COMPARATIVE STUDY OF GENERATIVE AI TOOLS IN VISUAL COMMUNICATION DESIGN EDUCATION: CREATIVITY AND USABILITY

Yana Erlyana[1*]; M Garry Saputra[2]

Visual Communication Design[1]
Interactive Design[2]
Universitas Bunda Mulia, Jakarta, Indonesia[1,2]
https://www.ubm.ac.id/[1,2]
yerlyana@bundamulia.ac.id*, msaputra@bundamulia.ac.id*

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *The integration of generative artificial intelligence (AI) into visual communication design education presents opportunities to enhance creativity and usability in learning. This study compares the effectiveness of three generative AI tools (MidJourney, DALL·E, and Adobe Firefly) in supporting students' creative outcomes and perceived usability, while also examining their broader pedagogical role in design education. A quasi-experimental design was conducted with 30 undergraduate students who each produced two poster designs: one manually and one with AI assistance. Creativity was evaluated using the Consensual Assessment Technique (CAT) by expert judges, and usability was measured using the System Usability Scale (SUS). Results showed that AI-assisted designs achieved significantly higher creativity scores (M = 4.3 vs. 3.2, p < 0.05) and usability ratings (range M = 74–82) compared to manual designs, with MidJourney rated highest in creativity and Adobe Firefly in usability. These findings provide empirical evidence that generative AI can act as a catalyst for creativity and usability in design education, offering theoretical insights into human–AI co-creation and practical implications for curriculum integration. Limitations include the small sample size and the study's focus on a single academic program, which may affect generalizability.*

Keywords: CAT-SUS, Creativity Assessment, Education, Generative AI, Usability.

Intisari— *Integrasi kecerdasan buatan generatif (AI) ke dalam pendidikan desain komunikasi visual membuka peluang untuk meningkatkan kreativitas dan kegunaan dalam proses pembelajaran. Penelitian ini membandingkan efektivitas tiga alat AI generatif (MidJourney, DALL·E, dan Adobe Firefly) dalam mendukung hasil kreativitas mahasiswa dan persepsi kegunaan, sekaligus menelaah peran pedagogisnya yang lebih luas dalam pendidikan desain. Desain kuasi-eksperimental dilakukan dengan melibatkan 30 mahasiswa sarjana yang masing-masing menghasilkan dua desain poster: satu secara manual dan satu dengan bantuan AI. Kreativitas dievaluasi menggunakan Consensual Assessment Technique (CAT) oleh panel ahli, sedangkan kegunaan diukur dengan System Usability Scale (SUS). Hasil penelitian menunjukkan bahwa desain berbasis AI memperoleh skor kreativitas yang lebih tinggi secara signifikan (M = 4,3 dibanding 3,2; p < 0,05) serta penilaian kegunaan yang lebih baik (rata-rata M = 74–82) dibandingkan desain manual. MidJourney mendapatkan skor tertinggi pada aspek kreativitas, sedangkan Adobe Firefly unggul dalam aspek kegunaan. Temuan ini memberikan bukti empiris bahwa AI generatif dapat berfungsi sebagai katalis kreativitas dan kegunaan dalam pendidikan desain, sekaligus menawarkan wawasan teoretis mengenai kolaborasi manusia–AI serta implikasi praktis bagi integrasi kurikulum. Keterbatasan penelitian ini mencakup jumlah sampel yang kecil dan konteks terbatas pada satu program studi, sehingga generalisasi hasil perlu dilakukan dengan hati-hati.*

Kata Kunci: CAT-SUS, Evaluasi Kreativitas, Generative AI, Pendidikan, Usability.

INTRODUCTION

The rapid development of generative artificial intelligence (AI) has significantly transformed creative industries and educational practices, offering opportunities to automate, accelerate, and expand human creativity [1], [2]. Generative AI refers to algorithms and models capable of producing new content such as images, audio, and text that closely resembles human-generated material [3], [4]. In visual communication design, these technologies enable students to create sophisticated outputs from simple prompts, shifting attention from manual execution to conceptual refinement. Popular tools such as Midjourney, DALL·E, and Adobe Firefly employ advanced diffusion or transformer-based models, producing high-quality images that have become widely adopted in both professional and educational contexts [5], [6], [7].

Visual communication design education emphasizes both conceptual thinking and technical execution [8]. Traditional methods such as sketching, prototyping, and digital rendering cultivate craftsmanship and critical thinking but are time-consuming and may limit rapid iteration. Generative AI offers a paradigm shift by allowing students to generate multiple design alternatives within minutes, thereby supporting contemporary educational goals that emphasize adaptive learning and creative problem-solving [9], [10], [11], [12]. However, challenges remain. Concerns have been raised regarding authenticity, originality, and ethical issues in AI-generated outputs [13], [14]. Students may also face usability barriers due to unfamiliar interfaces, inconsistent quality of outputs, and the need for precise prompts [15], [16], [17] Moreover, reliance on AI could foster creative dependency and reduce intrinsic ideation skills if not carefully integrated into curricula [1], [18]. These tensions highlight the need to better understand both the benefits and constraints of AI tools in design education.

Existing research on AI in education predominantly focuses on text-based applications such as automated essay grading, chatbots for learning support, and adaptive testing [19], [20], [21]. Empirical studies focusing on visual creativity and usability remain limited. Research employing the Consensual Assessment Technique (CAT) has demonstrated its value in evaluating originality, execution, and aesthetic quality [22], [23], while the System Usability Scale (SUS) has been widely used to measure user perceptions of system efficiency and satisfaction [24], [25], [26], [27]. Yet, few studies have conducted comparative analyses of multiple AI tools in design education, leaving an important research gap. Midjourney, DALL·E, and Adobe Firefly differ substantially in interface design, quality of outputs, and integration within creative workflows [28], [29], [30]. These variations suggest that students' experiences may diverge across platforms, affecting both creativity and usability. By systematically examining these differences, educators can make informed decisions about tool adoption and curriculum design.

This study addresses these gaps through a comparative evaluation of the three tools in an Indonesian higher education setting. Thirty undergraduate visual communication design students were tasked with creating manual and AI-assisted poster designs. Creativity was assessed using CAT [31], and usability was measured with SUS [32]. Beyond comparing tool performance, this study contributes theoretically by applying a dual framework of creativity and usability assessment, offering a model for future AI-in-design research. The findings are expected to provide empirical evidence on how generative AI functions as a catalyst for creativity and usability in design education while informing curriculum development and AI literacy in Indonesia.

MATERIALS AND METHODS

This research adopted a quasi-experimental design to compare creative outcomes and perceived usability of generative artificial intelligence tools with manual design methods in visual communication education [11], [22], [33], [34]. The focus was to evaluate three widely used AI platforms, namely Midjourney, DALL·E, and Adobe Firefly, in producing poster designs based on a common theme. The quasi-experimental approach was selected because it allowed controlled comparisons between different design methods while maintaining ecological validity in an academic classroom setting [35]. Thirty undergraduate students enrolled in a visual communication design program at a private university in Indonesia participated in this study. Participants were selected using purposive sampling to ensure baseline competency: all had completed at least one foundational course in digital design covering composition, color theory, and typography [36]. While this approach ensured technical feasibility and comparable skill levels, it limits generalizability to broader populations beyond the study site, such as students from public universities or other disciplines.

The data for this study comprised two main components: creative outputs generated by participants and usability responses provided after interacting with each AI tool. Creative outputs consisted of posters designed both manually and using AI assistance. Usability data were gathered using a standardized questionnaire administered immediately after each AI design task. The study was conducted in a controlled laboratory setting equipped with identical laptops to avoid variations in performance that might influence results. Data collection consisted of two primary components: (1) creative outputs generated through manual and AI-assisted poster design, and (2) usability evaluations of the three AI tools. The research was conducted in a controlled laboratory environment using identical laptops to minimize hardware-related performance variations.

The procedure comprised two stages. In the first stage, participants manually designed an environmental-awareness poster using Adobe Photoshop or Illustrator within three hours. In the second stage, they completed three additional AI-assisted poster designs, one using each AI tool with one hour allocated per tool. Prior to the AI sessions, a 45-minute introductory workshop was conducted to familiarize participants with prompt formulation, image generation processes, and post-processing adjustments. Minor refinements to layout and typography were permitted after AI generation. Although this design enables direct comparison between manual and AI-assisted methods, the unequal task duration (3:1 ratio) constitutes a significant limitation that may have influenced creative performance outcomes. The time discrepancy should therefore be interpreted as a major constraint in the comparative analysis. Future replications are encouraged to standardize task duration or normalize creative workload across methods to improve internal validity.

Creative quality of the poster designs was evaluated using the Consensual Assessment Technique (CAT) [31]. Three expert judges participated in the evaluation. Each judge held at least a Master's degree in Visual Communication Design and had more than five years of academic and professional experience. The judges independently rated all poster outputs using three standardized criteria: originality, aesthetic value, and relevance to theme [34], [37]. Each criterion was evaluated on a five-point Likert scale ranging from one (very low) to five (very high). The detailed criteria are presented in Table 1. Before the formal evaluation, the panel participated in a calibration session to align their understanding of the scoring criteria and to review several sample works. Inter-rater reliability was later examined using Cronbach's alpha and produced a coefficient of 0.87, which indicates a high level of consistency among evaluators.

Table 1. Adapted from Amabile's Consensual Assessment Technique (CAT) for creativity evaluation in design education

Criterion	Description	Scale Type	Evaluator
Originality	Degree to which the poster presents unique concepts, innovative ideas, and avoids clichés	1–5 Likert	Expert Panel
Aesthetic Value	Overall visual appeal including composition, color harmony, typography, and balance	1–5 Likert	Expert Panel
Relevance	How well the poster aligns with the given theme (environmental awareness) and communicates message	1–5 Likert	Expert Panel

Source : (Research Results, 2025)

The usability of the three AI tools was assessed using an adapted version of the System Usability Scale (SUS) [32]. he instrument consisted of ten statements rated on a five-point Likert scale. The questions covered three major dimensions of usability, namely satisfaction, effectiveness, and efficiency. All items were contextually adapted to suit the evaluation of generative AI tools in visual communication design tasks, as shown in Table 2. After completing each AI-assisted design, participants completed the SUS questionnaire corresponding to the specific AI tool they had just used. Although the adaptation of the SUS followed previous usability studies on digital design tools [24], [25], formal re-validation for generative AI contexts remains a necessary step for future research.

Table 2. Adapted from SUS Instrument, applied to generative AI tools in visual communication design

ID	Questionnaire	Usability	Sentiment
Q1	I think that I would like to use this AI tool frequently when creating poster designs	Satisfaction	Positive

ID	Questionnaire	Usability	Sentiment
Q2	I find this AI tool to be more complicated than it should be	Satisfaction	Negative
Q3	I thought the AI tool was easy to use	Overall	Positive
Q4	I think that I would need the support of a technical person to be able to use this AI tool	Effectiveness	Negative
Q5	I found the various functions in this AI tool were well-integrated	Effectiveness	Positive
Q6	I thought there was too much inconsistency in this AI tool	Effectiveness	Negative
Q7	I would imagine that most people would learn to use this AI tool very quickly	Efficiency	Positive
Q8	I find this AI tool to be time-consuming	Efficiency	Negative
Q9	I felt very confident using this AI tool	Effectiveness	Positive
Q10	I needed to learn a lot of things before I could get going with this AI tool	Efficiency	Negative

Source : (Research Results, 2025)

Data were analyzed using SPSS statistical software. Descriptive statistics were calculated to summarize participant demographics, creativity scores, and usability ratings [38]. nferential analyses included paired-sample t-tests to compare manual and AI-assisted design outputs, and one-way analysis of variance (ANOVA) followed by Tukey's HSD post hoc tests to identify significant differences among the three AI tools [39], [40]. Effect sizes, expressed as Cohen's d and eta squared (η^2), were computed to assess the magnitude of differences. Figure 1 illustrates the overall research workflow, from manual design to AI-assisted design and subsequent evaluation.

All participants provided written informed consent before participation and were assured of confidentiality and voluntary involvement. The study was conducted in accordance with the institutional ethical policy of Universitas Bunda Mulia (UBM). The Research Ethics Committee of UBM issued a favorable opinion for this project under registration number **015/Dir.SDM&PEN.01/IV/2025**, certifying that the study met the required ethical standards and that the foreseeable risks for participants were justified. The research did not involve any medical, psychological, or invasive procedures, and no personally identifiable or sensitive information was collected. All data were anonymized prior to analysis, and participants retained the right to withdraw at any stage without penalty. These measures ensured full compliance with international ethical principles, including the Declaration of Helsinki and COPE guidelines.
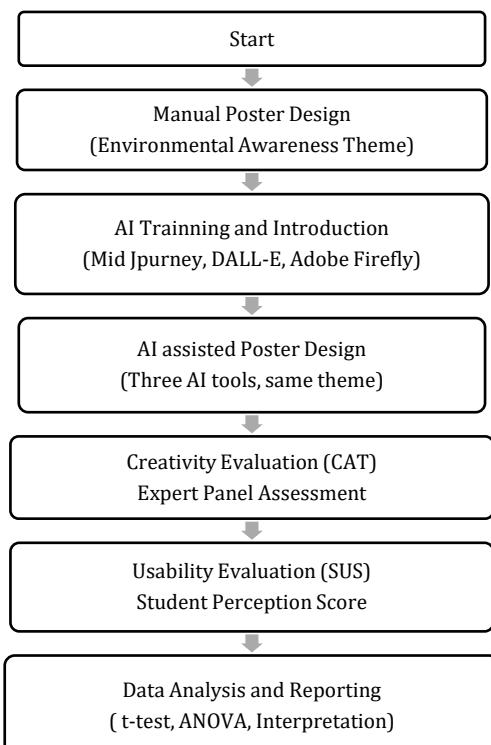
Source : (Research Results, 2025)
Figure 1. Procedure Flowchart

RESULTS AND DISCUSSION

Participant Demographics

Thirty undergraduate students from the visual communication design program participated in this study. Their demographic distribution is presented in Table 3. Most participants were female (60%) and aged between 21–23 years (66.7%). Only 30% reported prior experience using AI-based creative tools, indicating that the majority were first-time users of generative AI in design tasks.

Table 3. Demographic characteristics of participants (N = 30)

Variable	Category	n	%
Gender	Male	12	40
	Female	18	60
Age	18–20 years	10	33.3
	21–23 years	20	66.7
Prior AI Experience	Yes	10	33.3
	No	20	66.7

Source : (Research Results, 2025)
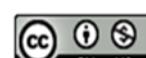
Creativity Evaluation Results

Creative quality of the poster designs was evaluated using the Consensual Assessment Technique (CAT), which considered originality, aesthetic value, and thematic relevance. Descriptive statistics are presented in Table 4. AI-assisted designs achieved higher creativity scores than manual designs, with manual posters receiving an average score of 3.2 (SD = 0.7, 95% CI [2.9, 3.5]) and AI-assisted designs producing substantially higher mean values: Midjourney (M = 4.5, SD = 0.4, 95% CI [4.3, 4.6]), Adobe Firefly (M = 4.3, SD = 0.5, 95% CI [4.1, 4.5]), and DALL·E (M = 4.1, SD = 0.6, 95% CI [3.9, 4.4]).

Table 4. Creativity scores of manual and AI-assisted designs (N = 30)

Design Method	M	SD	95% CI (Lower–Upper)
Manual	3.2	0.7	2.9 – 3.5
Midjourney	4.5	0.4	4.3 – 4.6
Adobe Firefly	4.3	0.5	4.1 – 4.5
DALL·E	4.1	0.6	3.9 – 4.4

Source : (Research Results, 2025)
Note: Creativity assessed using CAT; higher scores indicate greater originality, aesthetic value, and relevance

Inferential analysis confirmed these patterns. A paired-sample t-test showed that AI-assisted outputs were statistically more creative than manual outputs, t(29) = 6.74, p < 0.001, with a large effect size (Cohen's d = 1.23). One-way ANOVA revealed a significant effect of AI tool type, $F(2, 87)$ = 5.87, p < 0.01, partial η^2 = 0.12, indicating a medium effect. Post hoc Tukey tests showed that Midjourney produced significantly higher creativity scores than DALL·E (mean difference = 0.4, p < 0.01), while the difference between Midjourney and Adobe Firefly was not significant. The summary of effect sizes for these comparisons is presented in Table 5.

Table 5. Effect sizes for key statistical comparisons (Creativity-related rows only)

Comparison	Statistic	Effect Size
AI-assisted vs Manual (Creativity)	$t(29)$ = 6.74	Cohen's d = 1.23 (large)
Midjourney vs DALL·E (Creativity)	Tukey HSD	Mean diff = 0.4, p < .01
Midjourney vs Firefly (Creativity)	Tukey HSD	ns (not significant)
AI Tools (Creativity, ANOVA)	$F(2, 87)$ = 5.87	η^2 = 0.12 (medium)

Source : (Research Results, 2025)
Note: Cohen's d and η^2 indicate magnitude of differences. ns = not significant.

Table 5 presents the magnitude of the differences across statistical comparisons. The contrast between AI-assisted and manual designs produced a large effect, while the differences among AI tools were of medium magnitude. These results suggest that although AI-assisted outputs generally performed better, the strength of the differences varied depending on the platform. Creativity-score distributions with 95 percent confidence intervals are illustrated in Figure 2.
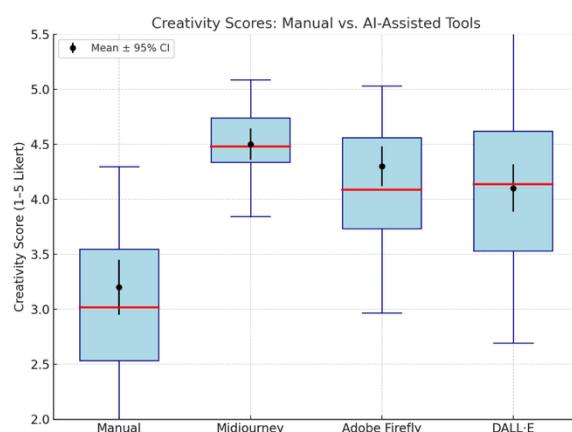
Source : (Research Results, 2025)
Figure 2. *Creativity scores comparison between manual and AI-assisted designs (with boxplots and 95% CI error bars*

Usability Evaluation Results

Usability perceptions of the three AI tools were assessed using the adapted System Usability Scale (SUS). Descriptive statistics are displayed in **Table 6**, showing that Adobe Firefly achieved the highest usability score (M = 82, SD = 6.3, 95 percent CI [80, 85]), followed by Midjourney (M = 78, SD = 7.1, 95 percent CI [76, 81]) and DALL·E (M = 74, SD = 8.4, 95 percent CI [71, 77]).

Table 6. Usability scores of AI-assisted tools (N = 30)

AI Tool	M	SD	95% CI (Lower–Upper)
Midjourney	78	7.1	76 – 81
Adobe Firefly	82	6.3	80 – 85
DALL·E	74	8.4	71 – 77

Source : (Research Results, 2025)
Note: Usability measured using adapted SUS; higher scores indicate greater ease of use, efficiency, and satisfaction.

A one-way ANOVA revealed significant differences in usability among the three tools, F(2,

87) = 4.62, p = 0.013, partial η^2 = 0.10, representing a medium effect. Post hoc Tukey tests showed that Adobe Firefly was rated significantly higher than DALL·E (mean difference = 8, p < 0.05), while the difference between Firefly and Midjourney was not statistically significant (p = 0.12). The magnitude of these effects is summarized in Table 7 (usability-related comparisons). As shown, Adobe Firefly's advantage over DALL·E was meaningful, whereas its difference from Midjourney did not reach significance.

Table 7. Effect sizes for key statistical comparisons (Usability-related rows only)

Comparison	Statistic	Effect Size
AI Tools (Usability, ANOVA)	$F(2, 87)$ = 4.62	η^2 = 0.10 (medium)
Firefly vs DALL·E (Usability)	Tukey HSD	Mean diff = 8, $p < .05$
Firefly vs Midjourney (Usability)	Tukey HSD	ns (not significant)

Source : (Research Results, 2025)
Note: η^2 indicates effect size magnitude. ns = not significant.

These findings are visualized in Figure 3, which presents boxplots with error bars representing 95 percent confidence intervals for each tool. The figure illustrates that Adobe Firefly achieved the highest usability scores with relatively low variability, Midjourney obtained moderately high scores with consistent results, and DALL·E displayed the lowest mean usability with the widest response spread, indicating more inconsistent user experiences.
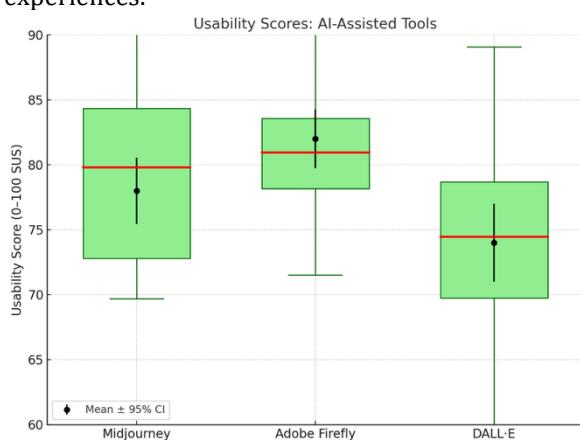
Source : (Research Results, 2025)
Figure 3. Usability scores of Midjourney, Adobe Firefly, and DALL·E with boxplots and 95% CI error bars. Adobe Firefly showed the highest ratings, significantly above DALL·E.

Inter-Rater Reliability

The reliability of expert evaluations using the CAT method was analyzed using Cronbach's alpha. The resulting value was 0.87, which is higher than the commonly accepted threshold of 0.70 for creativity assessments [41]. This result indicates high consistency among the evaluators and supports the validity of the creativity scores obtained in this study.

Discussion

The findings demonstrate that generative AI tools can meaningfully support creative performance and learning engagement in visual communication design education. Students' AI-assisted poster designs received higher ratings in originality, aesthetic quality, and thematic relevance than manually created designs. These results align with previous studies showing that generative AI systems can enhance creative ideation by providing rapid and diverse visual alternatives that stimulate conceptual exploration [42], [43], [44]. The strong performance of Midjourney may be attributed to its advanced diffusion-based architecture, which enables the production of nuanced and stylized imagery that effectively translates abstract ideas into visual form [1], [11], [45].

Nevertheless, this outcome must be interpreted with caution. The unequal task duration, with three hours for manual work and one hour per AI-assisted design, poses a significant internal validity threat that may have amplified the observed creativity difference. This imbalance limits the strength of causal inference regarding whether AI directly enhances creativity. It is therefore more accurate to interpret AI as a context-dependent facilitator rather than a universal catalyst for creativity. The comparative performance among the three AI tools, Midjourney, Adobe Firefly, and DALL·E, provides a more reliable basis for understanding AI's educational role than the contrast with the manual condition.

Differences among the AI tools highlight the importance of balancing creativity and usability when selecting technologies for educational use. Midjourney achieved the highest creativity scores, while Adobe Firefly received the highest usability ratings. This pattern reflects a pedagogical trade-off. Tools that produce more novel and stylized outputs may present usability challenges, whereas those integrated into familiar ecosystems such as Adobe Creative Cloud can lower cognitive barriers and improve user experience but may yield less distinctive creative outcomes [7], [46], [47], [48],
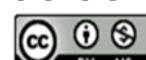
[49]. DALL·E, although accessible and versatile, scored lowest in both creativity and usability, possibly due to limited post-generation editing options and interface constraints.

Demographic factors, while not analyzed statistically due to the small sample size, could also moderate these relationships. Prior experience with AI, reported by 30 percent of participants, may have shaped usability perceptions, confidence, and learning outcomes. Differences in age or gender could further influence how students engage with AI-assisted creativity, as indicated in related studies on digital learning adoption [41], [42]. Future research with larger and more diverse samples should test these potential moderating variables to identify which learner groups benefit most from AI integration.

Several methodological limitations must be acknowledged. The small sample size of thirty participants, the single-institution context, and purposive sampling reduce generalizability. Although inter-rater reliability for CAT was high (α = 0.87), the assessment remains partly subjective. Moreover, the adaptation of the SUS instrument was not revalidated specifically for AI-assisted design tools. The controlled laboratory environment, though beneficial for internal control, differs from real classroom settings where resources, devices, and time constraints vary. These limitations reinforce the need for cautious interpretation and encourage replication with standardized task durations and broader participant profiles.

The discussion also requires acknowledgment of key limitations. The small sample size (n = 30), the single-institution context, and purposive sampling reduce generalizability. The reliance on CAT, while supported by high inter-rater reliability (α = 0.87), remains subjective, and the adaptation of the SUS instrument was not independently validated for AI tools. In addition, the controlled laboratory environment using identical laptops differs from real classroom contexts where resources and conditions vary. These constraints underscore the need for cautious interpretation and further replication before drawing broader conclusions.

Despite these constraints, the study provides several pedagogical implications. Integrating generative AI in design curricula can shift the learning emphasis from manual execution to conceptual evaluation, allowing students to focus on decision-making, iteration, and critique rather than production speed [18], [22], [45]. Prompt engineering also emerges as a new form of creative literacy, as the quality of AI-generated outcomes depends on how students articulate design intent and refine prompts [9], [11], [15]. Structured discussions about ethics, originality, and authorship are essential to maintain academic integrity and critical awareness in AI-mediated learning environments [13], [14], [47]. Balancing AI use with manual skill development is equally important to prevent overreliance on automation and to ensure that students continue building foundational creative competencies such as sketching, composition, and ideation [43], [48], [49].

Theoretically, this research contributes to the growing body of literature that combines creativity assessment (CAT) and usability measurement (SUS) within a dual-framework model [31], [32]. Practically, it provides evidence-based insights for educators and institutions seeking to integrate AI tools responsibly in creative disciplines. By comparing Midjourney, Adobe Firefly, and DALL·E under controlled conditions, the study demonstrates that different AI tools serve different pedagogical functions. Midjourney functions as a generator of high-concept visual ideas, Firefly as a user-friendly integrator, and DALL·E as an accessible but less specialized platform. Consequently, AI should be understood not as an inherent catalyst of creativity, but as a situational enabler whose educational value depends on context, task design, and reflective guidance from instructors [9], [12], [18], [42].

CONCLUSION

This study provides empirical evidence that generative AI tools have the potential to enrich design education by enhancing both creativity and usability compared to manual methods. By applying a dual evaluation framework that integrates the Consensual Assessment Technique (CAT) and the System Usability Scale (SUS), the research extends existing models of evaluation in design education and demonstrates how creativity and usability can be examined simultaneously in AI-assisted learning. The comparative results highlight that Midjourney performed best in terms of creative output, while Adobe Firefly received the highest usability ratings, indicating that different AI platforms offer distinct contributions to the educational context.

These findings must, however, be interpreted with caution. The study involved a relatively small sample of 30 students from a single private university in Indonesia, and the design task was limited to poster creation. Purposive sampling and unequal task durations may also have influenced

the results, while creativity evaluation through CAT remains partially subjective despite strong inter-rater reliability. These limitations restrict the generalizability of the findings and suggest that the conclusions should not be applied beyond the specific context of this study.

Despite these constraints, the study offers several practical implications for curriculum design. Educators are encouraged to introduce prompt engineering as a formal skill, to use different AI tools according to their strengths, such as Midjourney for creative exploration and Firefly for usability-focused workflows, to embed explicit discussion of authorship, originality, and ethics into classroom activities, and to maintain a balance between AI-assisted and manual design tasks so that students continue to develop core creative competencies.

Future research should build on this study by involving larger and more diverse samples across multiple universities and cultural contexts. Longitudinal studies are needed to examine how sustained use of AI affects students' manual design skills, originality, and critical thinking over time. Methodological refinement could include the use of mixed-methods approaches that combine quantitative measures with qualitative insights into students' creative processes. Additional investigations should also explore how demographic factors such as age, gender, and prior AI experience influence outcomes, and should test a broader range of design tasks beyond poster creation.

In conclusion, this study contributes to theory by demonstrating that CAT and SUS can be combined as a framework for evaluating creativity and usability in AI-assisted design education. It also contributes to practice by providing evidence that AI can be adopted in a context-sensitive manner to complement, rather than replace, manual skills. Careful integration of generative AI into design curricula will help ensure that students acquire both emerging AI literacy and enduring creative competencies.

REFERENCE

[1] Y. Liang, "Balancing: The Effects of AI Tools in Educational Context," *Frontiers in Humanities and Social Sciences*, vol. 3, no. 8, p. 7, 2023, doi: 10.54691/fhss.v3i8.5531.

[2] G. Franceschelli and M. Musolesi, "DeepCreativity: measuring creativity with deep learning techniques," *Intelligenza Artificiale: The international journal of the* *AIxIA*, vol. 16, no. 2, pp. 151–163, Dec. 2022, doi: 10.3233/IA-220136.

[3] V. Kujore, A. Adebayo, O. Sambakiu, and B. S. Segbenu, "Transformative role of generative AI in marketing content creation and brand engagement strategies," *GSC Advanced Research and Reviews*, vol. 23, no. 3, p. 1, 2025, doi: 10.30574/gscarr.2025.23.3.0145.

[4] Y. Li, "Advancing Design With Generative AI: A Case of Automotive Design Process Transformation," in *Proceedings of DRS*, 2024. doi: 10.21606/drs.2024.1260.

[5] D. Hanna, "The Use of Artificial Intelligence Art Generator 'Midjourney' in Artistic and Advertising Creativity," *Journal of Design Sciences and Applied Arts*, vol. 4, no. 2, p. 42, 2023, doi: 10.21608/jdsaa.2023.169144.1231.

[6] L. Tan and M. Luhrs, "Using Generative AI Midjourney to enhance divergent and convergent thinking in an architect's creative design process," *The Design Journal*, vol. 27, no. 4, p. 677, 2024, doi: 10.1080/14606925.2024.2353479.

[7] L. Tian, Q. Wang, B. Zhang, and L. Bo, "EMO: Emote Portrait Alive – Generating Expressive Portrait Videos with Audio2Video Diffusion Model under Weak Conditions," *ArXiv*, 2024, doi: https://doi.org/10.48550/arXiv.2402.17485.

[8] M. Chandra and Y. Erlyana, "Pilah sampah melalui media edukasi kreatif: pendekatan design thinking," *Jurnal Desain*, vol. 12, no. 3, p. 811, Jul. 2025, doi: 10.30998/jd.v12i3.24282.

[9] Y.-M. Fang, "The role of generative AI in industrial design: enhancing the design process and education," in *IET Conference Proceedings*, 2024, p. 135. doi: 10.1049/icp.2024.0303.

[10] M. L. Haryanti and R. Purbojo, "Peran Kepemimpinan Transformasional dan Literasi Digital terhadap Kompetensi Pedagogis Guru dalam Pengintegrasian Teknologi," *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 1, pp. 103–112, Jun. 2024, doi: 10.29408/edumatic.v8i1.25444.

[11] P. Cotroneo and J. Hutson, "Generative AI tools in art education: Exploring prompt engineering and iterative processes for enhanced creativity," *Metaverse*, vol. 4, no. 1, p. 14, 2023, doi: 10.54517/m.v4i1.2164.

[12] D. Satria, "Implementasi User Centered Requirements Engineering dalam Pencarian
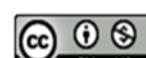
Kebutuhan Sistem eLearning," *Journal of Information System Research (JOSH)*, vol. 3, no. 4, pp. 437–445, Jul. 2022, doi: 10.47065/josh.v3i4.1618.

[13] K. A. Bartlett and J. D. Camba, "Generative Artificial Intelligence in Product Design Education: Navigating Concerns of Originality and Ethics," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 5, p. 55, 2024, doi: 10.9781/ijimai.2024.02.006.

[14] J. Liang, "The application of artificial intelligence-assisted technology in cultural and creative product design," *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-82281-2.

[15] M. Abdaljaleel *et al.*, "A multinational study on the factors influencing university students' attitudes and usage of ChatGPT," *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-52549-8.

[16] Y. Choi, "Integrating ChatGPT into the Design of 5E-Based Earth Science Lessons," *Educ. Sci. (Basel).*, vol. 15, no. 7, p. 815, 2025, doi: 10.3390/educsci15070815.

[17] M. G. Saputra and I. R. Andelina, "Analysis of Interface Design in the Game Honkai Star Rail Using the Usability Heuristics Principles of UI," *Gameology and Multimedia Expert*, vol. 2, no. 1, pp. 7–16, Jan. 2025, doi: 10.29103/game.v2i1.19957.

[18] B. Jose, J. Cherian, A. M. Verghis, S. M. Varghise, S. Mumthas, and S. Joseph, "The cognitive paradox of AI in education: between enhancement and erosion," *Front. Psychol.*, vol. 16, 2025, doi: 10.3389/fpsyg.2025.1550621.

[19] M. I. Baig and E. Yadegaridehkordi, "Factors influencing academic staff satisfaction and continuous usage of generative artificial intelligence (GenAI) in higher education," *International Journal of Educational Technology in Higher Education*, vol. 22, no. 1, 2025, doi: 10.1186/s41239-025-00506-4.

[20] S. Chardonnens, "Adapting educational practices for Generation Z: integrating metacognitive strategies and artificial intelligence," *Front. Educ. (Lausanne).*, vol. 10, 2025, doi: 10.3389/feduc.2025.1504726.

[21] R. Chugh *et al.*, "The Promise and Pitfalls: A Literature Review of Generative Artificial Intelligence as a Learning Assistant in ICT Education," *Computer Applications in Engineering Education*, vol. 33, no. 2, 2025, doi: 10.1002/cae.70002.

[22] J. Kim and H. Kim, "Generative AI in creative workflows: A case study on Midjourney and insights for future image tools," *Edelweiss Applied Science and Technology*, vol. 9, no. 4, p. 54, 2025, doi: 10.55214/25768484.v9i4.5936.

[23] R. Marrone, V. Taddeo, and G. Hill, "Creativity and Artificial Intelligence—A Student Perspective," *J. Intell.*, vol. 10, no. 3, p. 65, 2022, doi: 10.3390/jintelligence10030065.

[24] J. Fernandes Andry, M. Clara, W. Chandra, M. Antonio, and D. Yurisca Bernanda, "User Interface Analysis of PeduliLindungi Application to Improve User Experience with The Heuristic Evaluation Method," *Journal Research and Development (ITJRD)*, vol. 8, no. 2, 2024, doi: 10.25299/itjrd.2024.12295.

[25] J. Wu and J. Dong, "Impact of Augmented Reality Characteristics on Retail Brand Equity," 2022. doi: 10.2991/assehr.k.220107.042.

[26] N. Hiu and Y. Erlyana, "Redesigning User Interface of Datascripmall Mobile Apps Using User Centered Design Method," *Teknika*, vol. 13, no. 2, pp. 283–292, Jul. 2024, doi: 10.34148/teknika.v13i2.854.

[27] M. Gomez-Hernandez, X. Ferre, C. Moral, and E. Villalba-Mora, "Design Guidelines of Mobile Apps for Older Adults: Systematic Review and Thematic Analysis," *JMIR Mhealth Uhealth*, vol. 11, p. e43186, Sep. 2023, doi: 10.2196/43186.

[28] M. Basalla, J. Schneider, and J. vom Brocke, "Creativity of Deep Learning: Conceptualization and Assessment," in *Proceedings of the 14th International Conference on Agents and Artificial Intelligence*, 2022. doi: 10.5220/0010783500003116.

[29] J.-F. Chen, C.-C. Ni, P.-H. Lin, and R. Lin, "Designing the Future: A Case Study on Human-AI Co-Innovation," *Creat. Educ.*, vol. 15, no. 3, p. 474, 2024, doi: 10.4236/ce.2024.153028.

[30] Y. Meron, "Graphic design and artificial intelligence: Interdisciplinary challenges for designers in the search for research collaboration," in *Proceedings of DRS*, 2022. doi: 10.21606/drs.2022.157.

[31] P. Barth and G. Stadtmann, "Creativity Assessment over Time: Examining the

Reliability of CAT Ratings," *J. Creat. Behav.*, vol. 55, no. 2, pp. 396–409, Jun. 2021, doi: 10.1002/jocb.462.

[32] N. Pratama and R. Anrahvi, "Application of the System Usability Scale (SUS) Method in Measuring Student Satisfaction with the Academic Directory Website," *IJBEM: Indonesian Journal of Business Economics and Management*, vol. 3, no. 2, pp. 74–80, Jun. 2024, doi: 10.57152/ijbem.v3i2.2024.

[33] S. B. Kim and D. Oh, "Evaluating Creativity: Can LLMs Be Good Evaluators in Creative Writing Tasks?," *Applied Sciences*, vol. 15, no. 6, p. 2971, 2025, doi: 10.3390/app15062971.

[34] L. Hu, M. Boggess, and M. M. Shepley, "Comparing Expert, Quasi-Expert, and Novice Evaluations of Award-Winning Design Products Using the Consensual Assessment Technique," *Creat. Res. J.*, vol. 35, no. 4, pp. 730–748, Oct. 2023, doi: 10.1080/10400419.2022.2157591.

[35] B. Capili and J. K. Anastasi, "An Introduction to Types of Quasi-Experimental Designs," *American Journal of Nursing*, vol. 124, no. 11, pp. 50–52, Nov. 2024, doi: 10.1097/01.NAJ.0001081740.74815.20.

[36] Y. Erlyana and Ressiani, *Basic of Packaging: Belajar Kemas Kemasan*, 1st ed. Yogyakarta: Graha Ilmu, 2020.

[37] L. J. Ball and F. Vallée-Tourangeau, *The Routledge International Handbook of Creative Cognition*. London: Routledge, 2023. doi: 10.4324/9781003009351.

[38] R. B. Ohnson and L. B. Christensen, *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, 7th ed. California: SAGE Publications, Inc, 2020.

[39] A. Garg, K. Nisumba Soodhani, and R. Rajendran, "Enhancing data analysis and programming skills through structured prompt training: The impact of generative AI in engineering education," *Computers and Education: Artificial Intelligence*, vol. 8, p. 100380, Jun. 2025, doi: 10.1016/j.caeai.2025.100380.

[40] A. Nanda, Dr. B. B. Mohapatra, A. P. K. Mahapatra, A. P. K. Mahapatra, and A. P. K. Mahapatra, "Multiple comparison test by Tukey's honestly significant difference (HSD): Do the confident level control type I error," *International Journal of Statistics and Applied Mathematics*, vol. 6, no. 1, pp. 59–65,

Jan. 2021, doi: 10.22271/maths.2021.v6.i1a.636.

[41] P. Barth and G. Stadtmann, "Creativity Assessment over Time: Examining the Reliability of CAT Ratings," *Journal of Creative Behavior*, vol. 55, no. 2, pp. 396–409, Jun. 2021, doi: 10.1002/jocb.462.

[42] S. Guo *et al.*, "Vinci: An Intelligent Graphic Design System for Generating Advertising Posters," *Proceedings of CHI*, vol. 1, 2021, doi: 10.1145/3411764.3445117.

[43] J. Hartmann, Y. Exner, and S. Domdey, "The power of generative marketing: Can generative AI create superhuman visual marketing content?," *International Journal of Research in Marketing*, 2024, doi: 10.1016/j.ijresmar.2024.09.002.

[44] J. Lin *et al.*, "AutoPoster: A Highly Automatic and Content-aware Design System for Advertising Poster Generation," in *Proceedings of CHI*, 2023, p. 1250. doi: 10.1145/3581783.3611930.

[45] S. Chompunuch and T. Lubart, "AI as a Helper: Leveraging Generative AI Tools Across Common Parts of the Creative Process," *J. Intell.*, vol. 13, no. 5, p. 57, 2025, doi: 10.3390/jintelligence13050057.

[46] A. Ansone, Z. Zālīte-Supe, and L. Daniela, "Generative Artificial Intelligence as a Catalyst for Change in Higher Education Art Study Programs," *Computers*, vol. 14, no. 4, p. 154, 2025, doi: 10.3390/computers14040154.

[47] Q. Jiang, Y. Zhang, W. Wei, and C. Gu, "Evaluating Technological and Instructional Factors Influencing the Acceptance of AIGC-Assisted Design Courses," *Computers and Education Artificial Intelligence*, vol. 7, p. 100287, 2024, doi: 10.1016/j.caeai.2024.100287.

[48] M. U. Kahraman, Y. Şekerci, M. Develier, and F. Koyuncu, "Integrating Artificial Intelligence in Interior Design Education: Concept Development," *Journal of Computational Design*, vol. 5, no. 1, p. 31, 2024, doi: 10.53710/jcode.1418783.

[49] K. Seo, J. Tang, I. Roll, S. Fels, and D. Yoon, "The impact of artificial intelligence on learner–instructor interaction in online learning," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, 2021, doi: 10.1186/s41239-021-00292-9.