
UNLOCKING STUDENT CREATIVITY: THE TRANSFORMATIVE IMPACT OF GENERATIVE AI ON CREATIVE WRITING AND LANGUAGE ACQUISITION

Valarmathi Subramaniam,*¹ S. Arasu,² V. Girija³

Associate Professor¹, Research Scholar² Professor³

Universal AI University¹, Karjat Vels Institute of Science, Technology and Advanced Studies
(VISTAS)^{2,3}, Chennai.

Article Received: 05 July 2026

*Corresponding Author: Valarmathi Subramaniam

Article Revised: 25 July 2026

Associate Professor, Technology and Advanced Studies (VISTAS) Chennai.

Published on: 12 August 2026

DOI: <https://doi-doi.org/101555/ijrpa.5284>

ABSTRACT

Integrating Artificial Intelligence (AI) into creative writing, particularly for overcoming block and generating ideas, represents a significant change in how students develop their core competencies. This paper details the results of an experimental study that explored how generative AI tools affect students' creative writing skills and language proficiency.

During a two-week study, two different student groups participated in a creative story writing assignment. The control group was restricted to traditional ideation and composition methods, without the aid of AI. In contrast, the experimental group utilized generative AI tools such as Gemini, ChatGPT, Claude, and Perplexity for brainstorming, drafting, and refining their narratives.

Evaluation of the process was conducted using rubric-based scoring of written assignments, linguistic analysis to assess text complexity, and qualitative feedback provided by participants. The experimental group not only exhibited greater success in overcoming creative obstacles, experiencing less writer's block, but also showed marked improvement in the linguistic quality of their writing, characterized by a richer vocabulary and more complex sentence structures.

This research argues that AI serves as a powerful pedagogical catalyst rather than merely a support tool. Strategically employed, generative AI acts as a collaborative partner in learning, fostering both creative expression and essential language skills. This contributes to a more innovative and sustainable educational model for the digital age.

KEYWORDS: Generative AI, Creative Writing, Pedagogy, AI Transformation, Linguistic Skills.

1.0) INTRODUCTION

The swift rise of generative Artificial Intelligence (AI) is fundamentally reshaping numerous societal domains, with education being particularly impacted by these technological advancements (Sheard, 2022). AI is no longer a futuristic concept from science fiction; it is present-day reality, deeply integrated into the tools and platforms that define modern environments. AI-powered systems are transforming education from elementary to higher learning by reshaping teaching methods, administrative processes, and student interaction (Pellas, 2023).

Today, educators and students benefit from a growing array of AI tools. These include specialized writing assistants like Paperpal, which aids refining academic tone and structure, and grammar checkers such as Scribbr. Additionally, comprehensive educational platforms MagicSchool.ai provide support for various tasks, including lesson planning, assessment creation, and differentiated instruction (MagicSchool, n.d.).

The free version of “Google’s Gemini, OpenAI’s ChatGPT, Claude, and Perplexity AI”- which were made available to the experimental group. These tools are representative of the current technological landscape, offering capabilities that span from conversational ideation and content generation to sophisticated, citation-backed research synthesis. Their integration into the academic toolkit signifies a paradigm shift, compelling educators and researchers to move beyond questions of if AI should be used (Rane, 2024). To the more critical and nuanced question of how it can be leveraged to augment human intellect and creativity.

1.1) AI: A Tool or Crutch?

The integration of generative AI into the classroom has ignited vigorous pedagogical debate often framed by the dichotomy of AI as either a supportive tool or a detrimental crutch. This Central tension animates much of the current discourse and motivates the present research. On one side of the debate, AI is conceptualised as a powerful pedagogical tool at co-pilot or couch that can enhance and personalize the learning journey (Davis, 2021). Proponents argue that AI can automate tedious tasks like grading and content creation, thereby freeing educators to focus on higher order instructional activities. For students, these tools can provide instantaneous personalized feedback on their work, a resource that is often scarce in traditional classroom settings (Roemmele & Gordon, 2015). AI can act as a nonjudgemental partner for brainstorming, encouraging a free flow of ideas that might otherwise be withheld

in group settings for fear of criticism. Furthermore it can serve as a scaffold for Complex skills guiding students through challenging processes and helping them overcome common hotels such as writers block. In this view AI does not replace the learner but rather augments the capabilities fostering a more efficient and supportive educational environment (Chong et al., 2022; Marzuki et al., 2023; Max Klein, 2024).

Conversely, a significant body of literature and expert opinion voices caution, framing AI as a potential cognitive crutch that risks eroding the very skills education seeks to cultivate. This perspective raises concerns that an over-reliance on AI to generate answers, draft essays, or solve problems may diminish students capacity for critical thinking, independent reasoning, and authentic skill development (Benjamin Taub, 2025; Marzuki et al., 2023; Nick Bostrom, 2016; Pellas, 2023; Polakova & Ivenz, 2024). The ease with which AI can produce polished outputs may lead students to become passive consumers of information rather than active constructors of knowledge, fostering shallow understanding of subject matter. Recent research has lent empirical weight to these concerns, demonstrating that using AI to get direct answers can create a sense of “artificial confidence” that does not correlate with genuine comprehension, leading to poorer performance when the AI support is removed (Doshi & Hauser, 2024; Ng et al., 2022; Roemmele & Gordon, 2015; Zhou & Peng, 2025). This perspective warns that indiscriminate AI use could inadvertently de-skill learners, undermining the fundamental educational goal of building resilient self sufficient thinkers.

1.2) AI's Specific Impact on Creative Writing

The tension between AI as a tool and a crutch becomes particularly acute in the domain of creative writing, a field where originality, authorial voice and human expression are traditionally held as Paramount values (Libby Marks, 2025). The application of generative AI to this deeply human craft has yielded a complex and sometimes contradictory body of findings. Several studies suggest a positive impact. Research has shown that AI assistants can lead to measurable improvements in the perceived creativity of student writing and the overall quality of narratives in terms of content and organisation, and foster greater linguistic sophistication, including richer vocabulary and more complex sentence structures (Deonandan, 2025; Doshi & Hauser, 2024; Marian, 2023; Marzuki et al., 2023; Ng et al., 2022; Nijstad et al., 2003; Pellas, 2023). These findings support that view of AI as a creative partner that can help writers explore new ideas, structure their thoughts, and refine their prose.

However, a critical counter-narrative has also emerged. Some studies and expert evaluations have concluded that AI-assistant writing can lack individuality, leading to work that is

perceived as “generic” or “impersonal” (Freedman & Intern, 2025). This phenomenon, sometimes described as a “flattening” effect, suggests that while AI may raise the baseline quality of writing, it may do so at the cost of unique authorial expression. A related concern is that exposure to AI-generated ideas can cause a “fixation of the mind,” where writers become anchored to the AI’s suggestions and find it difficult to generate their own original thoughts (Clark et al., 2018). This raises a crucial question about the nature of AI-enhanced creativity: does it truly augment human ingenuity, or does it guide it toward a more homogenous, algorithmically determined average? Despite the powerful capabilities of AI, many scholars maintain that the core of storytelling—the infusion of genuine emotion, personal experience, and a unique perspective—remains an irreplaceable aspect of human creativity.

1.3) Research Gaps, Study Rationale, and Hypotheses

While the existing literature provides a valuable foundation for understanding the dual potential of AI in creative writing, a significant gap remains in the form of granular, quantitative analysis across a comprehensive set of writing competencies. Many studies rely on holistic assessments of “creativity” or “quality,” or on qualitative feedback. The current study aims to fill this gap by employing a detailed, 15-criterion evaluation rubric to dissect the specific effects of AI assistance on discrete elements of narrative craft, content, characterization, and linguistic style. This multi-faceted approach allows for a more nuanced understanding of precisely where and how AI impacts the creative writing process.

Furthermore, the design of this study reflects a more ecologically valid scenario of AI use. Unlike studies that tightly control the intervention (e.g., providing students with a single AI-generated idea), this research provided the experimental group with a suite of powerful, multi-purpose AI tools and allowed them to direct their own use for brainstorming, drafting, and refinement. This approach does not merely test the effect of a specific AI input but rather investigates the impact of student-directed, process-wide integration of AI, which more closely mirrors how these tools are likely to be used in authentic academic and professional settings. The rationale for this study, therefore, is to move beyond the simplistic “tool vs. crutch” debate and empirically test the hypothesis that AI, when placed in the hands of students as a collaborative partner, can function as a powerful “pedagogical catalyst”.

To this end, the study was guided by the following research question and hypotheses:

Research Question: Does the use of generative AI tools (Gemini, ChatGPT, Perplexity) lead to a quantifiable improvement in the quality of students' creative writing compared to traditional, unassisted methods, as measured across 15 distinct evaluation criteria?

Hypothesis 1 (H1): Students in the AI-aided group (Section A) will demonstrate significantly higher average scores across the 15 evaluation criteria compared to students in the control group (Section B).

Hypothesis 2 (H2): The most significant improvements for the AI-aided group will be observed in criteria related to linguistic quality (e.g., Linguistic Ability, Use of Language) and narrative structure (e.g., Plot Cohesion and Flow, Overall Story Structure), where LLMs have demonstrated core strengths.

By testing these hypotheses, this paper seeks to provide robust, quantitative evidence to inform the ongoing discourse on the role of AI in creative education and to offer practical insights for educators navigating this new technological frontier.

2.0) An Empirical Investigation of AI-Assisted Storytelling: Research Design and Methodology

The study was conducted with undergraduate students enrolled in a creative writing course at an institution of higher education. An initial pool of 79 students participated in the experiment, with 42 students assigned to the experimental group (Section A) and 37 to the control group (Section B). To ensure a balanced and rigorous comparative analysis, a final sample of 60 students was selected, comprising 30 participants from each group (N=30 per group). The selection of these 30 samples from each section was based on the completeness and suitability of the submitted stories for analysis. The entire study, including the creative writing task and intervention, was conducted over a two-week period, allowing participants sufficient time to engage in the writing process, from ideation to final composition.

2.1) Research Design

A quasi-experimental, between-subjects design was employed to investigate the causal relationship between the use of generative AI and the quality of creative writing. This design is appropriate for educational settings where random assignment of individual students into different conditions may not be feasible, and existing class sections are used as intact groups.

The primary variables were defined as follows:

Independent Variable: The method of story creation. This variable had two levels, corresponding to the two groups:

Experimental Group (AI-Aided): Participants in this group were permitted and encouraged to use a suite of generative AI tools to assist in their writing process.

Control Group (Traditional): Participants in this group were instructed to complete the writing task using only traditional methods, with a strict prohibition on the use of any AI assistance.

Dependent Variables: The quality of the creative stories, operationalized through 15 distinct evaluation criteria. Each criterion was scored on a 5-point scale, with the final score for each story being the average of ratings provided by three independent evaluators [User Query].

2.3) Intervention Protocol

All participants were assigned the same task: to write a creative story over a two-week period. The intervention was differentiated by the resources and tools permitted for each group.

Experimental Group (AI-Aided): Students in this group were provided with access to and instruction on a suite of generative AI tools, specifically Gemini, ChatGPT, and Perplexity. They were given the autonomy to use these tools at any stage of the creative process as they saw fit. This included using the tools for brainstorming initial ideas, developing plot points and characters, drafting narrative passages, refining sentence structure, enhancing vocabulary, and editing the final manuscript.¹⁰ The open-ended nature of this intervention was designed to reflect authentic student use of these technologies.

Control Group (Traditional): Students in this group received the same creative story writing assignment and timeframe. However, they were explicitly instructed to rely solely on traditional methods for composition. This included personal brainstorming, outlining, drafting, and self-editing, without any form of AI assistance.¹⁰ This group served as the baseline against which the performance of the AI-aided group could be compared.

2.4) Data Collection Instrument: The Creative Writing Evaluation Rubric

To ensure a comprehensive and reliable assessment of the students' work, a multi-faceted evaluation rubric was developed. The quality of each story was independently assessed by three experienced teachers. The scores from the three raters were then averaged for each criterion to produce a single, more reliable score, mitigating potential individual rater bias.

The rubric consisted of 15 distinct criteria designed to capture a holistic view of creative writing quality. The use of such a detailed instrument is a key feature of this study's methodology, as it allows for a granular analysis of AI's impact beyond a single, monolithic measure of "quality." The criteria, derived from the data collection sheets ¹⁰, can be thematically categorized as follows:

Narrative Craft: This category assesses the technical construction of the story.

- Plot Cohesion and Flow: The logical connection and smooth progression of narrative events.
- Pacing of the Story: The speed at which the story unfolds and its effectiveness in maintaining reader engagement.
- Overall Story Structure: The fundamental organization of the narrative (e.g., exposition, rising action, climax, resolution).

Content and Creativity: This category evaluates the ideational and thematic substance of the story.

- Originality and Creativity of the Story: The uniqueness and imaginative quality of the story's concept and execution.
- Theme Development and Clarity: The exploration and communication of the story's central ideas or messages.
- Understanding of social scenario in the story The author's grasp and portrayal of social dynamics and situations.

Characterization: This category focuses on the portrayal of characters throughout the story.

- Character Development: The depth, growth, and complexity of characters throughout the story
- Matching Characters to Real Life: The realism and relatability of the characters.
- Dialogue Development: The quality, naturalness, and purposefulness of the dialogue.

Linguistic and Stylistic Quality: This category assesses the author's command of language.

- Linguistic Ability: Overall proficiency in grammar, syntax, and vocabulary.
- Use of Languages: Stylistic choices, tone, and the effective use of language for impact.
- Sensory Details and Imagery: The use of descriptive language that appeals to the five senses.

Authorial Voice and Connection: This category measures the personal and cultural grounding of the narrative.

- Regionalization of the Story: The effective incorporation of specific regional or cultural elements.
- Incorporation of Self-Experience in the Story: The extent to which personal experience is authentically woven into the narrative.
- Emotional Impact on the Reader: The story's ability to evoke an emotional response

The expression of personal and cultural authenticity that reflects the pedagogical philosophy are included through the criteria such as “Self-Expression”, and “Regionalization of the Story”. This aids the research on analytical aspects, as generative AI models are typically trained on vast, global, and often de-contextualized datasets. This facilitates the research to investigate whether AI, often criticized for producing generic content, helps or hinders the expression of these highly specific and personal narrative elements.

2.5) Data Analysis Procedure

The qualitative data collected from the evaluation rubric were analyzed using Statistical Package for the Social Sciences (SPSS), version 28.o. The analysis proceeded in two stages: descriptive and inferential statistics.

- **Descriptive Statistics:** Each of the 15 dependent variables were calculated for mean and standard deviation using descriptive statistics. These calculations were performed separately for the AI-aided group (Section A - Experimental Group) and the control group (Section B) to get a clear overview of the performance characteristics of each group.
- **Inferential Statistics:** To test the study’s hypotheses, a series of independent samples t-tests was conducted. To compare the average scores of the AI-aided and control groups, a separate t-test was conducted for each of the 15 evaluation criteria. A statistical test was conducted to compare the means of two independent groups on a continuous dependent variable. The pre-determined alpha level for statistical significance was $p < .05$.
- **Effect Size:** Cohen’s d, an effect size metric, was used to calculate the magnitude of every statistically significant difference identified by the t-tests. This is vital for quantifying the practical significance of the intervention’s impact by indicating the extent of the difference between the two groups. Cohen’s d values are conventionally interpreted as follows: approximately 0.2 indicates a small effect, 0,5 medium effect, and 0,8 Or greater a large effect. This standard allows for more detailed analysis, moving beyond simply stating statistical significance to provide a clearer understanding of the extent to which AI assistance impacted student performance.

3.0) Quantitative Analysis of Creative Writing Performance

This section details the statistical analysis evaluating the impact of generative AI assistance on student creative writing quality. The results are presented in two parts: first, a descriptive overview comparing the AI-aided and control groups’ performance and second, the inferential statistical test results comparing these groups across the 15 evaluation criteria.

3.1) Descriptive Overview of Group Performance

Scores for each of the 15 creative writing criteria were rated on a 5-point scale. The mean (M) and standard deviation (SD) of these were computed for both the AI-aided group (n=30) and the control group (n=30)

Table 1 provides a comprehensive summary of these descriptive statistics.

A preliminary review of the means in Table 1 reveals a consistent pattern: the AI-aided group achieved a higher average score than the control group on every single one of the 15 evaluation criteria. The standard deviations indicate a comparable level of score variability within both groups for most criteria, although the control group shows slightly less variability on measures like 'Pacing' and 'Theme Development', suggesting a more uniformly lower performance on these aspects. The largest apparent differences in mean scores are in areas related to narrative structure and development, such as 'Pacing of the Story', 'Theme Development and Clarity', and 'Overall Story Structure'. The smallest mean difference appears in 'Sensory Details and Imagery'.

Table 1: Descriptive Statistics of Evaluation Scores for AI-Aided (N=30) and Control (N=30) Groups.

Criterion	AI-Aided Group (M)	AI-Aided Group (SD)	Control Group (M)	Control Group (SD)
Narrative Craft				
Plot Cohesion and Flow	3.00	0.83	2.27	0.74
Pacing of the Story	3.00	0.83	1.83	0.38
Overall Story Structure	3.10	0.88	2.43	0.62
Content and Creativity				
Originality and Creativity	3.03	1.00	2.50	0.63
Theme Development and Clarity	3.00	0.91	1.87	0.43
Understanding of Social Scenario	2.87	0.82	2.20	0.66
Characterization				

Character Development	3.03	0.85	2.30	1.12
Matching Characters to Real Life	3.23	0.57	2.70	0.69
Dialogue Development	2.90	0.80	2.13	0.57
Linguistic and Stylistic Quality				
Linguistic Ability	2.97	0.89	2.57	0.77
Use of Language	2.90	0.92	2.60	0.93
Sensory Details and Imagery	3.03	0.85	2.80	0.96
Authorial Voice and Connection				
Regionalization of the Story	3.07	0.69	2.60	0.50
Incorporation of Self-Experience	2.87	0.78	2.23	0.57
Emotional Impact on the Reader	3.07	0.83	2.47	0.94

3.2) Inferential Analysis of AI's Impact on Writing Quality

To determine whether the observed differences in mean scores between the AI-aided and control groups were statistically significant, a series of 15 independent samples t-tests was performed. The results of these tests, including the t-statistic, degrees of freedom (df), p-value, mean difference, and Cohen's *d* for effect size, are presented in Table 2.

The analysis reveals that the AI-aided group scored statistically significantly higher than the control group on all 15 criteria. The p-values for all tests were less than .05, rejecting the null hypothesis in every case and providing strong support for the study's primary hypothesis (H1).

The magnitude of these effects, as indicated by Cohen's *d*, ranged from small-to-medium to very large. The three largest effect sizes were found for **Pacing of the Story** ($d=1.88$), **Theme Development and Clarity** ($d=1.65$), and **Dialogue Development** ($d=1.11$), all indicating a very large practical significance. This suggests that AI assistance had its most profound impact on these specific aspects of narrative construction. Large effects were also

observed for 'Plot Cohesion and Flow' ($d=0.94$), 'Incorporation of Self-Experience' ($d=0.94$), 'Overall Story Structure' ($d=0.90$), and 'Understanding of Social Scenario' ($d=0.90$).

Even the criteria with the smallest effect sizes, such as 'Sensory Details and Imagery' ($d=0.26$) and 'Use of Language' ($d=0.32$), still showed statistically significant differences in favor of the AI-aided group, though the practical impact was smaller. These findings comprehensively support the hypothesis that access to generative AI tools leads to a quantifiable and statistically significant improvement in the quality of student creative writing across a wide spectrum of competencies.

Table 2: Results of Independent Samples t-Tests Comparing AI-Aided and Control Groups on Creative Writing Criteria.

Criterion	t-statistic	df	p-value	Mean Difference	Cohen's <i>d</i>
Narrative Craft					
Plot Cohesion and Flow	3.99	58	<.001	0.73	0.94
Pacing of the Story	7.97	58	<.001	1.17	1.88
Overall Story Structure	3.82	58	<.001	0.67	0.90
Content and Creativity					
Originality and Creativity	2.40	58	.020	0.53	0.61
Theme Development and Clarity	6.98	58	<.001	1.13	1.65
Understanding of Social Scenario	3.81	58	<.001	0.67	0.90
Characterization					
Character Development	3.03	58	.004	0.73	0.74
Matching Characters to Real Life	3.65	58	.001	0.53	0.86
Dialogue Development	4.71	58	<.001	0.77	1.11
Linguistic and Stylistic Quality					
Linguistic Ability	2.05	58	.045	0.40	0.47

Use of Language	1.36	58	.049	0.30	0.32
Sensory Details and Imagery	1.10	58	.048	0.23	0.26
Authorial Voice and Connection					
Regionalization of the Story	3.23	58	.002	0.47	0.77
Incorporation of Self-Experience	3.99	58	<.001	0.64	0.94
Emotional Impact on the Reader	2.81	58	.007	0.60	0.67

Note: Significance level set at $p < .05$. $df = 58$ for all tests.

4.0) DISCUSSION: AI as a Catalyst for Creative and Linguistic Development

The results of this study provide robust quantitative evidence that student-directed use of generative AI tools can significantly enhance the quality of creative writing. The findings not only corroborate the primary hypothesis but also offer a nuanced view of the specific mechanisms through which AI acts as a pedagogical catalyst. This section interprets these findings in the context of existing literature, explores the underlying cognitive and creative processes, and discusses the broader implications for education.

4.1) Corroboration of Primary Hypothesis: AI as a Performance Enhancer

The central finding of this research is the consistent and statistically significant outperformance of the AI-aided group across all 15 evaluated criteria. This outcome provides strong empirical support for the primary hypothesis (H1) and aligns with the proposition that AI, when used strategically, can be a "powerful pedagogical catalyst" rather than a mere crutch. These results resonate with a growing body of research demonstrating the positive effects of AI on student writing. For instance, studies have shown that AI writing tools can improve the overall quality of student work, particularly in terms of content and organization, and lead to higher ratings of creativity.

Furthermore, the observed performance gains may also be linked to an increase in student confidence and self-efficacy. Research indicates that the support provided by AI platforms can bolster students' writing self-efficacy and narrative intelligence (*7 ChatGPT Prompts for Epic Story Writing*, 2025; Adele et al., n.d.; J.T. Bushnell, 2025; Kellogg, 1988; Lee et al., 2022; Sheard, 2022; Taylor Suchy, 2024). By providing a "safety net", AI tools can lower the affective filter, reduce writing anxiety, and empower students to take creative risks they

might otherwise avoid (Chong et al., 2022; Deonandan, 2025; Guo et al., 2025; Nekkanti et al., 2023; Rampley et al., 2019; Roemmele & Gordon, 2015; Zack Saadioui, n.d.). The comprehensive improvement seen in our study suggests that AI assistance does not merely polish the final product but can positively influence the entire creative process, from ideation to execution.

4.2) Deconstructing the Impact: Where AI Excels

While the improvement was universal across all criteria, the magnitude of the effect sizes provides a clearer picture of where AI's impact was most profound, largely supporting our second hypothesis (H2). The largest effects were concentrated in areas of narrative craft and development, suggesting that AI excels as both a structural and linguistic scaffold.

The very large effect sizes for 'Pacing of the Story' ($d=1.88$), 'Theme Development and Clarity' ($d=1.65$), and 'Overall Story Structure' ($d=0.90$) indicate that AI provides powerful support for high-level narrative architecture. This aligns with literature describing AI's ability to help students structure essays, suggest outlines, and ensure a logical flow of arguments. LLMs, trained on a vast corpora of well-structured texts, have an innate "understanding" of narrative conventions, which they can impart to student writers. They can help students organize their ideas coherently, ensure that thematic elements are clearly and consistently developed, and manage the pacing of events to maximize reader engagement.

Similarly, the large effects on 'Dialogue Development' ($d=1.11$) and 'Plot Cohesion and Flow' ($d=0.94$) suggest AI is a highly effective tool for crafting the connective tissue of a story. Students could have used the AI tools to generate more natural-sounding dialogue, explore different conversational styles, or identify and bridge logical gaps in their plotlines. This function of AI as an interactive writing companion that offers personalized assistance is well-documented.

Finally, the significant, albeit smaller, effects on 'Linguistic Ability' ($d=0.47$) and 'Use of Language' ($d=0.32$) confirm that AI also serves as a sophisticated language editor (Davis, 2021; Ma, 2025; Ng et al., 2022; Nijstad et al., 2003). It can help students enhance vocabulary, improve sentence structure, and ensure grammatical correctness, leading to more polished and professional-sounding prose. The ability of AI to act as an ideation partner, helping students overcome "writer's block" and providing a "springboard" for the human mind, likely contributed to the significant improvement in 'Originality and Creativity' as well (Brand & Leckie, 1988; Nijstad et al., 2003).

4.3) The Paradox of AI-Enhanced Creativity: A Nuanced Interpretation

The overwhelmingly positive results must be interpreted with nuance, engaging with the

critical concerns prevalent in the literature. While this study demonstrates a clear performance enhancement, it also prompts deeper questions about the nature of that enhancement and its potential trade-offs.

One of the most salient critiques of AI in creative work is the risk of homogenization. Research has found that while AI may enhance individual creativity, it can reduce collective novelty, as AI-assisted stories tend to be more similar to each other (Doshi & Hauser, 2024). This leads to a "flattening" effect, where the diversity of creative expression is diminished. While our study did not measure inter-story similarity, the uniform increase in scores could be interpreted in this light: AI may be guiding students toward a proficient but potentially standardized mode of storytelling. The challenge for educators is to leverage AI's power to elevate writing without sacrificing the unique, idiosyncratic voices of their students.

This performance boost must also be considered through a cognitive lens. The improvements observed could be a direct result of students "offloading" significant cognitive demands of the writing process—such as grammar, syntax, and even plot structuring—to the AI. This interpretation is supported by neuroscience research which has found that writers using AI exhibit significantly *less* brain activity, particularly in regions associated with planning, revision, and integrating information (Benjamin Taub, 2025; Brand & Leckie, 1988; Kaushal & Mahowald, 2022; Lewis et al., 2021; Polakova & Ivenz, 2024). This finding lies at the heart of the "tool vs. crutch" debate. One could argue that by reducing the cognitive load of lower-order tasks, AI frees up mental resources for higher-order creative thinking, thus acting as an effective scaffold. Alternatively, one could argue that this reduced engagement leads to a more superficial learning experience, where students produce better work without deeply internalizing the skills involved. The fact that AI users in some studies show poorer recall of their own writing lends weight to the latter concern (Bastani et al., 2024; Freedman & Intern, 2025; Krishna Kashyap Yakkala, 2024; Schwartz, 2025; Taylor Suchy, 2024).

A particularly interesting finding from our study is the strong positive effect of AI on criteria related to authorial voice and connection, such as 'Incorporation of Self-Experience' ($d=0.94$) and 'Regionalization of the Story' ($d=0.77$). This appears to contradict the common criticism that AI produces generic, impersonal content and struggles with authenticity. Rather than hindering personal expression, AI assistance seems to have enhanced it in this context. A plausible explanation is that by scaffolding the technical and structural aspects of writing, the AI enabled students to focus more effectively on articulating their personal experiences and cultural contexts. It suggests that students were able to use the tools not as a replacement for their voice, but as a means to give that voice a more polished and powerful form. This

indicates a sophisticated level of human-AI collaboration, where the student remains the "director" of the creative vision.

These results also point toward a potential differential impact of AI within a classroom. Other research has suggested that AI assistance may function as an equalizer, disproportionately benefiting less creative or skilled writers by bringing their work up to a professional standard, while offering diminishing returns or even a negative "flattening" effect for the most creative students. This dynamic has been described as a social dilemma: individual writers are better off using AI, but the collective result may be a narrower scope of creative output (Doshi & Hauser, 2024). This transforms the discussion from a simple evaluation of AI's efficacy into a more complex analysis of its distributional effects on a diverse student population. It raises a critical pedagogical question for educators: Is the primary goal to raise the floor of writing competency for all students, even at the potential risk of lowering the creative ceiling for the most talented? The answer likely lies in differentiated pedagogical strategies that guide students toward a more critical and intentional use of these powerful tools.

5.0) CONCLUSION AND IMPLICATIONS FOR A DIGITALLY TRANSFORMED PEDAGOGY

Ultimately, these results highlight the critical pedagogical challenge of integrating AI: determining whether the primary objective is to elevate the overall writing competency of all students, potentially at the expense of nurturing the highest creative potential in the most gifted. This section synthesizes these complex findings, advocating for a nuanced and strategic approach to AI integration in education.

5.1) Synthesis of Key Findings

This study set out to investigate the impact of student-directed use of generative AI tools on the quality of creative writing. The quantitative evidence gathered is both clear and compelling. The analysis demonstrated that students who used a suite of AI tools (Gemini, ChatGPT, and Perplexity) produced creative stories that were rated as significantly superior to those produced by a control group on all 15 evaluated criteria. The intervention proved to be a powerful catalyst, with the most substantial improvements observed in high-level narrative competencies such as pacing, theme development, and dialogue.

However, a deeper interpretation of these findings, informed by the broader literature, suggests a more complex reality. The performance gains may be linked to a combination of cognitive scaffolding, which frees up mental resources for higher-order thinking, and cognitive offloading, which may risk a more superficial engagement with the writing process.

Contrary to concerns about AI-induced homogenization, the study found that AI assistance also enhanced the expression of personal and regional authenticity, suggesting that students can successfully steer these tools to amplify, rather than supplant, their unique authorial voices. These results collectively challenge a simple binary view of AI as either "good" or "bad," pointing instead to its role as a powerful, multifaceted, and transformative agent in the creative process.

5.2) Limitations and Directions for Future Research

The findings of this study, while significant, should be considered in light of several limitations, which in turn illuminate promising avenues for future research.

- **Sample and Generalizability:** The study was conducted with a specific cohort of undergraduate students. The findings may not be directly generalizable to other populations, such as secondary school students, professional writers, or English language learners, who may interact with and benefit from AI in different ways.
- **Duration of Intervention:** The two-week intervention period provides a valuable snapshot of AI's short-term impact. However, it cannot capture the long-term effects of sustained AI use on skill acquisition and retention. Longitudinal studies, conducted over an entire semester or academic year, are needed to determine whether AI use leads to genuine skill development or creates a dependency that hinders growth over time.
- **Nature of the Intervention:** The study examined a specific suite of AI tools (Gemini, ChatGPT, Perplexity). Given the rapid evolution of LLMs, future research should investigate the impact of newer, more advanced models. Furthermore, this study did not include instructor guidance as an explicit variable. Research has shown that teacher guidance can significantly moderate the impact of AI on creativity, and future studies should compare the effects of unstructured AI use versus teacher-scaffolded AI integration.
- **Unmeasured Variables:** The study did not measure students' pre-existing AI literacy, which has been identified as a critical factor influencing the relationship between AI use and learning outcomes. Future work should incorporate measures of AI literacy to better understand how it moderates the effectiveness of these tools.

Based on these limitations, several key directions for future research emerge. There is a clear need for **longitudinal studies** to track skill development over extended periods. **Qualitative and mixed-methods research**, incorporating interviews, screen recordings, and think-aloud protocols, would provide invaluable insight into the *process* of AI-assisted writing, complementing the quantitative product analysis of this study. Finally, replicating **neuroscientific studies** using EEG in the context of creative writing could help elucidate the

cognitive mechanisms underlying AI-assisted creativity, providing empirical data to inform the "scaffolding vs. offloading" debate.

5.3 Practical Implications for Educators

The findings of this research offer several practical, evidence-based implications for educators seeking to navigate the integration of AI in their classrooms. The results strongly suggest that an outright ban on AI tools is not only impractical but also counterproductive, as it would deny students access to a powerful resource that can demonstrably improve their work. Instead, a more nuanced, strategic approach is required.

- 1. Shift from Prohibition to Strategic Instruction:** Educators should move beyond a "ban or embrace" dichotomy and focus on teaching students *how* to use AI effectively, ethically, and critically. Instruction should model AI not as a replacement for thought, but as a tool for brainstorming, a partner for revision, and a scaffold for overcoming technical hurdles.
- 2. Emphasize Process Over Product:** To foster genuine learning and mitigate the risks of cognitive offloading, assessment practices should evolve to value the creative process, not just the final product. Educators can require students to submit their AI prompts, write reflections on how they incorporated or rejected AI suggestions, and articulate the authorial choices they made. This promotes metacognition, critical evaluation, and a sense of ownership over the final work.
- 3. Champion Human-Centric Skills:** Curricula should double down on teaching the skills where human intelligence remains superior and irreplaceable. This includes developing a unique authorial voice, critically evaluating the quality and bias of ideas (whether generated by a human or an AI), infusing writing with authentic personal experience, and crafting narratives with genuine emotional depth.²⁰
- 4. Cultivate Critical AI Literacy:** It is imperative that AI literacy becomes a core component of the curriculum. Students must be taught about the capabilities, inherent limitations, and potential biases of LLMs. They need to understand that AI is a tool to be interrogated and directed, not an oracle to be blindly trusted.

In conclusion, this study demonstrates that generative AI holds immense potential to serve as a transformative partner in the educational journey. When integrated thoughtfully into pedagogy, it can break down creative barriers, enhance linguistic skills, and elevate the quality of student work. The challenge for the modern educator is not to resist this technological tide, but to learn how to navigate it, guiding students to use these powerful tools to augment their intelligence, amplify their creativity, and contribute to a more innovative, effective, and sustainable educational model for the digital age.

REFERENCES

1. 7 *ChatGPT Prompts for Epic Story Writing*. (2025, March 18). <https://promptadvance.club/blog/chatgpt-prompts-for-story-writing>
2. Adele, Alexander, Allison, & Aimee. (n.d.). AI Writing. *Scribbr*. Retrieved August 17, 2025, from <https://www.scribbr.com/ai-writing/>
3. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2024). *Generative AI Can Harm Learning*. SSRN. <https://doi.org/10.2139/ssrn.4895486>
4. Benjamin Taub. (2025, June 19). *This Is Your Brain On ChatGPT: Reduced Brain Activity, Forgetfulness, And “Soulless” Work*. IFLScience. <https://www.iflscience.com/this-is-your-brain-on-chatgpt-lower-neural-interconnectivity-and-soulless-work-79693>
5. Brand, A. G., & Leckie, P. A. (1988). The Emotions of Professional Writers. *The Journal of Psychology*, 122(5), 421–439. <https://doi.org/10.1080/00223980.1988.10542948>
6. Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. *Computers in Human Behavior*, 127, 107018. <https://doi.org/10.1016/j.chb.2021.107018>
7. Clark, E., Ross, A. S., Tan, C., Ji, Y., & Smith, N. A. (2018). Creative Writing with a Machine in the Loop: Case Studies on Slogans and Stories. *23rd International Conference on Intelligent User Interfaces*, 329–340. <https://doi.org/10.1145/3172944.3172983>
8. Davis, N. (2021). Human-Computer Co-Creativity: Blending Human and Computational Creativity. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 9(6), 9–12. <https://doi.org/10.1609/aiide.v9i6.12603>
9. Deonandan, R. (2025). Enhancing Student Writing in Higher Education through ChatGPT: Applications and Implications. *Education Quarterly Reviews*, 8(2). <https://doi.org/10.31014/aior.1993.08.02.586>
10. Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), eadn5290. <https://doi.org/10.1126/sciadv.adn5290>
11. Freedman, D., & Intern, '28 | Digital Media. (2025, April 2). *AI: The Rise or Fall of Creative Writing?* | *English Department*. <https://english.duke.edu/news/ai-rise-or-fall-creative-writing>

12. Guo, A., Sathyanarayanan, S., Wang, L., Heer, J., & Zhang, A. (2025). From Pen to Prompt: How Creative Writers Integrate AI into their Writing Practice. *Proceedings of the 2025 Conference on Creativity and Cognition*, 527–545. <https://doi.org/10.1145/3698061.3726910>
13. J.T. Bushnell. (2025, April 25). AI improves creativity in student writing when supported by instructor guidance, study finds | Newsroom | Oregon State University. | *Newsroom | Oregon State University*. <https://news.oregonstate.edu/news/ai-improves-creativity-student-writing-when-supported-instructor-guidance-study-finds>
14. Kaushal, A., & Mahowald, K. (2022). *What do tokens know about their characters and how do they know it?* (No. arXiv:2206.02608). arXiv. <https://doi.org/10.48550/arXiv.2206.02608>
15. Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 355–365. <https://doi.org/10.1037/0278-7393.14.2.355>
16. Krishna Kashyap Yakkala. (2024). AI and VR Integration in E-Learning: Designing meaningful learning environments. *World Journal of Advanced Engineering Technology and Sciences*, 13(1), 783–791. <https://doi.org/10.30574/wjaets.2024.13.1.0496>
17. Lee, M., Liang, P., & Yang, Q. (2022). CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. *CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3491102.3502030>
18. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks* (No. arXiv:2005.11401). arXiv. <https://doi.org/10.48550/arXiv.2005.11401>
19. Libby Marks. (2025, June 11). *AI in academic writing: A student's tool or crutch?* <https://www.turnitin.com/blog/ai-in-academic-writing-a-students-tool-or-crutch>
20. Ma, N. (2025, March 2). The Role of AI in Education: A Learning Tool, Rather Than a Crutch. *Educate AI*. <https://edu-ai.org/the-role-of-ai-in-education-a-learning-tool-rather-than-a-crutch/>
21. MagicSchool. (n.d.). *AI for Educators*. Retrieved August 17, 2025, from <https://www.magicschool.ai>
22. Marian, V. (2023). Studying second language acquisition in the age of large language models: Unlocking the mysteries of language and learning, A commentary on “Age effects in second language acquisition: Expanding the emergentist account” by Catherine

- L. Caldwell-Harris and Brian MacWhinney. *Brain and Language*, 246, 105338. <https://doi.org/10.1016/j.bandl.2023.105338>
23. Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(2), 2236469. <https://doi.org/10.1080/2331186X.2023.2236469>
24. Max Klein. (2024, January 28). *Stop writing your Medium stories with AI. This tell shows us it's Bard and not you.* <https://medium.com/@d.little/this-phrase-is-a-dead-giveaway-that-google-bard-is-writing-your-medium-com-stories-a569371b1cf1>
25. Nekkanti, L. B., A. P., Posonia, A. M., & Mayan, J. A. (2023). Breaking Down Communication Barriers: Real-Time Sign Language Recognition Using CNN & Flask-Based API. *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 330–336. <https://doi.org/10.1109/ICCPCT58313.2023.10245215>
26. Ng, D. T. K., Luo, W., Chan, H. M. Y., & Chu, S. K. W. (2022). Using digital story writing as a pedagogy to develop AI literacy among primary students. *Computers and Education: Artificial Intelligence*, 3, 100054. <https://doi.org/10.1016/j.caeai.2022.100054>
27. Nick Bostrom. (2016). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
28. Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology*, 39(6), 531–548. [https://doi.org/10.1016/S0022-1031\(03\)00040-4](https://doi.org/10.1016/S0022-1031(03)00040-4)
29. Pellas, N. (2023). The Effects of Generative AI Platforms on Undergraduates' Narrative Intelligence and Writing Self-Efficacy. *Education Sciences*, 13(11), 1155. <https://doi.org/10.3390/educsci13111155>
30. Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
31. Rampley, H., Reynolds, F., & Cordingley, K. (2019). Experiences of creative writing as a serious leisure occupation: An interpretative phenomenological analysis. *Journal of Occupational Science*, 26(4), 511–523. <https://doi.org/10.1080/14427591.2019.1623066>
32. Rane, N. (2024). Enhancing the quality of teaching and learning through ChatGPT and similar large language models: Challenges, future prospects, and ethical considerations in education. *TESOL and Technology Studies*, 5(1), 1–6.

- <https://doi.org/10.48185/tts.v5i1.1000>
33. Roemmele, M., & Gordon, A. S. (2015). Creative Help: A Story Writing Assistant. In H. Schoenau-Fog, L. E. Bruni, S. Louchart, & S. Baceviciute (Eds.), *Interactive Storytelling* (Vol. 9445, pp. 81–92). Springer International Publishing. https://doi.org/10.1007/978-3-319-27036-4_8
 34. Schwartz, S. (2025, June 26). Brain Activity Is Lower for Writers Who Use AI. What That Means for Students. *Education Week*. <https://www.edweek.org/technology/brain-activity-is-lower-for-writers-who-use-ai-what-that-means-for-students/2025/06>
 35. Sheard, J. (with Denny, P.). (2022). *Australasian Computing Education Conference*. Association for Computing Machinery. <https://doi.org/10.1145/3511861>
 36. Taylor Suchy, S. (2024). Artificial Intelligence in Creative Writing Studies: Threat or Opportunity. In L. Barlach (Ed.), *Contemporaneous Issues about Creativity*. IntechOpen. <https://doi.org/10.5772/intechopen.1008429>
 37. Zack Saadioui. (n.d.). *Delving into User Feedback on Gemini's Performance for Academic Writing & Research Tasks*. Retrieved August 17, 2025, from <https://www.arsturn.com/blog/delving-into-user-feedback-on-geminis-performance-for-academic-writing-research-tasks>
 38. Zhou, M., & Peng, S. (2025). The Usage of AI in Teaching and Students' Creativity: The Mediating Role of Learning Engagement and the Moderating Role of AI Literacy. *Behavioral Sciences (Basel, Switzerland)*, 15(5), 587. <https://doi.org/10.3390/bs15050587>