MESR
منار الشرق للدراسات والأبحاث
Manar Elsharq for Studies & Research
Middle East Journal
of Educational and Psychological Sciences
Homepage:
http://meijournals.com/ar/index.php/mejeps/index
مجلة الشرق الأوسط
للعلوم التربوية والنفسية
ISSN 2789-1828

التحول التعليمي من خلال الذكاء الاصطناعي التوليدي متعدد الوسائط: نظام جيميني بنانا

Educational transformation through multimedia generative artificial intelligence: The Gemini Banana System

د. شذى صخر[1]، د. لؤي حميدي قطيش الفواعرة[2]

Dr. Shatha Sakher[1] & Dr. Loiy Hamidi Qutaish Alfawa'ra[2]

كلية الطب، جامعة البلقاء التطبيقية، السلط، الأردن[1] ، قسم اللغة الإنجليزية – جامعة عليكرة الإسلامية[2]

Faculty of Medicine, Al-Balqa Applied University, As-Salt, Jordan[1] & Department of English – Aligarh Muslim University[2]

استلام البحث: 04/01/2026 مراجعة البحث: 18/01/2026 قبول البحث: 09/02/2026

ملخص الدراسة:

يقدّم هذا البحث توليفًا للآثار التربوية المترتبة على دمج الذكاء الاصطناعي (AI) ضمن إطارَي Gemini Banana (إنتاج الفيديو) و Nano Banana (معالجة الصور) في تصميم التعليم المعاصر. وباستخدام منهجية التحليل البعدي النوعي، يقوم البحث بتحليل الوثائق التقنية والأدبيات التجريبية المنشورة بين عامي 2023 و2025 لتقييم إمكانات وتحديات الذكاء الاصطناعي التوليدي متعدد الوسائط.

وتشير النتائج الرئيسة إلى أن هذه الأدوات تُسهم بشكل ملحوظ في تحسين كفاءة التعليم وإتاحته، من خلال أتمتة إنشاء المحتوى، ودعم التعليم الحساس للغة والثقافة، وتمكين النمذجة السريعة للمواد التعليمية. وفي حين تُظهر المعلمون المُولَّدون بالذكاء الاصطناعي فاعلية في ترسيخ المعرفة، ولا سيما في مهام التعلم الإجرائي وتعلّم اللغات (مبدأ التكافؤ)، فإنهم يفتقرون إلى الحضور الاجتماعي والتفاعل التعاطفي اللذين يميّزان المعلم البشري.

ويخلص البحث إلى اقتراح إطار PATH المرتكز على الإنسان (الترويج، والتطوير، والتدريب، والتوظيف) من أجل الدمج الأخلاقي لهذه التقنيات، ويقدّم توصيات استراتيجية، مؤكدًا أن الذكاء الاصطناعي التوليدي متعدد الوسائط ينبغي أن يكون أداة تمكين لقدرات المعلم البشري، لا بديلًا عنه.

الكلمات المفتاحية: الذكاء الاصطناعي التوليدي، جيميني، جيميني بنانا، نانو بنانا، التصميم التعليمي، التعلم متعدد الوسائط، إطار عمل PATH، التكنولوجيا التعليمية، التدريس المستجيب ثقافيًا ولغويًا (CLRT).

Abstract

This research presents a synthesis of the educational implications of integrating artificial intelligence (AI) within the Gemini Banana (video production) and Nano Banana (image processing) frameworks in contemporary instructional design. Using a qualitative meta-analysis methodology, this study analyzes technical documentation and empirical literature published between 2023 and 2025 to assess the potential and challenges of multimedia generative AI. The main findings indicate that these tools significantly improve the efficiency and accessibility of education by automating content creation, supporting language- and culture-sensitive instruction, and enabling rapid prototyping of educational materials. While AI-generated teachers demonstrate effectiveness in knowledge consolidation, particularly in procedural and language learning tasks (equivalence principle), they lack the social presence and empathetic interaction characteristic of human teachers. The study concludes by proposing a human-centered PATH framework (promotion, advancement, training, and harnessing) for ethical integration, and offers strategic recommendations, arguing that MGAI should serve as an enabler, not a replacement, for human teacher capacity.

Keywords: Generative AI, Gemini, Gemini Banana, Nano Banana, Instructional Design, Multimodal Learning, PATH Framework, Educational Technology, Culturally and Linguistically Responsive Teaching (CLRT).

Introduction

Generative artificial intelligence (GAI) is changing how we teach. Instead of providing the student with a static teaching method, (GAI) provides teachers with tools to create personalized, co-created, and dynamic learning experiences for their students to work together to create a more holistic learning experience. This new type of learning experience has changed how educators view their role in the educational process; educators now have access to a suite of tools that were previously only available to those with specialized training in media production. In this context, Gemini Banana and Nano Banana provide educators with powerful tools and knowledge for the development of engaging and motivating educational experiences. In this paper, the authors provide an overview and synthesis of how these two systems work together to enhance learning outcomes, enhance student motivation, reduce teacher workload, and present ethical considerations for educators.

1.1 The Evolution of Gemini and the Nano Banana Engine

The journey of Gemini began as Google's ambitious response to the rapidly evolving landscape of generative AI. Officially introduced by Google DeepMind in late 2023 and early 2024, Gemini was designed from the ground up to be multimodal, meaning it could understand and operate across text, code, audio, image, and video. It eventually succeeded in "Bard," marking a shift toward a more integrated and powerful ecosystem of models designed to handle everything from complex reasoning to mobile-based tasks. As the platform matured, the need for specialized visual creativity led to the integration of dedicated generative tools. This is where the Nano Banana model comes into play. While "Gemini" serves as the primary conversational interface and cognitive engine, Nano Banana acts as the sophisticated image-generation powerhouse under the hood. It represents a state-of-the-art leap in text-to-image and image-editing capabilities, allowing for high-fidelity rendering, complex compositions, and even the ability to process text within images with remarkable accuracy. Note: While "Gemini" is the name you see on the front end, "Nano Banana" is the specialized creative engine that brings your visual ideas to life (Google DeepMind,2023).

1. Methodology

Utilizing qualitative research design, this study takes a meta-synthesis of existing research as one of its methodologies. Developed through a 3-phase process:

Phase 1: Gather Data. The data gathered consists of recently published academic studies and reports that are peer-reviewed articles, conference papers, systematic reviews, and reports by governmental bodies regarding Text-to-Video (T2V) and Text-to-Image (T2I) generative Artificial Intelligence Tools in the Classroom (Gemini Banana Ecosystems, Nano Banana Ecosystems).

Phase 2: Thematic Analysis. Inductive coding allowed for identification and analysis of unique and recurring themes. Major themes found were: (1) technological benefits- subjects were kept constant, as well as quality of image/video were held to a standard, and (2) impacts on pedagogy-intangible effects were cognitive loads faced, feeling of connectedness, creative confidence, instructional parity), (3) ethics of deploying these tools, and (4) structure of using these tools as part of integrated curriculum.

Phase 3. Synthesis All technical specifications and capabilities from Gemini Banana Ecosystem were combined with previously established approaches (CLT, CLRT, proposed PATH Model) into a practical model of creating evidence-based practices to support teacher's use of T2V and T2I tools with students.

2. Technological Architecture and Multimodal Affordances

The Gemini Banana ecosystem is designed on top of Gemini's transformer-based models (Google DeepMind, 2025), Gemini 2,5 Flash and Gemini 3 Pro. Both Gemini 2.5 and Gemini 3 Pro have been developed to perform deep visual reasoning and perform complex multimodal processing with the use of text prompts and reference images to produce coherent and high-fidelity images as outputs.

An important feature of using the Gemini Banana system for educational purposes is the ability to maintain the "subject consistency across scenes," where the same character or object maintains its visual identity throughout an entire story. By maintaining this identity, educators can decrease the amount of cognitive load placed on educational participants when they are exposed to visual narratives, thereby allowing them to focus more effectively on their learning experience (Yufeng, 2025).

Table 1: Technical Specifications and Pedagogical Implications of the Gemini Banana Ecosystem

Feature Component	**Technical Specification**	**Pedagogical Implication**
Model Foundation	Gemini 2.5 Flash / 3 Pro	Enables high-speed generation for real-time classroom visual aids and interactive content.

Subject Consistency	Up to 95% accuracy	Allows creation of consistent virtual instructors or pedagogical mascots across lessons.
Resolution Output	2K to 4K Ultra HD	Ensures clarity for large-scale displays, digital whiteboards, and high-quality prints.
Multilingual Support	Diverse font & voice styles	Supports linguistically responsive teaching for diverse student populations.
Multi-Image Reference	Accepts up to 14 images	Facilitates the creation of complex technical illustrations, timelines, and concept maps.

Models on the Gemini 3 level represent a significant upgrade in reasoning capacity with the introduction of features such as "Chain of Thought" reasoning and variable depths of thought. Using these capabilities enables the AI to address and complete very complicated tasks involving logical reasoning, science, and/or programming/computer coding, which will ultimately enable these AI models to produce good representations of highly complex relationships between concepts within educational materials. (Google DeepMind, 2025)

2.1 The Nano Banana Pro Framework: Precision and Control

The Google Image Generation Models for Professional Gemini Applications, called "Nano Banana Pro," have been developed to provide users with more advanced features than those offered with the Basic Tools model. Basic Tools allow simple instructions when creating new images, while Nano Banana Pro provides users with a way to create an image that uses more advanced photo editing techniques, including changing lighting, camera angles, and composition or layout after an image has been created, simply by typing in a request (QudAcademy, 2025). For example, this could allow an Instructor Designer to improve an educational graphic based on comments made by an educator regarding its content's accuracy and visual appeal (QudAcademy, 2025).

2.2 Iterative Design and Creative Control

The approach that Nano Banana uses is based on iterative design. Users who are educators can upload images of different items as well as use text prompts to build complete visual images that are matched well together. For example, if a teacher has an image of a technical device they want to put into a lab setting, they will be able to upload that technical device and request that it be given warm and natural lighting and is done with the highest level of fidelity without needing to have extensive computer software knowledge (as is true for programs like Photoshop) (Artist, 2025). This kind of flexible use provides educators with the power to create educational material that has "local" cultural or contextual relevance to their students (Yufeng, 2025) and at the same time, provide a way to easily accomplish this using spatial decomposition methods, such as those in models such as Unitrans Fer. Spatial decomposition methods allow educational videos/movies, and/or videos, to be separated into three areas or parts: the subjects in the video's foreground, its background, and the motion path (or flow) that are moving through the video (Liu et al., 2025). This feature also gives educators the ability to switch out specific pieces of an image (i.e. a virtual instructor's appearance to be better matched to student demographics etc.) without needing to change the educational content in the background.

3. Pedagogical Effectiveness of AI-Generated Instructional Videos (AIGIV)

The introduction of AIGIV has generated some academic discussion about how effective they may be compared to RVs created by humans. Between 2024-2025, research demonstrated what the authors referred to as the equivalence principle, with no statistically significant difference in the academic success of students taught by an AIGIV compared to being taught by a human instructor through RV format (Arkün-Kocadere & Özhan, 2024).

3.1 Comparing Engagement and Academic Performance

Even though students achieved equivalent cognitive outcomes, focus group interviews revealed that students frequently reported feeling less ""socially present" with virtual instructors than they would if they were being taught by a traditional instructor. According to the research conducted by Arkün-Kocadere and Özhan (2024), the greater level of human video engagement stemmed from subtle emotional cues and natural human movements that generated avatars lack. On the contrary, the lower levels of ""social noise"" associated with the use of avatar technology could have an unintended benefit; that is, by reducing the number of emotional distractions, students may experience less extraneous cognitive load, enabling them to concentrate on factual or technical information (Yufeng, 2025).

Table 2: Comparison of Human vs. AI-Generated Instructors

Criterion	Human Instructor (RV)	AI-Generated Instructor (AIGIV)
Social Presence	High, emotionally impactful.	Low, often perceived as "artificial."
Academic Performance	Consistent improvement from pre- to post-test.	Comparable improvement, sometimes superior in retention tasks.
Visual Engagement	High due to natural facial expressions.	Lower, potentially due to the "uncanny valley" effect.
Cost & Time	High requires recording equipment and setup.	Very low, driven by text-based prompts.
Cognitive Load	May increase due to extraneous social cues.	May decrease due to singular focus on content.

According to Xu's research (2025), AIGIVs provide learners with superior results on vocabulary tests compared to the traditional video method, indicating that a synthetic voice's clarity and uniformity positively influence a learner's auditory memory when learning a new language.

5. Culturally and Linguistically Responsive Teaching (CLRT) through AI

The Gemini Banana ecosystem is a powerful tool for enacting CLRT. In linguistically diverse

classrooms facing a shortage of bilingual teachers, T2V tools can help bridge the gap between student needs and teacher capacity (Rosa, 2025).

5.1 Linguistic Customization and Cultural Representation

The benefits of multilingual voice generation and automatic translation allow for educators to build “read-aloud” lessons that are accompanied by captions synchronized in English and a student’s native language (Yufeng, 2025). In addition, the ability to localize (change clothing, settings, landscapes) images using Nano Banana provides students with visual aids that are reflective of their cultural experiences (Yufeng, 2025). By creating personalized experience for them, students have a greater sense of who they are and where they belong, which is important for motivating them to learn deeply. There still exist challenges associated with how accurate translation can be done for languages that are less common, and the existence of stereotypes in certain types of default virtual assets requires that people continue to review and provide feedback (Rosa, 2025).

6. Developing Creative Confidence and Student Engagement

By incorporating these resources into the student generated content, it creates opportunities for developing not just the transfer of knowledge but also "Creative Confidence." Students learn to transition from being passive consumers to becoming "Smart Producers." They will learn about Prompt Engineering and Iterative Editing to create professional visual products from their abstract ideas (Yufeng, 2025). When students find that the outputs are not as they envisioned, it forces them to develop Critical Thinking and AI Literacy skills. In addition, the Gemini Banana ecosystem is a space that promotes Creative Risk Taking without the unnecessary financial or technical risks associated with illustrating a student's creative vision, thus enhancing Innovation and Reducing the Fear of Failure (Artist, 2025; Zielonka, 2025).

7. Ethical Implementation: The PATH Framework

To effectively and ethically integrate AI into education all over the globe, Brazil, Yang and Van Der Kleij (2025) created what they call a human-centered PATH Framework:

- Promote: The use of AI for any situation when the value of AI (e.g., efficiency, personalization) exceeds the negative impact that standard AI usage has on critical thinking and human connection.
- Advance: Supporting emotional and intellectual wellness of all people through human social interactions. By automating routine administrative responsibilities, more time will be available for mentoring.
- Train: Providing educators and students with sufficient knowledge to complete an adequate critique of the credibility, bias and limitations of generated content using AI.
- Harness: Aligning the use of AI with the ethical principles of Transparency, Privacy, and Fairness; and establishing a "Human-in-the-Loop" for the final evaluations, feedback and ethical oversight of AI-generated instructional materials.

The PATH Framework argues that instead of using AI as a replacement for teachers, AI should serve as a tool to alleviate some of the burdens on educators while maintaining the integrity and quality of education provided by teachers through supplemental human alternatives available to every educator (Brazil et al., 2025).

8. Practical Instructional Strategies

Incorporation of this ecosystem into the different steps of instructional design by teachers can take the following forms:

1. Lesson Design: A storyboard can be made to illustrate the learning pathway or to reconstruct Old World images to add depth to a history lesson.
2. Visualization of Technical and Scientific Principles: Converting a two-dimensional diagram into an artificially generated, three-dimensional image will help clarify a difficult-to-understand concept in STEM (science, technology, engineering, and mathematics).
3. Professional Simulation: Simulate activities in a virtual fitting room or create practice opportunities for students who are going into clothing design and environmental sciences. This experience will allow students to see what their creations would look like in a near-realistic way.
4. Interactive Learning Through Microlearning: Students can participate in short-form video production (such as Reels) as part of interactive microlearning. Micro-content that utilizes an algorithm to generate short-form videos is more engaging and facilitates better attention and effort against longer videos. This type of device helps reduce the cognitive load placed on students (Netland et al., 2025).

9. Governance and Public Perception Challenges

Growing concerns about "AI slop", or the low-quality and deceptively realistic content produced by AI, flood educational platforms (Timsah et al., 2025). The rise of powerful AI models such as OpenAI's Sora creates greater challenges regarding accountability in relation to privacy, the copyright of training data, and increasingly blurred lines between fact and fiction (Zhou et al., 2025). Experts suggest that establishing a regulatory environment, including mandatory digital watermarks for AI-generated content, enhancing laws concerning intellectual property, and public campaigns promoting understanding of AI among educators and learners, could help to alleviate these problems (Zhou et al., 2025). In terms of education, it is important that policymakers continue to focus on ensuring that students maintain an individual's ability to make decisions, and that the role of technology in education is to increase and not diminish students' creativity and critical thought processes.

10. Conclusion and Strategic Recommendations

The Gemini Banana/Nano Banana Ecosystems significantly enhance Education by integrating technology with educational delivery. Through research and analysis from the available literature

and data up to 2025, the following conclusions have been reached and have led to recommendations listed below:

1. Produced High Volume of Quality Resources/Production Efficiency: The introduction of Artificial Intelligence (A.I.) as a resource democratized the ability to produce quality educational resources therefore, allowing independent educators and smaller institutions to rapidly create engaging educational materials at a low cost (Netland et al., 2025).
2. Strategic Use of A.I.'s Equivalent Capabilities: To fully utilize the capabilities of AIGIV, instructional designers need to utilize A.I. for repetitive and/or technical content delivery, while still allowing the human instructor to provide emotional support, mentor complex topics, and develop social learning opportunities (Xu, 2025).
3. Learner-Focused Design: The effectiveness or success of these types of educational tools is mainly based on the quality of the instructional design and the degree of human supervision versus the technical attributes of the model (Netland et al., 2025).
4. Ethical Framework - PATH Framework: Educational Institutions, whether K-12 or College/University must create clear ethical guidelines defining the use of AI in a transparent and accountable manner and provide training to students and educators on Critical AI Literacy (Brazil et al., 2025).

As we look into the future of education, we will begin seeing "Generative Lectures" where learners will have access to AI replicas of their educators to get immediate, customized feedback in the form of questions or answers — resulting in a shift of the traditional educator's position from "providing information" to "creating ways to interactively learn" (Jo et al., 2025). The Gemini Banana ecosystem aims to provide the tools necessary to support and enhance human intellect rather than replace it in achieving greater levels of access, efficiency, and effectiveness in educating all people.

References

Al Kurdi, B., Al Anasweh, M. A., Alzoubi, H. M., Al Fouri, A., Sakher, S., & Alshurideh, M. T. (2023). The impact of digital education on acquiring cybersecurity skills among the students of the faculty of medicine at Al-Balqa Applied University. Migration Letters, 20(5), 370–387. https://doi.org/10.59670/ml.v20i5.3550

Arkün-Kocadere, S., & Özhan, S. Ç. (2024). Video lectures with AI-generated instructors: Low video engagement, same performance as human instructors. *The International Review of Research in Open and Distributed Learning, 25*(3), 350-369.

Artist, A. (2025). *Iterative visual design in education using generative AI*. Creative Tech Press.

Brazil, J., Yang, S., & van der Kleij, F. (2025). *AI systems in teaching and learning: Principles and practical examples*. Australian Council for Educational Research.

Fouri, A. A., aburumman, M. A., Siam, A., Yacoub, H., & Sakher, S. (2024). The impact of artificial intelligence applications on enhancing professional empowerment from the

perspective of medical sector workers. Journal of Law and Sustainable Development, 12(3), e3445. https://doi.org/10.55908/sdgs.v12i3.3445

Google DeepMind. (2023). Gemini: A Family of Highly Capable Multimodal Models. Technical Report. Retrieved from https://storage.googleapis.com/deepmind-media/gemini/gemini_1_report.pdf

Google DeepMind. (2025). *Gemini 2.5 & 3: Technical report on multimodal reasoning capabilities*. Google Research.

Jo, H.-Y., Kim, S., & Park, J. (2025). Generative lecture: Making lecture videos interactive with LLMs and AI clone instructors. *Computers & Education, 215*, 105001.

Liu, X., Chen, Y., & Wang, Z. (2025). *Spatial decomposition for editable educational video generation*. Proceedings of the International Conference on Machine Learning.

Netland, T., Haugland, V., & Olsen, K. (2025). The pedagogical effectiveness of AI-generated microlearning videos. *Computers & Education, 218*, 105123. QudAcademy. (2025). *Nano Banana Pro: A guide for educational content creators*. EdTech Publications.

Rosa, P. (2025). *Text-to-video AI for linguistically responsive teaching: Educators' use and reflections* (Doctoral dissertation). University of Nebraska-Lincoln. Timsah, Y., El-Din, M., & Farsi, N. (2025). Combating 'AI slop' in digital learning environments. *Journal of Educational Technology & Society, 28*(2), 89-104.

Sakher, S., Al Fouri, A., Al Fouri, S., & Alshurideh, M. T. (2023). The reality of artificial intelligence skills among eighth-grade students in public schools. In Proceedings of the 3rd International Conference on Artificial Intelligence and Computer Vision (AICV 2023) (pp. 91–106). Springer. https://doi.org/10.1007/978-3-031-27762-7_9

Xu, X. (2025). From recording to AI-generated instructional videos: A comparison of learning performance and cognitive load. *International Journal of Instructional Media, 52*(1), 45-62.

Yufeng, A. (2025). *Pedagogical applications of the Gemini Banana AI ecosystem: A qualitative study* [Unpublished manuscript]. Department of Educational Technology, Future University.

Zhou, K. Z., Zhang, L., & Zhao, H. (2025). 'Sora is incredible and scary': Public perceptions and governance challenges of text-to-video generative AI models. *Information Research, 30*(iConf), 508–522.

Zielonka, B. A. (2025). How Google's free Nano Banana AI powers visual learning and creative confidence. *AI in Education Review, 7*(4), 22-35.

ional, 23(4), 383–399