International Journal of Artificial Intelligence Research
Vol. 9, No. 1, June 2025 ISSN 2579-7298

Evaluating the Impact of Ai-Generated Outputs on Student Assessment: Educator's Perspective

Sammy Q. Dolba [a,1], Frederick A. Inoncillo [b,2], Jayrome L. Nuñez [c,3*]

[a] Philippine Normal University, City of Manila, Philippines
[b] Baliuag University, City of Baliwag, Philippines
[c] Tafe Arabia Education, Dammam City, Saudi Arabia

[1] dolba.sq@pnu.edu.ph; [2] frederickinoncillo@baliuagu.edu.ph; [3] nunez.j@tafear.edu.sa
* corresponding author

ARTICLE INFO

Article history
Received
Revised
Accepted

Keywords
artificial intelligence
AI-generated
teacher perspective
student assessment
evaluation

ABSTRACT

This study investigates educators' perceptions of the impact of artificial intelligence (AI)-)-generated outputs on student assessment in the Philippine educational context. With the rapid integration of AI technologies in education, understanding how educators view these tools is crucial for effective implementation. A descriptive quantitative research design was employed, utilizing a structured survey distributed to a diverse group of 93 educators across various disciplines. The findings reveal a generally positive perception of AI's role in enhancing teaching practices, with a mean score of M = 3.42 indicating high perceived value. However, concerns regarding the reliability and fairness of AI-generated output were noted, with mean scores of M = 3.30 and M = 3.28, respectively. Additionally, educators expressed moderate confidence in using AI tools, reflected by a mean score of M = 3.24. Qualitative responses highlighted ethical considerations and the need for continuous professional development to equip teachers with the necessary skills to effectively integrate AI into their assessments. This research underscores the potential benefits and challenges associated with AI in education, emphasizing the importance of addressing educators' concerns to maximize the advantages of AI technologies in student learning outcomes.

1. Introduction

In recent years, the ascent of artificial intelligence (AI) has been getting more and more traction in most fields, be in government, business, engineering, health care, hospitality, and education [6]. New generations of AI technologies appear in almost every sector of the world, and they seem to be no stopping. Its transformative power allowed everyone who has access to it to get full advantage of its features and countries that are early to adopt this have gotten the lead in the market of technological advancement [12]. As Ng (2017) mentioned in his talk on Stanford Graduate School of Business, artificial intelligence will be the new electricity of this age empowering the Fifth Industrial Revolution as one of its building blocks for it will amplify economic growth and development. AI incorporation is practically hyping each sector to cope with unending demand of people to fast, reliable, and accessible services 24/7.

In education, the presence of AI is undeniably being integrated in every aspect of teaching and learning to enhance and improve the delivery of lessons to the learners [1]. Studies show that the amalgamation of AI and traditional methods in educational setup enhanced the setup of services in administrative, giving feedback, accessibility, and personalization of learning [9] [14]. As the capacity of AI increases increasingly enhanced every single day, the rise of AI generated output among students and the incapacity of teachers, the illiteracy of teachers to detect the differences between AI generated and student-created output raised an alarm on the growing worry of the influence of this advancement [4]. In a study conducted by Gao et al. (2022), students' over-reliance on AI dialogue systems that closely mimic the human resulted to AI hallucinations and lack of validation by both teachers and learners [8]. Thus, the accountability of validation and verification should not be solely put to the learners utilizing the technology but also to the teachers and the educational institutions as well [21].

While AI promises significant advantages in teaching and learning, teachers' roles in these revolutions has been significantly redefined from mainly the provider and facilitator of learning to more integrated and student-involved collaborator and supporter of student-AI-aided-learning [20]. Many teachers still view AI as an adversary for learning as students become so dependent and self-actualization or academic hard work has been taken for granted due to the ease of access to AI technology, instead teachers should be able to adapt on how to take advantage of this technology to improve the learning process in classroom [13].

As AI takes a foothold in educational systems, the Philippines began formal and informal discussions about the subject's effects and influence on the country [7]. Researchers, Estrellado & Miranda (2023), mentioned in their papers a few challenges that the country may encounter as AI technology becomes established. One of which is the infrastructure and resources. Philippines still struggle with subpar digital infrastructure as it was amplified during the pandemic [5]. With the integration of AI in educational institutions, the digital divide among those who have and have notes will widen. They added [7], the training of teachers should be continuous as AI also continuously improves.

Moreover, one of the challenges that educators encounter with the integration of AI in the classroom is the evaluation and assessment. Without a progressive scale for evaluating AI generated or assisted outputs [16]. As AI has emerged as a vital tool in learning and skills development educational teachers should be able to recognize, get trained, and integrate it into their teaching skills [2]. Institutions have to make sure that educators are well-equipped with knowledge and usage of its addition to classes [3]. An example of an evaluation scale or criteria that can be used to assess AI generated output is shown in the paper of Jauhiainen & Guerra (2024). The authors highlighted three levels of evaluation of AI LLM (Large Language Models) into pre-evaluation, evaluation, and post-evaluation. In another paper written by Perkins et al. (2024), AI could be evaluated in different scales such as: No AI; AI-assisted idea generation and structuring; AI-assisted editing; AI task completion, human evaluation; and Full AI. This gives a structured approach to their assessment strategies. Each level specifies the range that AI is allowed to be used and the learner's responsibility.

Along with the challenges of AI utilization in the classroom is the perspective of the educators who are going to be in the front lines of its implementation. In a comprehensive study conducted among 74 educators in Turkey [19], it showed that there is general positive perception of AI integration however, the paper noted some concerns about privacy and ethical issues on its use. In a similar researched topic among 132 teachers in Russia, though it is promising and useful sometimes AI tools can be complex, risky, and not very smart. Even though the participants reported to have low competency on the use of AI, they are willing to undergo training and improve their skills [17]. One study in among 580 math teachers in the United Arab Emirates (UAE) recognizes the potential benefits of AI use in the classroom as it can cater to individual differences of learners and will help them improve at their own pace. However, the results also showed some concerns about teachers being pressured to use it and not being able to utilize it properly. Ethical and privacy issues were also on the list of concerns in the study [11] [18].

As AI technology continues to influence the education sector, this paper explores teachers' perceptions of students' use of AI-generated content in the Philippines. It is important to note that the

findings of this study are representative of the specific population sampled, rather than the entire archipelago.

2. Method

2.1. Research design

This study employs a descriptive quantitative research design to evaluate educators' perceptions and confidence regarding AI-generated outputs in student assessments. The design is aimed at collecting and analyzing numerical data to identify trends, patterns, and correlations related to the perspective of educators on AI-generated outcomes of the learners.

Using SPSS, the collected data were analyzed using the following methods:

1. Descriptive statistics. Percentages means, and standard deviations will be calculated to summarize the responses.
2. Correlation analysis. To assess the relationships between years of teaching experience, confidence, and perceived value of AI-generated outputs.

To ensure the validity of the questions, pilot testing was conducted on a small subset of participants to refine the tool and make sure that questions were aligned with the objectives of the study after rigorous experts' scrutiny. Reliability was ensured by calculating the internal consistency of the survey items using Cronbach's Alpha.

2.2. Research population

The participants in this study represented a diverse group of educators from the National Capital Region (NCR) of the Philippines. They teach various academic disciplines, including MAPEH (music, arts, and physical education), English, psychology, management, health sciences, social sciences, science education, mathematics, engineering, biology, computer engineering, history, nursing, business administration, accounting, economics, and general education.

Table 1. Distribution of respondents according to sex

College	Frequency	Percentage
Female	54	58.1
Male	39	41.9
Total	**93**	**100.0**

Table 1 showcases the distribution of respondents according to their sex. There are fifty-four (58.1%) of respondents are females and t (42.6 %) thirty-nine (41.9) % are males.

Table 2. Distribution of respondents according to department where they are teaching

Department	Frequency	Percentage
Elementary	15	16.1
Junior High School	19	20.4
Preparatory School	1	1.1
Senior High School	1	1.1
Tertiary	57	61.3
Total	93	100.0

Table 2 shows the distribution of respondent teachers according to three departments where they are teaching. There are fifteen (16.1 %) of respondents who are teaching Elementary, nineteen (20.4%) who are teaching Junior High School, one (1.1%) is teaching Preparatory School, one (1.1%) who is teaching senior High Schook, and fifty-seven (61.3%) who are teaching from Tertiary.

Table 3. Distribution of respondents by years of teaching experience

Years of Experience	Frequency	Percentage
0 - 6	19	20.4
7 - 12	24	25.8
13 - 19	18	19.4
20 - 25	17	18.3
26 - 32	11	11.8
32 – 38	3	3.2
39 -45	1	1.1

The frequency and percentage distribution of respondents according to their years of teaching experience is presented in Table 43. The majority of respondents (25.8%) have 7 to 12 years of teaching experience, followed by those with 0 to 6 years (20.4%). Respondents with 13 to 19 years and 20 to 25 years of experience comprise 19.4% and 18.3% of the sample, respectively. Fewer respondents reported longer teaching hours, with 11.8% having 26 to 32 years of experience, 3.2% with 32 to 38 years, and only 1.1% with 39 to 45 years of experience.

The distribution of teaching experience among respondents indicates a concentration of educators in the early to mid-career stages. Over 65% of the respondents have 25 years or less of teaching experience, suggesting that the majority are in the earlier phases of their professional journey. This trend may reflect recruitment patterns or retention challenges in the field, as fewer individuals remain in teaching roles for extended periods, with only 4.3% of respondents having more than 32 years of experience.

2.3. Data Collection Tool

A structured survey was used to gather data. The survey included both closed-ended questions (using Likert scales) and open-ended items, categorized under the following aspects:

1. Reliability and fairness of AI-generated outputs.
2. Effectiveness in accurately assessing student performance.
3. Alignment with learning objectives and knowledge representation.
4. Educators' confidence and readiness in using AI tools.
5. Perceived value and impact on teaching practices.
6. Contribution of AI tools to enhance learning experiences.

6.1. Ethical Considerations

Participants were briefed on the study's objectives, and their consent was secured before the questionnaire was distributed. Data handling was conducted with the highest level of security, ensuring that no personal information was used, in full compliance with the Data Protection Law of the Philippines.

3. Results and Discussion

The integration of artificial intelligence (AI) in education has prompted significant shifts in how educators perceive and utilize AI-generated outputs in student assessments. This study aimed to evaluate educators' perspectives on the effectiveness, reliability, and overall impact of AI tools on student performance. The results indicate a generally positive reception among educators regarding the role of AI in enhancing teaching practices, although concerns about its reliability and fairness remain prevalent.

Table 4. Analysis of Mean and Standard Deviation for Educators' Evaluation of AI-Generated Outputs' Impact on Student Assessment

	Score				
	N	**M**	**SD**	**Adjectival Rating**	**Interpretation**
Reliability and Fairness of AI-Generated Outputs	93	3.30	1.17	Neutral	Moderate
Effectiveness of AI in	93	3.40	.89	Neutral	Moderate

Student Assessment					
Alignment with Learning Objectives and Knowledge Representation	93	3.28	.83	Neutral	Moderate
Confidence and Preparedness in Using AI Tools	93	3.24	.87	Neutral	Moderate
Perceived Value and Impact of AI on Teaching	93	3.42	.84	Agree	High
Contribution to the Overall Learning Experience	93	3.31	.89	Neutral	Moderate

Table 4 offers a comprehensive overview of the mean and standard deviation of the educator's evaluation of the impact on AI-generated outputs on student assessment, dissected by its distinctive dimensions. The findings demonstrate that the respondents exhibit high levels of perceived value and impact AI on teaching while the rest of dimensions show moderate levels.

The degrees of educator's evaluation of the impact on AI-generated outputs on student assessment reported by the 93 respondents are effectively captured by the descending order of mean scores for each dimension, signifying varying levels of engagement. Leading the spectrum is perceived value and impact of AI on teaching with a high level, reflected by the highest mean score (M = 3.42, SD = 0.84), closely followed by effectiveness of ai in student assessment with moderate level has a mean score (M = 3, .40, SD= 0.89). Similarly, reliability and fairness of ai-generated outputs is rated moderately, with a mean score of (M = 3.30, SD = 0.87).

The dimensional alignment with learning objectives and knowledge representation follows closely at a moderate level (M = 3.28, SD = 0.83). Finally, confidence and preparedness in using AI tools concludes the list with a moderate level (M = 3.24, SD = 0.87). The presence of low standard deviations signifies that the scores closely align with the mean, underscoring a consistent level of educator's evaluation of the impact on AI-generated outputs on student assessment.

Educators have a moderately positive perception of AI's impact on student assessment, with the highest confidence in AI's value and impact on teaching. There is a consensus that while AI has potential, there are areas where further development and support are needed.

Table 5. Table of Reference. Likert Point Description

Numerical Scale	Weighted mean interval scale	Mean descriptive equivalent
5	4.21 – 5.00	Very high
4	3.41 – 4.20	High
3	2.61 – 3.40	Moderate
2	1.81 – 2.60	Low
1	1.00 – 1.80	Very Low

Table 5 presents the numerical scale, corresponding weighted mean interval scale, and their descriptive equivalents based on a Likert scale. The descriptions range from "Very High" (5) to "Very Low" (1), providing a clear interpretation of mean values within the specified intervals.

Table 6. Correlation analysis of teachers' experience and educators' perspectives on ai-generated outputs' impact on student assessment

Variables	Correlation Type	Correlation	Sig	N
TYE and RFA	Pearson's Product	.015	.887	93
TYE and EAS	Pearson's Product	.031	.765	93
TYE and ALO	Pearson's Product	.015	.890	93
TYE and CPU	Pearson's Product	.091	.384	93
TYE and PVI	Pearson's Product	.090	.390	93
TYE and COL	Pearson's Product	.069	.511	93
*p<.05				

Legend:

TYE	Teachers Years of Experience
RFA	Reliability and Fairness of AI-Generated Outputs
EAS	Effectiveness of AI in Student Assessment
ALO	Alignment with Learning Objectives and Knowledge Representation
CPU	Confidence and Preparedness in Using AI Tools
PVI	Perceived Value and Impact of AI on Teaching

COL Contribution to the Overall Learning Experience

Table 6 presents the weak correlation results between teachers' years of experience (TYE) and the dimensions of educators' perspectives on the impact of AI-generated outputs on student assessment. The findings indicate that there are no statistically significant correlations between TYE and any of the six dimensions studied. Specifically, the correlation coefficients ranged from .015 (for TYE and RFA, and TYE and ALO) to .091 (for TYE and CPU).

The significance values for all correlations exceeded the threshold of $p<.05$ with the lowest significance value being $p=.384$ (TYE and CPU). These findings suggest that teachers' years of experience do not significantly relate to their perceptions of AI's reliability, effectiveness, alignment with learning objectives, confidence in using AI tools, perceived value, or contribution to the overall learning experience.

The absence of statistically significant correlations between teachers' years of experience and the six dimensions of their perspectives on AI-generated outputs highlights that perceptions of AI in student assessment may not be influenced by tenure in the teaching profession. This finding suggests that educators across varying levels of experience share similar views regarding AI tools' reliability, fairness, effectiveness, and alignment with educational objectives.

4. Conclusion

This study underscores the transformative potential of AI in educational settings while simultaneously illuminating the challenges that accompany its integration into assessment practices. Educators recognize the significant advantages that AI can offer in enhancing teaching methodologies and student engagement; however, concerns regarding reliability, fairness, and ethical implications highlight the necessity for careful implementation and continuous training.

As educational institutions move forward with AI integration, it is imperative to establish robust frameworks for evaluating AI-generated outputs that prioritize both effectiveness and equity. Furthermore, ongoing professional development is essential to equip educators with the necessary skills to navigate this evolving landscape effectively. By addressing these challenges head-on, educational stakeholders can harness the full potential of AI technologies to foster meaningful learning experiences while maintaining academic integrity and promoting critical thinking among students.

Acknowledgment

We extend our heartfelt thanks to all the teachers who dedicated their time to participating in our survey on the perceptions of artificial intelligence in education.

Declarations

Funding statement. This research was entirely self-funded by the researchers, who covered all associated costs to maintain complete independence and objectivity throughout the study.

References

[1] Allam, H., Dempere, J., Akre, V., & Flores, P. (2023). Artificial Intelligence in Education (AIED): Implicationsand Challenges. *Proceedings of the HCT International General Education Conference (HCT-IGEC 2023)*. https://doi.org/10.2991/978-94-6463-286-6_10

[2] Bekdemir, Y. (2024). The Urgency of AI Integration in Teacher Training: Shaping the Future of Education. *Journal of Research in Didactical Sciences, 3*(1), 37-41 . https://doi.org/10.51853/jorids/15485

[3] Bellas, F., Naya-Varela, M., & Paz-Lopez, A. (2024). Education in the AI era: a long-term classroom technology based on intelligent robotics. *Humanities and Social Sciences Communications, 11*(1), 1-20. https://doi.org/10.1057/s41599-024-03953-y

[4] Bennett, L., & Abusalem, A. (2024). Artificial Intelligence (AI) and its Potential Impact on the Future of Higher Education. *Athens Journal of Education, 11*(3), 195-212. https://doi.org/10.30958/aje.11-3-2

[5] Cuisia-Villanueva, M. C., & Nuñez, J. L. (2021). A study on the impact of socioeconomic status on emergency electronic learning during the coronavirus lockdown. *FLDA Journal, 6*(1), 1-17. Retrieved from https://nsuworks.nova.edu/fdla-journal/vol6/iss1/6/

[6] Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management, 48*(1), 63-71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

[7] Estrellado, C. P., & Miranda, J. (2023). Artificial Intelligence in the Philippine Educational Context: Circumspection and Future Inquiries. *International Journal of Scientific and Research Publications, 13*(5), 16-22. https://doi.org/10.29322/IJSRP.13.04.2023.p13704

[8] Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *npj Digital Medicine*, 1-5. https://doi.org/10.1038/s41746-023-00819-6

[9] Göçen, A., & Aydemir, F. (2020). Artificial Intelligence in Education and Schools. *Research on Education and Media, 12*(1), 13-21. https://doi.org/10.2478/rem-2020-0003

[10] Jauhiainen, J. S., & Guerra, A. G. (2024). Generative AI in education: ChatGPT-4 in evaluating students' written responses. *Innovations in Education and Teaching International*, 1–18. https://doi.org/10.1080/14703297.2024.2422337

[11] Kalra , R. (2024). Exploring Teachers' Perceptions Toward the Integration of AI Tools in the Language Classroom. *NIDA Journal of Language and Communication, 29*(45), 21-36. Retrieved from https://lcjournal.nida.ac.th/main/public/abs_pdf/journal_v29_i45_2.pdf

[12] Khanal, S., Zhang, H., & Taeihagh, A. (2024). Development of New Generation of Artificial Intelligence in China: When Beijing's Global Ambitions Meet Local Realities. *Journal of Contemporary China*(April), 1-24. https://doi.org/10.1080/10670564.2024.2333492

[13] Koh, J., Cowling, M., Meena, J., & Sim, K. N. (2023). The Human Teacher, the AI Teacher and the AIed-Teacher Relationship. *Journal of Higher Education Theory and Practice, 23*(17), 199-211. https://doi.org/10.33423/jhetp.v23i17.6543

[14] Melchor, P. M., Lomibao, L. S., & Parcutilo, J. O. (2023). Exploring the Potential of AI Integration in Mathematics Education for Generation Alpha — Approaches, Challenges, and Readiness of Philippine Tertiary Classrooms: A Literature Review. *Journal of Innovations in Teaching and Learning, 3*(1), 39-44. https://doi.org/10.12691/jitl-3-1-8

[15] Ng, A. (2017, January 25). Artificial intelligence is the new electricity. Retrieved January 25, 2017, from https://www.youtube.com/watch?v=21EiKfQYZXc

[16] Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS): A Framework For Ethical Integration Of Generative AI In Educational Assessment. *Journal of University Teaching & Learning Practice, 21*(06), 1-19. https://doi.org/10.53761/q3azde36

[17] Sadykova, G., & Kayumova, A. (2024). Educators' Perception of Artificial Intelligence as Instructional Tool. *TEM Journal, 13*(4), 3194-3204. https://doi.org/10.18421/TEM134-54

[18] Tashtoush, M. A., Wardat, Y., Al Ali, R., & Saleh, S. (2024). Artificial Intelligence in Education: Mathematics Teachers' Perspectives, Practices and Challenges. *Iraqi Journal for Computer Science and Mathematics, 5*(1), 60-77. https://doi.org/10.52866/ijcsm.2024.05.01.004

[19] Uygun, D. (2024). Teachers' perspectives on artificial intelligence in education. *Advances in Mobile Learning Educational Research, 4*(1), 931-939. https://doi.org/10.25082/AMLER.2024.01.005

[20] Yolcu, H. H. (2024). Redefining the Teacher's Role in Education through Artificial General Intelligence (AGI). *Açıköğretim Uygulamaları Ve Araştırmaları Dergisi, 10*(1), 155-167. https://doi.org/10.51948/auad.1383166

[21] Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments, 11*(1), 1-38. https://doi.org/10.1186/s40561-024-00316-7