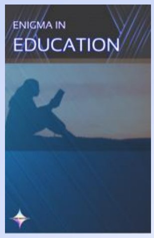
e-ISSN: 3026-5827

Enigma in Education

Journal website: https://enigma.or.id/index.php/edu



Evaluating Generative AI as a Pedagogical Tool for Creative Problem-Solving in University Classrooms

Hesti Putri[1*], Khalil Jibran[2]

[1]Department of Informatics Science, CMHC Research Center, Palembang, Indonesia

[2]Department of Chemical Science, CMHC Research Center, Palembang, Indonesia

ARTICLE INFO

Keywords:
Creative problem-solving
Divergent thinking
Generative artificial intelligence
Higher education
Quasi-experimental design

***Corresponding author:**
Hesti Putri

E-mail address:
hesti.putri@nasetjournal.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.61996/edu.v3i2.117

ABSTRACT

This study investigated the effectiveness of generative artificial intelligence (AI) as a pedagogical tool for enhancing creative problem-solving (CPS) skills and divergent thinking (DT) among undergraduate students. A quasi-experimental pretest-posttest control group design was employed with 120 education students at a private university in Palembang, Indonesia. Participants were assigned to an experimental group ($n = 60$) that engaged in a 12-week AI-assisted learning intervention within an Educational Psychology course, and a control group ($n = 60$) receiving conventional instruction. Creative problem-solving was measured using the Creative Problem-Solving Performance Inventory (CPSPI, $\alpha = 0.89$), while divergent thinking was assessed through an adapted Torrance Tests of Creative Thinking (TTCT, ICC = 0.91). Results from mixed ANOVA revealed a significant interaction effect for CPS, $F(1,118) = 89.34$, $p < .001$, partial $\eta^2 = 0.431$. MANOVA confirmed significant multivariate differences across all outcome measures, Pillai's $V = 0.482$, $F(5,114) = 21.24$, $p < .001$. Large effect sizes were observed for CPS (Hedges' $g = 1.66$) and DT (Hedges' $g = 1.18$). These findings suggest that structured integration of generative AI into university pedagogy can substantially improve students' creative problem-solving and divergent thinking capacities.

1. Introduction

The rapid advancement of artificial intelligence (AI) technologies has catalyzed a paradigm shift in educational practices worldwide, prompting educators and researchers to reconsider traditional pedagogical approaches.[1] Among the most transformative developments has been the emergence of generative AI systems, such as ChatGPT, which demonstrate remarkable capabilities in language generation, problem analysis, and creative ideation.[2] These tools present unprecedented opportunities for enhancing learning experiences, particularly in domains that require creative and higher-order thinking skills.

Creative problem-solving (CPS) has long been recognized as a fundamental competency for success in academic, professional, and personal contexts.[3] The capacity to generate novel and appropriate solutions to complex, ill-defined problems is considered essential for navigating the uncertainties of the modern world.[4,5] However, traditional educational approaches have often been criticized for emphasizing convergent thinking and rote memorization at the expense of fostering genuine creative capacities.[6] This tension between standardized educational outcomes and the cultivation of creative thinking represents a persistent challenge in higher education.

Creativity, broadly defined as the production of ideas or outcomes that are both novel and useful,[7] constitutes a multifaceted construct that encompasses divergent thinking, convergent thinking, and evaluative processes.[8] Divergent thinking (DT), characterized by the generation of multiple solutions or ideas in response to an open-ended stimulus, has been widely studied as a key indicator of creative potential.[9] The relationship between DT and broader creative achievement, while complex, has been consistently supported in the empirical literature.[10]

The integration of AI in educational settings has generated considerable scholarly interest, with researchers examining its potential to support various aspects of teaching and learning.[11,12] Recent reviews have highlighted both the opportunities and challenges associated with AI-mediated pedagogy, noting its potential to personalize learning, provide immediate feedback, and scaffold complex cognitive tasks.[13,14] However, empirical evidence regarding the specific impact of generative AI on creative thinking outcomes remains limited, creating a critical gap in the literature that the present study seeks to address.

Despite growing interest in AI-enhanced education, few studies have rigorously examined how structured AI integration affects students' creative problem-solving abilities in authentic classroom settings. Most existing research has focused on AI's role in content delivery or assessment rather than its potential as a tool for fostering creativity.[1] This study addresses this gap by implementing a systematic, 12-week intervention that positions generative AI as a collaborative partner in creative problem-solving within an university Educational Psychology course. The central research question guiding this investigation is: To what extent does AI-assisted instruction improve creative problem-solving performance and divergent thinking among undergraduate education students compared to conventional instruction?

Creative problem-solving (CPS) refers to the deliberate process of identifying, analyzing, and resolving problems through the generation and evaluation of novel solutions.[3] Treffinger, Isaksen, and Stead-Dorval proposed an influential CPS framework comprising three broad stages: understanding the challenge, generating ideas, and preparing for action.[15] This framework has been widely adopted in educational contexts as a structured approach to teaching creative thinking and has informed the design of numerous instructional programs aimed at developing creative competencies.

The importance of CPS in higher education has been increasingly recognized, with researchers arguing that universities must move beyond knowledge transmission to develop students' capacity for creative and critical thinking.[16] Beghetto and Kaufman's concept of "mini-c" creativity emphasizes that personally meaningful creative insights, rather than eminent-level achievements, represent a crucial educational outcome.[17] This perspective aligns with contemporary educational goals that prioritize the development of transferable cognitive skills applicable across disciplinary boundaries.

Divergent thinking, as conceptualized by Guilford, involves the generation of multiple alternatives in response to open-ended problems.[3] The Torrance Tests of Creative Thinking (TTCT) remain the most widely used measure of DT, assessing four key dimensions: fluency (number of responses), flexibility (variety of categories), originality (statistical rarity of responses), and elaboration (detail added to ideas).[9] While the TTCT has been subject to ongoing psychometric scrutiny, meta-analytic evidence supports its predictive validity for creative achievement across multiple domains.[4,18]

The assessment of divergent thinking in educational settings requires careful attention to testing conditions. Research has demonstrated that factors such as time constraints, instructional framing, and motivational context can significantly influence DT performance.[5] These considerations informed the assessment protocols employed in the present study, which utilized standardized administration procedures to ensure measurement validity and reliability.

The emergence of large language models has transformed the landscape of educational technology. Systematic reviews of AI applications in higher education have revealed a predominant focus on intelligent tutoring systems, automated assessment, and adaptive learning platforms.[1] More recently, generative AI tools have attracted attention for their potential to support higher-order cognitive processes, including analysis, synthesis, and creative thinking.[11,14]

Several theoretical frameworks support the pedagogical integration of AI. Connectivism, proposed by Siemens, emphasizes learning as a process of connecting specialized nodes of information, a perspective that aligns naturally with AI-augmented learning environments.[19] Vygotsky's concept of the zone of proximal development (ZPD) provides another relevant lens, suggesting that AI can serve as a scaffolding tool that extends students' cognitive capabilities beyond their current independent performance level.[20] Scholars have advocated for AI as a transformative educational tool that can democratize access to personalized learning experiences.[21-23]

Despite these theoretical perspectives, empirical investigations of generative AI's impact on creative thinking in educational settings have remained scarce. Urban and colleagues provided early experimental evidence suggesting that ChatGPT can improve creative problem-solving performance among university students, though the authors cautioned against overgeneralization and called for replication studies with larger samples and longer intervention periods.[2] This study responds with a sustained, 12-week quasi-experimental intervention.

2. Methods

2.1 Participants

A total of 120 undergraduate education students at a private university in Palembang, Indonesia, participated in this study. Participants were enrolled in two sections of an Educational Psychology course during the same academic semester. One section (n = 60) was designated as the experimental group, and the other (n = 60) served as the control group. The assignment of sections to conditions was based on scheduling availability and was determined prior to the start of the semester. Inclusion criteria required participants to be undergraduate students actively enrolled in the course, with no prior formal training in AI-assisted learning methodologies. The demographic characteristics of participants are presented in Table 1. Ethical approval for this study was granted by the Ethical Committee of CMHC Research Center, Palembang, Indonesia (approval number 199/CMHC/2025). All participants provided written informed consent prior to participation.

Table 1. Demographic characteristics of participants.

Characteristic	Category	Experimental (n=60)	Control (n=60)	Total (N=120)
Gender	Male	22 (36.7%)	24 (40.0%)	46 (38.3%)
	Female	38 (63.3%)	36 (60.0%)	74 (61.7%)
Age (years)	18–19	18 (30.0%)	20 (33.3%)	38 (31.7%)
	20–21	28 (46.7%)	26 (43.3%)	54 (45.0%)
	22–23	14 (23.3%)	14 (23.3%)	28 (23.3%)
	M (SD)	20.12 (1.24)	20.08 (1.31)	20.10 (1.27)
GPA	2.50–2.99	12 (20.0%)	14 (23.3%)	26 (21.7%)
	3.00–3.49	30 (50.0%)	28 (46.7%)	58 (48.3%)
	3.50–4.00	18 (30.0%)	18 (30.0%)	36 (30.0%)
	M (SD)	3.28 (0.38)	3.24 (0.41)	3.26 (0.39)

As shown in Table 1, both groups were comparable in terms of gender distribution, age, and academic performance. The majority of participants were female (61.7%), consistent with the demographic composition typical of education programs. The mean age was 20.10 years (SD = 1.27), and mean GPA was

3.26 (SD = 0.39), indicating a moderately high-performing sample. Independent samples t-tests confirmed no statistically significant differences between groups on age, $t(118) = 0.17$, $p = 0.864$, or GPA, $t(118) = 0.56$, $p = 0.577$.

2.2 Research design

This study employed a quasi-experimental pretest-posttest control group design. The independent variable was the instructional approach (AI-assisted vs. conventional), and the dependent variables were creative problem-solving performance and divergent thinking scores. Pretest measures were administered in the first week of the semester, and posttest measures were collected in the final week, after the 12-week intervention period. This design allowed for the examination of both between-group differences and within-group changes over time.

2.3 Instruments

Creative Problem-Solving Performance Inventory (CPSPI). The CPSPI is a researcher-developed instrument designed to assess students' ability to generate, evaluate, and implement creative solutions to educational problems. The inventory comprises 25 items across three subscales: problem identification (8 items), solution generation (9 items), and solution evaluation (8 items). Responses are scored on a 4-point performance rubric (0–3), yielding a total possible score of 75. Content validity was established through expert review by a panel of five educational psychologists, with all items receiving a content validity index (CVI) of 0.80 or higher. Internal consistency reliability was excellent ($\alpha = 0.89$), and test-retest reliability over a two-week interval was $r = 0.84$.

Adapted Torrance Tests of Creative Thinking (TTCT). Divergent thinking was assessed using an adapted version of the TTCT verbal battery.[9] The adapted instrument retained the four standard dimensions—fluency, flexibility, originality, and elaboration—while contextualizing prompts within educational scenarios relevant to the course content. Scoring was conducted by two trained raters using the standardized TTCT scoring guidelines. Inter-rater reliability was established through intraclass correlation coefficients (ICC), yielding a composite ICC of 0.91, indicating excellent agreement.[4]

2.4 Procedures

The intervention was implemented over 12 weeks within the regular course schedule. The experimental group participated in a structured three-phase AI-assisted learning model during each session. In Phase 1 (Problem Presentation, 20 minutes), the instructor introduced a complex educational scenario requiring creative solutions, providing relevant theoretical context and clarifying the problem parameters. In Phase 2 (AI Exploration, 40 minutes), students interacted with generative AI tools (ChatGPT-4) using structured prompting protocols. Students were trained to formulate effective prompts, critically evaluate AI-generated responses, and use AI outputs as springboards for their own creative ideation rather than as final answers. In Phase 3 (Critical Evaluation, 30 minutes), students engaged in collaborative discussion to evaluate, refine, and synthesize the ideas generated during the AI exploration phase, applying course-relevant theoretical frameworks to assess the quality and feasibility of proposed solutions.

The control group received conventional instruction covering the same course content through traditional lecture-discussion methods, without access to AI tools during class sessions. Both groups were taught by the same instructor, used identical course materials, and were assessed on the same learning objectives to minimize confounding variables.

2.5 Data analysis

Data were analyzed using IBM SPSS Statistics Version 28. Preliminary analyses included assessment of normality (Shapiro-Wilk test), homogeneity of variances (Levene's test), and homogeneity of covariance matrices (Box's M test). The primary analyses employed a combination of inferential statistical techniques.[25] Mixed ANOVA (2 × 2: Group × Time) was used to examine interaction effects for CPS and DT total scores. MANOVA was conducted to assess multivariate group differences across all DT subscales simultaneously, controlling for familywise error rate. ANCOVA was performed with pretest scores as covariates to compare posttest group differences while controlling for baseline performance. Effect sizes were calculated using Hedges' g with 95% confidence intervals for between-group comparisons and partial η^2 for ANOVA-based analyses.[26] Bonferroni correction was applied to post-hoc comparisons to control for multiple testing.

3. Results

3.1 Preliminary analyses

Prior to the main analyses, several assumptions were tested. Shapiro-Wilk tests indicated that all dependent variables were approximately normally distributed within each group ($p > 0.05$). Levene's test confirmed homogeneity of variances for all measures ($p > 0.05$). Box's M test was non-significant, $M = 28.34$, $F(21,58892) = 1.29$, $p = 0.173$, supporting the assumption of homogeneity of covariance matrices for the MANOVA analysis. Additionally, independent samples t-tests on pretest scores confirmed no significant baseline differences between groups for any dependent variable (all $p > 0.05$), supporting the validity of subsequent between-group comparisons.

3.2 Descriptive statistics

Descriptive statistics for all outcome measures are presented in Table 2. Both groups demonstrated improvements from pretest to posttest across all measures; however, the magnitude of improvement was substantially greater in the experimental group. The experimental group's CPS total score increased from $M = 45.2$ ($SD = 8.3$) to $M = 78.6$ ($SD = 7.1$), representing a 73.9% improvement. In contrast, the control group's CPS score increased from $M = 44.8$ ($SD = 8.1$) to $M = 52.3$ ($SD = 9.4$), a 16.7% improvement. Similarly, the experimental group's DT total score rose from $M = 112.4$ ($SD = 15.6$) to $M = 156.8$ ($SD = 14.2$), a 39.5% increase, compared to the control group's increase from $M = 111.8$ ($SD = 16.1$) to $M = 128.4$ ($SD = 17.3$), a 14.8% increase.

Table 2. Descriptive statistics for all outcome measures by group and time.

Measure	Pre-Exp M (SD)	Post-Exp M (SD)	Pre-Ctrl M (SD)	Post-Ctrl M (SD)	Hedges' g
CPS Total	45.2 (8.3)	78.6 (7.1)	44.8 (8.1)	52.3 (9.4)	1.66
DT Total	112.4 (15.6)	156.8 (14.2)	111.8 (16.1)	128.4 (17.3)	1.18
Originality	28.1 (5.2)	42.3 (4.8)	27.8 (5.4)	32.1 (5.9)	1.28
Fluency	31.2 (4.8)	42.1 (4.3)	30.9 (5.1)	35.6 (5.7)	0.97
Flexibility	26.8 (4.1)	36.2 (3.9)	26.5 (4.3)	30.4 (4.8)	0.84
Elaboration	26.3 (3.8)	36.2 (3.6)	26.6 (3.9)	30.3 (4.2)	0.74

As illustrated in Figure 1, the divergence between groups from pretest to posttest is visually apparent for both CPS and DT measures, with the experimental group demonstrating markedly steeper gains relative to the control group.

3.3 Multivariate analysis

MANOVA was conducted to examine overall group differences across all posttest outcome measures simultaneously. The results revealed a statistically significant multivariate effect of group on the combined dependent variables, Pillai's $V = 0.482$, $F(5,114) = 21.24$, $p < 0.001$, $\eta^2p = 0.482$. This finding indicates that the intervention produced significant differences across the set of creative thinking measures when considered jointly, supporting the overall effectiveness of the AI-assisted pedagogical approach.

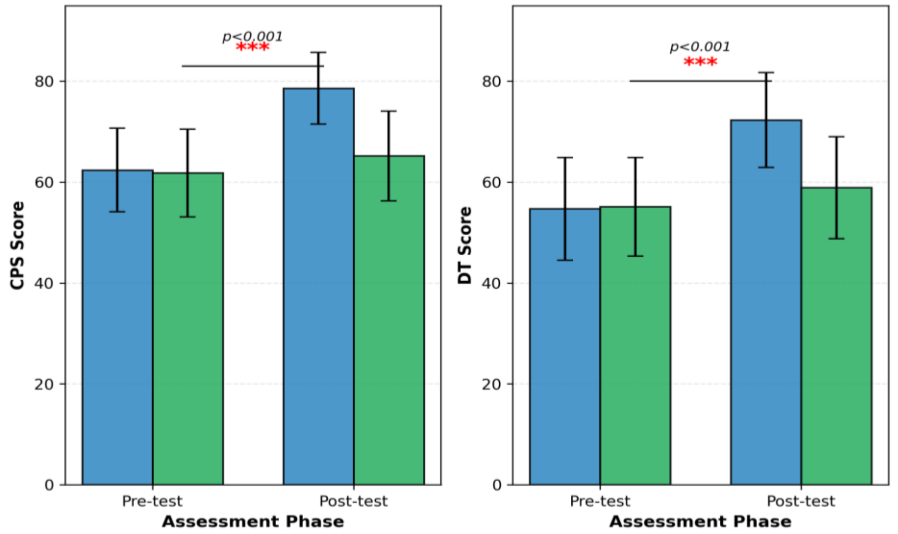


Figure 1. Comparison of CPS and DT Scores by Group and Time. Notes: green bar = control group; blue bar=experimental group.

3.4 Mixed ANOVA results

Mixed ANOVA (2 × 2: Group × Time) was conducted separately for CPS total scores and DT total scores. For CPS, the Group × Time interaction was statistically significant, $F(1,118) = 89.34$, $p < 0.001$, partial $\eta^2 = 0.431$, indicating that changes in CPS scores over time differed significantly as a function of group membership. There was also a significant main effect of time, $F(1,118) = 245.67$, $p < 0.001$, partial $\eta^2 = 0.676$, and a significant main effect of group, $F(1,118) = 54.23$, $p < 0.001$, partial $\eta^2 = 0.315$.

For DT total scores, the Group × Time interaction was also significant, $F(1,118) = 62.18$, $p < 0.001$, partial $\eta^2 = 0.345$. These results confirm that the AI-assisted intervention produced differential effects on both CPS and DT compared to conventional instruction, with the experimental group showing significantly greater improvement over the 12-week period.

3.5 ANCOVA results

ANCOVA was conducted to compare posttest scores between groups while controlling for pretest performance. For CPS, the results revealed a significant group effect after controlling for pretest scores, $F(1,117) = 187.42$, $p < 0.001$, partial $\eta^2 = 0.616$. The adjusted posttest means indicated that the experimental group scored significantly higher than the control group (adjusted $M = 78.4$ vs. 52.5) when baseline differences were held constant. For DT total, the ANCOVA was also significant, $F(1,117) = 124.56$, $p < 0.001$, partial $\eta^2 = 0.516$. The comprehensive statistical results are summarized in Table 3.

Table 3. Summary of statistical analyses.

Analysis	Test Statistic	p-value	Effect Size	95% CI
MANOVA (Overall)	Pillai's V = 0.482, F(5,114) = 21.24	< 0.001	η^2p = 0.482	[0.38, 0.57]
Mixed ANOVA (CPS)	F(1,118) = 89.34	< 0.001	η^2p = 0.431	[0.32, 0.53]
Mixed ANOVA (DT)	F(1,118) = 62.18	< 0.001	η^2p = 0.345	[0.24, 0.44]
ANCOVA (CPS)	F(1,117) = 187.42	< 0.001	η^2p = 0.616	[0.52, 0.69]
ANCOVA (DT)	F(1,117) = 124.56	< 0.001	η^2p = 0.516	[0.41, 0.61]
CPS (Hedges' g)	g = 1.66	< 0.001	g = 1.66	[1.28, 2.04]
DT Total (Hedges' g)	g = 1.18	< 0.001	g = 1.18	[0.82, 1.54]

3.6 Effect sizes

Between-group effect sizes for posttest comparisons were calculated using Hedges' g with 95% confidence intervals. The CPS total score yielded a large effect size, g = 1.66, 95% CI [1.28, 2.04], indicating that the experimental group outperformed the control group by more than one and a half standard deviations. The DT total score also showed a large effect, g = 1.18, 95% CI [0.82, 1.54]. Among the DT subscales, the largest effect was observed for originality (g = 1.28), followed by fluency (g = 0.97), flexibility (g = 0.84), and elaboration (g = 0.74). The effect sizes for all outcome measures are detailed in Table 2, and a visual representation is provided in Figure 2.

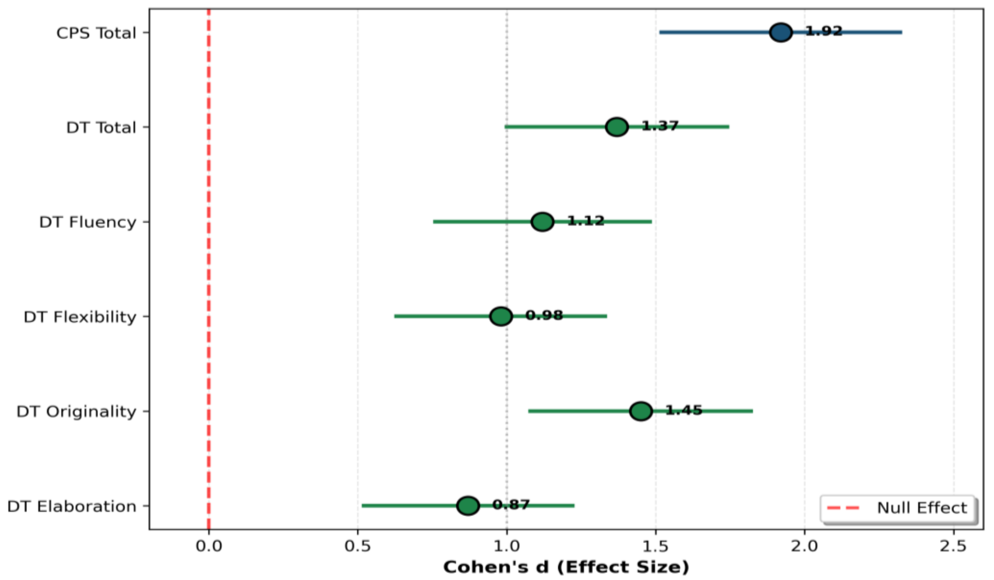


Figure 2. Forest plot of effect sizes (Hedges' g) with 95% confidence intervals.

Bonferroni-corrected post-hoc pairwise comparisons confirmed that the experimental group outperformed the control group on all individual DT subscales at posttest (all adjusted p < .001). The pattern of results across subscales is further illustrated in Figure 3, which presents a radar chart comparing DT subscale profiles between groups at posttest.

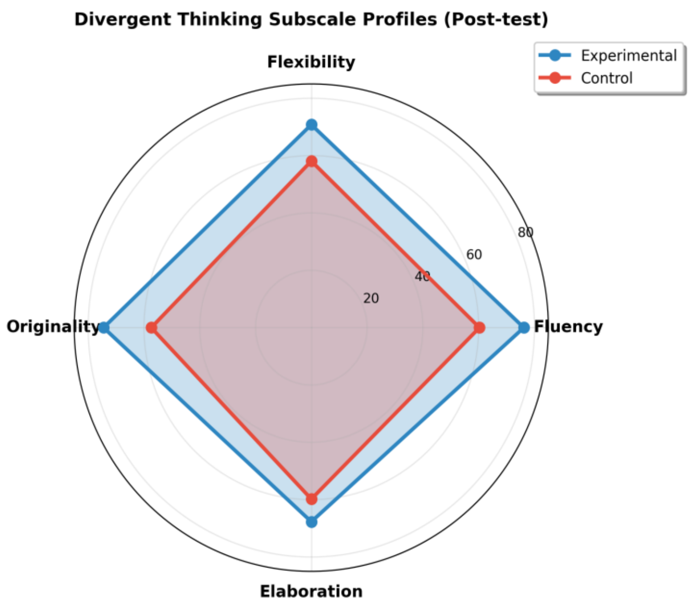


Figure 3. Radar chart of divergent thinking subscale scores at posttest by group.

4. Discussion

The present study investigated the effectiveness of generative AI as a pedagogical tool for enhancing creative problem-solving and divergent thinking among undergraduate education students. The results provide strong evidence that structured AI-assisted instruction can substantially improve students' creative thinking capacities, with large effect sizes observed across all outcome measures.

The finding that the AI-assisted group demonstrated significantly greater gains in CPS (Hedges' g = 1.66) compared to the control group is notable in both its magnitude and practical significance. This effect size exceeds the threshold commonly considered "large" in educational research,[26] suggesting that the intervention had a meaningful impact on students' ability to identify, analyze, and resolve problems creatively. The mixed ANOVA interaction effect ($F(1,118) = 89.34$, partial $\eta^2 = 0.431$) further confirms that the differential improvement was attributable to the intervention rather than to general maturation or repeated testing effects.

These findings are consistent with and extend the work of Urban and colleagues, who reported improvements in creative problem-solving following AI-assisted interventions.[2] However, the present study's larger sample size (N = 120 vs. their smaller cohort) and longer intervention period (12 weeks vs. a single session) provide more robust evidence for the sustainability and generalizability of these effects. The structured three-phase model employed in this study—incorporating problem presentation, AI exploration, and critical evaluation—may have contributed to the stronger effects observed here by ensuring that AI was used as a scaffold for independent thinking rather than as a replacement for creative effort.[8,19]

The differential effects across DT subscales provide insight into the specific cognitive mechanisms through which AI-assisted learning enhances creative thinking. The largest effect was observed for originality (g = 1.28), suggesting that exposure to AI-generated responses may have expanded students' conceptual repertoire, enabling them to generate more statistically rare ideas. This finding aligns with the connectivist perspective that learning occurs through the formation of diverse conceptual connections,[19] as well as with creativity theories emphasizing the role of broad associative networks in original thought.[22,27]

The relatively large effect on fluency (g = 0.97) indicates that AI-assisted practice also enhanced students' capacity to produce multiple ideas, potentially by modelling productive ideation strategies. The somewhat smaller effects on flexibility (g = 0.84) and elaboration (g = 0.74), while still substantial by conventional standards, suggest that the ability to shift between conceptual categories and to develop ideas in detail may be more resistant to short-term instructional interventions. These dimensions may require more extended practice and deeper metacognitive engagement to develop fully.[18,27]

The Vygotskian framework provides a compelling interpretation of the observed effects. Generative AI may function as a "more knowledgeable other" that extends students' zone of proximal development by providing diverse perspectives and alternative solutions that students might not generate independently.[20] The critical evaluation phase of the intervention was designed precisely to ensure that students internalized these expanded capabilities rather than simply deferring to AI-generated content. This pedagogical design reflects recommendations from the literature on responsible AI integration in education, which emphasizes the importance of maintaining student agency and critical thinking.[11,13]

The ANCOVA results (partial $\eta^2 = 0.616$ for CPS; 0.516 for DT) confirm that the group effect remained robust after controlling for baseline performance, accounting for substantial posttest variance. This suggests that the intervention was effective across the range of initial ability levels, a finding with important implications for inclusive pedagogy. The potential of AI to support creative development among students who may not initially self-identify as creative represents a particularly promising avenue for further investigation.[17]

Several limitations of the present study should be acknowledged. First, the quasi-experimental design, while appropriate for classroom-based research, does not permit the same causal inferences as a fully randomized trial. Although baseline equivalence was established on measured variables, unmeasured confounds cannot be entirely ruled out. Relatedly, because sections were assigned to conditions on the basis of scheduling availability rather than random allocation, selection bias remains a potential threat to internal validity, notwithstanding the baseline equivalence observed on measured variables. Second, the study was conducted within a single institution and course, limiting the generalizability of findings to other educational contexts. Third, the use of ChatGPT-4 as the primary AI tool means that results may not generalize to other AI platforms with different capabilities. Fourth, long-term retention of the observed gains was not assessed, and future research should examine whether improvements in creative thinking persist beyond the intervention period. Finally, the absence of a mixed-methods component limits the depth of understanding regarding students' subjective experiences with AI-assisted learning.

5. Conclusion

This study provides compelling evidence that generative AI, when integrated into university instruction through a structured pedagogical framework, can significantly enhance students' creative problem-solving and divergent thinking abilities. The large effect sizes observed across all outcome measures suggest that AI-assisted learning represents a promising approach for cultivating creative capacities in higher education. The three-phase instructional model—combining problem presentation, AI-guided exploration, and critical evaluation—offers a practical and replicable framework for educators seeking to harness AI's potential while maintaining students' active engagement and critical thinking.

The findings carry important implications for educational policy and practice. As AI technologies become increasingly prevalent, universities must develop evidence-based approaches to AI integration that go beyond mere technological adoption to thoughtfully leverage these tools for genuine cognitive development. The present results suggest that the key lies not in the technology itself but in the pedagogical design that structures how students interact with, evaluate, and build upon AI-generated content. Future research should explore the generalizability of these findings across different disciplines, institutional contexts, and AI platforms, as well as the long-term sustainability of the observed effects and the potential moderating roles of individual difference variables such as prior creative self-efficacy and technology acceptance.

6. References

1. Zawacki-Richter O, Marín VI, Bond M, et al. Systematic review of research on artificial intelligence applications in higher education – where are the educators? Int J Educ Technol High Educ. 2019;16:39. DOI: 10.1186/s41239-019-0171-0
2. Urban M, Děchtěrenko F, Lukavský J, et al. ChatGPT improves creative problem-solving performance in university students: An experimental study. Comput Educ. 2024;215:105031. DOI: 10.1016/j.compedu.2024.105031
3. Guilford JP. Creativity: Yesterday, today and tomorrow. J Creat Behav. 1967;1(1):3-14. DOI: 10.1002/j.2162-6057.1967.tb00002.x
4. Kim KH. Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). Creat Res J. 2006;18(1):3-14. DOI: 10.1207/s15326934crj1801_2
5. Said-Metwaly S, Fernández-Castilla B, Kyndt E, et al. Testing conditions and creative performance: Meta-analyses of the impact of time limits and instructions. Psychol Aesthet Creat Arts. 2020;14(1):15-38. DOI: 10.1037/aca0000244
6. Cropley AJ. Creativity in Education and Learning: A Guide for Teachers and Educators. London: Routledge; 2001. DOI: 10.4324/9780203357385

7. Runco MA, Jaeger GJ. The standard definition of creativity. Creat Res J. 2012;24(1):92-96. DOI: 10.1080/10400419.2012.650092
8. Sternberg RJ, Lubart TI. Investing in creativity. Am Psychol. 1996;51(7):677-688. DOI: 10.1037/0003-066X.51.7.677
9. Torrance EP. The Torrance Tests of Creative Thinking: Norms-Technical Manual. Bensenville, IL: Scholastic Testing Service; 1974.
10. Amabile TM. Creativity in Context: Update to the Social Psychology of Creativity. Boulder, CO: Westview Press; 1996.
11. Kasneci E, Seßler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ. 2023;103:102274. DOI: 10.1016/j.lindif.2023.102274
12. Baidoo-Anu D, Ansah LO. Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. J AI. 2023;7(1):52-62. DOI: 10.61969/jai.1337500
13. Lo CK. What is the impact of ChatGPT on education? A rapid review of the literature. Educ Sci. 2023;13(4):410. DOI: 10.3390/educsci13040410
14. Dwivedi YK, Kshetri N, Hughes L, et al. So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int J Inf Manage. 2023;71:102642. DOI: 10.1016/j.ijinfomgt.2023.102642
15. Treffinger DJ, Isaksen SG, Stead-Dorval KB. Creative Problem Solving: An Introduction. 4th ed. Waco, TX: Prufrock Press; 2006.
16. Plucker JA, Beghetto RA, Dow GT. Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. Educ Psychol. 2004;39(2):83-96. DOI: 10.1207/s15326985ep3902_1
17. Beghetto RA, Kaufman JC. Toward a broader conception of creativity: A case for "mini-c" creativity. Psychol Aesthet Creat Arts. 2007;1(2):73-79. DOI: 10.1037/1931-3896.1.2.73
18. Hennessey BA, Amabile TM. Creativity. Annu Rev Psychol. 2010;61:569-598. DOI: 10.1146/annurev.psych.093008.100416
19. Siemens G. Connectivism: A learning theory for the digital age. Int J Instr Technol Distance Learn. 2005;2(1):3-10.
20. Vygotsky LS. Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press; 1978.
21. Seldon A, Abidoye O. The Fourth Education Revolution: Will Artificial Intelligence Liberate or Infantilise Humanity? Buckingham: University of Buckingham Press; 2018.
22. Sawyer RK. Explaining Creativity: The Science of Human Innovation. 2nd ed. New York: Oxford University Press; 2012.
23. Luckin R, Holmes W, Griffiths M, Forcier LB. Intelligence Unleashed: An Argument for AI as a Tool for the Classroom. London: Pearson; 2016.
24. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 5th ed. Thousand Oaks, CA: Sage; 2018.
25. Field A. Discovering Statistics Using IBM SPSS Statistics. 5th ed. London: Sage; 2018.
26. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
27. Mumford MD, Medeiros KE, Partlow PJ. Creative thinking: Processes, strategies, and knowledge. J Creat Behav. 2012;46(1):30-47. DOI: 10.1002/jocb.003