ASEAN Journal of Educational Research and Technology

Journal homepage: https://ejournal.bumipublikasinusantara.id/index.php/ajert

Integrating Generative Artificial Intelligence (AI)-Based Multimodal Learning in Education to Enhance Literacy Aligned with Sustainable Development Goals (SDGs)

Asep Nurjamin[1], Lucky Rahayu Nurjamin[1], Yustika Nur Fajriah[1], Aisyah Khoerunnisa Nurjamin[2], Hijrin Ahmad Firdaus[1]*

[1]Institut Pendidikan Indonesia Garut, Indonesia

[2]KB Al-Uswah, Indonesia

*Correspondence: E-mail: asep5nurjamin@institutpendidikan.ac.id

ABSTRACT

This study aimed to investigate the integration of artificial intelligence (AI)-based multimodal learning in education to enhance students' literacy in alignment with Sustainable Development Goals. A quasi-experimental design was employed with elementary students who participated in lessons supported by teacher modelling and AI scaffolding. The intervention was implemented through literacy tasks focusing on descriptive writing and multimodal expression. The results indicated that students improved in their ability to construct coherent descriptive texts enriched with visual elements, demonstrating progress in vocabulary use and multimodal literacy. These findings are significant because the combination of teacher guidance and AI feedback strengthened the learning process by connecting linguistic concepts with meaningful tasks. The study highlights the impact of aligning educational innovations with Sustainable Development Goals (SDGs), showing that AI-assisted multimodal learning can foster student-centered pedagogy, improve literacy development, and provide sustainable pathways for integrating digital technologies into education.

© 2026 Bumi Publikasi Nusantara

ARTICLE INFO

Article History:

Submitted/Received 09 May 2024

First Revised 21 Jun 2025

Accepted 23 Aug 2025

First Available online 24 Aug 2025

Publication Date 01 Mar 2026

Keyword:

Deep learning,

Elementary students,

Generative AI,

Multimodal learning.

1. INTRODUCTION

With the rapid development of digital platforms, students are increasingly required to become literate in multimodal forms of communication (Fedorenko, 2019). In their daily lives, they encounter messages presented through various online platforms that combine text, images, audio, and video. Therefore, it is crucial to equip students with basic multimodal literacy skills that help them interpret and produce messages in meaningful ways (Shamim & Riaz, 2023; Chan & Sung, 2025).

In this regard, one of the effective methods is the integration of artificial intelligence (AI). Many reports regarding the use of AI have been well-documented (**Table 1**). Generative AI (GenAI)-based tools have emerged as a promising approach to support multimodal learning in schools (Mittal *et al*., 2017; Cheah *et al.*, 2025). Several recent studies have explored the use of GenAI in education, particularly in multimodal composition and language learning, as presented in **Table 2**. Existing studies demonstrate the growing interest in GenAI for educational purposes. However, most research has focused on higher education or advanced learners, with limited attention to elementary school contexts. There remains a gap in understanding how GenAI can be effectively utilized to support the development of early literacy and multimodal expression among young learners.

Table 1. Previous studies regarding the use of AI.

No	Title	Reference
1	A review of artificial intelligence in security and privacy: Research advances, applications, opportunities, and challenges	Al-Khassawneh, (2023)
2	Chatbot artificial intelligence as educational tools in science and engineering education: A literature review and bibliometric mapping analysis with its advantages and disadvantages	Al Husaeni *et al.* (2024)
3	How bibliometric analysis using VOSviewer based on artificial intelligence data (using ResearchRabbit Data): Explore research trends in hydrology content	Rochman *et al*., (2024)
4	Artificial intelligence (AI)-based learning media: Definition, bibliometric, classification, and issues for enhancing creative thinking in education	Solihat *et al*., (2024)
5	Smart electric resistance welding based on artificial intelligence (AI) based on real-time adaptive statistical features completed with bibliometric analysis	Fufon *et al*., (2025)
6	Primary education undergraduates' competency in the use of artificial intelligence for learning in Kwara State	Agarry *et al.* (2022)
7	Awareness and utilization of artificial intelligence-based intelligent tutoring systems (ITS) in enhancing chemistry education through information and communication technology (ICT)	Abdulmuhsin *et al*., (2024)
8	Trends in the use of artificial intelligence (AI) technology in increasing physical activity	Rahayu & Ismail, (2023)
9	Navigating the future of artificial intelligence: Innovations, ethical dilemmas, and societal impact	Mamaraimova & Safieva (2025)
10	Integrating artificial intelligence into education: Opportunities, challenges, and future prospects	Khalikova (2025)
11	Bibliometric analysis of research trends in conceptual understanding and sustainability awareness through artificial intelligence (AI) and digital learning media	Fiandini *et al*., (2023)
12	The future of learning: Ethical and philosophical implications of artificial intelligence (AI) integration in education	Nurhasanah & Nugraha, (2024)
13	Undergraduate perception, attitude, and utilization of artificial intelligence (AI) ChatGPT for learning: An educational technology perspective	Abdulmumin & Abdulrahman (2025)

Table 1 (continue). Previous studies regarding the use of AI.

No	Title	Reference
14	University students' awareness of, access to, and use of artificial intelligence for learning in Kwara State	Alimi *et al.* (2021)
15	The future of teaching: Artificial intelligence (AI) and artificial general intelligence (AGI) for smarter, adaptive, and data-driven educator training	Balasubramanian (2025)
16	Bibliometric analysis on artificial intelligence research in Indonesia vocational education	Rahmiyanti, (2024)
17	Accessibility and utilization of artificial intelligence (AI)-based intelligent tutoring systems (ITS) and information and communication technology (ICT) in enhancing biology education	Ibrahim *et al.*, (2025)
18	Leveraging artificial intelligence (AI) and geospatial technologies for community-centered urban expansion forecasting in Hyderabad	Nagaraju & Ramakrishna (2025)

Table 2. Previous studies on AI-based multimodal learning

No.	Title	Ref.
1.	Exploring Students' Multimodal Representations of Ideas About Epistemic Reading of Scientific Texts in Generative AI Tools	(Cheung *et al.*, 2025)
2.	Beyond Text-to-Text: An Overview of Multimodal and Generative Artificial Intelligence for Education Using Topic Modelling	(Tao *et al.*, 2025)
3.	When generative artificial intelligence meets multimodal composition: Rethinking the composition process through an AI-assisted design project	(Jiang, 2024)
4.	Multimodal Gen AI: Integrating Text, Image, and Video Analysis for Comprehensive Claims Assessment	(Reddy, 2024)
5.	Integrating generative AI into digital multimodal composition: A study of multicultural second-language classrooms	(Lin *et al.*, 2025)
6.	Multimodal composing with generative AI: Examining preservice teachers' processes and perspectives	(Smith *et al.*, 2025)
7.	Integrating artificial intelligence in literacy lessons for elementary classrooms: a co-design approach	(Kosmas *et al.*, 2025)
8.	Integrating Artificial Intelligence and Multimodality in Language Education: A Systematic Review of Emerging Trends and Practices	(Hamid, 2025)

Addressing this gap, the present study focuses on the utilization of a GenAI-based multimodal learning platform through a classroom intervention with elementary students. Powered by deep learning architectures, such a system enables real-time scaffolding through text generation, visualizations, and adaptive feedback. The purpose of this study is to examine how the combination of teacher modelling and AI scaffolding can improve students' ability to write descriptive texts and employ visualizations to support meaning-making.

In alignment with the current curriculum, which emphasizes student-centered, innovative, and meaningful learning, the intervention positions learners as active participants in constructing knowledge. Students are encouraged to engage with multimodal resources relevant to their daily experiences, while teachers act as facilitators who guide without dominating the process. The novelty of this research lies in demonstrating how advanced AI technologies can be pedagogically aligned with curriculum-oriented directions and in highlighting their contribution to Sustainable Development Goals (SDGs). This contribution shows that integrating AI with classroom instruction can enhance early literacy outcomes and provide sustainable pathways for digital innovation in education.

2. LITERATURE REVIEW

2.1. Generative AI Multimodal

Figure 1 elaborates the Dreamina workflow, which integrates several stages to foster early literacy and multimodal expression among students. The process begins with teacher modelling, a critical phase where teachers explicitly demonstrate how to execute a task, guiding students to recognize essential genre elements like nouns and adjectives. This provides the foundational knowledge students need to succeed (de Oliveira & dos Santos, 2025; Zhou, 2023). Next, students engage in independent creation, where they interact with the Dreamina platform to produce a multimodal output that includes text, images, and audio. As they work, the AI engine provides dynamic scaffolding by generating prompts, examples, and feedback that reinforce the generic features previously modelled by the teacher (Li & Wilson, 2025). The process is not linear but cyclical, with a strong emphasis on iterative refinement. This loop involves analysing student outputs and engagement to inform subsequent adjustments to both the AI model and the instructional strategies. This ensures that Dreamina remains pedagogically and developmentally appropriate. Thus, Dreamina workflow integrates teacher modelling, independent student creation, AI-driven scaffolding, and iterative refinement to foster early literacy and multimodal expression. This cyclical process adapts to student needs, ensuring effective and developmentally suitable learning experiences.

Ultimately, this integrated workflow cultivates desired learning outcomes, including a sophisticated understanding of genre conventions, enhanced descriptive writing skills, and a deeper competence in multimodal communication within a digital context.

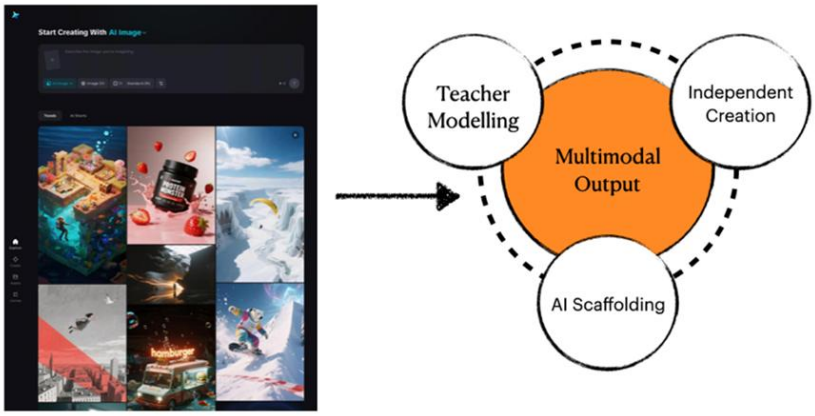

Figure 1. Gen. AI multimodal workflow.

2.2. Systemic Functional Linguistics - Genre-Based Approach (SFL-GBA)

The Systemic Functional Linguistics (SFL) Genre-Based Approach is a four-stage process designed to help students learn to write. This method moves from guided, collaborative activities to independent writing. The four stages are: Building Knowledge of Field (BKoF), Modelling of the Text (MoT), Joint Construction of the Text (JCoT), and Independent Construction of the Text (ICoT).

The SFL Genre-Based Approach begins with BKoF, a crucial phase where students develop a deep understanding of their writing topic. Teachers lead students to immerse themselves in the subject through reading, listening, and speaking activities, with a strong focus on building specialized vocabulary (Rahmawati & Fajriah, 2023). This stage lays the groundwork by having students analyse model texts, dissect new expressions, and explore the cultural

contexts surrounding the topic. Following this, the MoT stage introduces students to exemplary texts. Here, they go beyond simple reading, actively dissecting these models to internalize the structure and linguistic features of the target genre. Teachers guide them to answer comprehension questions and identify key structural elements and linguistic patterns, allowing students to grasp the core characteristics of the text type they are preparing to create.

JCoT, or Joint Construction of the Text, is a pivotal stage where students transition from analysis to action through collaborative writing (Lailatussyifa *et al*., 2025). This isn't just group work; it's a dynamic process where students, either in small groups or as a class, work together to create a text (Derewianka, 2003). Teachers act as essential facilitators, offering explicit guidance, immediate feedback, and careful monitoring to support this shared effort. JCoT is designed to build confidence and skill through a scaffolded, interactive environment. The final stage, ICoT, marks the culmination of the entire process. Here, students take everything they've learned and apply it autonomously (Alhammad, 2025; Nurjamin & Nurjamin, 2016). They are no longer simply imitating or collaborating; they are generating, organizing, drafting, and revising their own written work. While the teacher still provides support, the ultimate goal of ICoT is to cultivate student autonomy and to solidify their proficiency as independent writers.

Figure 2 shows that the cyclical pedagogical framework encompassing BKoF, MoT, JCoT, and ICoT provides a structured and scaffolded approach to writing instruction. The process begins with BKoF, where teachers establish a foundational understanding of the writing topic by engaging students in activities like reading, listening, and discussion to build relevant vocabulary and background knowledge. Next, the MoT stage introduces students to exemplary texts, enabling them to analyse the structure and linguistic features of a specific genre. Following this, the JCoT stage involves collaborative writing, where students work with their peers and teacher to create a text, receiving crucial guidance and feedback. The final stage, ICoT, culminates the process by requiring students to produce a text on their own, applying the skills and knowledge acquired in the preceding stages to foster autonomy and mastery. This cyclical, scaffolded approach empowers students to navigate the writing process effectively, transforming them into competent writers.

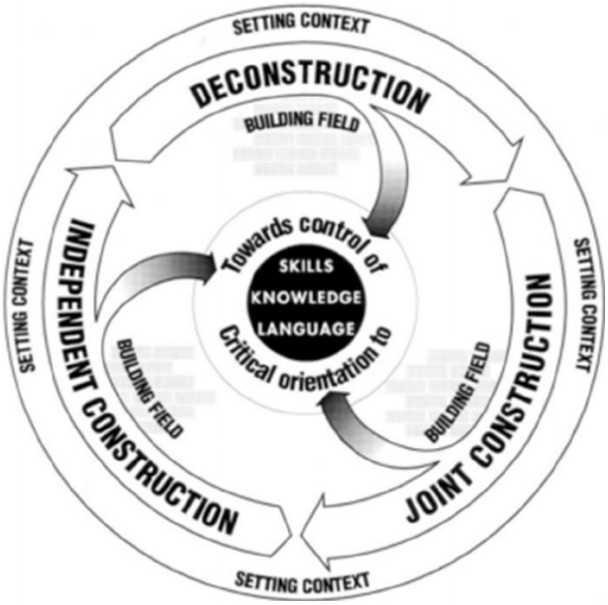

Figure 2. The teaching-learning cycle of GBA.

2.3. The Concept of How Image-Text Relation in Digital Storytelling is Taught

Involvement in a visual image is determined by the angle of representation, which positions the viewer as either a participant or an observer (**Table 3**). An inclusive relationship is created through a frontal angle, where the subject looks directly at the viewer, inviting them into the scene. Conversely, an exclusive relationship is conveyed with an oblique angle, where the subject is turned away, making the viewer a detached outsider (van Leeuwen, 2017). This angle of representation is a key tool for creators to control the viewer's perceived level of intimacy and connection to the image's narrative.

Visual angles also communicate power dynamics and social relationships. A vertical frontal angle, where the viewer looks down on the subject, can imply a sense of power or superiority. In contrast, a horizontal angle places the viewer on an equal level with the subject, signalling a balanced social dynamic and a sense of shared reality (Trisanti *et al*., 2024). This strategic use of vertical and horizontal angles allows the image to subtly communicate the power structures and social connections between the viewer and the subject.

Another crucial element is social distance, which is conveyed by the image shot. The framing of the subject, from a close-up to a long shot, determines the perceived intimacy between the viewer and the subject. A close-up shot suggests a high degree of intimacy and personal connection, while a wider shot creates a sense of detachment. By manipulating the shot type, photographers and artists can control the emotional and social proximity of the viewer to the image's content.

Table 3. Size of frame and social distance

Frame Size	Characteristics	Social Relation
very close up	less than the head and shoulders of the subject	intimate
close shot	head and shoulders of the subject	friendly or personal
medium close	cuts off the subject approximately at the waist	social or ' one of us'
medium shot	cuts off the subject approximately at knee level	'familiar' social
medium long	shows a full figure	general social
long shot	The human figure fills half the image height	public, largely impersonal
very long shot	anything beyond (wider) than half height	little or no social connection

The final aspect to be addressed is power. Power dynamics in visual communication are powerfully conveyed through the strategic use of camera angles. A horizontal angle places the viewer on an equal footing with the subject, suggesting a balanced relationship without dominance. In stark contrast, a high angle positions the viewer looking down on the subject, creating a sense of superiority and control, while a low angle forces the viewer to look up, implying that the subject holds the power or is in a position of authority. These deliberate choices in visual angles fundamentally shape how we perceive the relationship between ourselves and the subject in a photograph or image.

3. METHODS

This study employed a quasi-experimental design with a one-group pretest–posttest model. Detailed information regarding this method is explained elsewhere (Susilawati *et al*., 2025).

This method was used to examine the effectiveness of Dreamina, a generative AI-based multimodal learning model, in improving elementary students' ability to produce descriptive texts in Bahasa Indonesia. The research was conducted in five elementary schools located in Java and Bali, involving a total of 75 students (15 students per school). The intervention lasted

for five consecutive days, during which students engaged in activities such as analysing short descriptive passages, identifying nouns and adjectives, and constructing four simple descriptive sentences with AI-assisted visualization. This design was chosen to capture the short-term impact of multimodal learning on descriptive writing development in authentic classroom contexts without requiring random assignment.

Data were collected through pre-test and post-test writing tasks, where students composed short descriptive texts. The students' texts were scored using a rubric focusing on genre structure (identification and description), vocabulary use (especially nouns and adjectives), sentence accuracy, and the integration of multimodal elements (text-image alignment). Quantitative data were analysed using paired sample t-tests to determine significant improvement between pre-test and post-test scores. This statistical procedure ensured a rigorous examination of the effectiveness of Dreamina in enhancing descriptive text writing skills among elementary students. We analyzed statistics to get a better understanding of the results. Detailed information on how to analyze using statistical analysis is reported elsewhere (Fiandini *et al*., 2024; Rahayu *et al*., 2024; Afifah *et al*., 2022).

4. RESULTS AND DISCUSSION

4.1. Improvement in Writing Descriptive Texts

The improvement of students' writing of descriptive text is visualized in **Table 4**. The writing ability is important, while it is one of the difficult aspects (Nandiyanto *et al*., 2022; Andhini & Sakti, 2021; Olowoyeye *et al*., 2023; Damayanti & Santosa, 2024; Agustin *et al*., 2022). As paired-samples t-test using the combined scores of all 75 students across five schools indicated a significant improvement between the pre-test ($M = 30.5$, $SD = 7.4$) and post-test ($M = 71.3$, $SD = 8.1$), $t(74) = 12.6$, $p < 0.001$, with a large effect size (Cohen's $d = 1.8$).

When analyzed by aspect, **Table 4** shows all dimensions of descriptive writing consistent gains: Identification (+42; +150%), Description (+38; +119%), Nouns (+35; +87%), and Adjectives (+48; +218%). Detailed aspects are in the following:

(i) Adjectives (+48; +218%): As the most impressive gain, the +48 likely represents the average increase in the number of adjectives used by students after the intervention. The +218% is the percentage increase from their initial baseline performance. This dramatic jump suggests that the AI-assisted tool was particularly effective in teaching students how to use descriptive words to enhance their writing.

(ii) Identification (+42; +150%): This metric likely refers to the students' ability to identify the key elements or structure of a descriptive text. The significant increase suggests they gained a much better understanding of how a descriptive piece of writing is put together.

(iii) Description (+38; +119%): This likely measures the overall quality or richness of the descriptions students provided. The substantial gain here shows that they improved their ability to create vivid and detailed descriptions.

(iv) Nouns (+35; +87%): While the percentage gain is lower than the others, an 87% increase in the use of nouns is still a very strong result. This could indicate that students improved their ability to correctly use and identify the subjects or objects in their descriptions.

These results demonstrate that students not only gained awareness of the genre structure of descriptive texts but also improved in their use of linguistic features, particularly adjectives, to describe attributes such as colour, size, and shape. Overall, the findings confirm the effectiveness of the AI-assisted multimodal intervention in enhancing early literacy in Bahasa Indonesia.

The significant improvement observed in students' descriptive writing skills following the AI-assisted multimodal intervention aligns with contemporary research emphasizing the efficacy of multimodal learning strategies. Our study's results, including a large effect size (Cohen's d = 1.8) and substantial gains across all aspects of descriptive writing, corroborate findings from both international and local studies highlighting the positive impact of multimodal approaches on literacy outcomes. International studies demonstrate the potential of AI-assisted multimodal learning in enhancing writing skills. Students' generative AI-assisted writing processes reveal that AI tools support brainstorming, feedback, and scaffolding throughout the writing task explored. It was also found that picture-cued writing tasks combined with generative AI feedback improve descriptive writing skills and language proficiency (Zhuang *et al*., 2025). In addition, a report reviewed 43 studies on AI and multimodal learning analytics, highlighting consistent benefits of AI tools in supporting writing skills and learning efficiency (Arifin *et al*., 2025). Supporting these international findings, local research also indicates the effectiveness of multimodal learning, showing that multimodal learning strategies significantly enhance literacy and numeracy outcomes among prospective biology teachers in Indonesia, suggesting that integrating various instructional modalities can lead to improved learning outcomes across disciplines (Mohammadi, 2025). This also informs that adding technology can improve students' skills and ability, which is in good agreement with the literature (Al Husaeni *et al*., 2024).

Theoretical frameworks further support these findings. According to dual coding theory, processing information through both verbal and visual channels enhances learning. Our study's use of AI-assisted multimodal interventions, including visual aids and AI-generated feedback, likely facilitated deeper cognitive processing, leading to substantial improvements in students' descriptive writing skills. In conclusion, these results underscore the effectiveness of AI-assisted multimodal learning in enhancing descriptive writing skills among early learners, supported by both international and local research, and aligned with established cognitive learning theories.

Table 4. Descriptive writing improvement score

No	Descriptive writing Improvement score	Pre-test (M=30,5)	Post-test (M=71,3)	Gains
1.	Identification (introducing the object)	28	70	+42 (+150%)
2.	Description (providing details)	30	68	+38 (+119)
3.	Use of nouns (naming objects)	39	74	+35 (+87)
4.	Use of adjectives (colour, size, shape, quality)	22	70	+48 (+218)
5.	Multimodal integration (linking text with image idea)	26	74	+48 (+218)

4.2. Enhancement of Multimodal Literacy

Table 5 presents that this AI-assisted multimodal intervention proved successful in improving elementary students' descriptive writing skills. Students showed significant progress in integrating visual cues with textual descriptions, resulting in richer, more detailed, and coherent writing. They also demonstrated an improved use of descriptive vocabulary, incorporating adjectives and phrases that made their writing come to life. Additionally, students became more adept at developing context and environment within their texts,

adding details about location, time, and sensory elements. The simple sentences that were previously dominant evolved into more complex constructions, reflecting a substantial increase in their multimodal literacy.

These improvements align with various studies that support the effectiveness of multimodal learning strategies in elementary education. Research highlights that integrating sensory techniques is crucial for enhancing student engagement and comprehension (Guo, 2023). Similarly, multimodal literacy pedagogies have been shown to support young learners in developing critical thinking and expressive skills (Ntelioglou *et al*., 2014). Furthermore, advancements in AI-assisted learning tools have shown promise in improving students' writing outcomes. One study (Jiang, 2024) found that AI text generators can motivate students and improve their writing performance by providing personalized feedback and suggestions. These findings underscore the effectiveness of combining multimodal pedagogies with AI tools to foster comprehensive literacy development in elementary students.

Table 5. The students' writing improvement on multimodality

Type of improvement	Pre-test (ID/EN)	Post test (ID/EN)	Literacy aspects enhanced
Descriptive Vocabulary	*Kucing tidur* / "The cat sleeps."	*Kucing saya tidur di sofa yang empuk. Matanya biru dan bulunya lembut* / My cat sleeps on the soft sofa. Its eyes are blue and its fur is soft.	Added adjectives and descriptive phrases, richer vocabulary
Context / Setting /	*Kelinci makan* / "The rabbit eats."	*Kelinci putih kecil sedang makan wortel di kandang yang bersih*/ The small white rabbit is eating a carrot in a clean cage	Added location, time, and environmental details
Sensory Details	*Burung terbang* / "The bird flies."	*Burung merah itu terbang di langit pagi sambil berkicau* / "The red bird flies high in the morning sky while singing.	Added sensory information (sound, sight); richer engagement with text
Visual / Color Elaboration	*Kupu-kupu cantik* / "The butterfly is beautiful."	*Kupu-kupu biru dengan saya pindah hinggap di bunga merah*/ The blue butterfly with beautiful patterned wings lands on a red flower in the garden."	Added colours and visual details; enhanced alignment between text and imagery
Expansion of Simple Sentences	*Gajah besar* / "The elephant is big.	*Gajah abu-abu itu berjalan perlahan di hutan*/ The gray elephant walks slowly in the forest.	Extended simple sentences into complex ones; richer structure and detail

4.3. Teacher-AI Synergy

The findings, as mentioned in the previous part, highlight the importance of Teacher–AI Synergy in facilitating this progress. Teachers played a central role in guiding students to recognize literacy points, providing text examples, and modelling how verbal and visual elements can be integrated to convey meaning. Through teacher scaffolding, students learned to connect adjectives, nouns, and sentence structures with relevant visual cues. For instance, teachers demonstrated how a character's physical features or actions could be described accurately while corresponding to images, enabling students to internalize the principles of multimodal literacy. Importantly, this synergy ensured that AI was not merely a

tool for exposure; rather, the combined guidance of teacher and AI allowed students to engage more deeply with the learning material (Zhang & Zhang, 2024). Students progressed from producing disconnected sentences to creating coherent, multimodal texts, demonstrating that teacher-led scaffolding alongside AI feedback fosters more meaningful and developmentally appropriate learning in early literacy (see **Figure 3**) (Xu, 2024).

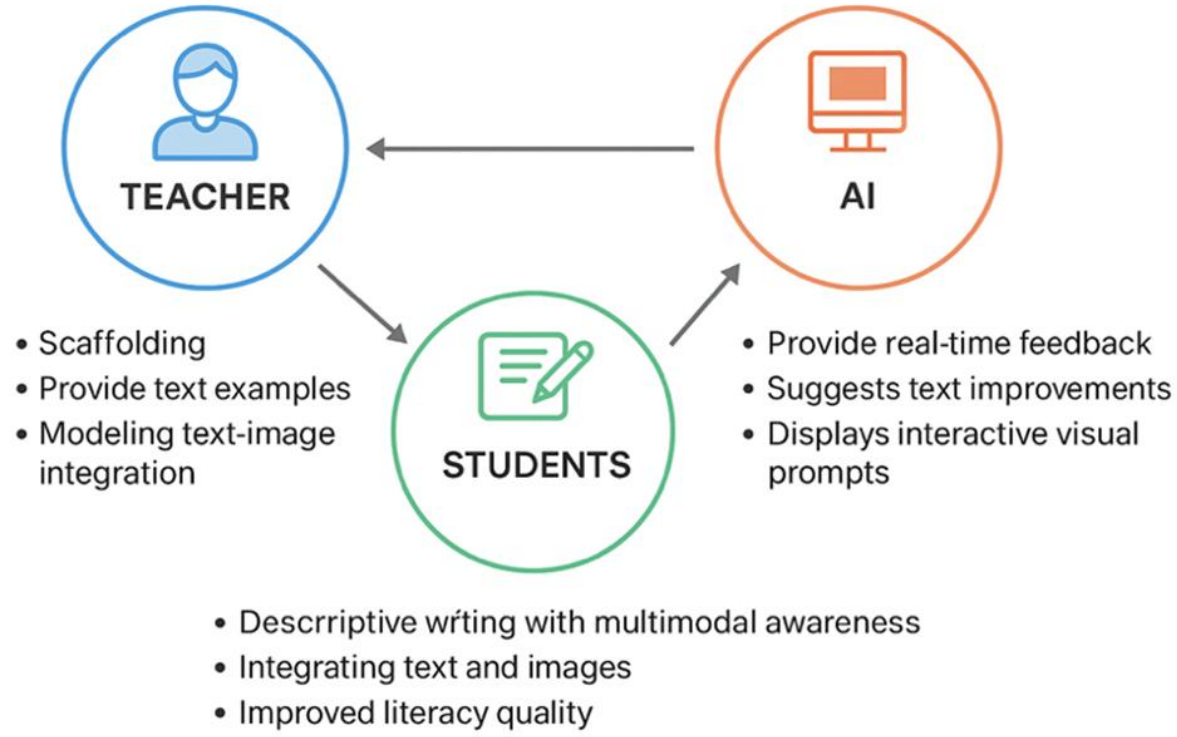

Figure 3. Visualization of Teacher-AI Sinergy.

The findings indicate that teacher guidance is pivotal in helping elementary students develop multimodal literacy. During pre-test activities, students often produced simple sentences with little connection to visual prompts, reflecting limited awareness of how text and images work together. Post-intervention, with teacher scaffolding, students began integrating visual cues with descriptive language, using adjectives and coherent sentences that aligned with the images (Dimitriadou & Lanitis, 2023; Qazi *et al*., 2024). This demonstrates that AI alone is insufficient for meaningful learning; teacher involvement ensures that students internalize the connections between verbal and visual modes (Purba *et al*., 2023).

AI tools, such as Dreamina, acted as facilitators of personalized learning, providing real-time feedback and adaptive scaffolds that guided students in aligning text with images. Studies have shown that AI can enhance literacy outcomes when integrated thoughtfully, offering individualized support that complements teacher instruction (Kosmas *et al*., 2025). By combining AI feedback with teacher modelling, students were able to move from producing minimal descriptions to creating coherent, multimodal texts that accurately represented the visual stimuli as an integration of multiple semiotic resources (Nurjamin *et al*., 2025). Finally, the synergy between teachers and AI underscores the importance of professional development in AI literacy. Educators who understand how to leverage AI tools can design learning experiences that are developmentally appropriate and pedagogically meaningful (Kelley & Wenzel, 2025). This approach not only strengthens early literacy skills

but also fosters students' awareness of how verbal and visual information can be coherently combined, preparing them for more advanced multimodal literacy tasks in the future.

4.4. Relevance to SDGs

The findings of this study also hold strong relevance to the global educational agenda outlined in the SDGs. **Figure 3**, which illustrates the synergy between teachers and AI in multimodal learning, can be further interpreted in the context of international commitments to equitable and quality education. This perspective allows the intervention not only to be viewed as a local classroom innovation but also as a contribution to the broader objective of sustainable human development. Detailed explanations are in the following:

(i) First, the results directly contribute to SDG 4: Quality Education, which aims to ensure inclusive and equitable education for all. The significant improvement in students' descriptive writing skills and multimodal literacy indicates that AI-assisted pedagogy can support foundational learning outcomes at the primary level. Because literacy is a cornerstone of lifelong learning, strengthening these skills in early education has long-term implications for students' future academic success. The intervention also demonstrated that innovative approaches can be embedded in regular classroom practice without replacing teachers, thereby ensuring that pedagogical integrity is maintained.

(ii) Second, the integration of advanced technologies into learning activities aligns with SDG 9: Industry, Innovation, and Infrastructure. The use of generative AI platforms in education illustrates how digital infrastructures can be adapted for pedagogical purposes, bridging the gap between technological development and classroom practice. This is significant because many discussions on AI remain focused on industrial or commercial applications, while this study provides evidence that educational contexts can also benefit from technological innovation. When schools adopt and adapt AI systems in developmentally appropriate ways, they contribute to fostering a generation that is not only literate but also digitally competent.

(iii) Third, the intervention contributes to SDG 10: Reduced Inequalities, particularly in terms of access to learning opportunities. Traditional classroom environments often struggle to provide individualized support due to large class sizes or limited teaching resources. The AI system in this study provided real-time, adaptive scaffolding that complemented teacher instruction. Because each student received personalized prompts and feedback, the intervention helped reduce disparities in participation and achievement. Such practices can be especially valuable in diverse classrooms, where students may come from different linguistic, cultural, or socio-economic backgrounds. By integrating AI into instruction, schools can reduce barriers to quality education and ensure that no child is left behind.

Furthermore, the intervention supports the holistic vision of the SDGs by fostering critical twenty-first-century skills such as creativity, communication, and digital literacy. Students who engaged with multimodal resources and visual prompts developed not only linguistic abilities but also multimodal awareness that reflects the demands of modern communication. This resonates with the broader goals of the SDGs, which seek to prepare individuals to participate actively in knowledge-based societies. In this sense, the study provides practical evidence that AI-assisted pedagogy contributes to building the human capital necessary for sustainable development.

The results also emphasize the need for teacher professional development in the era of AI. While the AI system provided immediate scaffolding, the role of teachers in guiding, contextualizing, and moderating the learning process remained central. This implies that to fully realize the potential of AI in achieving the SDGs, investment in teacher training and institutional support is essential. Without such efforts, the integration of advanced technologies may remain fragmented or uneven, limiting their capacity to reduce inequalities and improve the quality of education.

The intervention presented in this study demonstrates that AI-assisted multimodal learning is not only effective in improving literacy outcomes but also strongly aligned with the Sustainable Development Goals. It advances SDG 4 by strengthening early literacy, supports SDG 9 by integrating technological innovation into education, and contributes to SDG 10 by providing equitable access to individualized learning. These connections show that classroom-level innovations can have broader global significance, linking everyday teaching practices to the long-term vision of sustainable, inclusive, and high-quality education for all.

Finally, this study adds new information regarding SDGs as reported elsewhere (Djirong *et al*., 2024; Ragadhita *et al*., 2026; Kerans *et al*., 2024; Al Husaeni & Haristiani, 2025; Prasetyo *et al*., 2025; Makinde *et al*., 2024; Gemil *et al*., 2024; Keisyafa *et al*., 2024).

4. CONCLUSION

This study demonstrated that generative AI-based multimodal learning improved elementary students' descriptive writing and multimodal expression because the combination of teacher modelling and AI scaffolding enabled learners to connect linguistic concepts with visual representations. The results confirmed that technology does not replace teachers but enhances pedagogy by fostering student-centered and meaningful learning. These findings are significant for education as they align with Sustainable Development Goals, particularly quality education and reduced inequalities. The study provides evidence that integrating AI into classrooms creates sustainable pathways for literacy development in early education.

5. ACKNOWLEDGMENT

We extend our appreciation to the Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat (DRTPM) of the Ministry of Education, Culture, Research, and Technology of Indonesia under grant number 106/E5/PG.02.00.PL/2024 and to Lembaga Penelitian (LEMLIT) Institut Pendidikan Indonesia, Garut, under grant number 05/K-Penelitian/Lemlit-IPI/VI/2024.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

7. REFERENCES

Abdulmuhsin, I., Bolaji, H.O., Abdulazeez, B.T.A., and Salamia, A.O. Awareness and utilization of artificial intelligence-based intelligent tutoring systems (ITS) in enhancing chemistry education through information and communication technology (ICT). *ASEAN Journal of Educational Research and Technology*, *4*(2), 171-186.

Abdulmumin, I.A., and Abdulrahman, M.R. (2025). Undergraduate perception, attitude, and utilization of artificial intelligence (AI) ChatGPT for learning: An educational technology perspective. *Indonesian Journal of Multidiciplinary Research*, *5*(1), 175-186

Afifah, S., Mudzakir, A., and Nandiyanto, A.B.D. (2022). How to calculate paired sample t-test using SPSS software: From step-by-step processing for users to the practical examples in the analysis of the effect of application anti-fire bamboo teaching materials on student learning outcomes. *Indonesian Journal of Teaching in Science*, *2*(1), 81-92.

Agarry, R.O., Omolafe, E.V., Animashaun, V.O., and Babalola, E.O. (2022). Primary education undergraduates' competency in the use of artificial intelligence for learning in Kwara State. *ASEAN Journal of Educational Research and Technology*, *1*(2), 111-118.

Agustin, M.I., Muktiarni, M., and Mupita, J. (2022). Implementation of the reading and writing literacy program to elementary school students about the impact of the natural disaster of mount Merapi. *Indonesian Journal of Teaching in Science*, *2*(2), 99-106.

Al Husaeni, D.F., Al Husaeni, D.N., Nandiyanto, A.B.D., Rokhman, M., Chalim, S., Chano, J., Al Obaidi, A.S.M., and Roestamy, M. (2024). How technology can change educational research? Definition, factors for improving quality of education and computational bibliometric analysis. *ASEAN Journal of Science and Engineering*, *4*(2), 127-166.

Al Husaeni, D.F., Haristiani, N., Wahyudin, W., and Rasim, R. (2024). Chatbot artificial intelligence as educational tools in science and engineering education: A literature review and bibliometric mapping analysis with its advantages and disadvantages. *ASEAN Journal of Science and Engineering*, *4*(1), 93-118.

Al Husaeni, D.N., and Haristiani, N. (2025). What evidence supports the advancement of language learning through digital innovation? Toward achieving Sustainable Development Goals (SDGs) in the 21st century completed with bibliometric analysis. *ASEAN Journal of Science and Engineering*, *5*(2), 327-356

Alhammad, A. I. (2025). Implementing a genre-based approach to enhance essay writing skills in Saudi EFL learners. *Journal of Language Teaching and Research, 16*(1), 295-301.

Alimi, A.E., Buraimoh, O.F., Aladesusi, G.A., and Babalola, E.O. (2021). University students' awareness of, access to, and use of artificial intelligence for learning in Kwara State. *Indonesian Journal of Teaching in Scienc*e, *1*(2), 91-104.

Al-Khassawneh, Y.A. (2023). A review of artificial intelligence in security and privacy: Research advances, applications, opportunities, and challenges. *Indonesian Journal of Science and Technology*, *8*(1), 79-96.

Andhini, A.B., and Sakti, A.W. (2021). Impact of distance learning on reading and writing ability in elementary school students. *Indonesian Journal of Multidiciplinary Research, 1*(2), 393-398.

Arifin, K., Sirih, M., Munir, A., Jahidin, and Sabilu, M. (2025). The influence of multimodal learning strategies on prospective biology teachers' literacy-numeracy learning outcomes. *EURASIA Journal of Mathematics, Science and Technology Education, 21*(1), em2563.

Balasubramanian, K. (2025). The future of teaching: Artificial intelligence (AI) and artificial general intelligence (AGI) for smarter, adaptive, and data-driven educator training. *Indonesian Journal of Teaching in Science*, *5*(1), 49-54.

Chan, A. Y. W., and Sung, C. C. M. (2025). Enhancing students' digital literacy skills through their technology use in a course-based research project: A Hong Kong case study. *Asia Pacific Education Review, 2025*, 1-14.

Cheah, Y.H., Jingru, L., and Kim, J. (2025). Integrating generative artificial intelligence in K-12 education: Examining teachers' preparedness, practices, and barriers. *Computers and Education: Artificial Intelligence, 8*, 1-12.

Cheung, K. K. C., Pun, J., Kenneth-Li, W., and Mai, J. (2025). Exploring students' multimodal representations of ideas about epistemic reading of scientific texts in generative AI tools. *Journal of Science Education and Technology, 34*(2), 284-297.

Damayanti, K.D., and Santosa, M.H. (2024). The effect of prowritingaid writing tool on the writing skills of english as a foreign language (EFL) students: A systematic literature review. *Indonesian Journal of Multidiciplinary Research*, *4*(2), 469-478.

De Oliveira, L. C., and dos Santos, A. E. (2025). Using AI-text generated mentor texts for genre-based pedagogy in second language writing. *Journal of Second Language Writing, 67*, 101184.

Derewianka, B. (2003). Trends and issues in genre-based approaches. *RELC Journal, 34*(2), 133-154.

Dimitriadou, E., and Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments, 10*(1), 12.

Fedorenko, S.V. (2019). Experience of developing students' multimodal literacy in the digital learning environment of higher education institutions. *Information Technologies and Learning Tools, 69*(1), 12-20.

Fiandini, M., Nandiyanto, A.B.D., Al Husaeni, D.F., Al Husaeni, D.N., and Mushiban, M. (2024). How to calculate statistics for significant difference test using SPSS: Understanding students comprehension on the concept of steam engines as power plant. *Indonesian Journal of Science and Technology*, *9*(1), 45-108.

Fiandini, M., Nandiyanto, A.B.D., and Kurniawan, T. (2023). Bibliometric analysis of research trends in conceptual understanding and sustainability awareness through artificial intelligence (AI) and digital learning media. *Indonesian Journal of Multidiciplinary Research*, *3*(2), 477-486.

Fufon, K., Notesiri, U., Lekchaum, K., Chomsuwan, K., Tanitteerapan, T., and Puengsungewan, S. (2025). Smart electric resistance welding based on artificial intelligence (AI) based on real-time adaptive statistical features completed with bibliometric analysis. *ASEAN Journal of Science and Engineering*, *5*(1), 173-198.

Gemil, K.W., Na'ila, D.S., Ardila, N.Z., and Sarahah, Z.U. (2024). The relationship of vocational education skills in agribusiness processing agricultural products in achieving sustainable development goals (SDGs). *ASEAN Journal of Science and Engineering Education*, *4*(2), 181-192.

Guo, X. (2023). Multimodality in language education: Implications of a multimodal affective perspective in foreign language teaching. *Frontiers in Psychology, 14*, 1283625.

Hamid, S. (2025). Integrating artificial intelligence and multimodality in language education: A systematic review of emerging trends and practices. *Journal of Social and Amp; Organizational Matters, 4*(2), 400-416.

Ibrahim, A., Bolaji, H.O., and Abdulraheem, A.J. (2025). Accessibility and utilization of artificial intelligence (AI)-based intelligent tutoring systems (ITS) and information and communication technology (ICT) in enhancing biology education. *ASEAN Journal for Science Education*, *4*(2), 93-104.

Jiang, J. (2024). When generative artificial intelligence meets multimodal composition: Rethinking the composition process through an AI-assisted design project. *Computers and Composition, 74*, 102883.

Keisyafa, A., Sunarya, D.N., Aghniya, S.M., and Maula, S.P. (2024). Analysis of student's awareness of sustainable diet in reducing carbon footprint to support sustainable development goals (SDGs) 2030. *ASEAN Journal of Agricultural and Food Engineering*, *3*(1), 67-74.

Kelley, M., and Wenzel, S. (2025). Advancing teacher AI literacy: Professional partnerships for effective integration of AI in education. *Education Sciences, 15*(6), 659.

Khalikova, K.A. (2025). Integrating artificial intelligence into education: Opportunities, challenges, and future prospects. *Indonesian Journal of Educational Research and Technology*, *5*(2), 137-144.

Kosmas, P., Nisiforou, E. A., Kounnapi, E., Sophocleous, S., and Theophanous, G. (2025). Integrating artificial intelligence in literacy lessons for elementary classrooms: A co-design approach. *Educational Technology Research and Development, 2025*, 1-27.

Lailatussyifa, N.P., Al Baekani, A.K., and Hakim, P.K. (2025). Implementing genre-based approach in teaching reading: a case study at vocational high school students. *Journal of Educational Sciences, 9*(1), 120-128.

Li, M., and Wilson, J. (2025). AI-Integrated scaffolding to enhance agency and creativity in k-12 English language learners: A systematic review. *Information, 16*(7), 519.

Lin, C. H., Zhou, K., Li, L., and Sun, L. (2025). Integrating generative AI into digital multimodal composition: A study of multicultural second-language classrooms. *Computers and Composition, 75*, 102895.

Makinde, S.O., Ajani, Y.A., and Abdulrahman, M.R. (2024). Smart learning as transformative impact of technology: A paradigm for accomplishing sustainable development goals (SDGs) in education. *Indonesian Journal of Educational Research and Technology*, *4*(3), 213-224.

Mamaraimova, D., and Safieva, M. (2025). Navigating the future of artificial intelligence: Innovations, ethical dilemmas, and societal impact. *Indonesian Journal of Educational Research and Technology*, *5*(1), 71-76.

Mittal, U., Sai, S., Chamola, V., and Sangwan, D. (2017). A comprehensive review on generative AI for education. *IEEE Access. 12*, 142733-142759.

Mohammadi, M. (2025). Artificial intelligence in multimodal learning analytics: A systematic literature review. *Computers and Education: Artificial Intelligence, 8*, 100426.

Nagaraju, K., and Ramakrishna, S. (2025). Leveraging artificial intelligence (AI) and geospatial technologies for community-centered urban expansion forecasting in Hyderabad. *ASEAN Journal of Community Service and Education*, *3*(2), 135-146

Nandiyanto, A. B. D., and Azizah, N. N. (2022). Analysis of common mistake in writing scientific article in "engineering" undergraduate student. *Journal of Engineering Science and Technology*, *17*(1), 156-170.

Ntelioglou, B. Y., Fannin, J., Montanera, M., and Cummins, J. (2014). A multilingual and multimodal approach to literacy teaching and learning in urban education: A collaborative inquiry project in an inner-city elementary school. *Frontiers in Psychology, 5*, 533.

Nurhasanah, S., and Nugraha, M.S. (2024). The future of learning: ethical and philosophical implications of artificial intelligence (AI) integration in education. *Indonesian Journal of Multidiciplinary Research*, *4*(2), 341-352.

Nurjamin, A., and Nurjamin, L. R. (2016). The use of genre-based approach to teaching students' writing skill of hortatory exposition text in Bahasa Indonesia. *In Ninth International Conference on Applied Linguistics, 9*, 221-223.

Nurjamin, A., Nurjamin, L. R., Fajriah, Y. N., Nurjamin, A. K. Nugraha, I. (2025). Developing and evaluating an augmented reality (AR) digital storytelling video to foster multimodal literacy and narrative comprehension. *Journal of Engineering Science and Technology*, *20*(4) 919-956.

Olowoyeye, C.A.C., Deji-Afuye, O.O., and Aladesusi, G.A. (2023). Effect of multimedia instructional approach on English writing performance of pre-service technical teachers in South-Western Nigeria. *Indonesian Journal of Multidiciplinary Research, 3*(1), 21-30.

Prasetyo, I., Nugroho, A.A., Damayanto, A., and Fathoni, M.I. (2025). Developing an inclusive ICT-based academic information system using REST API to promote Sustainable Development Goals (SDGs) in higher education. *ASEAN Journal of Science and Engineering*, *5*(2), 395-416.

Purba, R., Herman, H., Fatmawati, E., Saputra, N., and Sabata, Y. N. (2023). The multimodal-based learning improves students' reading ability in perspective of systemic functional linguistics. *Journal of Education Technology, 7*(1), 61-70.

Qazi, S., Kadri, M. B., Naveed, M., Khawaja, B. A., Khan, S. Z., Alam, M. M., and Su'ud, M. M. (2024). AI-Driven learning management systems: Modern developments, challenges and future trends during the age of ChatGPT. *Computers, Materials and Continua, 80*(2), 3289-3314.

Rahayu, N.I., and Ismail, A. (2023). Trends in the use of artificial intelligence (AI) technology in increasing physical activity. *Indonesian Journal of Educational Research and Technology*, *3*(3), 295-304.

Rahayu, N.I., Muktiarni, M., and Hidayat, Y. (2024). An application of statistical testing: A guide to basic parametric statistics in educational research using SPSS. *ASEAN Journal of Science and Engineering*, *4*(3), 569-582.

Rahmawati, T., and Fajriah, Y.N. (2023). The implementation of building knowledge of the field (Bkof) in teaching narrative text. *English Education and Applied Linguistics Journal (EEAL Journal), 6*(1), 10-18.

Rahmiyanti, H. (2024). Bibliometric analysis on artificial intelligence research in Indonesia vocational education. *ASEAN Journal for Science Education*, *3*(2), 183-192.

Reddy, A. S. (2024). Multimodal Gen AI: Integrating Text, Image, and Video Analysis for Comprehensive Claims Assessment. *ESP International Journal of Advancements in Computational Technology, 2*(2), 133-141.

Rochman, S., Rustaman, N., Ramalis, T.R., Amri, K., Zukmadini, A.Y., Ismail, I., and Putra, A.H. (2024). How bibliometric analysis using VOSviewer based on artificial intelligence data (using ResearchRabbit Data): Explore research trends in hydrology content. *ASEAN Journal of Science and Engineering*, *4*(2), 251-294.

Shamim, F., and Riaz, M. N. (2023). Impact of multimodal digital media communication on generation Z's language use and literacy practices. *Academy of Education and Social Sciences Review, 3*(4), 398–409.

Smith, B. E., Shimizu, A. Y., Burriss, S. K., Hundley, M., and Pendergrass, E. (2025). Multimodal composing with generative AI: Examining preservice teachers' processes and perspectives. *Computers and Composition, 75*, 102896.

Solihat, A.N., Dahlan, D., Kusnendi, K., Susetyo, B., and Al Obaidi, A.S.M. (2024). Artificial intelligence (AI)-based learning media: Definition, bibliometric, classification, and issues for enhancing creative thinking in education. *ASEAN Journal of Science and Engineering*, *4*(3), 349-382.

Susilawati, A., Al-Obaidi, A.S.M., Abduh, A., Irwansyah, F.S., and Nandiyanto, A.B.D. (2025). How to do research methodology: From literature review, bibliometric, step-by-step research stages, to practical examples in science and engineering education. *Indonesian Journal of Science and Technology, 10*(1), 1-40

Tao, L., Deng, H., and Song, Y. (2025). Generative artificial intelligence in education: A topic-based bibliometric analysis. *Educational Technology and Society, 28*(2), p327.

Trisanti, N., Sukyadi, D., and Hermawan, B. (2024). University EFL teachers' stance and engagement in digital multimodal composing as multimodal literacy practices. *International Journal of Education, 17*(1), 61-70.

Van Leeuwen, T. (2017). Multimodal literacy. *Metaphor, 4*, 17-23.

Xu, Z. (2024). AI in education: Enhancing learning experiences and student outcomes. *Applied and Computational Engineering, 51*(1), 104-111.

Zhang, J., and Zhang, Z. (2024). AI in teacher education: Unlocking new dimensions in teaching support, inclusive learning, and digital literacy. *Journal of Computer Assisted Learning, 40*(4), 1871-1885.

Zhou, X. (2023). Teaching activities for English writing: based on a process genre. *English Language Teaching and Linguistics Studies, 5*(2), 208-222.

Zhuang, Y., Zhao, R., Xie, Z., and Yu, P. L. H. (2025). Enhancing language learning through generative AI feedback on picture-cued writing tasks. *Computers and Education: Artificial Intelligence, 9*, 100450.