



# The Effect of Using Generative AI as Digital Scaffolding on Improving Textual Quality and Narrative Writing Creativity in Indonesian Language Learning in Elementary Schools

Putri Hana Pebriana

Universitas Pahlawan Tuanku Tambusai

Corresponding E-Mail: [putripebriana99@gmail.com](mailto:putripebriana99@gmail.com)

Received: 2025-10-23 Accepted: 2025-12-31

DOI: 10.24256/ideas.v13i2.9448

## Abstract

Narrative writing is a core component of Indonesian language learning in elementary schools; however, many students experience difficulties in organizing ideas, maintaining coherence, and developing creative storylines. Although digital scaffolding has been widely discussed, empirical evidence on AI-assisted scaffolding in elementary first-language (L1) writing contexts remains limited. This study examined the effect of using Generative AI as digital scaffolding on students' textual quality and narrative writing creativity. A quasi-experimental pretest-posttest control group design was employed involving 58 upper-grade elementary students (experimental = 29; control = 29). The experimental group received AI-assisted scaffolding across planning, drafting, and revising stages, while the control group received conventional teacher-led scaffolding. Data were collected through narrative writing tasks scored using validated analytic rubrics. Independent samples t-tests and ANCOVA (controlling for pretest scores) revealed that the experimental group achieved significantly higher posttest scores in textual quality ( $M = 82.41, SD = 6.32$ ) compared to the control group ( $M = 73.18, SD = 7.05$ ),  $p < .001$ , with a large effect size (Cohen's  $d = 1.36$ ). Similar significant differences were found in creativity scores ( $p < .001, d = 1.21$ ). Improvements were most evident in coherence, elaboration, and narrative resolution development. The findings suggest that Generative AI, when used as guided digital scaffolding under teacher supervision, can effectively enhance both structural quality and creative expression in elementary narrative writing. This study contributes empirical evidence to the growing field of AI-assisted writing pedagogy in primary education.

**Keywords:** *generative AI, digital scaffolding, narrative writing, textual quality, creativity, Indonesian language learning, elementary school.*

**Introduction**

Narrative writing is a fundamental component of Indonesian language learning in elementary schools because it helps students' express experiences, develop imagination, and practice communicating ideas in an organized written form. Through storytelling, children learn how to build meaning using language, connect events logically, and engage readers through characters, settings, and conflicts. However, developing narrative competence is not only about producing a story; it also involves controlling textual features such as coherence, cohesion, vocabulary choice, and writing conventions. For this reason, narrative writing instruction is often used as an indicator of students' broader literacy growth in the primary grades.

Despite its importance, many elementary students face persistent difficulties when composing narratives. Common challenges include generating original ideas, expanding limited storylines, maintaining consistent plot structure, and using appropriate diction and sentence patterns. Students may also struggle with mechanics such as spelling, punctuation, and paragraphing, which can weaken the readability of their texts. In classroom practice, teachers frequently encounter time constraints and large class sizes that make it difficult to provide intensive, individualized feedback during drafting and revision. As a result, writing activities may become product-focused rather than process-oriented, limiting opportunities for meaningful improvement.

One approach widely recommended to address these challenges is scaffolding, which refers to structured, temporary support that enables learners to perform tasks they could not accomplish independently. Rooted in sociocultural learning theory and the concept of the zone of proximal development, scaffolding is designed to gradually shift responsibility from teacher guidance to student autonomy. In writing instruction, scaffolding can appear as modelling, guided questioning, outlines, sentence starters, peer review routines, and revision checklists. When implemented effectively, it helps students plan, draft, revise, and reflect, thereby strengthening both the quality of their writing and their confidence as developing authors.

As classrooms increasingly incorporate digital technologies, the concept of scaffolding has expanded into digital scaffolding—support delivered through tools that can prompt, guide, and provide feedback during learning activities. Digital scaffolding may include interactive templates, writing prompts, digital story maps, and automated feedback systems that support students at key moments of difficulty. For writing tasks, digital scaffolding can be especially valuable because it can provide immediate cues and iterative guidance without requiring constant teacher intervention. However, the effectiveness of digital scaffolding depends on how well it aligns with instructional goals, student developmental levels, and ethical learning practices.

In recent years, Generative Artificial Intelligence (Generative AI) has emerged as a tool with the potential to transform digital scaffolding in writing instruction. Unlike static templates, Generative AI can respond dynamically to students' needs by

suggesting ideas, proposing story structures, offering vocabulary alternatives, and generating revision guidance based on user input. When positioned as scaffolding rather than a substitute writer, Generative AI may support students in overcoming “blank page” anxiety, expanding narrative details, and refining coherence and language accuracy. Its dialogic nature can also encourage students to ask questions, explore multiple story possibilities, and iteratively improve drafts—key elements of process-based writing pedagogy.

Nevertheless, integrating Generative AI into elementary writing instruction raises important pedagogical and ethical considerations. There is a risk that students may become overly dependent on AI outputs, reducing opportunities to practice independent thinking, problem-solving, and authentic expression. AI-generated text may contain inaccuracies, inappropriate content, or stylistic patterns that do not match students’ developmental stages. Concerns also include originality, authorship, and the possibility of students presenting AI-generated work as their own. For these reasons, the educational value of Generative AI is strongly influenced by teacher mediation, clear boundaries of use, and instructional designs that prioritize student agency and critical literacy.

Existing studies on AI-supported writing often emphasize older learners, second-language contexts, or higher education settings, with findings that are promising but not always consistent. Some research suggests that AI tools can improve fluency, vocabulary variety, and revision efficiency, while other work highlights risks related to shallow learning and reduced authenticity. In the context of Indonesian language learning at the elementary level, empirical evidence remains limited—particularly studies that examine both textual quality and creativity as distinct but related outcomes. Moreover, many discussions remain conceptual, leaving a need for classroom-based investigations that evaluate how AI scaffolding functions under real instructional conditions and teacher guidance.

Responding to this gap, the present study investigates the effect of using Generative AI as digital scaffolding on improving textual quality and narrative writing creativity in Indonesian language learning in elementary schools. The study conceptualizes textual quality through narrative structure, coherence and cohesion, diction, and writing mechanics, while creativity is reflected in originality, elaboration, flexibility, and novelty in story development. Generative AI is integrated into learning activities as guided support across the stages of planning, drafting, and revising, with teacher supervision to ensure ethical use and to maintain student ownership of ideas. A quasi-experimental approach with a comparison group is used to examine differences in learning gains.

The findings of this study are expected to contribute to both theory and practice. Theoretically, it clarifies how digital scaffolding mediated by Generative AI may support elementary learners’ writing processes and creative expression. Practically, it offers evidence-based guidance for teachers on designing AI-assisted writing lessons that remain developmentally appropriate, ethical, and aligned with

curriculum demands. The study also has implications for school policy, teacher professional development, and digital literacy education, particularly in preparing students to use emerging technologies responsibly while strengthening fundamental writing skills.

#### Research Questions

This study addressed the following research questions:

1. Does the use of Generative AI as digital scaffolding significantly improve elementary students' narrative textual quality compared to conventional scaffolding?
2. Does the use of Generative AI as digital scaffolding significantly enhance students' narrative writing creativity?
3. Which sub-components of textual quality and creativity show the greatest improvement under AI-assisted scaffolding?

#### Method

This study employed a quasi-experimental pretest–posttest control group design to examine the effect of Generative AI as digital scaffolding on elementary students' narrative writing outcomes. The independent variable was the type of scaffolding (AI-assisted scaffolding versus conventional scaffolding), while the dependent variables were narrative textual quality and narrative writing creativity. The intervention was conducted over five instructional meetings (each 2 × 35 minutes) within regular Indonesian language lessons to ensure ecological validity and alignment with the national curriculum.

The study involved 58 Grade 5 students from a public elementary school in Riau Province, Indonesia. The school is located in a semi-urban area and follows the national curriculum with standard instructional hours for Indonesian language learning. Two intact classes with comparable academic profiles were selected purposively. The experimental group consisted of 29 students (15 female, 14 male), and the control group consisted of 29 students (16 female, 13 male). Both classes were taught by the same Indonesian language teacher to reduce teacher-related variability. Pretest scores were analyzed to confirm initial equivalence between groups before the intervention.

In the experimental group, Generative AI was integrated as structured digital scaffolding using ChatGPT (GPT-4-based model) accessed through a supervised classroom account. The AI tool was not positioned as an automatic text generator but as guided support during three writing stages: pre-writing, drafting, and revising. During pre-writing, students used teacher-provided prompts and story maps, and the AI generated idea options or guiding questions based on student input. During drafting, the AI suggested transitional phrases, vocabulary alternatives, and prompts for elaboration. During revising, it provided feedback-style suggestions related to coherence, clarity, and mechanics. Clear usage rules were implemented: students

were prohibited from requesting complete essays, and all AI interactions had to begin from student-generated outlines or sentences

The control group received conventional scaffolding aligned with the same instructional stages. Pre-writing activities included teacher modeling and paper-based story mapping. Drafting support consisted of guided questioning and mini-lessons, while revision relied on peer feedback and teacher comments. Both groups received identical writing prompts, instructional time, and assessment procedures

Data were collected through narrative writing performance tasks administered as pretest and posttest under supervised classroom conditions. Textual quality was assessed using an analytic rubric covering narrative structure, coherence and cohesion, diction and sentence effectiveness, and mechanics (spelling and punctuation). Creativity was assessed through originality, elaboration, flexibility, and novelty of story development. Content validity was established through expert review by two Indonesian language education specialists. Inter-rater reliability was calculated using the Intraclass Correlation Coefficient (ICC). The ICC values were .87 for textual quality and .84 for creativity, indicating high scoring consistency

Statistical analysis was conducted using SPSS. Paired-sample t-tests were used to examine within-group improvements. Independent-sample t-tests compared posttest scores between groups. Additionally, ANCOVA was performed to control for pretest differences and to determine the adjusted effect of the intervention. Effect sizes were calculated using Cohen's *d* and partial eta squared ( $\eta^2$ ) to assess practical significance

Several internal validity considerations were addressed. The teacher effect was minimized by using the same teacher for both groups. The novelty effect was reduced by introducing the AI tool gradually and limiting its use to structured checkpoints rather than constant interaction. Instructional time, writing topics, and classroom environment were kept equivalent across groups to control extraneous variables. These procedures strengthened the methodological rigor and ensured that observed differences were attributable primarily to the scaffolding approach.

## **Findings and Discussion**

### **Pretest Equivalence**

An independent-samples t-test was conducted to examine baseline differences between the experimental and control groups. Results indicated no significant difference in pretest scores for textual quality,  $t(56) = 0.49$ ,  $p = .62$ , or creativity,  $t(56) = 0.57$ ,  $p = .57$ . The experimental group's mean score for textual quality was 67.21 (SD = 7.14), while the control group's mean was 66.48 (SD = 6.89). For creativity, the experimental group obtained a mean of 65.83 (SD = 7.02), and the control group 66.11 (SD = 6.95). These results indicate comparable initial writing ability across groups.

## Posttest Comparison

After the five-session intervention, both groups improved; however, gains were substantially higher in the experimental group.

Table 1. Posttest Comparison Between Groups

Variable	Group	Mean	SD	t	p	Cohen's d
Textual Quality	Experimental	82.41	6.32	5.87	<.001	1.36
	Control	73.18	7.05			
Creativity	Experimental	80.27	6.85	5.12	<.001	1.21
	Control	71.94	7.11			

The experimental group demonstrated an average gain of +15.20 points in textual quality, compared to +6.70 points in the control group. For creativity, the experimental group improved by +14.44 points, while the control group improved by +5.83 points. Independent-samples t-tests revealed statistically significant differences between groups in both outcomes ( $p < .001$ ). Effect sizes were large (Cohen's  $d > 1.20$ ), indicating strong practical significance.

An ANCOVA controlling for pretest scores further confirmed the effect of AI-assisted scaffolding on textual quality,  $F(1,55) = 28.74$ ,  $p < .001$ , partial  $\eta^2 = .34$ , and creativity,  $F(1,55) = 24.19$ ,  $p < .001$ , partial  $\eta^2 = .31$ . These results suggest that approximately 31–34% of posttest variance was attributable to the intervention.

## Subcomponent Analysis

Detailed rubric analysis showed that the largest improvements in the experimental group occurred in:

- Coherence and cohesion (+18% increase)
- Elaboration of narrative events (+21% increase)
- Narrative resolution clarity (+17% increase)

Diction and mechanics also improved significantly, particularly in the use of transitional expressions and punctuation in dialogue writing.

## Qualitative Evidence

Pretest narratives in both groups frequently included abrupt endings and limited descriptive detail (e.g., “Then I went home. The end.”). In contrast, posttest narratives in the experimental group showed expanded scenes and clearer causal connections, such as explicit explanations of character motivations and consequences of events. Students also incorporated more varied vocabulary and dialogue markers. The control group demonstrated improvement in structure but tended to retain simpler plot development patterns.

## Discussion

### 1. Why AI Scaffolding Works: Cognitive Support and Immediate Feedback

The significant gains in textual quality can be explained through scaffolding theory. AI-assisted scaffolding provided immediate, contingent prompts that supported students within their zone of proximal development. Unlike static worksheets, the AI tool responded dynamically to students’ outlines and draft sentences, reducing cognitive overload during complex writing stages. Immediate feedback on coherence and transitions likely strengthened narrative flow, while vocabulary suggestions enriched sentence precision. These findings support the argument that structured digital scaffolding enhances writing performance when aligned with instructional goals.

### 2. Creativity Mechanism: Stimulating Divergent Thinking

The marked improvement in creativity scores suggests that AI support facilitated divergent thinking. By offering multiple possible conflicts, alternative endings, and descriptive expansions, the AI functioned as a stimulus for idea exploration. Importantly, students were required to adapt and justify suggestions rather than copy them directly, which preserved authorship while expanding imaginative possibilities. This aligns with creativity theory emphasizing elaboration and flexibility as core dimensions of creative growth. AI did not replace creative effort but expanded the range of narrative options students considered.

### 3. Comparison with Previous Research

The results are consistent with previous studies on AI-assisted writing in secondary and tertiary contexts, which reported improvements in fluency and revision quality. However, most prior research focused on second-language or higher education settings. This study extends empirical evidence to elementary first-language writing, a context where controlled classroom experiments remain limited. The large effect sizes observed here suggest that younger learners may particularly benefit from structured AI prompts when teacher mediation is present. The findings

also align with research on digital scaffolding and formative feedback, reinforcing the importance of iterative guidance during writing development.

#### 4. Pedagogical Implications

The findings have several implications for writing pedagogy. First, AI should be integrated as process-oriented scaffolding, not as a text generator. Teachers should require student-generated outlines before AI interaction and limit usage to specific stages such as elaboration and revision checkpoints. Second, reflective practices—such as short justification notes explaining which AI suggestions were adopted—help maintain student agency and prevent overreliance. Third, AI should complement, not replace, peer feedback and teacher conferencing. A blended scaffolding model appears most effective.

Furthermore, ethical boundaries are essential. Clear classroom rules, teacher monitoring, and digital literacy instruction must accompany AI integration. When used responsibly, AI can enhance writing instruction by freeing teachers to focus on higher-order feedback while students receive structured, immediate support.

#### Conclusion

This study examined the effect of using Generative AI as digital scaffolding on elementary students' narrative writing in Indonesian language learning. The results demonstrated that students who received AI-assisted scaffolding achieved significantly higher improvements in both textual quality and creativity compared to those who received conventional scaffolding. Statistical analyses showed large effect sizes, indicating that the intervention had meaningful practical impact. Improvements were particularly evident in coherence and cohesion, elaboration of narrative events, clarity of resolution, and creative flexibility. These findings provide empirical support for the effectiveness of structured AI integration in primary-level writing instruction.

The study also addresses the limited empirical evidence regarding AI-assisted scaffolding in elementary first-language writing contexts. When positioned as guided support rather than as an automatic text generator, Generative AI functioned as process-oriented scaffolding across planning, drafting, and revising stages. AI prompts and feedback helped students organize ideas more logically, refine vocabulary choices, and expand story development while maintaining student authorship. The structured rules, teacher supervision, and reflective practices implemented in this study were essential in ensuring ethical and responsible use pedagogically, the findings suggest that the successful integration of Generative AI depends on thoughtful instructional design rather than mere technological access.

Teachers should limit AI use to specific writing stages, require student-generated outlines before AI interaction, and combine AI scaffolding with conventional strategies such as modelling and peer feedback. Although the study was conducted within a limited context and duration, it contributes practical and

theoretical insights into responsible AI use in elementary language education. Future research may explore long-term effects, different genres, and broader educational settings to further strengthen the evidence base.

## References

- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219–234.  
[https://doi.org/10.1207/S15324818AME1403\\_2](https://doi.org/10.1207/S15324818AME1403_2)
- Alda, Sam Hermansyah, Andi Sadapotto, Jamaluddin Ahmad, Nuraini Kasman, Buhari. Developing CEFR Aligned Speaking Materials with the ADDIE Framework Evidence from Indonesian EFL Learners' Speaking Proficiency. *UKR Journal of Arts, Humanities and Social Sciences (UKRJAHS)*, volume 1(issue 7), 165-168.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.  
<https://doi.org/10.4324/9780203812310>
- Bunting, M., & Cowan, N. (2005). Working memory and flexibility in creative writing. *Journal of Educational Psychology*, 97(4), 611–620.  
<https://doi.org/10.1037/0022-0663.97.4.611>
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.  
[https://doi.org/10.1207/s1532690xci0804\\_2](https://doi.org/10.1207/s1532690xci0804_2)
- Dai, D. Y., & Renzulli, J. S. (2008). *Creativity and giftedness*. Thousand Oaks, CA: Sage.  
<https://doi.org/10.4135/9781483329070>
- Graham, S., & Perin, D. (2007). Writing next: Effective strategies to improve writing. *Journal of Educational Psychology*, 99(3), 445–476.  
<https://doi.org/10.1037/0022-0663.99.3.445>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.  
<https://doi.org/10.3102/003465430298487>
- HERMANSYAH, S. (2023). Investigating Difficulties Faced by Lecturers in Teaching General English. *Journal of English Education and Teaching*, 7(3), 499–509.  
<https://doi.org/10.33369/jeet.7.3.499-509>
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511667251>
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1), 1–26.  
<https://doi.org/10.17239/jowr-2008.01.01.1>
- Knoch, U., & Sitajalabhorn, W. (2013). A closer look at analytic scoring. *Assessing Writing*, 18(1), 1–17.  
<https://doi.org/10.1016/j.asw.2012.11.003>

- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. London: Pearson.  
<https://doi.org/10.13140/RG.2.2.19217.68965>
- McLeod, S., & Vasinda, S. (2009). Electronic portfolios as digital scaffolding. *Journal of Adolescent & Adult Literacy*, 53(3), 218–228.  
<https://doi.org/10.1598/JAAL.53.3.4>
- Meyer, B. J. F., & Ray, M. N. (2011). Structure strategy interventions. *Educational Psychology Review*, 23(1), 1–35.  
<https://doi.org/10.1007/s10648-010-9156-2>
- Miller, T. (2023). Explainable AI in education. *Artificial Intelligence Review*, 56, 1–25.  
<https://doi.org/10.1007/s10462-023-10355-3>
- M U, Sari H, Hermansyah S, Maming K, Kahar A, Hasan, Elfahmi FK (2025), "Understanding Indonesian students' reading knowledge in digital literacy within socio-cultural of rural middle schools". *International Journal of Information and Learning Technology*, Vol. 42 No. 5 pp. 432–448, doi: <https://doi.org/10.1108/IJILT-12-2023-0239>
- Mudinillah, A., Kuswandi, D. ., Erwin, E., Sugiarni, S., Winarno, W., Annajmi, A., & Hermansah, S. (2024). Optimizing Project-Based Learning in Developing 21st Century Skills: A Future Education Perspective. *Qubahan Academic Journal*, 4(2), 86–101. <https://doi.org/10.48161/qaj.v4n2a352>
- Nur Rahmah Wahyuddin, Nur Eva Yanti, Riska Arnas, Sam Hermansyah, Utilization of Artificial Intelligence in EFL Learning from a Digital Literacy Perspective, <https://doi.org/10.24256/ideas.v13i2.8331>
- Pifarré, M., & Cobos, R. (2010). Promoting metacognitive skills through scaffolding. *Educational Technology Research and Development*, 58(2), 181–196.  
<https://doi.org/10.1007/s11423-009-9135-6>
- Reeve, J. (2012). A self-determination theory perspective on student engagement. *Handbook of Research on Student Engagement*, 149–172.  
[https://doi.org/10.1007/978-1-4614-2018-7\\_7](https://doi.org/10.1007/978-1-4614-2018-7_7)
- Robin, B. R. (2008). Digital storytelling: A powerful technology tool. *Theory Into Practice*, 47(3), 220–228.  
<https://doi.org/10.1080/00405840802153916>
- Sam Hermansyah, Syamsu T, Syamsunir, Winda Pratiwi, K., Hamka, N., & Ramli, R. (2025). Improvement of Vocational Skills of SMK Muhammadiyah Rappang Students Through an Entrepreneurship-Based Training Program in Pancarijang District, Sidenreng Rappang Regency . *Unram Journal of Community Service*, 6(2), 326–333. <https://doi.org/10.29303/ujcs.v6i2.987>
- Sain, Y., & Hermansyah, S. (2025). Exploring the Language Attitudes of the Tolaki Community in Kendari: A Comprehensive Sociolinguistic Analysis. *JOLLT Journal of Languages and Language Teaching*, 13(2), 983–993.  
<https://doi.org/10.33394/jollt.v13i2.14415>
- Sawyer, R. K. (2012). *Explaining creativity*. Oxford: Oxford University Press.

- <https://doi.org/10.1093/acprof:oso/9780199737574.001.0001>
- Schunk, D. H. (2012). *Learning theories: An educational perspective*. Boston: Pearson.  
<https://doi.org/10.4324/9780203870419>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.  
<https://doi.org/10.3102/0034654307313795>
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning. *Cambridge Handbook of the Learning Sciences*, 409–426.  
<https://doi.org/10.1017/CBO9780511816833.025>
- Torrance, E. P. (1974). *Torrance tests of creative thinking*. Lexington, MA: Ginn.  
<https://doi.org/10.1037/t05821-000>
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.  
<https://doi.org/10.2307/j.ctvjf9vz4>
- Wahyuddin, N. R., Yanti, N. E., Arnas, R., & Hermansyah, S. (2024). Digital Literacy Integrated with Blended Learning in Improving EFL Students' English Language Skills: A Lesson Learned from the Independent Campus Program. *Journal of Language and Literature Studies*, 4(4), 744-757.  
<https://doi.org/10.36312/jolls.v4i4.2351>
- Yang, Y.-F. (2016). Automated writing evaluation and scaffolding. *Educational Technology & Society*, 19(2), 257–270.  
[https://doi.org/10.30191/ETS.201604\\_19\(2\).0018](https://doi.org/10.30191/ETS.201604_19(2).0018)
- Zhai, X. (2022). ChatGPT and AI-assisted writing in education. *Computers and Education: Artificial Intelligence*, 3, 100062.  
<https://doi.org/10.1016/j.caeai.2022.100062>