Adv Mobile Learn Educ Res, 2024, 4(2): 1197-1207
DOI: 10.25082/AMLER.2024.02.013

RESEARCH ARTICLE

Exploring user perceptions: The impact of ChatGPT on high school students' physics understanding and learning

Muhammad Aizri Fadillah[1*] **Usmeldi Usmeldi**[1] **Lufri Lufri**[1] **Mawardi Mawardi**[1] **Festiyed Festiyed**[1]

[1] Department of Science Education, Universitas Negeri Padang, Padang, Indonesia

Correspondence to: Muhammad Aizri Fadillah, Department of Science Education, Universitas Negeri Padang, Padang, Indonesia;
Email: m.aizrifadillah@student.unp.ac.id

Received: September 13, 2024;
Accepted: November 28, 2024;
Published: December 4, 2024.

Citation: Fadillah, M. A., Usmeldi, U., Lufri, L., Mawardi, M., & Festiyed, F. (2024). Exploring user perceptions: The impact of ChatGPT on high school students' physics understanding and learning. *Advances in Mobile Learning Educational Research*, 4(2), 1197-1207.
https://doi.org/10.25082/AMLER.2024.02.013

Abstract: Artificial intelligence (AI) in education is increasing, including ChatGPT as a learning tool in physics subjects. This study aims to analyze high school students' perceptions of using ChatGPT in physics learning, focusing on demographic factors such as gender, academic level, and duration of use. Involving 167 students, the study used a survey to evaluate students' views on various aspects of the learning experience with ChatGPT, including effectiveness, clarity, consistency of information, and the tool's ability to enrich understanding of physics concepts. Results showed that students' perceptions were positive overall, with ChatGPT perceived as helping to deepen concept understanding, improving the ability to correct misconceptions, and providing an enjoyable learning experience. Significant differences were found based on gender and academic level, while the duration of ChatGPT use showed no overall significant effect, though longer use enhanced specific benefits. These findings highlight the potential of ChatGPT to support physics education by addressing students' diverse needs and improving learning outcomes, offering insights for educators in integrating AI tools effectively into classrooms.

Keywords: ChatGPT, physics education, student perception, high school students, artificial intelligence

1 Introduction

AI has experienced rapid development in recent years and has become integral to various sectors, including education. AI technology offers various innovative solutions to support the teaching-learning process, such as automatically providing learning materials, assisting with assessments, and providing guidance that students can access at any time (Chen et al., 2020; Huang et al., 2021). One of the most prominent AI applications is ChatGPT, developed by OpenAI and launched on November 30, 2022. Within the first week after its launch, ChatGPT gained immense popularity by reaching over one million users, demonstrating the high public interest in AI's ability to respond and interact human-likely (Caldarini et al., 2022; Uğraş et al., 2024). The model was trained using an extensive dataset of human conversations, allowing ChatGPT to process and understand natural language by generating in-depth, relevant, and human-like responses (Susnjak & McIntosh, 2024). With 570 GB of data covering approximately 300 billion words and over 175 billion parameters, ChatGPT has tremendous capabilities in understanding and responding to questions and providing guidance directly to its users (Sabzalieva & Valentini, 2023).

The adoption of generative AI tools in education is proliferating. According to a recent survey by Wiley, 58% of college instructors reported using generative AI in their classrooms, and another third expressed openness to adopting it in the future (Aravantinos et al., 2024). Moreover, over 60% of instructors are somewhat or very familiar with generative AI tools, reflecting the increasing integration of this technology in educational settings (Colby, 2023). ChatGPT, as part of generative AI (GenAI) technology, has opened up new educational opportunities as a technical aid and an adaptive learning assistant in various fields of study. ChatGPT demonstrates its ability to help students develop critical thinking skills (Guo & Lee, 2023), solve complex problems (Urban et al., 2024), and deepen their understanding across multiple disciplines, including physics (Lavidas et al., 2024). The application of ChatGPT in physics education offers an opportunity for students to receive additional explanations outside of class, which can help overcome misconceptions or complex concepts that are often barriers to understanding physics (Festiyed et al., 2024; Kotsis, 2024; Liang et al., 2023). With its adaptive capabilities,

ChatGPT allows students to customize the interaction according to their needs and questions, creating a more dynamic and interactive learning experience (Mai et al., 2024). In addition, research shows that this technology increases students' intrinsic motivation and enriches their learning experience by presenting customized content to meet specific learning needs (Adiguzel et al., 2023; Baidoo-Anu & Ansah, 2023; Papadakis et al., 2023). In studies on its use, ChatGPT is effective as a tool to exercise students' critical and creative thinking skills, provide in-depth feedback on problem-solving-based tasks, and serve as an additional learning resource to reinforce understanding fundamental physics concepts (Krupp et al., 2024; Su et al., 2023).

Although the potential of ChatGPT in improving the quality of learning has been recognized in various fields, most research related to the use of ChatGPT in education has focused on student groups in other fields or at higher levels of education. For example, Shoufan (2023) explored the perceptions of senior students in a computer engineering program towards using ChatGPT, while Ngo (2023) examined university students' general perceptions regarding the benefits of ChatGPT in an educational setting. Chellappa & Luximon (2024) and Singh et al. (2023) also researched the perceptions of product design and computer science students but have yet to target the field of physics. Valova et al. (2024) and Xu et al. (2024) examined students' perceptions of ChatGPT without focusing on a specific discipline. Research more relevant to the physics context was conducted by Ding et al. (2023), which investigated students' perceptions of ChatGPT as a virtual tutor in physics classes. This study evaluated the accuracy of ChatGPT in answering physics questions, measuring the relationship between students' level of trust in ChatGPT and the accuracy of the answers provided, as well as the impact of that trust on students' perceptions. However, this study was limited to the role of ChatGPT as a technical tool in answering questions. It did not explore the broader student learning experience in understanding physics concepts using ChatGPT.

While prior studies have explored ChatGPT's applications in education, only some have examined its potential to enhance high school students' understanding of physics concepts (Hakiki et al., 2023). This lack of research highlights a gap in understanding how ChatGPT can specifically support physics education at the high school level, a critical phase for developing foundational knowledge in science. This study aims to fill this knowledge gap by analyzing high school students' perceptions of using ChatGPT to support physics learning (Abubakar et al., 2024). The research will look at critical aspects of students' experiences, including the effectiveness of ChatGPT in helping them understand physics concepts, the quality of information provided, and how students use ChatGPT to solve physics problems and deepen their understanding (Abubakar et al., 2024). By focusing on students studying physics, this study is expected to make new contributions regarding the application of AI technology in science education. This research can also guide educational technology developers to optimize ChatGPT features to support physics learning and improve students' outcomes.

2 Methodology

A quantitative research approach was used to comprehensively understand high school students' perceptions of using ChatGPT in physics learning. Based on the recommendations of Creswell and Creswell (2017), a survey strategy with questionnaires was chosen as it allowed for the collection of extensive data regarding students' general views on using ChatGPT. This strategy best suited the study's primary objective: thoroughly understanding students' perceptions (Petousi & Sifaki, 2020).

2.1 Questionnaire development

The questionnaire was developed based on the main focus of the study, which included six themes: understanding of physics concepts, effectiveness of problem solving and answers, clarification and correction of information, discussion and further exploration, trust in the accuracy of information, and experience and recommendation of use (Brandhofer & Tengler, 2024; Uygun et al., 2024). These themes were determined based on a review of the literature and relevance to the research objectives, namely exploring students' experiences in utilizing ChatGPT to improve their understanding of physics concepts, solve problems, and explore physics-related topics (Chan & Hu, 2023; Ding et al., 2023; Fadillah & Sahyar, 2023; Usmeldi, 2015).

The questionnaire comprised 18 items (Table 1), each designed to capture a specific aspect of students' experience using ChatGPT in physics learning. The questionnaire was pilot-tested on 50 senior high school students with ChatGPT experience in physics learning. Reliability tests

were conducted using Cronbach's alpha to ensure internal consistency. The results showed an alpha value of 0.969, which indicates excellent reliability (Tavakol & Dennick, 2011). After the pilot test, the questionnaire was used in the primary research phase.

Table 1 Questionnaire items (n =18)

Theme	Code	Items
Concept understanding	CU1	ChatGPT can provide relevant examples to explain physics concepts
	CU2	ChatGPT helps me relate physics concepts to everyday life.
	CU3	ChatGPT helps me find additional information that enriches my understanding of physics concepts.
	CU4	ChatGPT helps improve my understanding of physics.
	CU5	ChatGPT meets my needs in learning physics concepts
Effectiveness of question and answer completion	EC1	ChatGPT helps me solve physics problems better
	EC2	I feel that ChatGPT can answer my questions about physics concepts.
	EC3	ChatGPT provides consistent explanations of physics concepts.
	EC4	ChatGPT can answer advanced physics questions accurately
	EC5	The physics formulas provided by ChatGPT are always correct
Clarification and correction of information	CC1	ChatGPT can correct errors or provide clarification if I inquire about inaccurate information
	CC2	ChatGPT helped me identify and correct misconceptions about physics concepts.
Discussion and further exploration	DE1	I can ask ChatGPT follow-up questions to deepen my understanding of physics.
	DE2	I can explore practical applications of physics theories through discussions with ChatGPT.
Trust in the accuracy of information	TA1	I feel that ChatGPT sometimes provides ambiguous or confusing information
	TA2	I feel the need to verify information from ChatGPT with other sources
Experience and recommendations for use	ER1	Overall, my experience using ChatGPT in physics learning has been very positive
	ER2	I would recommend using ChatGPT to my friends who are learning physics

2.2 Data collection and sampling

The questionnaire was distributed to high school students in Indonesia using the convenience sampling method, which allows respondents to be selected based on their availability and willingness to participate. A 4-point Likert scale was used to reduce ambiguity due to neutral choices with options from "strongly disagree" (code 1) to "strongly agree" (code 4) (Taherdoost, 2022).

The questionnaire was divided into two sections: The first part contained demographic questions, while the second part contained 18 items to capture students' perceptions regarding using ChatGPT in physics learning. A screening question was asked at the beginning of the survey to ensure respondents had used ChatGPT for physics learning. After data collection, 117 valid responses were obtained. Data from the pilot study was incorporated into the analysis, resulting in a final sample size of 167 respondents.

A comparative analysis was conducted using a t-test between the first 20 respondents and the last 20 respondents to address the potential non-response bias that often arises in cross-sectional studies. The t-test results showed that there were no significant differences on all items between the initial and final respondents ($p > 0.05$), thus indicating that non-response bias was not a significant issue in this study (Baabdullah, 2024).

2.3 Data analysis

Data were analyzed using IBM SPSS Statistics version 26. Descriptive statistics (frequency, percentage, mean, and standard deviation) were used to summarize the survey data. In addition, inferential statistics (one-way ANOVA) were conducted to examine the effect of demographic variables, such as gender, academic level, and duration of ChatGPT use, on students' perceptions regarding the use of ChatGPT in physics learning. The choice of one-way ANOVA was based on the need to compare the mean perceptions among different groups based on the research objective of exploring the influence of demographics on students' perceptions.

Effect sizes were calculated using partial eta-squared ($\eta_p{}^2$) to interpret the practical significance of the findings. Partial eta-squared is an effect size measure that expresses the proportion of variance explained by one or more independent variables, commonly used alongside ANOVA. $\eta_p{}^2$ values can be interpreted as follows: 0.01 indicates a small effect, 0.06 represents a medium effect, and 0.14 suggests a significant effect (Norouzian & Plonsky, 2018). This additional analysis ensures a comprehensive understanding of the practical importance of the observed differences.

3 Results

3.1 Respondent characteristics

Table 2 displays the characteristics of the 167 respondents included in this study. As explained earlier, data from respondents in the pilot study were included in the analysis because they were considered valuable, so the final total sample included all responses obtained. By gender, most respondents were female (58.08%), while the rest were male (41.92%). Academically, the majority of respondents were in class X (52.69%), followed by class XI (28.74%) and class XII (18.56%). In terms of duration of ChatGPT usage, more than half of the respondents (62.87%) have been using ChatGPT for more than 6 months, 17.37% have been using it for more than a year, and another 19.76% have been using the platform for more than 1.5 years.

Table 2 Characteristics of respondents (n = 167)

Characteristic	Criteria	Frequency	Percentage
Gender	Male	70	41.92
	Female	97	58.08
Academic level	X (15 - 16 years old)	88	52.69
	XI (16 - 17 years old)	48	28.74
	XII (17 - 18 years old)	31	18.56
Duration of use of ChatGPT	Over 6 months	105	62.87
	Over 1 year	29	17.37
	Over 1.5 years	33	19.76

3.2 Student perception of ChatGPT

The survey results are presented in Table 3, showing that students' perceptions of ChatGPT in physics learning are generally positive, with some areas considered more prominent. Items TA2 and CU3 scored the highest, with an average of 3.20, indicating that although ChatGPT was perceived to enrich physics understanding, students still felt it was essential to verify information from other sources, reflecting their critical attitude. Items DE1 and ER1, with averages of 3.15 and 3.14, respectively, indicated that ChatGPT helped students ask follow-up questions to deepen understanding and provide an overall positive learning experience. In addition, ChatGPT was considered effective in correcting misconceptions (CC2) and clarifying information that may be inaccurate (CC1), scoring 3.14 and 3.13 respectively. However, some students also felt that the information provided by ChatGPT was sometimes ambiguous or confusing (TA1), indicating the need for more consistent explanations.

Table 3 Student perceptions of ChatGPT (data has been sorted from highest to lowest mean

Code	Item	Mean	SD
TA2	I feel the need to verify information from ChatGPT with other sources	3.20	0.83
CU3	ChatGPT helps me find additional information that enriches my understanding of physics concepts	3.20	0.77
DE1	I can ask ChatGPT follow-up questions to deepen my understanding of physics	3.15	0.80
ER1	Overall, my experience using ChatGPT in physics learning has been very positive	3.14	0.91
CC2	ChatGPT helped me identify and correct misconceptions about physics concepts	3.14	0.82
CC1	ChatGPT can correct errors or provide clarification if I inquire about inaccurate information	3.13	0.84
TA1	I feel that ChatGPT sometimes provides ambiguous or confusing information	3.13	0.87
EC1	ChatGPT helps me solve physics problems better	3.12	0.88
CU4	I feel that ChatGPT helps improve my understanding of physics	3.12	0.86
CU2	ChatGPT helps me relate physics concepts to everyday life	3.11	0.82
DE2	I can explore practical applications of physics theories through discussions with ChatGPT	3.09	0.83
EC2	I feel that ChatGPT can answer my questions about physics concepts	3.08	0.90
CU1	ChatGPT can provide relevant examples to explain physics concepts	3.07	0.87
EC3	ChatGPT provides consistent explanations of physics concepts	3.07	0.83
EC4	ChatGPT can answer advanced physics questions accurately	3.07	0.86
CU5	ChatGPT meets my needs in learning physics concepts	3.04	0.83
EC5	The physics formulas provided by ChatGPT are always correct	3.02	0.82
ER2	I would recommend using ChatGPT to my friends who are learning physics	3.00	0.96

Regarding supporting problem-solving, items EC1 and CU4, with an average score of 3.12, showed that ChatGPT helped students work through problems and improved their understanding of physics concepts. ChatGPT's ability to relate concepts to everyday life (CU2) and explore practical applications of physics theories (DE2) also received positive ratings, with scores of

3.11 and 3.09, respectively. On the other hand, although ChatGPT was perceived to provide clear answers (EC2, 3.08) and relevant examples (CU1, 3.07), the consistency of explanations (EC3, 3.07) as well as accuracy in answering advanced questions (EC4, 3.07) were still felt by students could be further improved. Item CU5 indicates that ChatGPT generally meets students' learning needs in understanding physics concepts (3.04), and item EC5 shows a reasonably positive perception regarding the accuracy of formulas provided by ChatGPT (3.02). Meanwhile, item ER (average 3.00) shows that some students are willing to recommend ChatGPT to friends, although this recommendation level is still relatively low compared to other items.

3.3 Student perceptions of ChatGPT based on variables

Table 4 presents student perceptions of ChatGPT, showing variations based on gender, academic level, and duration of use. These differences illustrate how demographic characteristics and usage experience contribute to students' views of ChatGPT. In the analysis of perceptions by gender, female students showed higher mean scores than male students on most items. For example, on item CU3, which measured how often students used ChatGPT to search for additional information, female students obtained a mean score of 3.36 (SD = 0.74) compared to male students, who had a mean score of 2.97 (SD = 0.76). This trend was also seen in item DE1, which assessed the specific benefits of ChatGPT, where female students recorded an average of 3.30 (SD = 0.71), higher than male students who only recorded 2.94 (SD = 0.87). This difference could indicate that female students may use the app more frequently in academic activities or appreciate the features offered by ChatGPT more. On the other hand, on some items such as EC4, which assessed students' views on ChatGPT's ability to explain complex concepts, male and female students' scores did not show significant differences, with mean scores of 2.97 (SD = 0.85) for males and 3.13 (SD = 0.86) for females respectively. It could indicate a similarity in perception between the two groups in certain aspects.

Table 4 Student perceptions of ChatGPT based on variables

Code	Gender				Academic level						Duration of use of ChatGPT					
	Male		Female		X		XI		XII		Over 6 months		Over 1 year		Over 1.5 years	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TA2	3.1	0.85	3.28	0.8	3.25	0.81	3.04	0.87	3.32	0.79	3.26	0.8	3.21	0.82	3.03	0.92
CU3	2.97	0.76	3.36	0.74	3.34	0.77	2.85	0.74	3.32	0.65	3.25	0.76	3.24	0.58	3	0.94
DE1	2.94	0.87	3.3	0.71	3.31	0.72	2.9	0.83	3.1	0.87	3.18	0.74	3.1	0.86	3.09	0.91
ER1	2.91	1	3.31	0.8	3.21	0.92	3.02	0.89	3.16	0.9	3.22	0.85	2.97	1.05	3.06	0.93
CC2	2.93	0.84	3.29	0.78	3.31	0.81	2.9	0.81	3.03	0.8	3.2	0.75	3.1	0.82	2.97	1.02
CC1	2.96	0.84	3.25	0.83	3.26	0.85	2.85	0.82	3.16	0.78	3.19	0.77	3.14	0.79	2.91	1.07
TA1	3	0.82	3.22	0.9	3.17	0.94	2.88	0.76	3.39	0.76	3.17	0.84	3.17	0.89	2.94	0.97
EC1	2.97	0.92	3.23	0.85	3.23	0.88	3	0.8	3	1	3.15	0.84	3.1	0.94	3.03	0.98
CU4	2.93	0.89	3.26	0.81	3.27	0.8	2.88	0.87	3.07	0.93	3.14	0.81	3.14	0.95	3.03	0.92
CU2	2.89	0.88	3.28	0.73	3.23	0.81	2.85	0.8	3.19	0.79	3.18	0.73	3.07	0.96	2.94	0.93
DE2	2.94	0.85	3.2	0.8	3.25	0.83	2.77	0.81	3.13	0.72	3.11	0.79	3.14	0.74	2.97	1.02
EC2	2.86	0.97	3.24	0.81	3.19	0.9	2.88	0.82	3.07	1	3.12	0.83	3.1	0.98	2.91	1.04
CU1	2.84	0.93	3.24	0.79	3.19	0.88	2.81	0.79	3.13	0.88	3.13	0.83	3.07	0.88	2.88	0.96
EC3	2.89	0.86	3.21	0.78	3.27	0.78	2.81	0.79	2.9	0.87	3.11	0.8	3.17	0.8	2.85	0.91
EC4	2.97	0.85	3.13	0.86	3.22	0.82	2.85	0.8	2.97	0.98	3.07	0.82	3.28	0.84	2.88	0.96
CU5	2.87	0.83	3.17	0.81	3.23	0.74	2.73	0.84	3	0.93	3.09	0.79	3.03	0.91	2.91	0.91
EC5	2.81	0.86	3.17	0.77	3.28	0.74	2.67	0.78	2.81	0.87	3.13	0.73	2.86	0.88	2.79	0.99
ER2	2.81	1.04	3.13	0.87	3.06	0.91	2.92	0.9	2.97	1.17	3.11	0.87	2.93	1.19	2.73	0.98

Students' perceptions of ChatGPT also varied according to their academic level. Grade X students tended to have more positive perceptions on some aspects than Grade XI and XII students. For example, on item CU3, grade X students recorded the highest mean of 3.34 (SD = 0.77), compared to grade XI students with a lower mean of 2.85 (SD = 0.74). The grade XII students' mean of 3.32 (SD = 0.65) was close to the grade X students' mean, suggesting that early and late grades may be more open to new information from this app than intermediate students. On item TA1, which measures ChatGPT's ability to provide transparent information, grade XII students obtained the highest mean score of 3.39 (SD = 0.76), while grade XI students had a lower score of 2.88 (SD = 0.76). Grade XII students may value the quality and clarity of information provided by ChatGPT.

Students' perceptions of ChatGPT also changed based on the duration of use of the application. Students who have used ChatGPT for more than six months generally have more positive perceptions than those who are new to the app or have used it for more than 1.5 years. For

example, on item TA2, which evaluates general perceptions of the benefits of ChatGPT, students with a duration of use of more than six months obtained an average score of 3.26 (SD = 0.80). In contrast, those who had used ChatGPT for over 1.5 years had a lower score of 3.03 (SD = 0.92). A similar trend was also seen in item ER1, where students who used ChatGPT for more than six months recorded an average of 3.22 (SD = 0.85), while those who had used ChatGPT for more than 1.5 years recorded an average of 3.06 (SD = 0.93). However, there were some aspects where the perceptions of students who used ChatGPT for more than a year were more favourable than those who were new to the app. On item EC4, which assessed ChatGPT in helping to explain complex concepts, students who had used the app for more than a year had the highest mean score of 3.28 (SD = 0.84) compared to students who had only used it for six months who had a mean score of 3.07 (SD = 0.82). Using ChatGPT for longer allows students to see more specific benefits from this tool.

3.4 One-way ANOVA for students' perception of ChatGPT

Before conducting the one-way Analysis of Variance (ANOVA) test, the test assumptions were fulfilled to ensure the reliability of the analysis. The normality of the data was checked using the skewness and kurtosis values, given that the sample in this study amounted to more than 50. Based on the suggestion of Hong et al. (2023), using Kolmogorov-Smirnov or Shapiro-Wilk tests is usually less effective in large samples, and skewness and kurtosis values can be an adequate alternative. According to Byrne (2013) and Hair et al. (2010), skewness values between -2 to +2 and kurtosis values between -7 to +7 indicate a near-normal distribution. In this study, the test results showed skewness values ranging from -1.046 to -0.699 and kurtosis ranging from -0.224 to 0.851, which met the normality criteria and allowed the data to be further analyzed using ANOVA. Table 5 shows the complete distribution of normality test results.

Table 5 Normality test results using skewness and kurtosis values

Code	Skewness	Kurtosis
TA2	-1.046	0.851
CU3	-0.916	0.848
DE1	-0.928	0.810
ER1	-0.976	0.264
CC2	-0.922	0.637
CC1	-0.850	0.286
TA1	-0.964	0.457
EC1	-0.874	0.144
CU4	-0.933	0.477
CU2	-0.951	0.799
DE2	-0.880	0.554
EC2	-0.861	0.109
CU1	-0.699	-0.157
EC3	-0.850	0.509
EC4	-0.764	0.084
CU5	-0.779	0.321
EC5	-0.751	0.314
ER2	-0.793	-0.224

In addition to normality, homogeneity of variance among groups was tested using Levene's test to ensure relatively consistent data variability in each group. Table 6 shows the results of Levene's test by gender group, academic level, and duration of ChatGPT use. Most items met the assumption of homogeneity of variance ($p > 0.05$). However, some items showed significant results in the Levene test, namely items CU3 and TA1 for the gender group and items CC1 and ER2 for the duration of the ChatGPT use group, with $p < 0.05$. A standard ANOVA was inappropriate because the assumption of homogeneity was not met for these items. Instead, Welch's test was used for these items, as this method is more reliable in situations of variance inhomogeneity (Field, 2024; Moder, 2010).

The results of the ANOVA test are summarized in Table 7. Significant differences were seen in gender and academic level for some items, highlighting variations in students' perceptions of ChatGPT in physics learning. Significant differences were found between genders for items CU3, DE1, ER1, CC2, CU4, CU2, EC2, CU1, EC3, CU5, EC5, and ER2 (all $p < 0.05$), with medium effect sizes ($\eta_p{}^2$) ranging from 0.03 to 0.06. Students' academic levels significantly influenced their perceptions for items CU3, DE1, CC2, CC1, TA1, CU4, CU2, DE2, CU1, EC3, EC4, CU5, and EC5 ($p < 0.05$), with effect sizes ($\eta_p{}^2$) ranging from 0.04 to 0.12. In contrast, the duration of ChatGPT usage had no significant impact on most items, as indicated by

Table 6 Homogeneity test results using Levene's value

Code	Levene Statistic		
	Gender	Academic level	Duration of use of ChatGPT
TA2	0.587	0.505	0.937
CU3	5.786*	1.954	0.263
DE1	0.038	0.371	0.616
ER1	0.539	1.330	0.505
CC2	0.556	1.432	0.164
CC1	2.296	1.344	0.044*
TA1	4.551*	3.000	0.800
EC1	0.907	2.748	0.928
CU4	0.453	0.108	0.655
CU2	0.026	0.355	0.330
DE2	0.833	0.389	0.386
EC2	0.596	2.919	0.238
CU1	1.173	0.773	0.503
EC3	0.000	0.529	0.736
EC4	1.726	2.250	0.166
CU5	0.086	0.828	0.617
EC5	0.068	0.195	0.087
ER2	3.426	2.838	0.013*

Note: $^{*}p < 0.05$.

nonsignificant F-values and very small effect sizes ($\eta_p{}^2 < 0.03$). While some minor differences were noted (*e.g.*, for EC5), these were not substantial enough to indicate a broader effect. The results indicate that students' perceptions of ChatGPT in physics learning differ between genders and also vary according to their academic level. However, the duration of use did not have a major effect in shaping students' perceptions of ChatGPT.

Table 7 ANOVA test results

Code	Gender		Academic level		Duration of use of ChatGPT	
	F	$\eta_p{}^2$	F	$\eta_p{}^2$	F	$\eta_p{}^2$
TA2	1.909	0.01	1.392	0.02	0.948	0.01
CU3	11.019**	0.06	7.204**	0.08	1.360	0.02
DE1	8.499**	0.05	4.393*	0.05	0.218	0.00
ER1	8.039**	0.05	0.642	0.01	1.061	0.01
CC2	8.163**	0.05	4.379*	0.05	1.019	0.01
CC1	4.918*	0.03	3.767*	0.04	1.405	0.02
TA1	2.527	0.02	3.598*	0.04	0.937	0.01
EC1	3.446	0.02	1.383	0.02	0.243	0.00
CU4	6.202*	0.04	3.538*	0.04	0.223	0.00
CU2	9.890**	0.06	3.523*	0.04	1.152	0.01
DE2	3.867	0.02	5.538**	0.06	0.440	0.01
EC2	7.560**	0.04	1.975	0.02	0.729	0.01
CU1	8.776**	0.05	3.147*	0.04	1.080	0.01
EC3	6.323*	0.04	5.956**	0.07	1.572	0.02
EC4	1.463	0.01	3.083*	0.04	1.665	0.02
CU5	5.197*	0.03	5.96**	0.07	0.565	0.01
EC5	7.642**	0.04	11.172***	0.12	2.893	0.03
ER2	4.639*	0.03	0.352	0.00	2.071	0.03

Note: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

As the assumption of homogeneity of variance was not met on items CU3, TA1, CC1, and ER2, Welch's test was used instead of standard ANOVA for these four items. Table 8 displays Welch's test results, which show a significant difference only in item CU3 ($p < 0.01$), meaning that there is variation in perceptions between gender groups regarding the use of ChatGPT in obtaining additional information for physics. However, Welch's test on items TA1, CC1, and ER2 showed no significant difference between groups.

Table 8 Welch test results

Gender		Duration of use of ChatGPT	
Code	Welch statistics	Code	Welch statistics
CU3	10.913**	CC1	0.969
TA1	2.611	ER2	2.04

Note: $^{**}p < 0.01$.

4 Discussion

The results of this study show that students' perceptions of the use of ChatGPT in physics learning are generally very positive. High scores on items TA2 and CU3, which measure ChatGPT's ability to enrich understanding and encourage critical thinking, confirm that this technology is a complementary learning tool. Nevertheless, students still verify the information obtained from ChatGPT through other sources, which reflects their critical evaluation skills. This finding supports the recommendation of Javaid et al. (2023) to encourage an evidence-based approach to utilizing educational technology.

One significant finding was the difference in perceptions based on gender. Female students showed significantly more positive perceptions than male students on some items, such as CU3 (using ChatGPT to seek additional information, $\eta_p{}^2 = 0.06$) and DE1 (perceived benefits of ChatGPT, $\eta_p{}^2 = 0.05$). These findings support Shoufan's (2023) view that gender differences can influence the adoption and perception of technology tools in education. However, such differences were not seen in advanced problem-solving ability (*e.g.*, EC4, $\eta_p{}^2 = 0.01$), indicating that the application of ChatGPT for complex cognitive tasks was perceived similarly across genders, thus highlighting the universal potential of this technology in supporting deep learning (Shoufan, 2023). In addition, variations in perceptions by grade level were also found. Grade X students reported higher positive perceptions on aspects such as seeking additional information (CU3, $\eta_p{}^2 = 0.12$) and clarity of information provided (TA1, $\eta_p{}^2 = 0.09$) than higher-grade students. It could indicate tremendous enthusiasm among younger students in adopting new learning tools, consistent with Susnjak and McIntosh's (2024) opinion that early exposure to technology can facilitate better integration into educational practices.

Interestingly, the duration of ChatGPT use did not significantly affect students' overall perceptions (*e.g.*, EC5, $\eta_p{}^2 = 0.02$). It could imply that the benefits of this technology can already be perceived within a short period. However, it contradicts the findings of Al-sa'di and Miller (2023), who stated that long-term use of AI tools tends to provide a deeper understanding of their functionality. This discrepancy may be due to the physics learning context, where advanced problem-solving requires more complex assistance beyond ChatGPT's capabilities. Another interesting result was the moderate effectiveness of ChatGPT's information verification item (TA2, $\eta_p{}^2 = 0.03$) despite its high mean score (3.20). This finding highlights the importance of maintaining critical engagement when using AI-based tools, as emphasized by Kasneci et al. (2023). The data also showed that although ChatGPT was effective in correcting misconceptions (CC2, $\eta_p{}^2 = 0.05$), occasional response inconsistency (EC3, $\eta_p{}^2 = 0.002$) was one of the limitations identified by students, thus reinforcing the need to verify information independently.

These findings have important implications for educators and policymakers. Differences in perceptions by gender and grade level suggest the need for tailored strategies to optimize the integration of ChatGPT in the physics curriculum. For example, specific training programs could be designed to meet the needs of male or upper-level students who tend to show lower engagement with ChatGPT. In addition, students' critical behaviour in verifying information emphasizes the importance of positioning ChatGPT as a complementary, rather than primary, tool. The integration of AI literacy in the curriculum can help students evaluate and utilize information from ChatGPT more effectively, in line with the framework of Javaid et al. (2023). ChatGPT can be a cost-effective tool for policymakers to reduce the resource gap in underprivileged schools. Given its moderate effect in supporting clarification (CC1, $\eta_p{}^2 = 0.03$) and correcting misconceptions (CC2, $\eta_p{}^2 = 0.05$), this technology can improve physics conceptual understanding, especially for students with limited access to traditional learning resources.

This study has limitations that need to be noted. The sample of high school students in one region limits the generalizability of the findings to other educational levels or contexts. In addition, the survey approach relied on subjective reports, which may need to fully reflect the impact of ChatGPT on physics learning. Future research could use experimental designs to elucidate causal relationships and evaluate the role of external factors, such as teacher facilitation, in optimizing the use of ChatGPT. Furthermore, physics as a scientific discipline requires a high level of conceptual understanding, which sometimes goes beyond the current capabilities of ChatGPT. The item related to advanced questions (EC4) showed moderate perception ($\eta_p{}^2 = 0.02$), which hints at the need for a hybrid approach that combines AI tools with direct guidance from teachers. Future research could also explore how the integration of ChatGPT in a blended learning environment impacts profound learning outcomes and the development of critical thinking skills. In closing, given the evolution of AI tools such as ChatGPT, further research needs to extend the scope to other STEM subjects or cross-disciplinary learning. Longitudinal

and exploratory studies in diverse education systems would provide greater insight into the long-term value of this technology in education.

5 Conclusion

This study investigated students' perceptions of using ChatGPT in physics learning, focusing on its understanding of physics concepts, effectiveness of problem solving and answers, clarification and correction of information, discussion and further exploration, trust in the accuracy of information, and experience and recommendation of use. The main findings show that students have generally positive perceptions of this technology, with the highest scores on ChatGPT's ability to clarify concepts and provide additional relevant information. Differences in perceptions based on gender and grade level were also found, with female students and grade X students showing more positive responses than other groups. Furthermore, while ChatGPT effectively supported concept clarification and misconception correction, its limitations in providing consistent responses highlight the need for self-verification of information. This study contributes to the development of literature in the field of educational technology, particularly on the integration of AI-based tools such as ChatGPT in physics learning. The findings support the use of AI technology as a complementary tool that can enhance student's learning experience, especially in terms of conceptual understanding. In addition, the identification of differences in perceptions by gender and grade level provides new insights into how technology adoption can be optimized through a more personalized and contextualized approach. The results of this study provide an empirical foundation for educators and policymakers to design more inclusive and effective strategies for integrating AI technologies into learning practices.

Conflicts of interest

The authors declare that they have no conflict of interest.

References

Abubakar, U., Falade, A. A., & Ibrahim, H. A. (2024). Redefining student assessment in Nigerian tertiary institutions: The impact of AI technologies on academic performance and developing countermeasures. Advances in Mobile Learning Educational Research, 4(2), 1149–1159. https://doi.org/10.25082/amler.2024.02.009

Abubakar, U., Onasanya, S. A., & Ibrahim, H. A. (2024). Student perspectives and impact of AI integration in pedagogical practices in Nigerian tertiary institutions. Advances in Mobile Learning Educational Research, 4(2), 1135–1148. https://doi.org/10.25082/amler.2024.02.008

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. Contemporary Educational Technology, 15(3), ep429. https://doi.org/10.30935/cedtech/13152

Al-sa'di, A., & Miller, D. (2023). Exploring the Impact of Artificial Intelligence language model ChatGPT on the User Experience. International Journal of Technology, Innovation and Management (IJTIM), 3(1), 1–8. https://doi.org/10.54489/ijtim.v3i1.195

Aravantinos, S., Lavidas, K., Voulgari, I., Papadakis, S., Karalis, T., & Komis, V. (2024). Educational Approaches with AI in Primary School Settings: A Systematic Review of the Literature Available in Scopus. Education Sciences, 14(7), 744. https://doi.org/10.3390/educsci14070744

Baabdullah, A. M. (2024). The precursors of AI adoption in business: Towards an efficient decision-making and functional performance. International Journal of Information Management, 75, 102745. https://doi.org/10.1016/j.ijinfomgt.2023.102745

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. Journal of AI, 7(1), 52–62. https://doi.org/10.61969/jai.1337500

Brandhofer, G., & Tengler, K. (2024). Acceptance of artificial intelligence in education: opportunities, concerns and need for action. Advances in Mobile Learning Educational Research, 4(2), 1105–1113. https://doi.org/10.25082/amler.2024.02.005

Byrne, B. M. (2013). Structural Equation Modeling with Mplus. Routledge. https://doi.org/10.4324/9780203807644

Caldarini, G., Jaf, S., & McGarry, K. (2022). A Literature Survey of Recent Advances in Chatbots. Information, 13(1), 41. https://doi.org/10.3390/info13010041

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20(1). https://doi.org/10.1186/s41239-023-00411-8

Chellappa, V., & Luximon, Y. (2024). Understanding the perception of design students towards ChatGPT. Computers and Education: Artificial Intelligence, 7, 100281. https://doi.org/10.1016/j.caeai.2024.100281

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. IEEE Access, 8, 75264–75278. https://doi.org/10.1109/access.2020.2988510

Colby. (2023). Generative AI already being used in majority of college classrooms, according to instructors in new wiley survey. Wiley. https://newsroom.wiley.com

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. International Journal of Educational Technology in Higher Education, 20(1). https://doi.org/10.1186/s41239-023-00434-1

Fadillah, M. A., & Sahyar, S. (2023). Development of Higher Order Thinking Skills (HOTS) Test Instruments on Parabolic and Circular Motion Materials in High Schools. Berkala Ilmiah Pendidikan Fisika, 11(3), 329. https://doi.org/10.20527/bipf.v11i3.16697

Festiyed, F., Tanjung, Y. I., & Fadillah, M. A. (2024). ChatGPT in Science Education: A Visualization Analysis of Trends and Future Directions. JOIV: International Journal on Informatics Visualization, 8(3-2), 1614-1624. https://doi.org/10.62527/joiv.8.3-2.2987

Field, A. (2024). Discovering statistics using IBM SPSS statistics. Sage Publications Limited.

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for Enhancing Critical Thinking Skills. Journal of Chemical Education, 100(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

Hair, J. F., Black, B., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis: Global edition (7th Editio). Harlow: Pearson Education.

Hakiki, M., Fadli, R., Samala, A. D., Fricticarani, A., Dayurni, P., Rahmadani, K., & Astiti, A. D. (2023). Exploring the impact of using Chat-GPT on student learning outcomes in technology learning: The comprehensive experiment. Advances in Mobile Learning Educational Research, 3(2), 859–872. https://doi.org/10.25082/amler.2023.02.013

Hong, W. C. H., Chi, C., Liu, J., Zhang, Y., Lei, V. N.-L., & Xu, X. (2022). The influence of social education level on cybersecurity awareness and behaviour: a comparative study of university students and working graduates. Education and Information Technologies, 28(1), 439–470. https://doi.org/10.1007/s10639-022-11121-5

Huang, J., Saleh, S., & Liu, Y. (2021). A Review on Artificial Intelligence in Education. Academic Journal of Interdisciplinary Studies, 10(3), 206. https://doi.org/10.36941/ajis-2021-0077

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 3(2), 100115. https://doi.org/10.1016/j.tbench.2023.100115

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kotsis, K. T. (2024). Correcting Students' Misconceptions in Physics Using Experiments Designed by ChatGPT. European Journal of Contemporary Education and E-Learning, 2(2), 83–100. https://doi.org/10.59324/ejceel.2024.2(2).07

Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K. E., Lukowicz, P., Kuhn, J., Küchemann, S., & Karolus, J. (2024). Unreflected Acceptance – Investigating the Negative Consequences of ChatGPT-Assisted Problem Solving in Physics Education. HHAI 2024: Hybrid Human AI Systems for the Social Good. https://doi.org/10.3233/faia240195

Lavidas, K., Voulgari, I., Papadakis, S., Athanassopoulos, S., Anastasiou, A., Filippidi, A., Komis, V., & Karacapilidis, N. (2024). Determinants of Humanities and Social Sciences Students' Intentions to Use Artificial Intelligence Applications for Academic Purposes. Information, 15(6), 314. https://doi.org/10.3390/info15060314

Liang, Y., Zou, D., Xie, H., & Wang, F. L. (2023). Exploring the potential of using ChatGPT in physics education. Smart Learning Environments, 10(1). https://doi.org/10.1186/s40561-023-00273-7

Mai, D. T. T., Da, C. V., & Hanh, N. V. (2024). The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1328769

Moder, K. (2010). Alternatives to F-test in one-way ANOVA in case of heterogeneity of variances (a simulation study). Psychological Test and Assessment Modeling, 52(4), 343.

Ngo, T. T. A. (2023). The Perception by University Students of the Use of ChatGPT in Education. International Journal of Emerging Technologies in Learning (IJET), 18(17), 4–19. https://doi.org/10.3991/ijet.v18i17.39019

Norouzian, R., & Plonsky, L. (2017). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. Second Language Research, 34(2), 257–271. https://doi.org/10.1177/0267658316684904

Papadakis, S., Kravtsov, H. M., Osadchyi, V. V., Marienko, M. V., Pinchuk, O. P., Shyshkina, M. P., ... & Striuk, A. M. (2023). Revolutionizing education: using computer simulation and cloud-based smart technology to facilitate successful open learning. https://doi.org/10.31812/123456789/7375

Petousi, V., & Sifaki, E. (2020). Contextualising harm in the framework of research misconduct. Findings from discourse analysis of scientific publications. International Journal of Sustainable Development, 23(3/4), 149. https://doi.org/10.1504/ijsd.2020.115206

Sabzalieva, E., & Valentini, A. (2023). ChatGPT and artificial intelligence in higher education: Quick start guide.

Shoufan, A. (2023). Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey. IEEE Access, 11, 38805–38818. https://doi.org/10.1109/access.2023.3268224

Singh, H., Tayarani-Najaran, M.-H., & Yaqoob, M. (2023). Exploring Computer Science Students' Perception of ChatGPT in Higher Education: A Descriptive and Correlation Study. Education Sciences, 13(9), 924. https://doi.org/10.3390/educsci13090924

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. Assessing Writing, 57, 100752. https://doi.org/10.1016/j.asw.2023.100752

Susnjak, T., & McIntosh, T. (2024). ChatGPT: The End of Online Exam Integrity? Education Sciences, 14(6), 656. https://doi.org/10.3390/educsci14060656

Taherdoost, H. (2022). What is the best response scale for survey and questionnaire design; Review of different lengths of rating scale/attitude scale/likert scale. International Journal of Academic Research in Management, 8(1), 1–10.

Festiyed, F., Tanjung, Y. I., & Fadillah, M. A. (2024). ChatGPT in Science Education: A Visualization Analysis of Trends and Future Directions. JOIV: International Journal on Informatics Visualization, 8(3-2), 1614-1624. https://doi.org/10.62527/joiv.8.3-2.2987

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Uğraş, H., Uğraş, M., Papadakis, S., & Kalogiannakis, M. (2024). ChatGPT-Supported Education in Primary Schools: The Potential of ChatGPT for Sustainable Practices. Sustainability, 16(22), 9855. https://doi.org/10.3390/su16229855

Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. Computers & Education, 215, 105031. https://doi.org/10.1016/j.compedu.2024.105031

Usmeldi, U. (2015). The effectiveness of physics-based research in learning engineering physics. Jurnal Pendidikan IPA Indonesia, 4(1), 79–85. https://doi.org/10.15294/jpii.v4i1.3505

Uygun, D., Aktaş, I., Duygulu, İ., & Köseer, N. (2024). Exploring teachers' artificial intelligence awareness. Advances in Mobile Learning Educational Research, 4(2), 1093–1104. https://doi.org/10.25082/amler.2024.02.004

Valova, I., Mladenova, T., & Kanev, G. (2024). Students' Perception of ChatGPT Usage in Education. International Journal of Advanced Computer Science and Applications, 15(1). https://doi.org/10.14569/ijacsa.2024.0150143

Xu, X., Su, Y., Zhang, Y., Wu, Y., & Xu, X. (2024). Understanding learners' perceptions of ChatGPT: A thematic analysis of peer interviews among undergraduates and postgraduates in China. Heliyon, 10(4), e26239. https://doi.org/10.1016/j.heliyon.2024.e26239