

A Comparative Study on Visual Aesthetic Quality and Originality of Children's Book Illustrations: Midjourney Generation vs. Manual Illustrators

1st Moch Bagoes Pakarti, 2nd Tigus Juni Betri

^{1,2}Informatika, Universitas Islam Negeri Raden Mas Said Surakarta

¹moch.bagoes@staff.uinsaid.ac.id, ²tigusjuni.betri@staff.uinsaid.ac.id

Abstract—The emergence of Generative AI, such as Midjourney, has sparked significant disruption in the children's book publishing industry by offering efficiency and rapid creation of visual works, while simultaneously raising fundamental debates regarding aesthetic quality and authenticity. This study aims to comparatively evaluate the visual aesthetic quality and originality levels between Midjourney-generated illustrations and those created by manual illustrators. Utilizing an explanatory sequential mixed-methods approach, data were collected through a structured evaluation rubric assessed by a panel of 15 experts (visual communication design academics, children's book editors, professional illustrators, and child psychologists). Quantitative data were analyzed using the Mann-Whitney U test, complemented by descriptive qualitative analysis. Findings reveal that Midjourney excels in technical rendering precision and execution speed, yet exhibits critical flaws in character consistency across story pages and narrative emotional expression. Conversely, manual illustrators scored significantly higher in originality, signature visual style, and narrative visual sensitivity appropriate for children's emotional development. This study concludes that AI cannot currently replace the intuitive role of manual illustrators in producing complete, impactful visual narratives for children's books. The findings imply that publishers should position AI merely as an initial ideation tool while underscoring the necessity of preserving the originality and ethical value of human illustrators' work. Please write in English. This document gives formatting instructions for authors preparing papers for publication in the IJCIS. The authors must follow the instructions given in the document for the papers to be published. You can use this document as both an instruction set and as a template into which you can type your own text. Abstract should be no longer than 250 words. It gives a brief summary of the content of the paper and point out the main objective, the methods employed, the results obtained and major conclusions.

Keywords : Children's Book Illustration, Manual Illustration, Midjourney, Visual Aesthetics, Originality.

I. INTRODUCTION

1.1 Background

The rapid advancement of Generative Artificial Intelligence (Gen-AI) in recent years has fundamentally altered the landscape of the creative industry, particularly in the realm of digital visual arts. One of the most prominent text-to-image platforms, Midjourney, leverages deep learning architectures and diffusion models to process simple text prompts into highly detailed, richly textured, and aesthetically sophisticated visual art within seconds [1]. This speed and accessibility massively lower the technical barriers of visual art production while democratizing image creation for independent publishers, authors, and designers (Epstein et al., 2023). This technological acceleration offers unprecedented time and cost efficiency, making it an increasingly attractive option to displace traditional design workflows.

Conversely, illustrations in children's books play a role far more complex than mere decorative elements accompanying text. Children's illustrations serve as crucial pedagogical and psychological instruments; they act as primary narrative conduits, spatial imagination builders, and triggers for emotional perception and empathy in early readers [2]. Through visual-text interplay, young children learn to decipher expressions, identify character emotions, and intuitively comprehend story arcs before they can read text independently [3]. Consequently, the quality criteria for children's illustrations are measured not only by superficial visual beauty, but also by emotional sensitivity, clarity of visual communication, and educational value.

The intersection of instant Generative AI efficiency and the complex functions of children's book illustrations triggers a

critical debate regarding the boundaries of originality, technical consistency, and the presence of "human touch" in artwork. Although Midjourney produces visually striking images, algorithmic synthesis frequently falls into generic visual patterns that threaten the uniqueness of an artist's signature style [4]. Furthermore, children's book illustrations demand strict character consistency and narrative continuity across pages—a major technical vulnerability in generative AI algorithms, which tend to yield randomized variations across prompt iterations [5]. The inability of AI to deeply comprehend story context risks generating "cold" and soulless illustrations. Given this dilemma between technological efficiency and human aesthetic-originality values, a comparative study evaluating Visual Aesthetic Quality and Originality between Midjourney generations and manual illustrators is urgently needed.

1.2 Research Questions

1. What are the differences in visual aesthetic quality (comprising composition, coloring, visual narrative, and character consistency) between children's book illustrations generated by Midjourney and those created by manual illustrators?
2. What are the levels of originality and visual style uniqueness in children's book illustrations produced by Midjourney compared to those by manual illustrators?

1.3 Research Objectives

To objectively analyze and compare the visual aesthetic quality and originality levels of both illustration mediums based on standard assessment rubrics and expert panel evaluations. Any compared to those by manual illustrators?

1.4 Research Significance

- Academic: Expands literature on the integration of Artificial Intelligence within Visual Communication Design (VCD) and child media studies.
- Practical: Provides objective guidelines for publishers, illustrators, and educators regarding the adoption and ethical boundaries of AI tools in children's media.

1.5 Principles of Children's Book Illustration

Illustrations in children's books serve a pedagogical and narrative function that transcends ornamentation. According to visual reading theory, children read pictures intuitively before mastering verbal literacy [2]. In this context, color psychology plays a vital role; vibrant colors and warm contrasts stimulate cognitive perception and foster emotional connections in early readers [6]. Visual storytelling relies on organizing image elements to guide a child's logical thinking without complete dependence on text. Another critical factor is character design consistency. Main characters must maintain uniform physical traits, proportions, and visual identities across every page turn to enable recognition and sustained emotional engagement [3].

1.6 Generative AI in Visual Arts (Midjourney)

Generative AI in the visual arts operates via deep learning diffusion models trained on billions of text-image pairs [7]. Midjourney functions through prompt engineering, wherein users provide text instructions to guide automated image synthesis. Its primary strength lies in generating highly detailed, dramatically lit, and textured visuals within seconds [1]. Nevertheless, the platform exhibits fundamental limitations in translating narrative nuances and psychological depth. Because it relies on statistical probabilities and data reconstruction, Midjourney often struggles to grasp abstract story contexts or maintain consistent character forms across varying angles and expressions[5].

1.7 Parameters of Visual Aesthetics

The aesthetic quality of visual artwork is evaluated through art elements (line, shape, space, color, texture) and design principles (balance, proportion, rhythm, emphasis) [8]. In children's book illustration, aesthetic parameters are judged not merely by technical polish, but by how effectively these principles yield balanced, uncluttered compositions easily processed by a child's developing visual perception[9].

1.8 Concepts of Originality and Authenticity

In the digital age, originality has shifted from "pure lack of influence" to the transformative execution of new ideas. However, AI introduces ethical concerns regarding unauthorized data scraping of artists' works without consent, attribution, or compensation [10]. Human illustrators infuse their work with a signature style—an authentic imprint shaped by personal experience, emotional intuition, and artistic expression [4]. Conversely, AI outputs are statistical compilations from pre-existing datasets, rendering their authenticity questionable as they reproduce generic patterns lacking genuine artistic intent.

1.9 State of the Art

Prior studies have examined Generative AI's broader impact on creative industries, such as public perceptions of AI art attribution [5], creativity in prompt engineering [1], and computational aesthetic evaluation [4]. However, a literature gap remains regarding comparative evaluations specifically focused on children's book illustrations, where character consistency, educational narrative value, and child emotional sensitivity are paramount. The novelty of this study lies in its comparative experimental design measuring visual aesthetics and originality specifically tailored for children's picturebooks. objectively analyze and compare the visual aesthetic quality and originality levels

II. RESEARCH METHODS

1.1 Research Design

This study employs an explanatory sequential mixed-methods design. Quantitative data are used to statistically test score differences between the two mediums, while qualitative data explain the underlying artistic, narrative, and psychological rationales..

1.2 Objects and Samples

The study uses 2 sample groups of children's book illustrations based on a standardized short narrative script requiring 3 key sequential scenes with character and emotional consistency:

1. Group A (AI): Images generated via Midjourney v6 using optimized prompt engineering techniques.
2. Group B (Manual): Artwork created by professional/early-career manual illustrators based on the identical narrative brief.

1.3 Subjects (Expert Evaluation Panel)

Using purposive sampling, 15 expert panelists were selected across four domains:

- VCD / Fine Arts Academics (4): Experts in design principles, visual elements, and originality theory.
- Children's Book Editors / Publishers (4): Experts in industry standards, market viability, and story readability.
- Professional Illustrators (4): Experts in technical execution, signature style, and character consistency.
- Child Psychologists / Early Childhood Educators (3): Experts in emotional sensitivity and visual appropriateness for young children.

1.4 Evaluation Rubric (Matrix Instrument)

Variable 1: Visual Aesthetic Quality (X ₁).			
No	Indicator	Operational Criteria & Description	Rating Scale (1 - 5)
1.1	Technical Precision	Sharpness of line art, cleanliness of rendering details,	1 = Very Poor (Severe artifacts)

	& Rendering	smoothness of gradients, and absence of visual artifacts or glitches.	3 = Fairly Neat 5 = Highly Precise & Flawless
1.2	Composition & Visual Elements	Balance, spatial utilization, and clarity of focal points that facilitate visual reading for children.	1 = Cluttered/Unbalanced 3 = Fairly Balanced 5 = Harmonious & Well-Focused
1.3	Color Harmony & Psychology	Color palette cohesion, contrast levels, and alignment with children's emotional responses and age levels.	1 = Dull/Inappropriate 3 = Moderately Appealing 5 = Highly Engaging & Child-Friendly
1.4	Visual Storytelling & Narrative	Ability of the illustration to convey plot arcs, emotional expressions, and scene moods without relying entirely on text.	1 = Passive/Fails to Tell Story 3 = Moderately Communicative 5 = Highly Expressive & Narrative
1.5	Character Consistency	Uniformity of main character proportions, physical features, costumes, and style across different story scenes.	1 = Highly Inconsistent 3 = Moderately Consistent 5 = Perfectly Identical & Consistent

Variable 2: Originality and Authenticity (X₂)

No	Indicator	Operational Criteria & Description	Rating Scale (1 - 5)
2.1	Signature Style Uniqueness	Distinctiveness of visual style, free from generic or repetitive	1 = Highly Generic/Derivative 3 = Moderately Distinct

		algorithmic visual patterns.	5 = Highly Unique & Characterful
2.2	Conceptual Novelty & Creativity	Boldness of visual ideas, creative camera angles, and fresh presentation of narrative concepts.	1 = Highly Cliche 3 = Moderately Creative 5 = Highly Innovative & Fresh
2.3	Depth of Emotional Touch ("Human Touch")	Presence of artistic soul, emotional warmth, and intuitive gestures that evoke empathy in young readers.	1 = Rigid/Emotionally Cold 3 = Moderately Warm 5 = Deeply Moving & Empathetic

$$\text{Class Interval} = \frac{\text{Max Score} - \text{Min Score}}{\text{Categories}} = \frac{5 - 1}{5} = 0.80$$

Score Categorization:

- 1.00–1.80: Very Low;
1.81–2.60: Low;
2.61–3.40: Moderate;
3.41–4.20: High;
4.21–5.00: Very High

2.5 Data Collection

1. Experimental Production: Execution of optimized Midjourney prompts vs. manual illustration creation following the exact same script.
2. Blinded Expert Evaluation: Unlabeled artwork presented anonymously to the 15 expert panelists to avoid bias toward AI or human media.
3. In-Depth Interviews & Focus Group Discussions (FGD): Qualitative sessions to capture expert rationales behind numerical rubric ratings.

2.6 Data Analysis

- Quantitative Analysis: Non-parametric Mann-Whitney U Test ($\alpha = 0.05$) to evaluate statistical score differences between Midjourney and Manual groups.
- Qualitative Analysis: Visual Criticism Analysis through data reduction, coding (e.g., character consistency, emotional resonance), and thematic synthesis.

III.RESULT AND ANALYSIS**3.1 Quantitative Evaluation Summary**

Table 1. Summary of Expert Panel Evaluation Results (Likert Scale 1–5)

Variable	Evaluation Indicator	Midjourney (AI)	Manual Illustration	Difference (Δ)	Mann-Whitney U (p -value)					
Visual Aesthetics (X_1)	1.1 Technical Precision & Rendering	4.67	4.20	+0.47	$p < 0.05$	2.3	2.40	4.80	-2.40	$p < 0.001$
	1.2 Composition & Elements	4.00	4.33	-0.33	$p > 0.05$	Emotional Touch (<i>Human Touch</i>)				
	1.3 Color Harmony & Psychology	4.13	4.40	-0.27	$p > 0.05$	Mean X_2	2.87	4.53	-1.66	Significant
	1.4 Visual Storytelling & Narrative	3.20	4.60	-1.40	$p < 0.001$	3.2 Comparative Analysis of Visual Aesthetic Quality (X_1)				
	1.5 Character Consistency	2.13	4.73	-2.60	$p < 0.001$	3.2.1 Technical Precision vs. Narrative Limitations				
	Mean X_1	3.63	4.45	-0.82	Significant	Midjourney significantly outperformed manual work in Technical Precision & Rendering (AI: 4.67 vs. Manual: 4.20), demonstrating mastery over lighting, micro-textures, and crisp digital lines within seconds. However, AI performance dropped sharply in Visual Storytelling & Narrative (AI: 3.20 vs. Manual: 4.60). Expert editors noted that Midjourney's over-rendered visual details frequently create visual noise, distracting young readers from the primary story focus. Manual illustrators excel at establishing focal hierarchies that guide a child's eyes toward character expressions.				
Originality (X_2)	2.1 Signature Style Uniqueness	2.80	4.53	-1.73	$p < 0.001$	3.2.2 The Critical Flaw of Character Consistency				
	2.2 Conceptual Novelty	3.40	4.27	-0.87	$p < 0.05$	The most severe divergence appeared in Character Consistency (AI: 2.13 vs. Manual: 4.73). Despite using prompt parameter references (--cref), Midjourney repeatedly introduced random variations in facial structure, clothing details, and body proportions across different scenes. This inconsistency is fatal for children's books, where young readers rely on visual stability to build familiarity with characters.				
						3.3 Comparative Analysis of Originality and Authenticity (X_2)				
						3.3.1 Generic Visual Patterns vs. Signature Style				
						Manual illustrators dominated the Originality variable ($\bar{X} = 4.53$ vs. AI $\bar{X} = 2.87$). Panelists highlighted that Midjourney outputs exhibit generic visual patterns—homogenized lighting, digital brush styles, and eye renderings recognizable as "typical AI art." Manual illustrators displayed distinct signature styles, organic brush textures, and personal color choices.				
						3.3.2 Absence of Human Touch and Intuitive Emotion				
						The indicator for Depth of Emotional Touch revealed a large score gap ($\Delta = -2.40$). While AI can synthesize generic facial expressions (e.g., smiles, tears), it fails to capture micro-expressions—such as hesitant glances or subtle bodily warmth. Manual illustrators leverage human intuition and lived experience to infuse characters with empathy, vital for early childhood emotional development.				
						3.4 Industry and Ethical Implications				
						1. AI as an Ideation Tool: Midjourney is currently unsuitable as a standalone final generator for children's				

picturebooks due to narrative and consistency failures. However, it offers valuable efficiency during pre-production (moodboarding, color exploration, character ideation).

2. Preserving Human Illustrators' Value: Results demonstrate that AI cannot replace human artistic intuition and emotional storytelling. Over-reliance on AI by publishers risks degrading narrative and educational standards.
3. Copyright and Ethical Considerations: Low originality scores reinforce concerns over unlicensed data scraping, reinforcing the need for publishers to protect human illustrators' IP and fair labor.

VI. CONCLUSION

4.1 Conclusion

1. Visual Aesthetic Quality: Manual illustrators achieved significantly higher overall aesthetic quality ($\bar{X} = 4.45$) compared to Midjourney ($\bar{X} = 3.63$). While AI excels in technical rendering speed and surface polish, it suffers from critical limitations in visual storytelling and character consistency across pages.
2. Originality and Authenticity: Manual artwork demonstrated superior originality ($\bar{X} = 4.53$) over Midjourney ($\bar{X} = 2.87$). AI outputs remain constrained by generic visual patterns and lack the emotional depth, subtle human touch, and distinct signature styles essential for early childhood engagement.

In summary, Generative AI offers powerful technical efficiency but cannot replace the holistic, empathetic, and narrative capabilities of human illustrators in children's publishing.

4.2 LIMITATIONS

For Publishers: Avoid replacing human illustrators entirely with AI to preserve story quality, emotional resonance, and ethical standards.

- For Illustrators: Adopt AI tools collaboratively during early pre-production (ideation, moodboards) while retaining manual execution for final artwork.
- For Future Research: Conduct direct audience testing with young children using eye-tracking and behavioral observation to measure real-time emotional engagement across AI and human illustrations.

4.3 RECOMMENDATIONS

- For Publishers: Avoid replacing human illustrators entirely with AI to preserve story quality, emotional resonance, and ethical standards.
- For Illustrators: Adopt AI tools collaboratively during early pre-production (ideation, moodboards) while retaining manual execution for final artwork.
- For Future Research: Conduct direct audience testing with young children using eye-tracking and behavioral observation to measure real-time emotional engagement across AI and human illustrations.

THANK-YOU NOTE

Titles for thanks and references are not numbered. Thank you to the IJCIS Team for taking the time to create this template.

REFERENCES

- [1] J. Oppenlaender, *The Creativity of Text-to-Image Generation*. 2022. doi: 10.1145/3569219.3569352.
- [2] M. Nikolajeva and C. Scott, "How picturebooks work," *How Picturebooks Work*, pp. 1–293, Jan. 2013, doi: 10.4324/9780203960615.
- [3] M. Salisbury and M. Styles, *Children's Picturebooks: The Art of Visual Storytelling*. Laurence King Publishing, 2012. [Online]. Available: <https://books.google.co.id/books?id=QXpEXwAACAAJ>
- [4] E. Cetinic and J. She, "Understanding and Creating Art with AI: Review and Outlook," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 18, no. 2, Feb. 2022, doi: 10.1145/3475799.
- [5] O. Kaynar and M. Tuna, "Generative AI In Media, Communication and Visual Arts," 2025, pp. 295–325.
- [6] J. Albers, *Interaction of Color: 50th Anniversary Edition*, 50th anniversary. New Haven, CT: Yale University Press, 2022.
- [7] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, and M. Chen, *Hierarchical Text-Conditional Image Generation with CLIP Latents*. 2022. doi: 10.48550/arXiv.2204.06125.
- [8] D. A. Lauer and S. Pentak, *Design Basics*. Cengage Learning, 2011. [Online]. Available: <https://books.google.co.id/books?id=dW8hSkqO9oUC>
- [9] C. Malamed, *Visual Design Solutions: Principles and Creative Inspiration for Learning Professionals*. 2015. doi: 10.1002/9781119153801.
- [10] S. Cui, F. Liu, B. Li, and X. Yuan, *Sharing instead of being infringed: How to build a responsible community for online artwork sharing?* 2024. doi: 10.21606/drs.2024.417.