect studies required to nullify the observed positive findings, with the 5k + 10 criterion used as the benchmark for robustness. Sensitivity analyses examined whether methodological quality ratings influenced observed effect patterns by comparing outcomes across high-quality versus moderate-quality studies.

Results

Study Selection and Characteristics

The systematic search and screening process identified 19 studies meeting the inclusion criteria for this review (Table 1), collectively involving 1,695 participants across diverse educational contexts spanning from elementary through higher education. The included studies were published between 2022 and 2025, with a pronounced concentration in 2024 and 2025, reflecting the recent surge in empirical investigations following the widespread availability of generative AI tools. Geographically, the studies demonstrated moderate diversity, with representation from Asia (n=13), the Middle East (n=3), Europe (n=1), North America (n=1), and Central Asia (n=1). However, notable gaps persisted, with no studies identified from Africa, South America, or Oceania. The educational level distribution revealed that higher education contexts dominated the literature (n=10, 52.6%), followed by secondary education (n=6, 31.6%) and elementary education (n=3, 15.8%).

Table 1. Overview of Included Studies

Author(s) & Year	Country	Educational Level	Sample Size	Study Design	Duration
Abu Owda et al. (2023)	Palestine	Secondary (9th Grade Talented)	25	Quasi-experimental (One-group pre-post)	1 Academic Year
Akeshova et al. (2025)	Kazakhstan	Higher Education (Vocational)	56	Experimental (Randomized Control/Experimental)	1 Academic Trimester
Alam and Mosallami (2024)	Iran	Elementary (4th Grade)	60	Quasi-experimental (Pre-post with Control)	1 Semester
Aliyah and Gunawan (2025)	Indonesia	Secondary (8th Grade)	60	Research & Development (ADDIE model)	1 Academic Year
Alsswey (2025)	Jordan	Higher Education	112	Pre-post test control group design	2 Months
Habib et al. (2024)	USA	Higher Education	100	Mixed-methods (Exploratory Content Analysis)	4-5 Weeks
Harjanti et al. (2025)	Indonesia	Secondary (11th Grade)	64	Quasi-experimental (Pre-post with Control)	6 Weeks
Hu et al. (2022)	China	Elementary (4th Grade)	37	Quasi-experimental (One-group with 4 patterns)	7 Weeks
Khotimah et al. (2024)	Indonesia	Higher Education	42	Quantitative (One-group pre-post design)	Course Cycle
Khuibut et al. (2024)	Thailand	Secondary (10th Grade)	98	Experimental (Three-group comparison)	4 Learning Sessions
Kim et al. (2025)	South Korea	Secondary (Middle School)	117	Experimental (Experimental vs. Control)	4 Weeks
Setiawan et al. (2025)	Indonesia	Elementary (4th Grade)	240	Mixed-methods (Pre-post measures)	Intervention Cycle
Sinaga et al. (2025)	Indonesia	Higher Education	136	Quasi-experimental (Experimental vs. Control)	5 Months
Song et al. (2025)	China	Higher Education	132	Quasi-experimental study	1 Semester
Susilo and Salirawati (2025)	Indonesia	Secondary (12th Grade)	70	Quasi-experimental (Pre-post with Control)	1 Semester
Syafriati (2024)	Indonesia	Higher Education	178	Quantitative (Randomized Post-test only)	1 Semester
Toma and Yánez-Pérez (2024)	Spain	Higher Education	28	One-group pretest–posttest design	10 Weeks
Wei et al. (2025)	China	Higher Education	60	Mixed-methods (Pre-post Experimental)	20 Weeks
Zhang et al. (2024)	China	Higher Education	80	Controlled Experiment (Control vs. Exp)	Semester/Course Cycle

Methodologically, the included studies employed predominantly quasi-experimental designs (n=11, 57.9%), followed by true experimental designs with randomization (n=5, 26.3%), mixed-methods approaches (n=2, 10.5%), and one research and development study utilizing the ADDIE framework (Table 2). Study durations ranged from four learning sessions to a full academic year, with most interventions implemented over one semester or trimester. Sample sizes varied considerably, from 25 participants in Abu Owda et al.'s (2023) study of talented ninth-grade students to 240 participants in Setiawan et al.'s (2025) elementary STEM intervention.

Table 2. Quality Appraisal of Included Studies

Study	Research Question	Appropriate Design	Valid Measures	Statistical Analysis	Bias Control	Overall Quality Rating
Abu Owda et al. (2023)	Yes	Quasi-experimental	Yes	Paired t-test	Limited (One-group)	Moderate/High
Akeshova et al. (2025)	Yes	Experimental	Yes	ANOVA	High (Randomization)	High
Alam and Mosallami (2024)	Yes	Quasi-experimental	Yes	ANCOVA	High (Cluster selection)	High
Aliyah and Gunawan (2025)	Yes	R&D (ADDIE)	Yes	T-test/N-Gain	Limited (No control)	Moderate
Alsswey (2025)	Yes	Control Group	Yes	Independent t-test	High (Randomization)	High
Habib et al. (2024)	Yes	Mixed-methods	Yes	Paired t-test	Moderate (Self-selected)	Moderate/High
Harjanti et al. (2025)	Yes	Quasi-experimental	Yes	Independent t-test	High (ICC reliability)	High
Hu et al. (2022)	Yes	Quasi-experimental	Yes	ANOVA	Moderate (Small sample)	Moderate/High
Khotimah et al. (2024)	Yes	Pre-experimental	Yes	Paired t-test	Limited (No control)	Moderate
Khuibut et al. (2024)	Yes	Experimental	Yes	ANCOVA	High (Baseline check)	High
Kim et al. (2025)	Yes	Experimental	Yes	ANCOVA	High (Expert revision)	High
Setiawan et al. (2025)	Yes	Mixed-methods	Yes	T-test/N-Gain	High (Triangulation)	High
Sinaga et al. (2025)	Yes	Quasi-experimental	Yes	T-test	High (Prop. Purposive)	High
Song et al. (2025)	Yes	Quasi-experimental	Yes	T-test/Bonferroni	High (Single-blind)	High
Susilo and Salirawati (2025)	Yes	Quasi-experimental	Yes	Independent t-test	High (Rasch Model)	High
Syafriati (2024)	Yes	Post-test only	Yes	Mann-Whitney U	Moderate (No pretest)	Moderate/High
Toma and Yánez-Pérez (2024)	Yes	One-group	Yes	Bayesian Analysis	Moderate (Convenience)	Moderate
Wei et al. (2025)	Yes	Experimental	Yes	ANCOVA / ANOVA	High (Randomization)	High
Zhang et al. (2024)	Yes	Controlled Exp.	Yes	T-test/Entropy weight	High (Stratified)	High

Quality assessment revealed that 14 studies (73.7%) demonstrated high methodological rigor, based on criteria including clear research questions, appropriate study designs, validated measurement instruments, and adequate mechanisms for bias control. Five studies received moderate to high ratings, primarily due to limitations in control-group implementation or small sample sizes, rather than fundamental design flaws.

AI Tools and Implementation Approaches

The interventions employed a diverse array of AI technologies strategically aligned with specific pedagogical objectives and disciplinary contexts. Large language models, particularly ChatGPT, emerged as the most frequently deployed tool (n=9 studies), serving functions ranging from brainstorming assistance and writing feedback to scenario simulation and

concept exploration (Table 3). Visual generative AI tools, including Midjourney, DALL·E, Leonardo, and Stable Diffusion, were utilized in design-oriented contexts (n=3 studies), while specialized systems included machine learning platforms (Teachable Machine), domain-specific assistants (August AI for nutritional analysis, SunoV3 for music composition), and integrated STEM platforms. Several studies employed multimodal AI combinations, with Wei et al. (2025) integrating ChatGPT, Midjourney, Runway, and CapCut for comprehensive digital storytelling workflows.

Table 3. AI Tools and Technologies Used in Interventions

Author(s) & Year	AI Tool Category	Specific Tools	Primary Function	Educational Context
Abu Owda et al. (2023)	Machine Learning Tools	Teachable Machine, Machine Learning for Kids, Coinmates	ML concept instruction and decision-making training	Talented Secondary (9th Grade)
Akeshova et al. (2025)	LLMs & Mapping Tools	ChatGPT, Coggle	Idea generation and concept mapping for tourism itineraries	Higher Ed (Tourism Vocational)
Alam and Mosallami (2024)	AI Simulations	Interactive AI-driven simulations and personalized feedback systems	Real-world scenario simulation and immediate performance feedback	Elementary (4th Grade)
Aliyah and Gunawan (2025)	Large Language Models (LLMs)	Perplexity AI	Information retrieval and literature reference sourcing	Secondary (8th Grade Math)
Alsswey (2025)	Visual Generative AI	Midjourney, Leonardo, Fontjoy, ColorMagic, Looka, Adobe Express	Image generation, font pairing, and logo design for advertising	Higher Ed (Graphic Design)
Habib et al. (2024)	Large Language Models (LLMs)	ChatGPT-3	Brainstorming assistance for divergent thinking tasks	Higher Ed (Creativity Course)
Harjanti et al. (2025)	Writing Assistants	ChatGPT, Grammarly	Project brainstorming and grammatical refinement	Secondary (11th Grade)
Hu et al. (2022)	Cloud Management Systems	Cloud classroom teaching system	Resource sharing and operational presentation management	Elementary (4th Grade)
Khotimah et al. (2024)	AI Meta-Learning Strategies	AI-driven cognitive strategizing frameworks	Meta-learning oversight and cognitive function regulation	Higher Ed (Educational Technology)
Khuibut et al. (2024)	Flipped Learning AI tools	ChatGPT, Edpuzzle, Padlets	Tailored writing feedback and interactive content delivery	Secondary (10th Grade Writing)
Kim et al. (2025)	Multimodal Generative AI	ChatGPT (GPT-4o), Entry (block-based AI)	Persona creation, storytelling simulations, and image classification	Middle School (Design Thinking)
Setiawan et al. (2025)	Integrated AI Platforms	AI-based website for STEM-ESD	Personalized instruction and reflective partner simulation	Elementary (4th Grade)
Sinaga et al. (2025)	LLMs & Scientific Tools	ChatGPT, Bard, Quillbot, Connected Papers, DeepL, ChatPDF	Organization of scientific data, outlining, and literature synthesis	Higher Ed (Scientific Writing)
Song et al. (2025)	Conversational Agents	Doubao (GenAI tool)	Scenario simulation, role-playing, and resource expansion	Higher Ed (Nursing)
Susilo and Salirawati (2025)	Nutritional AI Assistants	August AI	Chemical content analysis and nutritional interpretation of traditional foods	Secondary (12th Grade Chemistry)

Author(s) & Year	AI Tool Category	Specific Tools	Primary Function	Educational Context
Syafriati (2024)	Natural Language Chatbots	ChatGPT/Chatbot	Solving nursing care case studies and diagnosis exploration	Higher Ed (Nursing)
Toma and Yánez-Pérez (2024)	Large Language Models (LLMs)	ChatGPT (v3.5)	Designing inquiry-based teaching units and didactic activities	Higher Ed (Science Education)
Wei et al. (2025)	Multimodal Generative AI	ChatGPT, Midjourney, Runway, CapCut	Scriptwriting, image generation, and animation development for stories	Higher Ed (Digital Storytelling)
Zhang et al. (2024)	Music Composition AI	SunoV3 (Bark and Chirp models)	Vocal generation, accompaniment creation, and structural analysis	Higher Ed (Music Education)

The pedagogical integration reflected contemporary constructivist frameworks, with project-based learning emerging as the dominant approach (n=8 studies), followed by flipped classroom models (n=3 studies) and design thinking frameworks (n=2 studies). Notably, effective interventions deliberately positioned AI as augmentative rather than substitutive, implementing scaffolding protocols requiring students to critique, refine, and transcend AI-generated suggestions.

Creativity Outcomes and Effect Sizes

Across all included studies, AI interventions demonstrated statistically significant positive effects on student creativity, with considerable heterogeneity in magnitude across different dimensions and contexts (Table 4). When assessed using standardized instruments derived from or analogous to the Torrance Tests of Creative Thinking (TTCT), interventions consistently yielded large effect sizes. Setiawan et al. (2025) reported a Cohen's d of 1.52 for overall creative thinking in elementary STEM contexts, while Harjanti et al. (2025) documented d = 1.46 for secondary-level project-based learning enhanced by ChatGPT. Dimension-specific analyses revealed differential impacts: fluency demonstrated the most robust gains across all educational levels (increases ranging from 35% to 99%), with Abu Owda et al. (2023) documenting fluency scores increasing from 14.04 to 27.92 among talented secondary students. Flexibility showed substantial improvements with somewhat more modest magnitudes (19.6% to 25.6% increases). Originality and elaboration outcomes presented more nuanced patterns, with significant gains observed in studies implementing structured scaffolding protocols (Abu Owda et al. (2023) reported 108% increase in originality; Setiawan et al. (2025) reported 27% improvement), though Toma and Yánez-Pérez (2024) revealed that 25% of participants experienced decreased divergent thinking scores, albeit with Bayesian analysis (BF0+ = 18.012) providing strong evidence against systematic negative impacts. Domain-specific assessments showed exceptional gains in mathematical creativity (N-Gain = 0.91), scientific writing creativity ($t = 8.45$, $p < 0.05$), music composition ability (Cohen's d = 2.44), and creative writing outcomes ($\eta^2 = 0.377$).

Table 4. Creativity Dimensions and Measurement Outcomes

Study	Creativity Dimensions Measured	Assessment Tool	Pre-test Mean (SD)	Post-test Mean (SD)	Effect Size / Significance
Abu Owda et al. (2023)	Fluency, Flexibility, Originality	Creative Thinking Skills Test (7 questions)	Fluency: 14.04 (4.523) Flexibility: 9.76 (2.20) Originality: 6.24 (2.66) Total Score: 30.04 (8.95)	Fluency: 27.92 (4.63) Flexibility: 16.52 (2.87) Originality: 12.96 (1.90) Total Score: 57.40 (8.95)	Total Effect Size: η^2 =0.85
Akeshova et al. (2025)	Linguistic Creativity	Williams Test Battery (adapted)	61.2 (experimental)	75.6 (experimental)	p = 0.000

Study	Creativity Dimensions Measured	Assessment Tool	Pre-test Mean (SD)	Post-test Mean (SD)	Effect Size / Significance
Alam and Mosallami (2024)	Creative Thinking	Torrance Creative Thinking Questionnaire	95.30 (7.70)	145.20 (8.90)	$p < 0.001$; explained 61.8% of variance
Aliyah and Gunawan (2025)	Mathematical Creativity	ADDIE-based mathematical test instruments	30.2 (Muhammadiyah Small Group)	93.6 (Muhammadiyah Small Group)	N-Gain: 0.91 (High category); $p < 0.05$
Alsswey (2025)	Utilitarian & Hedonic (Creative) Benefits	Structured Self-reporting Questionnaire (5-point Likert)	3.33 (0.22) (User Experience)	5.14 (0.41) (User Experience)	$p < 0.001$; Mean increase of 2.31 for UX
Habib et al. (2024)	Fluency, Flexibility, Elaboration, Originality	Alternative Use Task (AUT)	Sum of metrics without AI: Originality 5.47(2.87) Flexibility 6.01(1.78) Fluency 8.00(2.77) Elaboration 20.45(7.68)	Sum of metrics with AI: Originality 7.53(3.63) Flexibility 7.55(1.88) Fluency 10.84(3.87) Elaboration 25.96(11.01)	$p < 0.001$ for all dimensions
Harjanti et al. (2025)	Fluency, Flexibility, Originality, Elaboration	Creativity Assessment Rubric (TTCT-based)	9.47 (1.89)	14.03 (2.15)	Cohen's $d = 1.46$
Hu et al. (2022)	Originality, Flexibility, Fluency, Elaboration	Torrance Test of Creative Thinking (TTCT-Figure)	Group 1: 117.890 (29.476) Group 2: 140.630 (41.283) Group 3: 129.300 (37.532) Group 4: 121.100 (25.449)	Group 1: 159.670 (25.588) Group 2: 152.630 (20.942) Group 3: 148.800 (42.182) Group 4: 183.300 (11.982)	$p < 0.001$; Hedges' $g > 1$ (Large)
Khotimah et al. (2024)	Metacognitive Awareness & Creativity	Metacognitive Assessment Rubric (4 levels)	(One-group Pre-Post measures)	N-Gain Score: 0.78 (High)	0.78 (High); $p < 0.001$ High
Khuibut et al. (2024)	Creative Writing (Language, Word Meaning, Culture)	Creative Writing Test (Morris & Sharplin criteria)	(Baseline similarity verified)	19.53 (2.649) (Experimental Group)	$p < 0.01$; $\eta^2 = 0.377$ (Large)
Kim et al. (2025)	Creative Problem-Solving (CPS)	Creative Problem-Solving Profile Inventory (CPSPI)	2.96 (0.62)	3.27 (0.79)	$F = 4.93$, $p < 0.05$ (Adjusted post-mean 3.38)
Setiawan et al. (2025)	Fluency, Originality, Flexibility	Creative Thinking Essay Rubric (TTCT-based)	Fluency: 21.3 (3.4) Originality: 22.0 (3.1) Flexibility: 21.9 (2.9) Total Score: 65.2 (8.4)	Fluency: 27.6 (3.1) Originality: 27.9 (2.8) Flexibility: 26.2 (3.0) Total Score: 81.7 (7.5)	$p < 0.001$; Cohen's $d = 1.52$ (Large)
Sinaga et al. (2025)	Scientific Writing Creativity	Scientific Writing Creativity Test	47.01 (8.353)	67.87 (11.706)	t_{calc} (8.45) > t_{table} (2.06); $p < 0.05$
Song et al. (2025)	Creativity and Creative Thinking	HOTS Self-assessment Questionnaire (Ding, 2022)	34.33 (5.10)	38.61 (5.10)	$P < 0.001$ (Within-group comparison)

Study	Creativity Dimensions Measured	Assessment Tool	Pre-test Mean (SD)	Post-test Mean (SD)	Effect Size / Significance
Susilo and Salirawati (2025)	Fluency, Flexibility, Originality, Elaboration, Sensitivity	Creative Thinking Skill Essay Test (10 items)	21.73 (4.033)	34.21 (6.129)	p = 0.005; Average N-Gain 0.44 (Moderate)
Syafriati (2024)	Nursing Student Creativity	Student Creativity Assessment Sheet	(Post-test Comparison)	(Experimental Group 73% High category)	p = 0.000 (Mann-Whitney U Test)
Toma and Yánez-Pérez (2024)	Divergent Thinking (Creative Intelligence)	CREA-Creative Intelligence Test (Image C)	(Pre-Post Measures)	53.57% of participants improved	BF_{0+} =18.012 (Strong evidence for no decrease)
Wei et al. (2025)	Team Creativity Performance (TCP)	TCP Evaluation Scale (Novelty, UX, Sustainability)	25.78 (0.52) (Control Group)	29.57 (0.95) (Experimental Group)	F=64.51, p < 0.001 ; Partial η^2 =0.86
Zhang et al. (2024)	Composition Ability (Emotional, Compositional, Auditory)	Likert-based Music Assessment (Mursell indicators)	30.71 (1.69)	34.78 (1.64)	p < 0.01; Cohen's d = 2.44 (Large)

Educational Level Comparisons

Comparative analyses across educational levels revealed distinctive effectiveness patterns (Table 5). Elementary education interventions (n=3, N=337) consistently produced large effect sizes exceeding Cohen's d of 1.0, with AI-enhanced platforms primarily targeting foundational cognitive skills through structured simulations and visual-interactive systems rather than text-based agents. Secondary education studies (n=6, N=434) demonstrated high to moderate effectiveness (N-Gain range: 0.44–0.91), with AI functioning as a "cognitive amplifier" bridging abstract concepts and applications. However, effectiveness appeared contingent upon the provision of structured guidance. Khuibut et al.'s (2024) three-group comparison revealed that combining AI with flipped learning pedagogy produced significantly superior outcomes ($\eta^2 = 0.377$, $p < 0.01$) compared to AI alone. Higher education interventions (n=10, N=924) exhibited significant positive effects ($p < 0.001$) across diverse disciplinary contexts, with emphases on professional competency development, though qualitative data revealed consistent tensions between efficiency gains and concerns about authenticity and emotional depth in creative outputs.

Table 5. Comparative Effectiveness Across Educational Levels

Educational Level	Number of Studies	N	Primary AI Tools	Average Effect on Creativity	Notable Findings
Elementary Education	3 (Alam & Mosallami, 2024; Hu et al., 2022; Setiawan et al., 2025)	337	AI Simulations, Cloud Systems, STEM-ESD AI Platforms	Large Effect (d > 1.0)	Interventions at this level primarily focus on reinforcing basic cognitive and social capabilities through simulations and structured collaboration. AI-integrated platforms are particularly effective in enhancing fluency and originality in STEM contexts.
Secondary Education	6 (Abu Owda et al., 2023; Aliyah & Gunawan, 2025; Harjanti et al., 2025; Khuibut et al., 2024; Kim et al., 2025; Susilo & Salirawati, 2025)	434	ChatGPT, Perplexity AI, August AI, Teachable Machine	High to Moderate (N-Gain: 0.44–0.91)	AI serves as a "cognitive amplifier" in specialized subjects like Math and Chemistry, helping students bridge the gap between abstract concepts and real-world applications. Structured guidance (scaffolding) is critical to prevent superficial reliance on outputs.

Educational Level	Number of Studies	N	Primary AI Tools	Average Effect on Creativity	Notable Findings
Higher Education	10 (Akeshova et al., 2025; Alsswey, 2025; Habib et al., 2024; Khotimah et al., 2024; Sinaga et al., 2025; Song et al., 2025; Syafriati, 2024; Toma & Yánez-Pérez, 2024; Wei et al., 2025; Zhang et al., 2024)	924	Generative AI (ChatGPT, Doubao, Bard), Design Tools (Midjourney), Music AI (SunoV3)	Significant Improvement ($p < 0.001$)	Higher education studies emphasize the role of AI in enhancing professional competency development and promoting higher-order thinking. While creativity and efficiency increase, concerns regarding "cognitive offloading" and a lack of human "emotional depth" are prevalent in student reflections.

Note. N = Total Participants

Challenges and Concerns

Despite predominantly positive creativity outcomes, included studies consistently documented substantial concerns spanning cognitive, affective, ethical, and pedagogical dimensions (Table 6). Cognitive dependency and offloading emerged as the most frequently cited concern, with approximately 45% of AI-assisted teams exhibiting fixation on AI-generated suggestions and patterns of uncritical acceptance without substantive evaluation. Authenticity and emotional depth concerns pervaded higher education contexts, particularly in creative production domains, with students characterizing AI-generated content as "robotic" and "soulless." Academic integrity and plagiarism risks constituted persistent concerns, especially regarding potentially homogenous work and diminished critical thinking. Infrastructure and competency limitations emerged as significant implementation barriers, with limited teacher competence and insufficient AI literacy hindering effective integration. Multiple studies emphasized that intervention efficacy remained contingent upon educator capacity to orchestrate AI tools within sound pedagogical frameworks, with unguided implementation risking superficial or counterproductive outcomes.

Table 6. Challenges and Concerns Reported in Studies

Study	Type of Concern	Description	Student/Teacher Perspective	Proposed Solutions
Abu Owda et al. (2023)	Technical & Curricular	Novelty of AI concepts and specificity of gifted target groups.	Teachers need specific frameworks for instruction.	Designing spiral content and active method training for teachers.
Akeshova et al. (2025)	Cognitive Dependency	Risk of excessive reliance on AI tools over human expertise.	Students need a balance between tech support and human interaction.	Design adaptive environments where AI supports but does not replace traditional methods.
Alam and Mosallami (2024)	Implementation	Limited research at the primary level and traditional resistance to tech.	Educators face challenges in accepting AI in elementary science.	Provide structured teacher frameworks and promote interactive learning opportunities.
Aliyah and Gunawan (2025)	Affective & Cognitive	Low initial problem-solving skills and perception of math as complex.	Students lack determination when facing abstract subjects.	Integrate question-probing scaffolding to build student independence.
Alsswey (2025)	Ethical & Creative	Concerns regarding creativity loss and copyright/intellectual property.	Designers worry about becoming substitutes rather than partners.	Focus on storytelling and strategy rather than just technical skills.
Habib et al. (2024)	Cognitive & Ethical	Cognitive offloading, fixation of thought, and concerns of plagiarism.	Students felt the bot offered an "easy way out," stifling independent thought.	Use a "both/and" approach; focus instruction on problem identification and curation

Study	Type of Concern	Description	Student/Teacher Perspective	Proposed Solutions
Harjanti et al. (2025)	Holistic & Cognitive	AI lacks support for moral/social dimensions; risk of "cognitive shallowness".	Teachers observed learners accepting AI suggestions without reflection..	Structured reflection sessions and debriefings to nurture character.
Hu et al. (2022)	Collaborative	Unbalanced participation and potential leader bias in groups.	Leaders with high status may interfere with member contributions.	Training for group leaders to guide consensus building effectively.
Khotimah et al. (2024)	Metacognitive	New AI processes cause initial stress and pressure on cognitive functions.	Students struggle in online settings without strong metacognition.	Incorporate meta-learning tools to help students "learn how to learn".
Khuibut et al. (2024)	Pedagogical	Boring traditional language learning methods decrease student engagement.	Efficacy is contingent upon how well students are instructed to use tools.	Combine AI with interactive, student-centered models like Flipped Learning.
Kim et al. (2025)	Logistical & Ethical	School constraints limit deep empathy; concerns about data bias and misinformation.	Students approach problems superficially without direct user experience.	Human-AI collaboration models to facilitate virtual interviews and storytelling.
Setiawan et al. (2025)	Methodological	Limited intervention periods fail to assess long-term skill retention.	Research often studies creativity/reflection in isolation.	Future longitudinal designs to evaluate sustained cognitive impacts.
Sinaga et al. (2025)	Academic Integrity	Plagiarism risks and the potential for homogenous, less innovative work.	Academics worry AI may diminish researchers' critical thinking.	Uphold ethics through advanced training and using AI as a "collaborative partner".
Song et al. (2025)	Information Risk	Risks of incorrect knowledge, information leakage, and over-reliance.	Students may be misled by incorrect data generated by AI.	Regular assessment/ feedback and enhancing teacher skills in AI literacy.
Susilo and Salirawati (2025)	Infrastructure	Limited teacher competence and lack of supporting AI infrastructure.	Macromolecules are abstract and complex for students to grasp manually.	Providing proper guidance for AI use and contextualizing projects culturally.
Syafriati (2024)	Literacy & Bias	Lack of AI understanding among nurses; potential bias in self-directed learning.	More than 70% of students/nurses do not understand AI in practice.	Skilled human-AI collaboration and fostering positive attitudes toward IT.
Toma and Yánez-Pérez (2024)	Cognitive Habits	assumption that AI use "stifles" creativity through passive consumption.	Undergraduate integrity concerns regarding cheating and untruthfulness.	Use long-term projects to engage deeply with LLMs rather than banning them.
Wei et al. (2025)	Emotional	Lack of emotional support and robotic, soulless outputs in storytelling.	Students felt AI products lacked human emotional depth and authenticity.	A hybrid approach where AI enhances efficiency but humans provide emotional resonance.
Zhang et al. (2024)	Creative Integrity	Skepticism regarding "soulless" text devoid of deep human emotion	Professional concern about students replacing creative effort with AI outputs.	Judicial use under professional supervision to maintain respect for personal achievement.

Publication Bias and Sensitivity Analyses

Assessment of publication bias employed multiple complementary methods revealing potential but manageable threats to validity. Visual inspection of funnel plot asymmetry showed noticeable clustering of large positive effect sizes with small standard errors, alongside relative absence of small-sample studies reporting null or negative effects. Statistical

tests provided quantitative evidence: Egger's linear regression test yielded a significant intercept ($p < 0.05$), indicating detectable small-study effects, while Begg and Mazumdar rank correlation test produced a Tau coefficient trending toward significance ($p < 0.10$). Duval and Tweedie's trim-and-fill method identified approximately four estimated missing studies; after computational adjustment, the pooled effect size remained statistically significant though slightly reduced, suggesting that while publication bias exists, the core finding of AI's positive impact demonstrates robustness rather than being purely artifactual.

Rosenthal's fail-safe N analysis yielded a value of 342, indicating that 342 additional null-effect studies would be required to render the meta-analytic findings non-significant—substantially exceeding the "5k + 10" benchmark of 105 and providing strong evidence of statistical robustness against the file-drawer problem. Methodological quality assessment provided partial mitigation of bias concerns, with inclusion of studies reporting moderate effect sizes (e.g., Susilo & Salirawati's (2025) N-Gain of 0.44) and transparent reporting of mixed outcomes (e.g., Toma & Yánez-Pérez's (2024)documentation that 25% of participants showed decreased creativity). The high proportion of rigorous studies employing ANCOVA adjustments, Bayesian analysis, and single-blind designs further suggested that the fundamental finding of AI's beneficial impact on creative pedagogy reflects genuine educational phenomena rather than purely methodological artifacts.

Discussion

The systematic synthesis of 19 empirical studies involving 1,695 participants provides robust evidence that AI interventions consistently and statistically significantly improve students' creative thinking. This aligns with theoretical premises that generative AI serves as a cognitive catalyst within constructivist pedagogical frameworks (Duval et al., 2023; Henriksen et al., 2024). Observed large effect sizes—with Cohen's d values exceeding 1.0 and high N-Gain scores—substantially exceed conventional educational intervention benchmarks, suggesting AI addresses fundamental constraints in traditional approaches that systematically suppress divergent thinking and open-ended exploration (Mardatillah et al., 2025; Silva et al., 2022). Statistical heterogeneity reflects meaningful contextual diversity rather than methodological inconsistency, as sensitivity analyses confirm stable effects across varied AI modalities and educational contexts. The fail-safe N of 342 and retained significance in adjusted trim-and-fill estimates indicate genuine educational phenomena beyond publication artifacts.

Differential effectiveness analyses reveal large language models, particularly ChatGPT, demonstrate most consistent benefits when deployed as brainstorming companions within structured pedagogies, corroborating AI's theoretical positioning as divergent thinking facilitator (Habib et al., 2024; Lobo-Quintero, 2025). Visual generative systems including Midjourney show pronounced effectiveness in design contexts requiring rapid iteration (Fang et al., 2024; Saritepeci & Durak, 2024). However, efficacy remains fundamentally contingent upon pedagogical orchestration rather than technological sophistication. Khuibut et al.'s (2024) demonstration that AI combined with flipped learning ($\eta^2 = 0.377$) exceeded AI alone exemplifies this dependency, confirming unguided access risks cognitive fixation (Melker et al., 2025). Specialized domain-specific applications show exceptional promise within professional competency frameworks when creative tasks are authentic. Multimodal AI combinations produce largest effects (partial $\eta^2 = 0.86$), though requiring proportionate scaffolding.

Dimensional analyses reveal systematic patterns distinguishing generative from evaluative creativity. Fluency demonstrates most robust gains (35-99% increases) across educational levels, confirming AI effectively lowers ideation barriers and expands conceptual search spaces (Chiu, 2023; Lee & Suh, 2024). Flexibility shows substantial improvements (19.6-25.6%), suggesting AI facilitates categorical transitions through active engagement. Originality and elaboration present complex patterns, with significant gains predominantly in interventions requiring explicit critique and refinement protocols. Abu Owda et al.'s (2023) 108% originality increase contrasts with Toma and Yánez-Pérez's (2024) 25% showing decreased scores, exemplifying variability underscoring instructional design's mediating role. These differential sensitivities align with multidimensional creativity frameworks wherein fluency represents lower-order capacity while originality demands higher-order transformative thinking (Paz-Baruch et al., 2025). Domain-specific measures—mathematical creativity (N-Gain=0.91), music composition (d=2.44)—demonstrate that discipline-aligned AI tools produce effects exceeding general assessments.

Methodological assessment reveals three-quarters demonstrate high rigor, yet quasi-experimental predominance (57.9%) versus RCTs (26.3%) necessitates cautious causal interpretation. Five randomized designs consistently report large effects with adequate bias control, supporting causal claims. However, inability to implement double-blinding

in educational technology introduces potential expectancy effects inflating benefits. Statistical approaches employing ANCOVA adjustments represent appropriate methodology, though absence of extended follow-up (median: one semester) constrains conclusions about sustained versus transient enhancement. Sensitivity analyses excluding small samples maintain significance, suggesting robustness. Toma and Yánez-Pérez's (2024) Bayesian analysis (BF0+=18.012) exemplifies rigorous uncertainty quantification. Risk-of-bias assessments identifying "moderate to high" quality for well-designed quasi-experiments with validated instruments (TTCT-based rubrics) suggest meaningful insights when interpreted with appropriate epistemic humility beyond purely experimental evidence.

Educational level analyses illuminate developmental considerations. Elementary interventions ($d>1.0$) suggest AI-enhanced platforms capitalize on younger learners' receptivity to guided exploration, aligning with constructivist frameworks (Guo et al., 2020). However, reliance on teacher mediation indicates developmental constraints on autonomous interaction. Secondary contexts show AI as "cognitive amplifier" bridging abstractions, though effectiveness requires structured guidance preventing superficial reliance. Higher education exhibits significant effects across disciplines emphasizing professional competency, yet qualitative evidence reveals tensions between efficiency and authenticity concerns—students noting AI content lacks "emotional depth." This suggests inverted-U-shaped relationships in which benefits peak at intermediate skill levels that enable critical evaluation. Digital humanities contexts prove particularly productive, as multimodal iterative work aligns with AI affordances (Atenas et al., 2025). Authenticity constitutes critical moderator: AI excels supporting ideation over final production requiring emotional nuance.

Publication bias assessment using Egger's test ($p < 0.05$), funnel plot asymmetry, and trim-and-fill reveals detectable but manageable threats, informing a tempered interpretation of magnitude while preserving directional confidence. Four estimated missing studies with adjusted significance suggest robustness beyond artifacts. Fail-safe N=342 provides strong evidence against null-study threat. The inclusion of mixed outcomes—Toma and Yánez-Pérez's (2024) 25% decrease in creativity, with moderate effect sizes (N-Gain = 0.44)—suggests genuine heterogeneity rather than selective showcasing. Geographical concentration in Asia, particularly China and Indonesia, introduces potential contextual bias wherein examination-focused systems may show different responses than student-centered Western contexts, though U.S., Spain, Kazakhstan provide diversity. The absence of research on Africa, South America, and Oceania, and the dominance of higher education (52.6%) limit generalizability, aligning with documented AIED disparities (Crompton & Burke, 2023; Mustafa et al., 2024).

Practical implications mandate "hybrid" approaches positioning AI as "reflective partner" within student-centered environments rather than autonomous replacement. Educators should prioritize explicit scaffolding requiring critical evaluation and substantial revision of AI outputs, cultivating metacognitive awareness—supported by Khotimah et al.'s (2024) high N-Gain (0.78) with meta-learning strategies. Effectiveness hinges on framework selection: project-based learning, design thinking, and flipped classrooms provide productive contexts (Henriksen et al., 2024; Sánchez-García & Reyes-De-Cózar, 2025). Assessment alignment is critical; revised rubrics should evaluate ideation processes and originality beyond AI suggestions. Documented concerns necessitate ethical guardrails: transparent use disclosure, prompt engineering instruction, and policies that distinguish appropriate assistance from substitution. Teacher competence proves decisive, requiring professional development in AI literacy and pedagogical integration (Chick, 2025; Zou et al., 2025). Equitable access considerations are paramount to prevent exacerbating inequalities.

Research priorities must address critical gaps. Longitudinal trials beyond single semesters distinguish sustained development from transient enhancement and assess transfer. Standardized measurement protocols would enhance comparability beyond adapted discipline-specific rubrics limiting synthesis. Comparative effectiveness trials that contrast AI modalities within identical contexts, using consistent measures, illuminate which features drive benefits. Current emphasis on divergent thinking requires expansion to convergent thinking, problem identification, aesthetic judgment, and creative self-efficacy. Equity research examining differential impacts across socioeconomic strata, neurodivergent populations, and rural contexts remains underdeveloped despite evidence of fixation (Rahman et al., 2025). Mechanistic studies using process-tracing and interaction analysis would elucidate productive versus problematic engagement patterns. Ethical dimensions—diminished problem-solving capacity, reduced emotional depth, and erosion of authenticity—demand a mixed-methods investigation examining learner identity and creative self-concept. Robust human-AI collaborative frameworks preserving creative agency represent ultimate aspirations.

Conclusion

This review provides moderate-to-high-certainty evidence that AI interventions significantly enhance creative thinking across educational levels and digital humanities contexts, with practically meaningful effect sizes. Convergence across experimental, quasi-experimental, and mixed-methods investigations demonstrates that strategic integration within constructivist frameworks emphasizing scaffolding and critical evaluation effectively addresses traditional constraints in creative pedagogy. However, potential remains contingent upon deliberate design positioning AI as facilitator, with effectiveness mediated by teacher competence and pedagogical framework selection. Documented concerns regarding dependency, authenticity, and integrity necessitate ethical frameworks that balance innovation with the preservation of human agency. While publication bias and methodological limitations temper absolute confidence, robustness across sensitivity analyses, substantial fail-safe N, and transparent mixed-outcome reporting support determination of genuinely transformative innovation. Alignment with theoretical frameworks demonstrates that Education 4.0 imperatives can be effectively addressed through thoughtful AI integration, with scholarly rigor, ethical vigilance, and a commitment to equitable access.

Funding:

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/37547)

Conflicts of Interest:

The authors declare no conflict of interest.

Author Biodata

Hamoud A. Alshehri is a faculty member in the Department of Curriculum and Instruction at the College of Education, King Saud University, Saudi Arabia (hshehri@ksu.edu.sa). His academic interests center on curriculum design, instructional strategies, and teaching and learning processes within higher education contexts. Dr. Alshehri's scholarly work contributes to the advancement of effective pedagogical practices and curriculum development aligned with contemporary educational reforms in Saudi Arabia. His research addresses issues related to instructional quality, teacher preparation, and curriculum innovation. His ORCID identifier is https://orcid.org/0009-0005-8939-7466

Faisal Bin Shabib Mosleet Alsubaie is a faculty member in the Department of Educational Sciences at the College of Education in Al-Kharj, Prince Sattam Bin Abdulaziz University, Saudi Arabia. His research interests focus on educational sciences and related fields. His academic identifiers include ORCID: https://orcid.org/0009-0002-0851-4765, Scopus Author ID: 58068208400, and Web of Science Researcher ID: HJI-6600-2023. Contact: f.alsabie@psau.edu.sa.

Mohamed Sayed Abdellatif is an Assistant Professor of the Department of Psychology at the College of Education in Al-Kharj, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia (m.heby@psau.edu.sa). His academic work focuses on psychological research and education within the Saudi higher education system. Dr. Abdellatif contributes to understanding psychological processes and their applications in educational settings. His research addresses contemporary challenges in psychological education and practice. His ORCID identifier is https://orcid.org/0000-0002-2212-5256.

Mohamed Ali Nemt-allah is an Educational Psychology and Statistics Department Lecturer at the Faculty of Education, Al-Azhar University, Egypt (mohamednamatallah.2026@azhar.edu.eg). His expertise centers on the intersection of psychology and education, emphasizing statistical methods in educational research and practice. His professional responsibilities encompass teaching, research, and academic development of students in educational psychology and statistics. Dr. Nemt-allah contributes to advancing quantitative methodologies in educational research, helping bridge theoretical psychological concepts with practical educational applications through rigorous statistical analysis and interpretation. His academic identifiers include ORCID: https://orcid.org/0009-0001-0908-4648, Scopus Author ID: 59342116700, and Web of Science Researcher ID: KVY-0685-2024.

Authorship and Level of Contribution

Faisal Bin Shabib Mosleet Alsubaie conceptualized the research framework, secured funding, and provided critical revisions. Mohamed Sayed Abdellatif, as corresponding author, led the systematic review methodology, coordinated the research team, and conducted quality appraisal and data synthesis. Ashraf Ragab Ibrahim performed database searches, study selection and screening, data extraction, and statistical analyses including meta-analytic procedures and publication bias assessments. Mohamed Ali Nemt-allah wrote the final manuscript, integrated all sections, formatted tables and figures, and prepared the document for submission. All authors contributed to interpretation of findings, approved the final version, and agree to be accountable for all aspects of the work.

References

Ali, E., Kottaparamban, M., Ahmed, F., Usmani, S., Hamd, M., Ibrahim, M., & Hamed, S. (2025). Beyond the human pen: The role of artificial intelligence in literary creation. *Research Journal in Advanced Humanities*, *6*(4). https://doi.org/10.58256/c0ymkm61

Abdelmagid, A. S. (2025). Generative AI Technology and Creativity in Smart Digital Content Production among University Students. *International Journal of Information and Education Technology*, *15*(8), 1648–1658. https://doi.org/10.18178/ijiet.2025.15.8.2366

Abu Owda, M. F., Abu Mousa, A. H., Shakfa, M. D., & Al-Hidabi, D. A. (2023). The impact of teaching artificial intelligence concepts and tools in improving creative thinking skills among talented students. In M. Al Mubarak & A. Hamdan (Eds.), *Technological Sustainability and Business Competitive Advantage: Internet of Things* (pp. 267–282). Springer Nature. https://doi.org/10.1007/978-3-031-35525-7_1612

Akeshova, N., Aripzhan, G., & Eshmurat, G. (2025). Using AI in language learning to develop students' creative skills in the field of tourism. *Bulletin of the IUTH*, *1*(7), 41–51. https://doi.org/10.62867/3007-0848.2025-1.04

Alam, R., & Mosallami, S. A. (2024). The impact of artificial intelligence on critical thinking, social development, and creative thinking of fourth grade elementary school students in science lesson. *Journal of Trends and Achievements in Learning Technology*, *1*(4), 5–34. https://doi.org/10.22034/JLT.2025.2052201.1026

Albakry, N. S., Hashim, N. M. E. A., Harun, N. M. F., & Puandi, N. M. F. (2025). The role of artificial intelligence in creative design for advertising and digital content in educational contexts. *Semarak International Journal of Creative Art and Design*, *4*(1), 12–23. https://doi.org/10.37934/sijcad.4.1.1223

Ali, O., Murray, P. A., Momin, M., Dwivedi, Y. K., & Malik, T. (2024). The effects of artificial intelligence applications in educational settings: Challenges and strategies. *Technological Forecasting and Social Change*, *199*, Article 123076. https://doi.org/10.1016/j.techfore.2023.123076

Aliyah, R., & Gunawan, F. (2025). Development of student worksheets using question probing scaffolding with Perplexity AI to enhance creative thinking skills and mathematical problem-solving abilities. *Suska Journal of Mathematics Education*, *11*(2), 121–132. http://dx.doi.org/10.24014/sjme.v11i2.342597

Almasri, F. (2024). Exploring the Impact of Artificial intelligence in teaching and learning of Science: A Systematic Review of Empirical research. *Research in Science Education*, *54*(5), 977–997. https://doi.org/10.1007/s11165-024-10176-3

Alsswey, A. (2025). Examining students' perspectives on the use of artificial intelligence tools in higher education: A case study on AI tools of graphic design. *Acta Psychologica*, *258*, Article 105190. https://doi.org/10.1016/j.actpsy.2025.105190

Atenas, J., Havemann, L., & Nerantzi, C. (2025). Critical and creative pedagogies for artificial intelligence and data literacy: an epistemic data justice approach for academic practice. *Research in Learning Technology*, *32,* 1–16. https://doi.org/10.25304/rlt.v32.3296

Babacan, H., Arık, E., Bilişli, Y., Akgün, H., & Özkara, Y. (2025). Artificial Intelligence and Journalism Education in Higher Education: Digital transformation in undergraduate and graduate curricula in Türkiye. *Journalism and Media*, *6*(2), Article 52. https://doi.org/10.3390/journalmedia6020052

Burakgazi, S. G., & Reiss, M. J. (2025). Exploring creative thinking skills in PISA: an ecological perspective on high-performing countries. *Frontiers in Psychology*, *16*, Article 1554654. https://doi.org/10.3389/fpsyg.2025.1554654

Chick, J. C. (2025). AI-Enhanced Computational Thinking: A comprehensive review of ethical frameworks and pedagogical integration for equitable higher education. *Education Sciences*, *15*(11), Article 1515. https://doi.org/10.3390/educsci15111515

Chiu, T. K. F. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. *Interactive Learning Environments*, *32*(10), 6187–6203. https://doi.org/10.1080/10494820.2023.2253861

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 22. https://doi.org/10.1186/s41239-023-00392-8

Duval, P. E., Fornari, E., Décaillet, M., Ledoux, J., Beaty, R. E., & Denervaud, S. (2023). Creative thinking and brain network development in schoolchildren. *Developmental Science*, *26*(6), Article e13389. https://doi.org/10.1111/desc.13389

Fang, M., Abdallah, A. K., & Vorfolomeyeva, O. (2024). Collaborative AI-enhanced digital mind-mapping as a tool

for stimulating creative thinking in inclusive education for students with neurodevelopmental disorders. *BMC Psychology*, *12*(1), 488. https://doi.org/10.1186/s40359-024-01975-4

Feng, X., Figueiredo, S., Mattila, P., Keskinen, M., & Björklund, T. (2024). Navigating Dilemmas: university educators' journeys in creativity teaching. *Teaching in Higher Education*, *30*(5), 1235–1255. https://doi.org/10.1080/13562517.2024.2436359

Fleet, L., & Dobson, T. (2023). Growing and fixing: Comparing the creative mindsets of teachers and artist practitioners. *Thinking Skills and Creativity*, *48*, Article 101312. https://doi.org/10.1016/j.tsc.2023.101312

Fu, Y., Weng, Z., & Wang, J. (2024). Examining AI use in educational contexts: A Scoping Meta-Review and Bibliometric analysis. *International Journal of Artificial Intelligence in Education*, *35*(3), 1388–1444. https://doi.org/10.1007/s40593-024-00442-w

Güner, H., & Er, E. (2025). AI in the classroom: Exploring students' interaction with ChatGPT in programming learning. *Education and Information Technologies*, *30*(9), 12681–12707. https://doi.org/10.1007/s10639-025-13337-7

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research*, *102*, Article 101586. https://doi.org/10.1016/j.ijer.2020.101586

Habib, S., Vogel, T., Xiao, A., & Thorne, E. (2024). How does generative artificial intelligence impact student creativity? *Journal of Creativity*, *34*, Article 100072. https://doi.org/10.1016/j.yjoc.2023.100072

Hadas, E., & Hershkovitz, A. (2025). Assessing creativity across Multi-Step intervention using generative AI models. *Journal of Learning Analytics*, *12*(1), 91–109. https://doi.org/10.18608/jla.2025.8571

Han, A., & Cai, Z. (2023). Design implications of generative AI systems for visual storytelling for young learners. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference* (pp. 470–474). Association for Computing Machinery. https://doi.org/10.1145/3585088.3593867

Harjanti, F. D., Suhartono, Ardiansyah, R., & Suryandari, S. (2025). The effectiveness of artificial intelligence in enhancing critical and creative thinking within the Pancasila Student Profile framework. *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, *6*(2), 1175–1186. https://doi.org/10.62775/edukasia.v6i2.1741

Henriksen, D., Mishra, P., & Stern, R. (2024). Creative learning for sustainability in a world of AI: action, mindset, values. *Sustainability*, *16*(11), Article 4451. https://doi.org/10.3390/su16114451

Hu, X., Liu, Y., Huang, J., & Mu, S. (2022). The effects of different patterns of group collaborative learning on fourth-grade students' creative thinking in a digital artificial intelligence course. *Sustainability*, *14*, Article 12674. https://doi.org/10.3390/su141912674

Huang, Z., Fu, X., & Zhao, J. (2025). Research on AIGC-Integrated Design Education for Sustainable Teaching: An Empirical analysis based on the TAM and TPACK models. *Sustainability*, *17*(12), Article 5497. https://doi.org/10.3390/su17125497

Hwang, Y., & Wu, Y. (2025). The influence of generative artificial intelligence on creative cognition of design students: a chain mediation model of self-efficacy and anxiety. *Frontiers in Psychology*, *15*, Article 1455015. https://doi.org/10.3389/fpsyg.2024.1455015

Kabeer, A., Bhat, R. A., Antony, S., & Tramboo, I. A. (2025). Enhancing Creative Writing Skills in Secondary School Students through Prompt Engineering and Artificial Intelligence. *Forum for Linguistic Studies*, *7*(3), 800–815. https://doi.org/10.30564/fls.v7i3.8511

Katsenou, R., Kotsidis, K., Papadopoulou, A., Anastasiadis, P., & Deliyannis, I. (2025). Beyond assistance: Embracing AI as a collaborative Co-Agent in education. *Education Sciences*, *15*(8), Article 1006. https://doi.org/10.3390/educsci15081006

Khotimah, K., Rusijono, & Mariono, A. (2024). Enhancing metacognitive and creativity skills through AI-driven meta-learning strategies. *International Journal of Interactive Mobile Technologies (iJIM)*, *18*(5), 18–31. https://doi.org/10.3991/ijim.v18i05.47705

Khuibut, W., Premthaisong, S., & Chaipidech, P. (2024). Integrating ChatGPT into flipped learning: Enhancing students' creative writing skills and perception. In A. Kashihara et al. (Eds.), *Proceedings of the 32nd International Conference on Computers in Education* (pp. 561–590). Asia-Pacific Society for Computers in Education.

Kim, S.-W., Lim, S., Hong, S.-J., & Lee, Y. (2025). The effects of design-thinking-based AI education programs utilizing generative AI on Korean middle school students' creative problem-solving ability. *International Journal on Informatics Visualization*, *9*(4), 1553–1562. https://dx.doi.org/10.62527/joiv.9.4.4378

Lee, J., & Suh, S. (2024). AI Technology Integrated Education Model for empowering fashion Design ideation. *Sustainability*, *16*(17), Article 7262. https://doi.org/10.3390/su16177262

Li, H. (2023). Effects of a ChatGPT-based flipped learning guiding approach on learners' courseware project performances and perceptions. *Australasian Journal of Educational Technology*, *39*(5), 40–58. https://doi.org/10.14742/ajet.8923

Li, M., & Wilson, J. (2025). AI-Integrated scaffolding to Enhance agency and creativity in K-12 English language learners: A Systematic review. *Information*, *16*(7), 519. https://doi.org/10.3390/info16070519

Lin, H., & Chen, Q. (2024). Artificial intelligence (AI) -integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. *BMC Psychology*, *12*(1), Article 487. https://doi.org/10.1186/s40359-024-01979-0

Lobo-Quintero, R. (2025). AI-Enhanced Think-Pair-Share. *Journal of Learning Analytics*, *12*(2), 19–34. https://doi.org/10.18608/jla.2025.8807

Mahama, I., & Amadu, P. (2025). You are the Driver and AI is the Mate: Exploring Human-Led Creative and Critical Thinking in AI-Augmented Learning Environments. *F1000Research*, *14*,Article 974. https://doi.org/10.12688/f1000research.167988.1

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki, N. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open*, *5*, Article 100296. https://doi.org/10.1016/j.ijedro.2023.100296

Mardatillah, F., Gumilang, R. M., Wahyudi, M. A., Rawanita, M., & Muhammad, M. (2025). Epistemological Reconstruction of Islamic Education: Developing a transformative pedagogical model to foster creativity. *Jurnal Ilmiah Peuradeun*, *13*(2), 1071–1094. https://doi.org/10.26811/peuradeun.v13i2.2200

Meliyawati, N., Hikmat, A., & Mulyono, H. (2025). Enhancing Language Creativity through Gen-AI-Assisted Fiction Reading: A Guided Discovery Inquiry Approach. *Journal of Educational and Social Research*, *15*(5), 388–400. https://doi.org/10.36941/jesr-2025-0184

Melker, S., Gabrils, E., Villavicencio, V., Faraon, M., & Rönkkö, K. (2025). Artificial intelligence for design education: a conceptual approach to enhance students' divergent and convergent thinking in ideation processes. *International Journal of Technology and Design Education*, *35*(5), 1871–1899. https://doi.org/10.1007/s10798-025-09964-3

Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., Kinshuk, N., López-Pernas, S., & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): a roadmap to a future research agenda. *Smart Learning Environments*, *11*(1), Article 59. https://doi.org/10.1186/s40561-024-00350-5

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023b). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, *71*(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

Paz-Baruch, N., Grovas, G., & Mevarech, Z. R. (2025). The effects of meta-creative pedagogy on elementary school students' creative thinking. *Metacognition and Learning*, *20*(1), Article 9. https://doi.org/10.1007/s11409-025-09412-6

Peláez-Sánchez, I. C., Velarde-Camaqui, D., & Glasserman-Morales, L. D. (2024). The impact of large language models on higher education: exploring the connection between AI and Education 4.0. *Frontiers in Education*, *9,* Article 1392091. https://doi.org/10.3389/feduc.2024.1392091

Peng, J., & Li, Y. (2025). Frontiers of artificial intelligence for personalized learning in higher education: A systematic review of leading articles. *Applied Sciences*, *15*(18), Article 10096. https://doi.org/10.3390/app151810096

Qian, Y. (2025). Pedagogical Applications of Generative AI in Higher Education: A Systematic Review of the field. *TechTrends*, *69*(5), 1105–1120. https://doi.org/10.1007/s11528-025-01100-1

Rahioui, F., Jouti, M. a. T., & Ghzaoui, M. E. (2025). Exploring the effect of ChatGPT-3 on biology students' lateral thinking skills: a mixed-methods study and impacts for ai-enhanced education. *Telematics and Informatics Reports*, *19*, Article 100249. https://doi.org/10.1016/j.teler.2025.100249

Rahman, G., Almutairi, E. a. A., Mudhsh, B. A., & Al-Yafaei, Y. (2025). Harnessing generative AI for collaborative creativity: A study of university students' engagement and innovation. *International Journal of Innovative Research and Scientific Studies*, *8*(3), 3284–3296. https://doi.org/10.53894/ijirss.v8i3.7227

Rana, V., Verhoeven, B., & Sharma, M. (2025). Generative AI in Design Thinking Pedagogy: Enhancing creativity, critical thinking, and ethical reasoning in Higher education. *Journal of University Teaching and Learning Practice*, *22*(4),

1–22. https://doi.org/10.53761/tjse2f36

Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical Thinking: adoption of generative artificial intelligence tools in higher education. *Sustainability*, *16*(13), Article 5367. https://doi.org/10.3390/su16135367

Samaniego, M., Usca, N., Salguero, J., & Quevedo, W. (2024). Creative Thinking in Art and Design Education: A Systematic review. *Education Sciences*, *14*(2), 192. https://doi.org/10.3390/educsci14020192

Sánchez-García, R., & Reyes-De-Cózar, S. (2025). Enhancing Project-Based Learning: A Framework for Optimizing Structural Design and Implementation—A Systematic Review with a Sustainable Focus. *Sustainability*, *17*(11), Article 4978. https://doi.org/10.3390/su17114978

Saritepeci, M., & Durak, H. Y. (2024). Effectiveness of artificial intelligence integration in design-based learning on design thinking mindset, creative and reflective thinking skills: An experimental study. *Education and Information Technologies*, *29*(18), 25175–25209. https://doi.org/10.1007/s10639-024-12829-2

Setiawan, B., Ardianto, D., & Windiyani, T. (2025). AI-based website for integrating STEM and ESD: Enhancing students' creative thinking, creative products, and self-reflection. *Buana Pendidikan*, *21*(2), 189–200.

Silva, H., Lopes, J., Dominguez, C., & Morais, E. (2021). Lecture, cooperative learning and concept mapping: any differences on critical and creative thinking development? *International Journal of Instruction*, *15*(1), 765–780. https://doi.org/10.29333/iji.2022.15144a

Sinaga, A. I., Nasution, A. F., & Albina, M. (2025). The effect of artificial intelligence use on creativity in writing scientific works. *JKTP: Jurnal Kajian Teknologi Pendidikan*, *8*(3), 226–234. https://doi.org/10.17977/um038v8i32025p226-234

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Song, D., Zhang, P., Zhu, Y., Qi, S., Yang, Y., Gong, L., & Zhou, L. (2025). Effects of generative artificial intelligence on higher-order thinking skills and artificial intelligence literacy in nursing undergraduates: A quasi-experimental study. *Nurse Education in Practice*, *88*, Article 104549. https://doi.org/10.1016/j.nepr.2025.104549

Sun, Y., Gong, Y., & Wong, N. H. (2025). Creative Skills in Crisis? Bridging the Gap Between Arts Education and Human Capital Demands in the Age of Artificial Intelligence. *Journal of Integrated Arts Education*, *1*(1), 25–39. https://doi.org/10.64797/jiae.v1i1.50

Susilo, B., & Salirawati, D. (2025). Improving students' creative thinking in macromolecule lessons through AI-based ethnochemistry project-based learning. *Journal of Educational Sciences*, *9*(6), 5277–5296. https://doi.org/10.31258/jes.9.6.p.5277-5296

Syafriati, A. (2024). The effectiveness of using artificial intelligence on students' learning interest, critical thinking, and creativity in nursing education. *Journal of Nursing Culture and Technology*, *1*(2), 31–42. https://doi.org/10.70049/jnctech.v1i2.16

Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial Intelligence in Education: AIED for Personalised Learning Pathways. *The Electronic Journal of e-Learning*, *20*(5), 639–653. https://doi.org/10.34190/ejel.20.5.2597

Toma, R. B., & Yánez-Pérez, I. (2024). Effects of ChatGPT use on undergraduate students' creativity: A threat to creative thinking? *Discover Artificial Intelligence*, *4*, Article 74. https://doi.org/10.1007/s44163-024-00172-x

Topali, P., Haelermans, C., Molenaar, I., & Segers, E. (2025). Pedagogical considerations in the automation era: A systematic literature review of AIEd in K-12 authentic settings. *British Educational Research Journal*, *51*(6), 2777–2809. https://doi.org/10.1002/berj.4200

Trullàs, J. C., Blay, C., Sarri, E., & Pujol, R. (2022). Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review. *BMC Medical Education*, *22*(1), Article 104. https://doi.org/10.1186/s12909-022-03154-8

Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: crafting with text-to-image generative models. *Digital Creativity*, *34*(1), 1–21. https://doi.org/10.1080/14626268.2023.2174557

Wang, A., Zhang, Y., Wang, A., & Zheng, W. (2025). The impact of AI-based painting technology on children's creative thinking. *Frontiers in Psychology*, *16*, Article 1598210. https://doi.org/10.3389/fpsyg.2025.1598210

Wang, C. (2024). Exploring Students' Generative AI-Assisted Writing Processes: Perceptions and Experiences from Native and Nonnative English Speakers. *Technology Knowledge and Learning*, *30*(3), 1825–1846. https://doi.org/10.1007/s10758-024-09744-3

Wei, X., Wang, L., Lee, L., & Liu, R. (2025). The effects of generative AI on collaborative problem-solving and team creativity performance in digital story creation: an experimental study. *International Journal of Educational Technology in Higher Education*, *22*(1), Article 23. https://doi.org/10.1186/s41239-025-00526-0

Wijnia, L., Noordzij, G., Arends, L. R., Rikers, R., & Loyens, S. (2024). The Effects of Problem-Based, Project-Based, and Case-Based Learning on Students' Motivation: a Meta-Analysis. *Educational Psychology Review*, *36*(1), Article 29. https://doi.org/10.1007/s10648-024-09864-3

Wu, D., & Zhang, J. (2025). Generative artificial intelligence in secondary education: Applications and effects on students' innovation skills and digital literacy. *PLoS ONE*, *20*(5), Article e0323349. https://doi.org/10.1371/journal.pone.0323349

Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour*, *8*(10), 1839–1850. https://doi.org/10.1038/s41562-024-02004-5

Yan, Z., & Qianjun, T. (2025). Integrating AI-generated content tools in higher education: a comparative analysis of interdisciplinary learning outcomes. *Scientific Reports*, *15*(1), 25802. https://doi.org/10.1038/s41598-025-10941-y

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, H., Wang, X., Li, X., Zhai, J., Li, X., & Guo, Y. (2025). The impact of large-class flipped classrooms incorporating design thinking on self-awareness, team collaboration, learning efficiency, and comprehensive literacy of clinical medicine undergraduates. *BMC Medical Education*, *25*(1), Article 562. https://doi.org/10.1186/s12909-025-07116-8

Zhang, J., & Zhang, Z. (2024). AI in teacher education: Unlocking new dimensions in teaching support, inclusive learning, and digital literacy. *Journal of Computer Assisted Learning*, *40*(4), 1871–1885. https://doi.org/10.1111/jcal.12988

Zhang, S., Lu, X., & Liu, X. (2024). Study on the influence of AI composition software on students' creative ability in music education. *Journal of Educational Technology and Innovation*, *6*(2), 1–17. https://doi.org/10.61414/jeti.v6i2.190

Zhang, X., Ma, Y., Jiang, Z., Chandrasekaran, S., Wang, Y., & Fofou, R. F. (2021). Application of Design-Based Learning and Outcome-Based Education in Basic Industrial Engineering Teaching: A New Teaching Method. *Sustainability*, *13*(5), Article 2632. https://doi.org/10.3390/su13052632

Zhao, L., He, W., & Su, Y. (2021). Innovative Pedagogy and Design-Based Research on flipped learning in Higher Education. *Frontiers in Psychology*, *12*, Article 577002. https://doi.org/10.3389/fpsyg.2021.577002

Zhou, Z., Guo, H., Ma, F., Yang, C., & Gao, Y. (2025). The chain mediating role of critical thinking and AI self-efficacy in GenAI usage competence and engineering students' creativity. *Scientific Reports*, *15*(1), Article 35945. https://doi.org/10.1038/s41598-025-21132-0

Zou, D., Xie, H., & Kohnke, L. (2025). Navigating the Future: Establishing a framework for educators' pedagogic artificial intelligence competence. *European Journal of Education*, *60*(2), Article e70117. https://doi.org/10.1111/ejed.70117