Preprints.org

Article

The Impact of ChatGPT on Science Education: Teachers' Perceptions Across Diverse Disciplines

Olukayode Apata [*]

Article

The Impact of ChatGPT on Science Education: Teachers' Perceptions Across Diverse Disciplines

Olukayode Apata

Texas A&M University

apata.olukayode@tamu.edu

Abstract

This study examines science teachers' perceptions of ChatGPT's impact on classroom assessments using the Technology Acceptance Model and Theory of Planned Behavior. Data from a publicly available survey of 19 science educators across life sciences, physical sciences, and technology/engineering disciplines were analyzed using one-way ANOVA and multiple regression. Results showed no statistically significant differences in perceptions across disciplines or based on teaching experience, age, and gender. These findings suggest uniform perceptions of ChatGPT, reflecting its perceived usefulness and ease of use, irrespective of teaching area or demographics. The study highlights ChatGPT's potential to enhance assessment practices while emphasizing the need for institutional support and professional development. Recommendations for equitable AI integration and future research are provided.

Keywords: ChatGPT; science education; teacher perceptions; technology acceptance model (tam); classroom assessment

Introduction

Artificial intelligence (AI) has made significant strides in recent years, introducing us to remarkable tools like ChatGPT. This generative AI model has been gaining traction, particularly in the field of education. Dempere and colleagues (2023) emphasized the need to comprehend the potential impacts of AI. For instance, AI can provide personalized support, automate administrative tasks, and introduce innovative teaching methods (Pokkakillath & Suleri, 2023; Apata et al., 2025). Research also indicates that AI can enhance student engagement and customize instruction to meet individual learning needs (Tan, 2023). However, there are valid concerns. Yang (2023) highlighted ethical issues, accuracy, and potential biases in AI systems that could affect fairness in education. Similarly, prior work has shown that factors such as teachers' qualifications and students' background in mathematics can significantly influence performance outcomes in science subjects (Apata, 2019). Therefore, to ensure the effective and equitable use of these AI tools, it is crucial to understand and value educators' perceptions and feelings about them. This understanding is key to addressing potential issues and maximizing the benefits of AI in education (Rueda et al., 2023).

As ChatGPT becomes more prevalent, educators' opinions are split. Some see these tools as groundbreaking, offering new ways to enhance teaching and make assessments more efficient and personalized. However, it is important to recognize concerns raised by some that the reliance on ChatGPT may diminish critical thinking and creativity, which are vital components of student development (Wairisal et al., 2023; Rueda et al., 2023). These concerns are particularly pressing for science teachers tasked with evaluating complex concepts (Pokkakillath & Suleri, 2023). Moreover, the unique challenges faced by educators in different scientific disciplines may influence how they view the integration of AI tools like ChatGPT into their classroom assessment (Altarawneh, 2023).

This study seeks to understand how science teachers across various fields perceive the impact of ChatGPT on their assessment methods. Additionally, it explores how factors such as teaching

experience, age, and gender might shape these perceptions. The research aims to provide insights that could help shape the future of teaching and assessment in science education (Lo, 2023). These insights are essential for developing effective strategies to integrate AI tools to enhance educational outcomes while addressing educators' unique challenges in different scientific disciplines (Rueda et al., 2023). Understanding these dynamics can also guide policy-making and professional development programs that support teachers in adapting to emerging educational technologies.

To guide this study, we focus on the following research questions:

1. How do perceptions of the impact of ChatGPT on assessment differ among science teachers across different science courses?
2. What is the association between science teachers' experience and their perceptions of the impact of ChatGPT on assessment, controlling for age and gender?

Theoretical Framework

Technology Acceptance Model and the Theory of Planned Behavior

We built our study on two theoretical frameworks to understand how teachers perceive and use ChatGPT: the Technology Acceptance Model (TAM) and the Theory of Planned Behavior (TPB). The Technology Acceptance Model (TAM), introduced by Davis in 1989, looks at two main factors: perceived usefulness and perceived ease of use. Perceived usefulness refers to how users believe ChatGPT makes educational tasks easy. Perceived Ease of Use refers to how effortlessly users interact with ChatGPT. On the other hand, TPB, developed by Ajzen in 1991, looks at the attitudes that shape our intentions to use technology.   It includes our attitudes towards the behavior, the social pressure we feel to engage in the behavior, and our confidence in our ability to perform the behavior. TPB is applicable to our study in the following ways: Teachers' attitudes show how they feel about using ChatGPT for teaching. Social pressure results from the influence of colleagues and the wider educational community. Finally, teachers' confidence in their abilities relates to their belief in their skills and the resources they have to use ChatGPT effectively.

Methodology

Participants

We obtained our dataset from a publicly available survey by seasoned researchers from Macquarie University, Australia, who conducted an online survey to determine what teachers think about ChatGPT and its relationships with their teaching and assessment. The original sample size comprises 318 teachers. However, for our analysis, we focused exclusively on science-related subjects. After filtering for relevant teaching areas, we are left with a final sample of 19 participants, comprising high school teachers and university professors from different countries. The dataset is available at the Macquarie University ChatGPT Survey Dataset.

Plan of Analysis

Our study used STATA 18 software (StataCorp, 2023) to analyze our data. First, we used a one-way ANOVA to understand how teachers/professors in different courses (such as physics, biology, and computer science) perceive the usefulness of ChatGPT on their assessments in the classroom. We then grouped these subjects into three: physical sciences (e.g., Physics, Chemistry, Mathematics), life sciences (e.g., Biology, Health, Physiology), and technology/engineering (e.g., Computer Science, Information Technology, Engineering). Physical science is the reference group. After that, we used

multiple regression analysis to looked closely at the association between science teachers' teaching experience and their views on ChatGPT's role in assessments. This approach also allowed us to account for other factors such as teaching area, age, and gender.

Analysis and Results

The results show no statistically significant differences in perceptions of ChatGPT's impact across science disciplines. Using the Technology Acceptance Model, this suggests that teachers perceive ChatGPT as similarly useful and easy to use, regardless of subject area. Teachers may be focused on the general functions of ChatGPT, such as grading support and assessment feedback, rather than discipline-specific applications, leading to uniform perceptions. The Theory of Planned Behavior helps explain this consistency because teachers' attitudes, social expectations, and confidence in using ChatGPT appear comparable across groups. The novelty of ChatGPT and its rapid emergence in education likely created a shared sense of experimentation, where adoption decisions are shaped more by collective norms and institutional factors than by differences in teaching discipline or demographics.

Research Question 1

Perceptions of the Impact of ChatGPT on Assessment in Science Courses

Before conducting the ANOVA analysis, we checked the normality of residuals and the homogeneity of variances (homoscedasticity) assumptions. The residuals from the histogram and q-q plot were normally distributed. Also, the result for Levene's test for homogeneity of variances assumption was not statistically significant (p = .40). So, the assumptions are met. Descriptive statistics for each group are presented in Table 1. Teachers in the life sciences group reported the highest perceived impact of ChatGPT on their classroom assessment (M = 3.75, SD = 0.50), followed by the physical sciences (M = 2.83, SD = 0.75), and the last was technology and engineering groups (M = 2.56, SD = 0.88). The ANOVA results showed no statistically significant differences, F (2, 16) = 3.25, p = .065, η^2 = .29 (see Table 2). This suggests that while there is a variation in how different science disciplines perceive the usefulness of ChatGPT, these differences are not strong enough to be considered statistically meaningful. The non-significant differences might also reflect the early stages of AI adoption in education, where teachers are still exploring and understanding the potential of these tools.

Table 1. Mean and Standard Deviation of Science Teachers' Perception of the Impact of ChatGPT on Classroom Assessment.

Teaching Area Group	N	M	SD
Physical Sciences	6	2.83	0.75
Life Sciences	4	3.75	0.50
Technology and Engineering	9	2.56	0.88
Total	19	2.89	0.88

Table 2. ANOVA Results for the Impact of Teaching Area on Science Teachers' Perceptions of ChatGPT.

Source of Variance	SS	df	MS	F	P
Between Groups	3.98	2	1.99	3.25	0.065
Within Groups	9.81	16	0.61		
Total	13.79	18			

Note. SS = Sum of Squares, df = Degrees of Freedom, MS = Mean Square, F = F-value.

Table 3. Multiple Regression Analysis Predicting Perceptions of ChatGPT's Impact on Assessment Practices.

Predictor	B	SE	T	P	95% CI
Teaching Experience	-0.012	0.030	-0.42	.684	[-0.076, 0.051]
ScienceGroup (Life Sciences)	0.914	0.554	1.65	.123	[-0.281, 2.110]
ScienceGroup (Technology & Engineering)	-0.232	0.462	-0.50	.624	[-1.229, 0.765]
Age	0.026	0.031	0.81	.430	[-0.044, 0.095]
Gender	-0.123	0.436	-0.28	.782	[-1.065, 0.819]
R²				.326	
Adjusted R²				.068	
F(5, 13)				1.26	

Note. B = Unstandardized regression coefficient; SE = Standard error; CI = Confidence interval.

Research Question 2

Teaching Experience and Perceptions of ChatGPT's Impact on Assessment Practices

Relevant assumptions were checked before conducting the multiple regression analysis. The residuals were approximately normally distributed, and the heteroskedasticity test was not significant, $\chi^2 (1) = 1.25$, $p = .263$. The multicollinearity assumption was also met because all Variance Inflation Factor (VIF) values were below 2 (see Table 4).

Table 4. Variance Inflation Factors (VIF) for Predictors.

Predictor	VIF	1/VIF
Teaching Experience	1.98	0.505
Science Group (Life Sciences)	1.36	0.738
ScienceGroup (Technology & Engineering)	1.41	0.708
Age	2.01	0.499
Gender	1.09	0.916
Mean VIF		1.57

The overall model was not statistically significant, $F (5, 13) = 1.26$, $p = .337$, $R^2 = .33$, Adjusted $R^2 = .07$, indicating that the predictors did not significantly explain the variance in teachers' perceptions of ChatGPT's impact on assessment practices. The result of our multiple regression analysis revealed unexpected non-significant results. Teaching experience ($B = -0.012$, $SE = 0.030$, $t(13) = -0.42$, $p = .684$), the Life Sciences group ($B = 0.914$, $SE = 0.554$, $t(13) = 1.65$, $p = .123$), the Technology & Engineering group ($B = -0.232$, $SE = 0.462$, $t(13) = -0.50$, $p = .624$), age ($B = 0.026$, $SE = 0.031$, $t(13) = 0.81$, $p = .430$), and gender ($B = -0.123$, $SE = 0.436$, $t(13) = -0.28$, $p = .782$) are not all statistically significant. Even though teachers in the life sciences group appeared to perceive a greater impact of ChatGPT than those in the physical sciences and technology & engineering groups, these differences did not reach statistical significance. This finding suggests that the impact of ChatGPT on assessment is not strongly influenced by demographic factors such as teaching experience, age, or gender. The novelty of ChatGPT may lead to a relatively uniform perception across different demographic groups, as all teachers are still adapting to this new technology. Moreover, the lack of significant predictors may indicate that other factors such as institutional support or personal attitudes toward technology, play a more crucial role in shaping these perceptions (Durff & Carter, 2019; Francom, 2020; Hew & Brush, 2007).

Contribution to the Teaching and Learning of Science

This study makes a significant contribution to the teaching and learning of science by providing empirical insights into how ChatGPT is perceived by science instructors in different disciplines.

Understanding these perceptions as AI continues to integrate into educational practices is crucial for developing effective strategies that enhance science education. The findings of this research highlight the varied impact of ChatGPT across different science disciplines, suggesting that ChatGPT can potentially support more personalized and efficient assessment practices. By identifying educators' specific needs and concerns within life sciences, physical sciences, and technology/engineering, this study offers valuable information for tailoring AI implementation in ways that resonate with the unique demands of each field.

However, it is essential to note that the small sample size of 19 participants may limit the generalizability of these findings. While the insights gained are valuable, they should be interpreted with caution. The study provides a preliminary understanding that can serve as a foundation for further research involving larger and more diverse samples. Despite this limitation, the research emphasizes the importance of teacher training and professional development, ensuring that educators can use these tools and maximize their potential in fostering student learning. These insights are expected to guide policymakers, curriculum developers, and educational institutions in making informed decisions about adopting and integrating AI technologies in science education, ultimately contributing to improving teaching practices and student outcomes in science disciplines.

Scholarly Significance of the Study

This study adds to the emerging research on the integration of generative AI tools in science education. While many studies examine AI in general educational contexts, limited attention has been given to how teachers in different science disciplines perceive its impact on assessment practices. This study focuses on perceptions across life sciences, physical sciences, and technology or engineering, offering discipline-specific insights that are often overlooked in AI adoption research.

The findings suggest that demographic factors such as teaching experience, age, and gender do not significantly influence perceptions of ChatGPT's impact, implying that adoption challenges and opportunities may be more closely linked to institutional and pedagogical contexts than personal characteristics. This insight contributes to theory by showing how perceptions of emerging technologies can transcend individual demographics, highlighting the need to consider contextual and systemic variables.

These findings align with the Technology Acceptance Model and the Theory of Planned Behavior. Teachers' consistent perceptions across disciplines suggest that perceived usefulness and ease of use, rather than teaching area or demographic factors, shape attitudes toward ChatGPT. Social norms and institutional contexts appear to influence adoption intentions more than individual characteristics. This highlights the importance of collective support structures and professional development when integrating AI tools.

Practically, the study informs professional development and policy decisions. Educational leaders and policymakers can use these findings to design training and support structures tailored to diverse science disciplines, ensuring equitable AI integration. Furthermore, the study identifies gaps in current adoption research, emphasizing the need for larger and more diverse samples. These contributions help build a foundational understanding of how generative AI like ChatGPT can influence teaching and assessment in science education, supporting informed decision-making and future empirical research.

References

Altarawneh, H. (2023). ChatGPT impact on student educational performance: A conceptual analysis. *EAI Endorsed Transactions on e-Learning, 13*(47), Article e4574. https://doi.org/10.4108/eetel.4574

Apata, O. E. (2019). Predictive validity of school and student factors on secondary school students' academic performance in physics. *IOSR Journal of Research & Method in Education, 9*(5), 34–42. https://www.iosrjournals.org/iosr-jrme/papers/Vol-9%20Issue-5/Series-3/B0905033442.pdf

Apata, O. E., Ajamobe, J. O., Ajose, S. T., Oyewole, P. O., Olaitan, G. I. (2025, March). *The role of artificial intelligence in enhancing classroom learning: Ethical, practical, and pedagogical considerations.* Proceedings of the 2025 ASEE Gulf-Southwest Annual Conference, Arlington, Texas, https://peer.asee.org/55084

Bower, M., Torrington, J., Lai, J., Petocz, P., & Alfano, M. (2023). *Public data files containing the data used for the ChatGPT survey (XLSX) and the survey containing variable selection codes (DOCX).* Macquarie University. Dataset. https://doi.org/10.25949/24123306.v1

Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. (2023). The impact of ChatGPT on higher education. *Frontiers in Education, 8.* https://doi.org/10.3389/feduc.2023.1206936

Durff, L., & Carter, M. (2019). Overcoming second-order barriers to technology integration in K–5 schools. *Journal of Educational Research and Practice*, 9(1), 246–260. https://doi.org/10.5590/JERAP.2019.09.1.18

Francom, G. M. (2020). Barriers to technology integration: A time-series survey study. *Journal of Research on Technology in Education*, 52(1), 1–16. https://doi.org/10.1080/15391523.2019.1679055

Hew, K., & Brush, T. A. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research and Development.* https://doi.org/10.1007/S11423-006-9022-5

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers, 12*(8), 153. https://doi.org/10.3390/computers12080153

Pokkakillath, S., & Suleri, J. (2023). ChatGPT and its impact on education. *Research in Hospitality Management, 13*(1), 31-34. https://doi.org/10.1080/22243534.2023.2239579

StataCorp. (2023). *STATA Statistical Software: Release 18.* StataCorp LLC.

Tan, X. (2023). The Impact of ChatGPT on Education and Future Prospects. *Highlights in Science, Engineering, and Technology, 61.* https://doi.org/10.54097/hset.v61i.10285

Wairisal, P. L., Elijonnahdi, E., Susanto, N., & Megahati, R. R. P. S. (2023). Freedom to learn and freedom to teach in science learning through ChatGPT: Systematic literature review. *Jurnal Penelitian Pendidikan IPA, 9*(10), Article 5089. https://doi.org/10.29303/jppipa.v9i10.5089

Yang, J. (2023). The impacts and responses of ChatGPT on education. In *Proceedings of the 7th International Conference on Education and Multimedia Technology (pp. 69–73).* https://doi.org/10.1145/3625704.3625732