PANTAO
(International Journal of the Humanities and Social Sciences)
Volume 5, Issue 4
ISSN 3028-0877 | e-ISSN 3155-6043 | DOI 10.69651/2022PIJHSS

Assessing students' perception on teacher-guided AI chatbots and their academic performance in Applied Social Sciences

Aje B. Barretto*
Cebu Technological University – Main Campus, Cebu City, Philippines

ABSTRACT

Artificial intelligence (AI) chatbots are increasingly present in Philippine senior high school classrooms, yet evidence on which of their functions support learning remains limited. This study assessed the perception of Grade 12 Humanities and Social Sciences (HUMSS) students on teacher-guided AI chatbots and determined whether this perception predicted their academic performance in Applied Social Sciences. A descriptive research design employing multiple linear regression analysis was used, involving 153 Grade 12 HUMSS students of the University of Cebu Lapu-Lapu and Mandaue, Mandaue City, Cebu, identified through complete enumeration. Data were gathered using an adapted 20-item perception survey and a researcher-constructed 50-item summative test, and were analyzed using frequency, percentage, weighted mean, standard deviation, and multiple linear regression. The respondents' perception was Positive in terms of clarifying complex social concepts, prompting deeper follow-up questions, identifying potential bias in information, and perceived teacher guidance ($M = 3.74$, $SD = 0.94$), while their academic performance was Very Satisfactory ($M = 39.63$, $SD = 10.59$). Perception significantly predicted academic performance, $F(4, 148) = 4.54$, $p = .002$, $R^2 = .109$. Concept clarification ($\beta = .279$, $p = .020$) and teacher guidance ($\beta = .216$, $p = .030$) were positive predictors, follow-up questioning was a negative predictor ($\beta = -.349$, $p = .007$), and bias identification was not significant ($\beta = .101$, $p = .408$). The contribution of AI chatbots to learning depends on the specific function used and on active teacher mediation. The study proposes an Instructional Enhancement Plan for function-specific, teacher-mediated AI chatbot use in Applied Social Sciences instruction.

Keywords: Academic performance, AI chatbots, Applied Social Sciences, HUMSS, teacher guidance.

Date Submitted: October 20, 2026
Date Accepted: November 05, 2026
Date Published: November 15, 2026

INTRODUCTION

Social studies education is central to developing learners who can think critically, interpret social issues, and arrive at well-reasoned conclusions (González-Valencia et al., 2022). Inquiry-based instruction in the field asks students to formulate questions, evaluate evidence from diverse sources, weigh competing perspectives, and communicate well-

*Corresponding author / Email: king.aje1234@gmail.com
DOI: http://doi.org/10.69651/PIJHSS05041670

Recommended citation:
Barretto, A. B. (2026). Assessing students' perception on teacher-guided AI chatbots and their academic performance in Applied Social Sciences. Pantao (International Journal of the Humanities and Social Sciences) 5 (4), 1509-1518. http://doi.org/10.69651/PIJHSS05041670

supported conclusions (National Council for the Social Studies [NCSS], 2013). This orientation has converged with the rise of generative artificial intelligence (AI), particularly AI chatbots, which can provide immediate feedback, clarify difficult concepts, and sustain learner interaction. Chatbots have been shown to provide rapid academic assistance and support personalized learning (Okonkwo & Ade-Ibijola, 2021), and AI tools help learners generate ideas and ask follow-up questions (Tlili et al., 2023). At the same time, AI-generated responses may contain inaccuracies, oversimplifications, or biases when used without proper supervision (Kasneci et al., 2023).

The instructional value of chatbots appears to depend less on the technology than on the structure through which students use it. Okonkwo and Ade-Ibijola (2021) found that chatbots produce educationally significant results only when embedded within a structured pedagogical framework, and Kasneci et al. (2023) cautioned that, without teacher scaffolding, students tend to consume AI outputs uncritically, a risk with particular weight in social studies, where AI-generated content may omit contested perspectives. Meta-analytic evidence reports positive effects of ChatGPT on academic performance and higher-order thinking (Deng et al., 2025), yet Weidlich et al. (2025) warned that many of the underlying experiments do not adequately describe the treatment and outcomes being compared.

In the Philippines, curricular policies prescribe inquiry-based and constructivist approaches in Social Studies (Pecson, 2026), and DepEd Order No. 003, series of 2026, permits AI use in basic education provided that it remains ethical, human-centered, and subordinate to teacher judgment, with AI functioning only as an auxiliary tool that never replaces the teacher (Department of Education [DepEd], 2026). Philippine studies have documented the benefits and risks of AI tool use among senior high school students (Estrellado & Miranda, 2023) and a significant relationship between AI chatbot benefits and student engagement (Pimentel et al., 2024), while Filipino students have been found to value ChatGPT as a convenient academic guide at the risk of relying on it at the expense of essential learning (Acena, 2025). Milloria et al. (2024), in a quasi-experimental study of Grade 11 HUMSS students in Mandaue City, Cebu, found that AI-integrated instruction using ChatGPT significantly improved academic performance in Social Sciences. That study, however, introduced ChatGPT as an experimental treatment and measured general performance, without isolating whether concept clarification, follow-up questioning, bias identification, or teacher mediation relates to learning outcomes, and without examining chatbot use as it already occurs in classrooms.

This gap has practical consequences at the University of Cebu Lapu-Lapu and Mandaue (UCLM), where Grade 12 HUMSS students take Disciplines and Ideas in the Applied Social Sciences (DIASS) within a program that espouses inquiry-driven instruction. The researcher's classroom observations indicate that some students consult AI chatbots for quick answers instead of working through inquiry tasks themselves, while some teachers hesitate to use these tools in the absence of school-sanctioned guidelines. With the institution's policy on classroom AI use still under deliberation, decisions about AI chatbots are being made without local evidence on how guided use relates to performance, in a strand whose analytical goals have been described as only unevenly realized in practice (Estrera, 2020).

The study is anchored on three frameworks. Connectivism holds that learning in a digital era is the formation of connections among distributed nodes such as people, tools, and networked systems (Siemens, 2005, as cited in Alam, 2023), and generative AI fosters knowledge connectivity when students use it as an interactive agent rather than consume its output passively (Liang & Bai, 2025); clarifying concepts, prompting questions, and identifying bias represent the ways students draw meaning from, question, and evaluate the chatbot as an information node. Human-Centered AI theory holds that responsible AI should amplify human capability and preserve human agency (Shneiderman, 2022), a principle supported by evidence that teacher presence increases students' engagement with AI chatbots

(Li et al., 2026), which grounds perceived teacher guidance as a distinct dimension. The Intelligent TPACK framework stresses the teacher's knowledge of integrating AI-based tools ethically with content and pedagogy (Celik, 2023), linking these perceptions to academic performance in Applied Social Sciences measured across the levels of Bloom's revised taxonomy (Anderson & Krathwohl, 2001).

Statement of the problem

This study assessed the perception of Grade 12 HUMSS students of the University of Cebu Lapu-Lapu and Mandaue on teacher-guided AI chatbots and its influence on their academic performance in Applied Social Sciences for the school year 2026–2027 as basis for a proposed Instructional Enhancement Plan. Specifically, it answered the following questions:

1. What is the level of the respondents' perception on teacher-guided AI chatbots in learning in terms of clarifying complex social concepts, prompting deeper follow-up questions, identifying potential biases in information, and perceived teacher guidance?
2. What is the level of academic performance of the respondents in Applied Social Sciences?
3. Does the respondents' perception on teacher-guided AI chatbots in learning significantly predict their academic performance in Applied Social Sciences?
4. Based on the findings of the study, what Instructional Enhancement Plan can be proposed?

The null hypothesis that the respondents' perception did not significantly predict their academic performance was tested at the 0.05 level of significance.

METHODOLOGY

The study used a descriptive research design employing multiple linear regression analysis under the quantitative approach, which describes the current status of variables and determines whether some variables significantly predict another without manipulating them or establishing cause and effect (Creswell & Creswell, 2023). No intervention was administered, since teacher-guided AI chatbot use was already part of the respondents' regular DIASS instruction. The study followed an input-process-output approach, in which the respondents' perception and academic performance served as the input, instrument development, data gathering, and statistical analysis formed the process, and an Instructional Enhancement Plan was the output.

The study was conducted at UCLM, a private, non-sectarian satellite campus of the University of Cebu system along A.C. Cortes Avenue, Barangay Looc, Mandaue City, Cebu, Philippines, whose Senior High School program generally serves 1,500 to 2,500-plus students. The campus was purposively selected because it offers DIASS specifically to Grade 12 HUMSS students, has the infrastructure and internet connectivity needed for teacher-guided chatbot use, and was accessible to the researcher as a faculty member.

The Grade 12 HUMSS population for the school year 2026–2027 consisted of five sections with a total enrollment of 200 students. Complete enumeration was employed to maximize coverage and support the stability of the regression estimates. Of the 200 enrollees, 153 students (76.5%) participated, composed of 35 males, 106 females, and 12 who identified as LGBTQ+; the remaining 47 were not included due to absence during data gathering, incomplete consent or assent forms, or unavailability.

Two instruments were used. The first was an adapted 20-item survey measuring perception across four dimensions of five items each, rated on a five-point Likert scale from 5 (Strongly Agree) to 1 (Strongly Disagree). Items were adapted from Davis (1989), Al-Abdullatif (2023), and Tlili et al. (2023). A pilot test with 30 non-respondent Grade 12 HUMSS students yielded Cronbach's alpha values of 0.750 for clarifying complex social concepts, 0.872 for prompting deeper follow-up questions, 0.794 for identifying potential bias in information, and 0.786 for perceived teacher guidance, all above the 0.70 threshold (Taber, 2018). The second was a researcher-constructed 50-item multiple-choice summative test aligned with the Table of Specifications drawn from the approved DIASS curriculum guide, covering the competencies of the first three weeks of the first quarter, with ten items for each of the levels of Remembering, Understanding, Applying, Analyzing, and Evaluating (Anderson & Krathwohl, 2001). Its content validity was established by expert validators, and pilot testing with 30 non-respondent students from a different school yielded a Kuder-Richardson 20 coefficient of 0.89, after which 17 items were revised.

Written permission was secured from the School Director and the Senior High School Principal of UCLM. Informed consent was obtained from parents or guardians and informed assent from the students, who were told that participation was voluntary, that they could withdraw without penalty, and that participation had no bearing on their grades. The summative test and the survey were administered after the relevant instructional units had been completed and were retrieved on the same day. Identification codes replaced names, and all procedures complied with the Data Privacy Act of 2012.

Frequency and percentage described the distribution of respondents across performance levels; weighted mean and standard deviation determined the levels of perception and performance; and multiple linear regression determined whether the four perception dimensions predicted the total test score, with R^2, the F-statistic, and standardized beta coefficients examined at the 0.05 level. Perception means were interpreted using equal-interval bands (Joshi et al., 2015): 4.21–5.00, Very Positive; 3.41–4.20, Positive; 2.61–3.40, Neutral; 1.81–2.60, Negative; and 1.00–1.80, Very Negative. Test scores were interpreted using the descriptive labels of DepEd Order No. 8, s. 2015 (DepEd, 2015), applied to equal intervals: 41–50, Outstanding; 31–40, Very Satisfactory; 21–30, Satisfactory; 11–20, Fairly Satisfactory; and 0–10, Poor.

RESULTS AND DISCUSSION

Clarifying complex social concepts

This dimension obtained the highest aggregate mean, 3.88 (SD = 0.81), interpreted as Positive. The highest-rated statement was "The AI chatbot helped me clarify terms and ideas that I found confusing" (M = 4.07, SD = 0.82), and the lowest was "The AI chatbot helped me connect classroom lessons to real-life social situations" (M = 3.56, SD = 0.91). The respondents valued the chatbot most in the moment of confusion, when it restated a term, offered an example, or simplified an abstract idea, and less for linking the lesson to the wider social world. Its standard deviation was also the smallest of the four dimensions, indicating that the benefit was experienced broadly. This is consistent with Okonkwo and Ade-Ibijola (2021) and Milloria et al. (2024), although a favorable rating reflects perceived usefulness rather than demonstrated mastery, and the benefit depends on users understanding the tool's limitations (Kasneci et al., 2023). Teachers can build on this strength by using the chatbot to define, exemplify, and restate concepts while using class discussion to extend those ideas to the real-life situations that the lowest-rated statement shows are still underdeveloped.

ISSN 3028-0877 | e-ISSN 3155-6043 | DOI 10.69651/2022PIJHSS

Prompting deeper follow-up questions

This dimension obtained the lowest aggregate mean, 3.61 (SD = 0.98), still interpreted as Positive. The highest-rated statement was "The AI chatbot helped me develop researchable questions during inquiry activities" (M = 3.74, SD = 0.92), and the lowest was "The AI chatbot helped me think more deeply about social issues discussed in class" (M = 3.50, SD = 1.00). The respondents credited the chatbot with helping them produce questions, a visible act, but were least convinced that it deepened their thinking, an internal one, which suggests that it supplied the wording of questions more readily than the reasoning behind them. Tlili et al. (2023) similarly noted concerns about whether AI-prompted questioning fosters genuine independent thinking, and Gerlich (2025) linked heavier reliance on AI tools to weaker critical thinking through cognitive offloading. Each chatbot-generated question should therefore be treated as the start of a teacher-directed discussion or written reflection rather than as a finished learning product.

Identifying potential bias in information

This dimension obtained an aggregate mean of 3.63 (SD = 0.97), interpreted as Positive, with statement means ranging narrowly from 3.56 to 3.65. The lowest-rated statement was "The AI chatbot encouraged me to verify information before accepting it as true" (M = 3.56, SD = 1.03). The narrow range suggests a general impression that the chatbot made the respondents somewhat more careful readers, while the low rating for verification, the most effortful behavior in this dimension, indicates that awareness was perceived more strongly than action. Because DIASS content involves contested social and political issues, students who compare viewpoints through the chatbot rely on its own framing of those viewpoints, which supports the view that the competencies needed to evaluate AI output must be deliberately developed (Kasneci et al., 2023; Shneiderman, 2022). Explicit instruction, such as comparing a chatbot's answer with a primary source, identifying whose viewpoint is missing, and writing a short verification note before accepting an answer, can move students from awareness to practice.

Perceived teacher guidance

This dimension obtained an aggregate mean of 3.85 (SD = 1.01), interpreted as Positive. The highest-rated statement was "My teacher helped us use the AI chatbot responsibly and ethically" (M = 3.99, SD = 0.96), while the lowest, both at 3.73, were "My teacher gave clear instructions on how to use the AI chatbot during class" (SD = 1.07) and "My teacher's guidance made the use of the AI chatbot more helpful for my learning" (SD = 1.05). Guidance was thus experienced mainly as ethical and evaluative direction and less as procedural coaching, and its standard deviation, the largest of the four, suggests that it may not have been delivered uniformly across sections. This aligns with Li et al. (2026) and with Shneiderman's (2022) principle that AI functions best under reliable human oversight, and the weakest statements identify where guidance can be strengthened first, in line with the auxiliary role assigned to AI by DepEd Order No. 003, series of 2026. Across the four dimensions, the grand mean was 3.74 (SD = 0.94), interpreted as Positive, consistent with the favorable attitudes toward chatbots reported by Al-Abdullatif (2023).

Level of academic performance in Applied Social Sciences

The respondents' mean score was 39.63 (SD = 10.59), which falls within the Very Satisfactory range. Of the 153 respondents, 95 (62.09%) scored Outstanding, 29 (18.95%) Very Satisfactory, 14 (9.15%) Satisfactory, and 15 (9.80%) Fairly Satisfactory, while none scored Poor. In all, 81.04% performed at the Very Satisfactory level or higher, although the large standard deviation shows that a minority of 29 respondents performed well below the group mean. The high scores may partly reflect the measurement, since the test covered three weeks of recently taught content and two-fifths of its items assessed remembering and understanding. The result is consistent with Milloria et al. (2024) and with the positive average effects reported by Deng et al. (2025), although the correlational design does not allow the scores to be attributed to the chatbot.

Perception as a predictor of academic performance

The regression model was significant, $F(4, 148) = 4.54$, $p = .002$, $R^2 = .109$, adjusted $R^2 = .085$, so the null hypothesis was rejected. Clarifying complex social concepts ($B = 4.44$, $\beta = .279$, $p = .020$) and perceived teacher guidance ($B = 2.65$, $\beta = .216$, $p = .030$) were positive predictors, prompting deeper follow-up questions was a negative predictor ($B = -4.43$, $\beta = -.349$, $p = .007$), and identifying potential bias in information was not significant ($B = 1.26$, $\beta = .101$, $p = .408$). The modest R^2 is acceptable in social science research when most predictors are significant (Ozili, 2023), and it was also bounded by the concentration of scores in the Outstanding range. Holding the other dimensions constant, each one-point increase in the clarification rating was associated with 4.44 more test points and each one-point increase in teacher guidance with 2.65 more points, while each one-point increase in follow-up prompting was associated with 4.43 fewer points.

The positive effect of clarification is coherent with a test that rewarded understanding of a bounded set of recent competencies, and it shows that the benefit reported in earlier studies (Deng et al., 2025; Milloria et al., 2024) attaches specifically to the clarification function. Bastani et al. (2025), however, found that unrestricted chatbot access improved practice performance but lowered performance once the tool was withdrawn, so the result supports guided clarification rather than unsupervised answer-seeking. The positive effect of teacher guidance suggests that clear direction and accuracy checks keep students' attention on lesson objectives, consistent with Li et al. (2026), Celik (2023), and the guarded tutoring design of Bastani et al. (2025), although the study measured perceived rather than actual teacher practices. The non-significant effect of bias identification may reflect a mismatch with a test that assessed content rather than evaluative skill, while the near-identical statement means gave the model little discriminating information; this skill should therefore be taught explicitly and assessed directly rather than assumed to follow from chatbot use.

The negative effect of follow-up prompting, the largest in the model, shows that a positive perception of a chatbot function does not guarantee a positive relationship with performance. Plausible explanations, none tested directly, include cognitive offloading, as students valued generating questions ($M = 3.74$) more than the reflection that should follow ($M = 3.50$); misalignment, as exploring beyond the lesson may divert attention from tested content; and reverse direction, as struggling students may rely more on the function. The result parallels Gerlich (2025) and Bastani et al. (2025) and contrasts with meta-analyses reporting cognitive benefits of overall ChatGPT use (Deng et al., 2025; Heung & Chiu, 2025), possibly because the present study isolated a single function with the others held constant. In practice, chatbot-generated questions should serve as prompts for human thinking, paired with teacher-

led discussion, written reflection, or a requirement that students answer the question in their own words before consulting the chatbot again.

Taken together, the results show that the relationship between perception and performance depends on the function perceived: concept clarification and teacher guidance, in which understanding and direction are central, were associated with higher scores; bias identification, which depends on the student's own follow-through, showed no relationship; and question prompting, the function most open to passive use, was associated with lower scores. This ordering supports the premise, drawn from Connectivism, Human-Centered AI, and Intelligent TPACK, that teacher mediation and not the chatbot alone shapes whether AI use becomes learning. These findings formed the basis of the proposed Instructional Enhancement Plan, which comprises four components: reinforcing AI-assisted concept clarification, formalizing teacher-guided ethical and responsible AI use, explicit critical-evaluation and bias-identification instruction, and structuring follow-up-question prompting. The findings should be read in light of the single-institution sample, self-reported perceptions, a single summative test covering three weeks of content, and the correlational design.

CONCLUSION

Grade 12 HUMSS students perceived teacher-guided AI chatbots positively and demonstrated a Very Satisfactory level of academic performance in Applied Social Sciences, but a favorable perception of the tool did not by itself guarantee stronger performance. The rejection of the null hypothesis confirms that perception matters, yet its contribution depends on the function involved: AI chatbots support learning when they clarify concepts under active teacher guidance, but may hinder it when students rely on them to generate questions without structure. The ability to evaluate AI-generated information does not develop through chatbot use alone and must be deliberately taught. It is the teacher, and not the technology, that determines whether AI chatbot use translates into meaningful learning.

AI chatbot use in DIASS should therefore be planned function by function, with the teacher actively mediating each activity, which supports the teacher-centered safeguards of DepEd Order No. 003, series of 2026. The study recommends the adoption of the proposed Instructional Enhancement Plan by the Applied Social Sciences Department of UCLM, implemented over one grading quarter and evaluated through the Classroom Observation Checklist and the Teacher Self-Assessment Survey. Future research may involve several schools and subjects, measure teachers' actual guidance practices, test the explanations for the negative effect of follow-up prompting through experimental or longitudinal designs, and include direct measures of critical-evaluation skills.

REFERENCES

Acena, J. G. (2025). Student's experiences using ChatGPT for academic purposes. Journal of Interdisciplinary Perspectives, 3(1), 146–153. https://doi.org/10.69569/jip.2024.0615

Al-Abdullatif, A. M. (2023). Modeling students' perceptions of chatbots in learning: Integrating technology acceptance with the value-based adoption model. Education Sciences, 13(11), Article 1151. https://doi.org/10.3390/educsci13111151

Alam, M. A. (2023). Connectivism learning theory and connectivist approach in teaching and learning: A review of literature. Bhartiyam International Journal of Education & Research, 12(2), 1–15. https://www.researchgate.net/publication/369734538

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. Proceedings of the National Academy of Sciences, 122(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Computers in Human Behavior, 138, Article 107468. https://doi.org/10.1016/j.chb.2022.107468

Creswell, J. W., & Creswell, J. D. (2023). Research design: Qualitative, quantitative, and mixed methods approaches (6th ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/research-design/book270550

Data Privacy Act of 2012, Republic Act No. 10173 (2012). https://privacy.gov.ph/data-privacy-act/

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. Computers & Education, 227, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

Department of Education. (2015). Policy guidelines on classroom assessment for the K to 12 Basic Education Program (DepEd Order No. 8, s. 2015). https://www.deped.gov.ph/2015/04/01/do-8-s-2015-policy-guidelines-on-classroom-assessment-for-the-k-to-12-basic-education-program/

Department of Education. (2026). Foundational guidelines on artificial intelligence (AI) in basic education (DepEd Order No. 003, s. 2026). https://www.deped.gov.ph/2026/02/20/february-20-2026-do-003-s-2026-foundational-guidelines-on-artificial-intelligence-ai-in-basic-education/

Estrellado, C. J. P., & Miranda, J. C. (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.13.05.2023.p13704

Estrera, E. B. (2020). A perception-based curricular review on the K to 12 HUMSS strand curriculum. IAFOR Journal of Education, 8(4), 25–44. https://doi.org/10.22492/ije.8.4.02

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), Article 6. https://doi.org/10.3390/soc15010006

González-Valencia, G., Massip Sabater, M., & Santisteban Fernández, A. (2022). Critical global citizenship education: A study on secondary school students. Frontiers in Education, 7, Article 867113. https://doi.org/10.3389/feduc.2022.867113

Heung, Y. M. E., & Chiu, T. K. F. (2025). How ChatGPT impacts student engagement from a systematic review and meta-analysis study. Computers and Education: Artificial Intelligence, 8, Article 100361. https://doi.org/10.1016/j.caeai.2025.100361

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. British Journal of Applied Science & Technology, 7(4), 396–403. https://doi.org/10.9734/BJAST/2015/14975

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, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Li, Y., Wu, Y., & Chiu, T. K. F. (2026). How teacher presence affects student engagement with a generative artificial intelligence chatbot in learning designed with first principles of instruction. Journal of Research on Technology in Education, 58(5), 1097–1113. https://doi.org/10.1080/15391523.2025.2493942

Liang, E. S., & Bai, S. (2025). Generative AI and the future of connectivist learning in higher education. Journal of Asian Public Policy, 18(2), 329–351. https://doi.org/10.1080/17516234.2024.2386085

Milloria, B. R. B., Marzon, A. M. D., & Derasin, L. M. C. (2024). Investigating AI-integrated instruction in improving academic performance of senior high school students in the Philippines. Journal of Harbin Engineering University, 45(6), 61–66. https://harbinengineeringjournal.com/index.php/journal/article/view/3090

National Council for the Social Studies. (2013). The College, Career, and Civic Life (C3) Framework for Social Studies State Standards: Guidance for enhancing the rigor of K–12 civics, economics, geography, and history. https://www.socialstudies.org/standards/c3

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence, 2, Article 100033. https://doi.org/10.1016/j.caeai.2021.100033

Ozili, P. K. (2023). The acceptable R-square in empirical modelling for social science research. In C. Saliya (Ed.), Social research methodology and publishing results: A guide to non-native English speakers (pp. 134–143). IGI Global. https://doi.org/10.4018/978-1-6684-6859-3.ch009

Pecson, R. R. (2026). Pedagogical approaches in social studies in the Philippine basic education. Interdisciplinary Research Review, 21(1), 28–41. https://doi.org/10.14456/irr.2026.4

Pimentel, J. V., Malaluan, K. D., Daluperi, F. N. M., De Vera, J. L. S., Sunga, J. E. T. A., Marin, A. J. B., Aguilar, J. M. D., Valdez, E. J. T., & Limos-Galay, J. A. (2024). Benefits of AI chatbots and engagement of senior high school students in Divine Word College of San Jose. International Journal of Research Studies in Educational Technology, 8(4), 77–83. https://doi.org/10.5861/ijrset.2024.8036

Shneiderman, B. (2022). Human-centered AI. Oxford University Press. https://doi.org/10.1093/oso/9780192845290.001.0001

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. Research in Science Education, 48(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10(1), Article 15. https://doi.org/10.1186/s40561-023-00237-x

Weidlich, J., Gašević, D., Drachsler, H., & Kirschner, P. A. (2025). ChatGPT in education: An effect in search of a cause. Journal of Computer Assisted Learning, 41(5), Article e70105. https://doi.org/10.1111/jcal.70105

BIONOTE

Aje B. Barretto is a Licensed Professional Teacher and a college instructor at the University of Cebu Lapu-Lapu and Mandaue, Mandaue City, Cebu, Philippines. He is completing his Master of Arts in Education, major in Teaching Social Studies, at Cebu Technological University-Main Campus, Cebu City. This article is based on his master's thesis.