

REVIEW

Open Access



# Google Gemini as a next generation AI educational tool: a review of emerging educational technology

Muhammad Imran<sup>1\*</sup>  and Norah Almusharraf<sup>1</sup>

\*Correspondence:  
mimran@psu.edu.sa

<sup>1</sup> Education Research Lab,  
College of Humanities  
and Science, Prince Sultan  
University, Rifa Street,  
11586 Riyadh, Saudi Arabia

## Abstract

This emerging technology report discusses Google Gemini as a multimodal generative AI tool and presents its revolutionary potential for future educational technology. It introduces Gemini and its features, including versatility in processing data from text, image, audio, and video inputs and generating diverse content types. This study discusses recent empirical studies, technology in practice, and the relationship between Gemini technology and the educational landscape. This report further explores Gemini's relevance for future educational endeavors and practical applications in emerging technologies. Also, it discusses the significant challenges and ethical considerations that must be addressed to ensure its responsible and effective integration into the educational landscape.

**Keywords:** Multimodal AI, Google Gemini, Educational technology, Smart learning, Personalized learning, Intelligent tutoring systems

## Introduction

The past year has witnessed massive growth in artificial intelligence (AI) systems and their unprecedented impact on human creativity and productivity (Ali et al., 2023; Badshah et al., 2023). OpenAI's development of large language models (LLMs) like GPT-3 paved the way for explosive growth in innovative AI chatbots, such as ChatGPT-3.5. However, LLMs have made significant progress and have gone beyond the unimodal input methods, where they only perform a particular task like text or speech recognition. At present, multimodal AI tools and language models have the ability to interact with and identify various inputs interleaved with a variety of text, images, audio, video, and PDFs. These multimodal are ChatGPT-4 or ChatGPT-4V, Inworld AI, Meta ImageBind, Runway Gen-2, and Google DeepMind Gemini, the most used ones. The present study discusses Google Gemini as a multimodal AI tool because it is the latest and most novelty-based LLM multimodal that can perform multiple tasks simultaneously. Despite being a user-friendly and efficient AI tool, Gemini revolutionized the way to access and interact with various information by providing advanced, more accurate, and contextually relevant responses. According to a Google team report (Team et al., 2023), Gemini's

visual coding is inspired by their own foundational work on Flamingo, CoCa, and PaLI (Alayrac et al., 2022; Yu et al., 2022; Chen et al., 2022) with this distinction that such multimodal models from the beginning can natively output images using discrete image tokens (Team et al., 2023).

Gemini, the latest multimodal artificial intelligence (AI) tool launched on December 06, 2023, is a Google's DeepMind AI model with Visual Language Model (VLM) technology that directly competes with OpenAI's ChatGPT, GPT-4, and GPT-4 with vision (Coles, 2023; Perera & Lankathilaka, 2023). This AI tool features multiple large language models (LLMs) and natural language processing (NLP) technologies (Farrokhnia et al., 2023). It comprises various LLM sizes and encompasses three versions: Gemini Nano, Gemini Pro, and Gemini Ultra. These three versions are designed according to the users' needs and demands. Nano is designed for 'on-device' efficient and more accessible use on smartphones. Ultra is the most powerful version of these three and is used to the fullest extent of Google's AI capabilities. However, the Pro version is as balanced in terms of use and AI capabilities as it is a combined version of the other two (Team et al., 2023).

This Gemini tool is helpful for tackling reinforcement learning, deep learning, and problems and tasks related to digital education. Its interdisciplinary use will help in integrating AI tools in different fields for future technology integration, collaboration, and innovations, particularly for researchers, educators, and digital content creators. It further assists in finding diverse responses and seeking help to provide solutions for future learning innovation through generative AI and its incorporation in education, health, management, climate change, etc. (Ali et al., 2023; Imran & Almusharraf, 2023a, 2023b). Through this emerging technology report, the researchers intend to explore the merits and drawbacks of Google's Gemini as a next-generation AI educational tool and consider its use and role in educational technology.

### **Gemini features**

Gemini, GenAI's most advanced and capable tool, offers a wide range of features that distinguish it in the AI landscape. It has the most robust general capabilities across modalities as well as cutting-edge understanding and reason performance in each domain (Team et al., 2023).

### **Multimodal capabilities**

Gemini's most striking feature is its ability to understand and work with different data types, such as text, images, audio, PDFs, and videos. This versatility allows Gemini to generate more complete answers that fit the context, making it helpful for various tasks and applications. Moreover, Gemini is a potential source of educational technology advancement and practical applications beyond its theoretical framework (Lee et al., 2023a, 2023b). Unlike ChatGPT, Google Gemini is not limited to text-based tasks; instead, it can process various types of inputs, including audio, visual, and video data, and generate the output based on the data received (Portakal, 2023; Koubaa et al., 2023).

### **Advanced performance**

The Gemini 1.0 Ultra model stands out because of its exceptional performance across domains. Its multimodality proves helpful for those with limited access to digital

learning tools and AI platforms to interact with diverse, rich and sustained learning environments. Through its diverse functions, an individual can benefit from learning language, object recognition, responses supported with multiple input options, and making a real-time conversation on any topic (Nyaaba, 2023). Gemini excels in various tasks such as analyzing text, helping with programming, using logic, reading comprehension, solving mathematical problems, and code generation. According to a Google report, Gemini is trained to mitigate risks of harmful response generation. The Google DeepMind Team enumerates about twenty types of harmful clues and phrases, such as suggestions regarding dangerous behavior, hate speech, security issues, medical advice, etc. Therefore, Gemini's responses are based on a dataset free of potential harm-inducing inputs and queries (Team et al., 2023).

### **Generative AI**

Gemini, powered by GenAI, is a compelling AI model that excels at generating new content based on the input it receives. With its impressive capabilities, Gemini can create a wide range of data types, including text, code, images, and more. Therefore, there is no doubt that Gemini is the ultimate tool for creative tasks, content creation, and problem-solving. Unlike previous models trained on unchanging datasets, Gemini has the possibility to tap into Google Search to acquire and process real-world information (Portakal, 2023). This allows Gemini to tailor its responses to ongoing events, ensuring they reflect the latest developments.

### **Versatility in communication**

Another feature of Gemini lies in its ability to handle different communication tasks and styles, potentially adapting its responses to be informative, comprehensive, or even causal and engaging depending on the need and situation. It further offers interactive simulations and learning environments. By combining audio, video, image, and text, Gemini creates immersive educational experiences that bring abstract concepts to life (Team et al., 2023). Among all its competitor AI educational tools (Bing chat, Claude 2.0, Ernie, ChatGPT, etc.), Gemini, with its ability to understand and interpret various input data, emerges as a powerful contender in meeting the needs for personalized, accessible, and dynamic learning experiences necessitates innovative solution in the educational landscape (Perera & Lankathilaka, 2023). Moreover, this Gemini technology is advanced in providing personalized feedback and explanations for various tasks and prompts (Saeidnia, 2023). It has the ability to analyze students' responses and provide tailored feedback, including explanations through visualized concepts, natural responses, and relevant examples.

### **Feedback and assessment**

Gemini uses its advanced understanding of language and logic for more systematic assessment and offers efficient and consistent feedback and grading of coded and written assignments and tasks (Team et al., 2023). It is equally beneficial for educators, learners, and professionals. For educators, Gemini can generate thought-provoking prompts and scenarios, encouraging their students to think critically, analyze things logically, develop hypotheses, and explore solutions. According to Saeidnia (2023), Google Gemini

facilitates knowledge exchange and communication across diverse learning communities and promotes a collaborative learning environment. This technology is designed to offer a contextualized conversational experience for personalized, accurate, and relevant responses in a user-friendly way.

#### **Better understanding across modalities**

Gemini 1.5 Pro is trained to perform highly sophisticated reasoning tasks for different modalities. According to Nyaaba (2023), Gemini proved to be a more elaborative and informed view of the nature of science than other LLMs' supported tools like ChatGPT. Therefore, Gemini helps in understanding complex, multidimensional, and evolving concepts of scientific theories, methods, approaches, knowledge, and alignment with the latest scientific understandings (Knight, 2023; Nyaaba, 2023). For learning and instruction, Gemini emphasizes systematic inquiry and evidence-based reasoning. This feature makes it versatile among competitors in providing clear and concise contracts among various responses dealing with subjects such as science, religion, and philosophy. It has the capacity to break down the experimental components and discuss research-based priorities and methodologies (Nyaaba, 2023).

#### **Problem solving**

Gemini 1.5 Pro is designed to provide problem-solving with larger blocks of code. For instance, as the Google Team report (Team et al., 2023) claimed, it can reason across 100,000 lines of code, giving helpful solutions, modifications, and explanations. Moreover, Gemini is the first model to outperform human experts in Massive Multitask Language Understanding (MMLU). MMLU is among the most popular methods to test the knowledge and problem-solving abilities of AI models.

#### **Emerging technology in practice**

The traditional definition of AI systems centers on replicating human intelligence and problem-solving abilities. However, in educational and learning environments, the focus shifts from AI's human-like flexibility to its interaction with learners and outcomes (Lee et al., 2023a, 2023b). This shift results in more dynamic, inclusive, and valuable AI-tailored tools to meet the changing needs of educators and students, albeit deviating from traditional AI goals. Therefore, AI technology and tools based on LLMs, GPTs, and generative systems are getting worldwide attention and becoming popular in every field of life (Aktay et al., 2023; Imran & Almusharraf, 2023a). Unlike Recurrent Neural Networks (RNNs), these AI systems do not face any limitations in handling long-term dependencies. Hence, OpenAI's ChatGPT-4 and Google's DeepMind Gemini have introduced efficient ways and effectiveness that often increase with model and training corpus size (Lee et al., 2023a, 2023b; OpenAI, 2023).

In the context of Gemini in future education, Lee et al., (2023a, 2023b) examined the significant role that multimodal AI approaches have been contributing towards the realization of generative AI in educational perspectives. They further delved into the crucial aspects of AI systems, including advancing knowledge representation, strategic planning, cognitive frameworks, adaptive learning mechanisms, integration of diverse data sources, and sophisticated language processing (Lee et al., 2023a, 2023b, p. 1). In another

study, Perera and Lankathilaka (2023) presented a case study of various participants, including government officials, educators, learners, and researchers, to prepare a series of recommendations and proposals to ensure the effective and ethical use of AI systems and tools. This study concluded that policymakers and experts can have the opportunity to cultivate an environment that is both transformative and ethical, helpful in harnessing the complete capabilities of generative AI products like Google's Gemini and Open AI's ChatGPT while ensuring the protection of learners' welfare and upholding academic integrity.

In contrast to Lee et al., (2023a, 2023b), Perera and Lankathilaka (2023), Nyaaba (2023) utilized VNASO based questionnaire for the in-service and pre-service educators to compare human and AI (Gemini and GPT-4) understanding of the nature of science. Nyaaba tested the same questionnaire with AI technologies and human beings and summarised the findings, stating that AI tools have more informed views on the nature of science and scientific knowledge than humans. Human responses were mixed with informed and naïve views, whereas both AI tools, GPT-4 and Gemini, offered consistent informed views. Moreover, Gemini tended to be more comprehensive and elaborative in analyzing and answering queries among AI tools. The following are the advantages of Gemini for learners and teachers, as well as for creating educational content.

#### **Gemini for learners**

Learners can benefit from its various functions, such as using Gemini as a study buddy for personalized learning. It is a tool that can help in finding answers to questions related to any subject in a clear and informative way, tailored and customized to an individual's specific requirements and understanding. Moreover, it can adapt its explanations to the learner's learning style, level, and subject, providing targeted support (Saeidnia, 2023; Team et al., 2023). Gemini also assists learners struggling with accessibility and exploration of any concept. For instance, it can generate various representations of the understudy topic/concept through visualization, diagrams, simulations, or even creative narratives/stories to help learners grasp the concept from an innovative angle. Kinesthetic and visual learners can benefit the most from this Gemini feature. At a more advanced level, Gemini can help learners in research and analysis tasks. Being a powerful research assistant, Gemini helps learners and researchers to find relevant resources, conceive innovative ideas, synthesize information, and even identify various patterns and trends in any field of study.

#### **Gemini for teachers**

For teachers and instructors, Gemini has made several teaching and assessment tasks easy to accomplish. They can leverage its' capabilities to create engaging materials, differentiation, and rapid assessment and feedback. For instance, for an interactive lesson plan, Gemini helps teachers in generating worksheets, quizzes, personalized learning paths for students, interactive exercises, etc. According to the Google Team report (Team et al., 2023), Gemini has the potential to cater to the diverse learning needs of the students in a classroom, which was a challenging task before. This AI tool has the ability to create differentiated materials, design multiple activities for different levels of students, or provide additional explanations for those who need extra support or face

different learning challenges (Imran & Almusharraf, 2024). Gemini is very supportive in providing effective feedback in real-time, which is crucial for learners' growth. It can analyze learners' work and offer personalized feedback, identifying areas for further improvement and suggesting resources for further learning and practice.

### **Gemini for educational content generation**

Being a powerful multimodal AI tool, Gemini can work as an educational content/material generator in various capacities because of its exceptional GenAI capabilities to organize study materials, help in creating outlines, draft lesson plans, add visual effects, and other teaching resources such as worksheets, puzzles, creative ideas, fiction and non-fiction, error analysis, and most importantly output in multiple shapes beyond only text such as image, video, and graphs (Nyaaba, 2023). Furthermore, it helps in dealing with multilingual tasks by breaking down language barriers. Gemini's potential for multilingual communication could be an advantage for creating educational materials that cater to diverse student populations, particularly in a multilingual community with diverse linguistic needs. Interactive learning objects play a crucial role in promoting successful classroom practices. Therefore, Gemini, being a multitasking tool, has the ability to provide interactive learning elements that can boost engagement within less time. Hence, this study encourages teachers to use Gemini to develop daily classroom tasks and assignments. It can help in developing simulations, quizzes, or other interactive features to make learning materials more engaging and dynamic.

### **Significant challenges**

Despite Google Gemini's promises and opportunities, significant challenges also exist. The major challenge in using generative AI and other AI technologies is the lack of ethical guidelines and policies for its fair use in educational environments (Perera & Lankathilaka, 2023). Since December 2022, after the launch of ChatGPT-3.5, a serious discussion regarding AI and its potential use in academia has been started, particularly for academic integrity and ethical concerns (Imran & Almusharraf, 2023a, 2023b). Therefore, proper ethical guidelines and policies are necessary not only for the education sector but also for all other departments and sectors where AI is potentially being used. Recently, all leading publication houses, including Sage, Springer, Tylor and Francis, and Oxford University Press, to name a few, have issued guidelines according to their journals' authorship criteria and use of AI for content production (Garnier, 2023).

Lee et al. (2023a) performed a comparative study comparing Gemini and other multimodal tools like ChatGPT-4V and concluded that Gemini has certain limitations in completing various tasks. They highlighted a few examples, such as, "Gemini fails to retrieve a random example from few-shot examples" (p. 10). In another test, Lee et al. (2023a) highlighted Gemini Pro's limited performance on the automatic scoring task. For example, the authors "tried to get automatic scoring results for one student drawn image by providing it with..., Gemini Pro did not return any scoring result (Lee et al., 2023a, p. 13)." A few other studies (Akter et al., 2023; Fu et al., 2023) also pointed out similar concerns and difficulties Gemini faced in dealing with complex visual contexts. Fu et al. (2023) highlighted Gemini's failure to meet the challenges of interpreting images with a large number of elements and to utilize few-shot visual examples effectively because of



its concision approach. Similarly, Lee et al. and's (2023a, 2023b) study provides a detailed examination of these specific shortcomings, corroborating previous research's observations and providing a concrete comparison of how different AI models handle complex educational tasks (p. 15).

Regardless of any particular AI tool, Academia and other stakeholders like publishing houses, creative centers, and educational material developers strongly emphasize acknowledging the AI's role and limitations. In this context, Levene (2023), the operations manager of the Committee on Publication Ethics (COPE), published a detailed report regarding challenges in AI-produced content. This report highlighted that AI tools of any kind have no reliability, replicability, or truth in their responses when getting written responses. Instead, these tools collect and summarize a large number of fetched data and statements in their repository and return a tailored response for which they are tailored. Moreover, an AI chatbot does not consider and care whether the utility of the produced response is true or false, and the cases and issues become more diverse in different subjects and disciplines. COPE further mentioned AI technology users' experiences, including various examples of fabricated citations, false references, and misattributed suggestions.

## Conclusion

Indeed, Google Gemini is a valuable addition to educational technologies; however, its algorithms must be tested and monitored rigorously to ensure that they are free from biases that could disadvantage certain groups of learners and cause fairness challenges. Like other AI chatbots and tools, Gemini also has to safeguard data privacy and security by protecting the users' data and ensuring ethical data collection and practices, which are crucial to responsible AI integration in education. In conclusion, Google Gemini represents a powerful transformative force in educational technology competition. Its multimodal capabilities, reasoning, and generation skills offer unprecedented opportunities for personalized learning, engaging instruction, and dynamic assessment. However, careful consideration of ethical challenges, responsible development, and transparent implementation are essential to harnessing the full potential of this generative AI technology. Hence, by prioritizing human-centered design, addressing biases, and upholding ethical standards, Google Gemini can pave the way for a future where technology empowers personalized learning experiences for all.

## Abbreviations

AI	Artificial Intelligence
VLM	Visual language model
NLP	Natural language processing
RNNs	Recurrent neural networks
COPE	Committee on Publication Ethics

## Acknowledgements

The authors would like to thank the Education Research Lab, Prince Sultan University, for technical and financial (APC) support.

## Author contributions

Both authors equally contributed in this manuscript. Dr. Imran worked on initial idea generation, data collection, and initial write-up. Dr. Almusharraf worked on final draft, conclusion, technological aspects, and overall supervision of this study.

## Funding

This study received funding from Prince Sultan University, Riyadh, Saudi Arabia.

**Availability of data and materials**

No data is associated with this study.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

The authors agreed to publish this paper.

**Competing interests**

The authors declare no competing interests.

Received: 30 January 2024 Accepted: 13 May 2024

Published online: 23 May 2024

**References**

- Aktay, S., Seçkin, G., & Uzunoğlu, D. (2023). ChatGPT in education. *Türk Akademik Yayınlar Dergisi (TAY Journal)*, 7(2), 378–406.
- Akter, S. N., Yu, Z., Muhamed, A., Ou, T., Bäuerle, A., Cabrera, Á. A., et al. (2023). An in-depth look at Gemini's language abilities. [arXiv:2312.11444](https://arxiv.org/abs/2312.11444)
- Alayrac, J. B., Donahue, J., Luc, P., Miech, A., Barr, I., Hasson, Y., et al. (2022). Flamingo: A visual language model for few-shot learning. *Advances in Neural Information Processing Systems*, 35, 23716–23736.
- Ali, A. H., Alajlanbi, M., Yaseen, M. G., & Abed, S. A. (2023). Chatgpt4, DALL· E, Bard, Claude, BERT: Open Possibilities. *Babylonian Journal of Machine Learning*, 2023, 17–18.
- Badshah, A., Rehman, G. U., Farman, H., Ghani, A., Sultan, S., Zubair, M., & Nasralla, M. M. (2023). Transforming educational institutions: Harnessing the power of internet of things, cloud, and fog computing. *Future Internet*, 15(11), 367.
- Chen, X., Wang, X., Changpinyo, S., Piergiovanni, A. J., Padlewski, P., Salz, D., et al. (2022). Pali: A jointly-scaled multilingual language-image model. [arXiv:2209.06794](https://arxiv.org/abs/2209.06794)
- Coles, G. (2023). Google DeepMind Gemini AI release date: The world's new most powerful language model. PCguide.com. Retrieved December 25, 2023, from <https://www.pcguide.com/apps/google-deepmind-gemini-release-date/>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of Generative AI: Implications for educational practice and research. *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2023.2195846>
- Fu, C., Zhang, R., Lin, H., Wang, Z., Gao, T., Luo, Y., et al. (2023). A challenger to GPT-4V? Early explorations of Gemini in visual expertise. [arXiv:2312.12436](https://arxiv.org/abs/2312.12436)
- Garnier, M. (2023). True or false: Publishers' policies around ChatGPT and other AI tools. Writefull.com. Retrieved December 20, 2023, from <https://blog.writefull.com/true-or-false-publishers-policies-around-chatgpt-and-other-ai-tools/>
- Imran, M., & Almusharraf, N. (2023a). Review of Teaching innovation in university education: Case studies and main practices. *The Social Science Journal*. <https://doi.org/10.1080/03623319.2023.2201973>
- Imran, M., & Almusharraf, N. (2023b). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), 464.
- Imran, M., & Almusharraf, N. (2024). Digital learning demand and applicability of quality 4.0 for future education: a systematic review. *International Journal of Engineering Pedagogy (IJEPP)*, 14(4), 38–53. <https://doi.org/10.3991/ijepv14i4.48847>
- Knight, W. (2023). Google just launched Gemini, its long-awaited answer to ChatGPT. <https://www.wired.com/story/google-gemini-ai-model-chatgpt>
- Koubaa, A., Boulila, W., Ghouti, L., Alzahem, A., & Latif, S. (2023). Exploring ChatGPT capabilities and limitations: a survey. *IEEE Access*, 11, 118698–118721. <https://doi.org/10.1109/ACCESS.2023.3326474>
- Lee, G. G., Latif, E., Shi, L., & Zhai, X. (2023a). Gemini pro defeated by GPT-4V: Evidence from education. [arXiv:2401.08660](https://arxiv.org/abs/2401.08660)
- Lee, G. G., Shi, L., Latif, E., Gao, Y., Bewersdorf, A., Nyaaba, M., et al. (2023b). Multimodality of AI for education: Towards artificial general intelligence. [arXiv:2312.06037](https://arxiv.org/abs/2312.06037)
- Levene, A. (2023). Artificial intelligence and authorship. Committee on Publication Ethics (COPE). Publicationethics.org. Retrieved December 15, 2023, from <https://publicationethics.org/news/artificial-intelligence-and-authorship>
- Nyaaba, M. (2023). Comparing human and AI's (GPT-4 and Gemini) understanding of the nature of science. <https://ssrn.com/abstract=4661602> or <https://doi.org/10.2139/ssrn.4661602>
- OpenAI. (2023). GPT-4 Technical report. [arXiv:2303.08774](https://arxiv.org/abs/2303.08774). <https://api.semanticscholar.org/CorpusID:257532815>
- Perera, P., & Lankathilaka, M. (2023). Preparing to revolutionize education with the multi-model GenAI tool Google Gemini? A journey towards effective policy making. *Journal of Advances in Education and Philosophy*, 7(8), 246–253.
- Portakal, E. (2023). Google's Gemini AI review. Textcortex.com. Retrieved December 23, 2023, from <https://textcortex.com/post/gemini-ai-review>
- Saeidnia, H. R. (2023). Welcome to the Gemini era: Google DeepMind and the information industry. *Library Hi Tech News*. <https://doi.org/10.1108/LHTN-12-2023-0214>
- Team, G., Anil, R., Borgeaud, S., Wu, Y., Alayrac, J. B., Yu, J., et al. (2023). Gemini: A family of highly capable multimodal models. [arXiv:2312.11805](https://arxiv.org/abs/2312.11805)
- Yu, J., Wang, Z., Vasudevan, V., Yeung, L., Seyedhosseini, M., & Wu, Y. (2022). Coca: Contrastive captioners are image-text foundation models. [arXiv:2205.01917](https://arxiv.org/abs/2205.01917)

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.