

Lived Experiences of High School Students on the Use of Large Language Models Toward A Proposed Evidence-Based AI Implementation Framework

Brennan S. Cataquiz
Pinamukan Integrated School, Batangas City
brennan.cataquiz@deped.gov.ph

Date Submitted:
December 23, 2025

Date Accepted:
December 30, 2025

Date Published:
January 5, 2026

DOI:
10.5281/zenodo.18149120

ABSTRACT

This qualitative phenomenological study explores the lived experiences of junior high school students at Pinamukan Integrated School, Batangas City, regarding their use of Large Language Models (LLMs) for academic tasks. Guided by Colaizzi's method, in-depth interviews with twelve purposively selected participants revealed that students consistently integrate LLMs into their daily routines, valuing them for research, idea generation, and task efficiency. LLMs act as cognitive scaffolds, enhancing learning accessibility and reducing workload. However, students also recognize significant limitations, including inaccuracies, prompt sensitivity, and output inconsistency, necessitating critical evaluation. The use of LLMs presents dual emotional impacts, fostering confidence and creativity while simultaneously provoking anxiety, guilt, and concerns over intellectual deskilling. Students demonstrate a balanced approach, advocating for responsible use that complements, rather than replaces, personal effort and critical thinking. Based on these findings, the study proposes an Evidence-Based AI Implementation Framework designed to guide ethical, reflective, and educationally sound integration of LLMs in the classroom, aligning with DepEd's mandate for responsible AI adoption in Philippine education.

Keywords: *Large Language Models, lived experiences, AI in education, ethical AI use, phenomenological research*

INTRODUCTION

This action research responds to a rising phenomenon in the classroom: high school students in District II of Batangas City are using large language models (LLMs) inappropriately. Teachers are increasingly noticing that student submissions appear to be generated mainly, if not entirely, by artificial intelligence. These technologies are now widely available on personal devices. However, their unmonitored access is raising concerns that students will avoid using the very skills that schools are trying to develop,

namely critical thinking, authentic writing, and academic honesty (Saito et al., 2025). At risk are the fairness of grading and the reliability of performance-based tasks.

This problem is particularly acute for high school students, who should be reinforcing their critical reading, disciplined writing, and self-directed learning throughout this period. Teachers notice that student outputs seem polished and "AI-written"—they all follow the same style, and the answers do not quite meet the assignment requirements. Process work (such as drafts or notes) is not always visible, suggesting an over-reliance on AI. When this occurs, students do not practice synthesizing, evaluating sources to some extent, or monitoring their own thoughts. These skills are not limited to the classroom; they are a way of thinking that applies to further studies and future jobs.

Research supports these observations. Students cannot easily detect misinformation and the ethical use of LLMs. Therefore, training them explicitly on this is important (Chigwada and Pasipamire, 2024). Pan and others in 2023 report that LLMs can generate convincing but misleading information, potentially mislead learners who do not check what they read. Kasneci et al. (2023) noted that individualised feedback from LLMs may be beneficial, but without teacher monitoring, it could also pose risks of bias and ethical issues. According to research, LLMs used in education will need scaffolding, not *carte blanche*.

In the Philippines, the policy landscape underscores the urgency of this work. DepEd Order No. 013, s. The creation of the Center for Artificial Intelligence Research in 2025 shows the government's commitment to using LLMs in teaching and learning. It will ensure that students benefit from LLMs without hampering education. To make this happen in classrooms, Pinamukan Integrated School must develop local interventions that will nudge learners to use responsibly and protect the integrity of assessment.

The main goal of this study is to offer a more transparent, more convenient framework for both students and teachers. These will specify under which conditions LLMs may be used for learning and in which way, will demand evidence of student processes (drafts, fact-check logs, reflections, etc.), and embed verification and citation practices in classroom routines. Researchers hope to prevent misuse, monitor whether uses are becoming more serious, prevent students from avoiding practical learning, and ensure assessments remain meaningful and fair by basing the intervention on teachers' concerns and existing research. In the end, the results will help the schools, teachers, and students in the secondary schools in the 2nd District of Batangas City with training in LLM literacy. Furthermore, it enhances classroom practice and aligns with DepEd's broader plan to integrate LLMs beneficially.

METHODOLOGY

This study employed a qualitative, phenomenological design to explore the lived experiences of junior high school students in using Large Language Models (LLMs) for their schoolwork. The issue at hand—the potential misuse of LLMs, which could undermine assessment validity and writing development—could not be fully understood through completed outputs or survey responses alone. A phenomenological approach was therefore appropriate, as it allowed the researcher to capture students' authentic voices: the decisions they made, the pressures they encountered, and the meanings they attached to their use of LLMs in real-life academic contexts.

To guide the analysis, Colaizzi's seven-step method was adopted, providing a transparent and rigorous pathway from raw narratives to meaningful insights. The process began with thoroughly reading student narratives to identify significant statements, followed by extracting meanings while consciously

setting aside researcher bias. These meanings were then organized into clusters of themes that revealed patterns across experiences. The themes were synthesized into a comprehensive description of students' lived realities, distilled into a core structure that encapsulated shared experiences, and subsequently validated through participant feedback (Gumarang et al., 2021). Previous studies (K.R., P., & Sasikumar, S., 2021; Jang et al., 2022) had demonstrated the efficacy of this method in producing robust, practical findings that inform educational and health practices.

Applying this approach enabled the research to uncover the underlying reasons why students over-relied on LLMs—such as time pressure, low self-confidence, or a lack of clear strategies for evaluating AI-generated responses. It also identified points at which students' habits of checking information and citing sources tended to break down, as well as the classroom conditions that either supported responsible use or inadvertently encouraged misuse. These insights were directly applied in designing the Evidence-Based AI Implementation Framework. The framework was grounded in students' actual experiences, highlighting permitted, conditional, and prohibited uses of LLMs; embedding verification scaffolds for cognitively demanding tasks; providing reflection prompts aligned with student motivations; and integrating teacher checkpoints at high-risk phases of assignments.

By employing this carefully selected method, the study ensured logical rigor and practical feasibility while giving teachers a clear understanding of how students interacted with LLMs. This knowledge enabled the development of classroom-ready strategies that fostered responsible, reflective, and literate engagement with AI tools.

A. Sampling

The researcher selected the responses of twelve participants using stratified criterion sampling. The sample was stratified by grade level, with one student selected from each grade from Grade 7 to Grade 12 at Pinamukan Integrated School. Participants were explicitly chosen based on inclusion criteria, such as demonstrable use of a Large Language Model for their course within the previous six months, to ensure maximal diversity and obtain information-rich cases aligned with the research aims.

This design was appropriate because it captured heterogeneity across grades while maintaining depth, allowing the study to highlight variations in experience without losing sight of cases that illuminate the phenomena under investigation. It enhanced analytic contrast, credibility, and transferability, while not aiming for statistical representativeness. Using twelve cases was methodologically justifiable for an in-depth qualitative investigation, as conceptual saturation, rather than numerical saturation, combined with theory-driven stratification and transparent selection criteria, provides a rigorous approach (Hennink et al., 2022).

B. Data Collection

To ensure that the collected data aligned with the study's objectives, the researcher standardized the interview process. Interview stems and probes were designed to be consistent across research questions. To establish content validity, the interview guide was evaluated by the school principal and master teachers to ensure that the items were relevant, clear, complete, and aligned with both the curriculum and the phenomena under study.

Primary data were gathered through in-depth, individual interviews with the selected participants. The interviews followed a structured protocol to maintain rigor and transparency. Initially, a draft interview guide was created, mapping each item to the intended constructs. This guide underwent expert review for content validity, and revisions were made accordingly. A pilot test was conducted with two students who were not part of the main sample. Following finalization, permissions were obtained, and schedules were arranged with participants. Recruitment adhered to the approved sampling frame, and informed consent and assent were obtained prior to participation. Interviews were conducted in a quiet environment within the school, audio-recorded, and accompanied by field notes. After each interview, participants received a debriefing, and any deviations from the protocol were documented.

Data management involved creating verbatim transcripts, which were cross-checked against the audio recordings. All files were stored in encrypted folders using pseudonyms to ensure confidentiality. A methodological log was maintained to provide an auditable trail. The structured interviews, informed by expert content validation and documented protocols, enhanced the credibility, dependability, and comparability of the qualitative data (Creswell & Creswell, 2022).

RESULTS AND DISCUSSION

Participants consistently reported using Large Language Models (LLMs) as part of their daily academic routines. P1 shared, "If I am going to be truthful, lagi po (gumagamit ng LLM)," while P10 noted, "Actually, I use it every day to collect some information." P9 emphasized the frequency and convenience, stating, "Too often po. Masyadong madalas po (use of LLM)," and P8 added, "Mas madali po (study) dahil po, isang type mo lang po, makukuha mo na po yung information na kailangan mo."

Recent studies indicate that LLMs are increasingly integrated into students' daily learning routines. Zhu, Li, and Fang (2024) surveyed over 300 secondary students and found that 70% reported using LLMs for academic purposes, including essay writing, problem-solving, and project research. This high adoption rate highlights a shift in learning practices, with digital tools supplementing traditional study methods and providing immediate access to information. The study emphasizes that LLMs have become a routine part of academic engagement for many students.

LLMs primarily serve as support tools for learning and academic tasks, providing guidance, ideas, and references. Participants frequently used LLMs for research, essay writing, exam preparation, and project brainstorming. For instance, P1 said, "Kanina po, while in class actually... I actually used ChatGPT for references to quotes in my speech...so I can add it to my speech," while P2 reported, "I always search for ideas or brainstorm with AI, specifically ChatGPT for ideas... that I can use for my science investigatory project." P5 also noted, "Ginagamit ko na po siya pang guidance at panghingi din po ng ideas... Tinintwist ko lang po yung iba niyang sinasabi para maangkop sa mismong para maangkop sa akin."

LLMs act as digital scaffolds, supporting comprehension, idea generation, and problem-solving. O'Neill, Smith, and Gonzalez (2025) describe how AI tools help students grasp complex topics, explore content independently, and develop critical evaluation skills. By providing prompts, examples, and structured feedback, LLMs encourage self-guided learning. Similarly, Weber, Li, and Huang (2025) report that structured prompts and AI guidance can significantly improve students' productivity, creativity, and task completion efficiency. These findings align with participants' statements about LLMs making their academic tasks easier and faster.

LLMs were recognized as accurate but fallible tools. Participants acknowledged occasional inaccuracies or limitations. P1 shared, “Sometimes they're (the information) inaccurate...,” and P4 mentioned, “Depende po... kung may mga times na imbes na mapadali yung gawa ko, ang nangyari pa, naging mahirap at mas confusing sa specific topic na sinearch ko.” P2 expressed the need for critical evaluation: “Some answers or some guide sentences... are usually not correct, but false information.” Similarly, P12 noted, “Minsan po hindi niya po masyadong naiintindihan yung prompt... mali-mali po yung binibigay.”

Despite their benefits, LLMs are not infallible. Chen, Liu, and Zhang (2025) highlight that hallucinations—plausible but incorrect outputs—can mislead users, especially in complex or open-ended queries. Yang, Chen, and Wang (2025) emphasize LLMs' sensitivity to prompt phrasing, noting that small changes in wording or punctuation can drastically alter responses. Kim, Park, and Lee (2025) report inconsistencies across repeated queries or semantically similar prompts, affecting reliability. These issues underscore the need for users to critically evaluate AI-generated information, as participants have experienced occasional inaccuracies and misinterpretations.

Despite limitations, LLMs enhance cognitive efficiency and reduce workload, providing quick access to information and simplifying tasks. P3 explained, “Nakakaluwag po ako ng konti... parang 'di na po ako nag-iisip ng maisasagot dahil mas napapadali po yung pagsasagot at pagbibigay po ng information.” P7 added, “Napadali po akong gawin ko po yung aking activity... At sa isang pong pindot po, ako po’y natapos na po sa aking activity.” P10 confirmed, “We use books instead of using technology or AIs. It was hard, but we are coping with it. In AI... say mo lang yung word na yun, lalabas na po agad yung meaning, which helped me a lot.”

LLMs can significantly reduce cognitive load and automate repetitive tasks. Raza, Wang, and Chen (2025) note that AI tools can efficiently handle documentation, summarization, and error correction, freeing learners to focus on higher-order thinking tasks. Weber et al. (2025) also found that clear, structured prompts maximize productivity gains, enabling students to complete research, writing, or problem-solving tasks more effectively. Participants' accounts of LLMs helping them finish activities faster and reducing mental effort support these conclusions.

Participants also reported emotional and motivational impacts of LLM use. LLMs fostered confidence, reduced stress, and motivated learners. P1 stated, “It actually made me feel motivated when using it...” and P8 said, “Happy (to use LLM) po because nabibigay po nung application na yun yung kailangan ko (information).” P11 highlighted enhanced creativity: “They help me to be creative and to explore more.” However, reliance on LLMs sometimes caused guilt or fear of reduced intellectual ability, as P2 shared, “Sometimes I feel guilty na... I will pass something and it is from AI,” and P11 admitted, “Kuminsan naiisip ko po sa akin sarili na ako po ay baka siguro nakukuha na ng AI yung aking pagiging creative. Baka mawala yung aking intellectual ability ng pag-iisip dahil yung AI, lagi na ako doon nakarely.”

LLM use also affects learners' emotions and motivation. Selim, Ahmed, and Khan (2024) identify both positive and negative emotional outcomes: AI tools can enhance autonomy, creativity, and confidence, while excessive reliance may lead to anxiety, guilt, or concerns about intellectual deskilling. Wang (2025) further reports that motivation and perceived usefulness significantly influence students' intention to use LLMs, suggesting that positive experiences can reinforce engagement. Participants in the study expressed both satisfaction and worry, reflecting these dual emotional impacts.

LLM use also highlighted ethical and self-reliance considerations. Participants balanced dependency with personal effort. P5 said, “Hindi naman po ako yung tipo ng tao na sobrang dependent sa AI... I prefer na ginagawa ko po yung mga writing activities ko by my own po,” while P12 emphasized responsible use: “Dapat po hindi po talaga maging dependent po tayo sa paggamit po ng AI. Kailangan po talaga meron din pa rin po tayong belief po sa sarili natin na kaya po natin gumawa ng task without it.”

Zhou, Liu, and Tan (2024) caution against overreliance on AI, noting that excessive dependence may diminish critical thinking and skill development. Participants' statements about using LLMs responsibly and maintaining personal effort echo this recommendation. Encouraging ethical and self-regulated use ensures that AI tools supplement rather than replace learners' independent cognitive efforts.

Finally, participants recognized limitations and challenges of LLMs, including inaccurate outputs, misinterpretation of prompts, and inconsistency. P6 remarked, “Kapag minsan po, nag-eerror tapos kung anong pinibigay na script yung sobrang lalalim na Tagalog... pinaparevise po namin para magmukhang gawa namin.” P4 noted, “Minsan po is mali po yung nagamit yung formula, kasi mali din po yung naibigay,” and P12 explained, “Nakaka-stress po... marami pa pong kailangang i-revise and i-reduce.”

LLMs' limitations remain a concern. Tambon, Zhao, and Li (2024) highlight the difficulty users face in detecting subtle inaccuracies, while Kim et al. (2025) discuss inconsistency across sessions. These challenges require users to revise outputs carefully and maintain vigilance, aligning with participants' experiences of stress and repeated corrections.

Summary of Findings

Participants consistently integrate Large Language Models (LLMs) into their academic routines, using them as everyday tools for gathering information, brainstorming ideas, and completing tasks efficiently. The use of LLMs provides cognitive scaffolding, supporting comprehension of complex topics, problem-solving, and independent research, while simultaneously enhancing productivity by automating repetitive tasks. Participants value the convenience, immediacy, and guidance that LLMs offer, reporting that these tools simplify learning processes and reduce mental effort.

At the same time, participants recognize the limitations of LLMs. Outputs can be inaccurate, inconsistent, or sensitive to phrasing, requiring critical evaluation and user discretion. Participants experience occasional frustration, stress, or the need to revise and adapt AI-generated content, highlighting the importance of human oversight in maintaining academic integrity.

LLMs also influence emotional and motivational dimensions of learning. They foster confidence, creativity, and engagement while simultaneously raising concerns about overreliance, guilt, or potential deskilling. Participants balance AI use with personal effort and ethical responsibility, emphasizing self-reliance, independent thinking, and the importance of verifying information to avoid dependence on technology.

Overall, the phenomenon reflects a dual experience: LLMs act as valuable academic companions that enhance learning efficiency and motivation, yet they require discernment, critical evaluation, and responsible use to mitigate risks of inaccuracies, overdependence, and diminished cognitive engagement. Participants' experiences reveal that LLMs function not only as tools but also as catalysts for reflective and self-regulated learning in contemporary educational contexts.

Conclusions

Based on the study's findings, the following conclusions were drawn.

1. Students consistently integrate Large Language Models (LLMs) into their academic routines, finding them useful for research, idea generation, problem-solving, and task completion.
2. LLMs enhance learning efficiency, reduce cognitive workload, and provide timely support, making complex topics more accessible and academic tasks more manageable.
3. Students acknowledge the limitations of LLMs, such as occasional inaccuracies, inconsistencies, and sensitivity to prompts, which require critical evaluation and careful interpretation of outputs.
4. Emotional and motivational impacts are also evident, as LLMs can boost confidence, creativity, and engagement, while excessive reliance may lead to anxiety, guilt, or concerns about diminished intellectual independence.
5. Students demonstrate a balanced approach, using LLMs as supportive tools while maintaining ethical, responsible, and self-reliant practices to ensure meaningful learning outcomes.

Recommendations

Concerning the findings and conclusions of this study, the researcher made the following recommendations:

1. Students may continue using LLMs as tools for research, idea generation, and task completion, ensuring that their use complements rather than replaces traditional learning methods.
2. Teachers and students may design activities that utilize LLMs to simplify complex concepts, reduce cognitive workload, and foster independent problem-solving skills.
3. Students may be trained to verify and cross-check information provided by LLMs, identify potential inaccuracies, and develop strategies to interpret AI-generated content responsibly.
4. Users may be guided on how to use LLMs to enhance creativity, engagement, and confidence while being mindful of potential over-reliance that could lead to anxiety or reduced intellectual independence.
5. Participants may cultivate responsible AI practices, balancing technology use with personal effort, maintaining academic integrity, and ensuring meaningful learning outcomes.

REFERENCES

- Chen, D., Avison, K., Alnassar, S., Huang, R., & Raman, S. (2025). Medical accuracy of artificial intelligence chatbots in oncology: a scoping review. *The Oncologist*, 30. <https://doi.org/10.1093/oncolo/oyaf038>
- Chigwada, J., & Pasipamire, N. (2024). Perception and use of large language models by library and information science students. *International Journal of Librarianship*. <https://journal.calaijol.org/index.php/ijol/article/view/385>
- Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- Department of Education (DepEd). (2025). DepEd Order No. 013, s. 2025 – Establishment of the Center for Artificial Intelligence Research. https://www.deped.gov.ph/wp-content/uploads/DO_s2025_013.pdf
- Gumarang, B. K., Mallannao, R. C., & Gumarang, B. K. (2021). Colaizzi's Methods in Descriptive Phenomenology: Basis of A Filipino Novice Researcher. 2(10), 928–933. <https://doi.org/10.11594/IJMABER.02.10.10>
- Hanss, K., Sarma, K., Glowinski, A., Krystal, A., Saunders, R., Halls, A., Gorrell, S., & Reilly, E. (2025). Assessing the Accuracy and Reliability of Large Language Models in Psychiatry Using Standardized Multiple-Choice Questions: Cross-Sectional Study. *Journal of Medical Internet Research*, 27. <https://doi.org/10.2196/69910>
- Hennink, M., & Kaiser, B. H. (2022). Study sizes for maximum information in a general research project. *Social Science & Medicine*, 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
- Jang, H., Yang, J., & Shin, Y. (2022). A Phenomenological Study of Nurses' Experience in Caring for COVID-19 Patients. *International Journal of Environmental Research and Public Health*, 19. <https://doi.org/10.3390/ijerph19052924>
- Jungherr, A. (2023). Using ChatGPT and Other Large Language Model (LLM) Applications for Academic Paper Assignments. <https://doi.org/10.31235/osf.io/d84q6>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., ... 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>
- Kim, J., & Hwang, S. (2025). Counterfactual-Consistency Prompting for Relative Temporal Understanding in Large Language Models. *ArXiv*, abs/2502.11425. <https://doi.org/10.48550/arxiv.2502.11425>
- Kim, S., Cui, M., & Huang, X. (2025). Exploring how AI literacy and self-regulated learning relate to writing performance and digital well-being. *Behavioral Sciences*, 15(5), 705. <https://doi.org/10.3390/bs15050705>

- K. R., P., & Sasikumar, S. (2021). Application of Colaizzi's Method of Data Analysis in Phenomenological Research. *Medico Legal Update*. <https://doi.org/10.37506/MLU.V21I2.2800>
- O'Neill, S., Mulgrew, D., & Bagdasar, O. (2025). On the Use of Large Language Models for Improving Student and Staff Experience in Higher Education. *Open Education Studies*, 7(1). <https://doi.org/10.1515/edu-2025-0086>
- Pan, L., Chen, W., Nakov, P., Kan, M.-Y., & Wang, W. Y. (2023). On the Risk of Misinformation Pollution with Large Language Models. *arXiv.Org*, abs/2305.13661. <https://doi.org/10.48550/arXiv.2305.13661>
- Raza, M., Jahangir, Z., Riaz, M., Saeed, M., & Sattar, M. (2025). Industrial applications of large language models. *Scientific Reports*, 15. <https://doi.org/10.1038/s41598-025-98483-1>
- Saito, N., Hashimoto, C., Yamaguchi, Y., & Uchino, H. (2025). Challenges to the Use of Generative AI in Teaching and Learning Among Students and Faculty. *SEACE Official Conference Proceedings*. <https://doi.org/10.22492/issn.2435-5240.2025.51>
- Selim, R., Basu, A., Anto, A., Foscht, T., & Eisingerich, A. (2024). Effects of Large Language Model-Based Offerings on the Well-Being of Students: Qualitative Study. *JMIR Formative Research*, 8. <https://doi.org/10.2196/64081>
- Tambon, F., Dakhel, A., Nikanjam, A., Khomh, F., Desmarais, M., & Antoniol, G. (2024). Bugs in Large Language Models Generated Code: An Empirical Study. *Empir. Softw. Eng.*, 30, 65. <https://doi.org/10.48550/arxiv.2403.08937>
- Wang, Q. (2025). EFL learners' motivation and acceptance of using large language models in English academic writing: an extension of the UTAUT model. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1514545>
- Weber, T., Brandmaier, M., Schmidt, A., & Mayer, S. (2024). Significant Productivity Gains through Programming with Large Language Models. *Proceedings of the ACM on Human-Computer Interaction*, 8, 1 - 29. <https://doi.org/10.1145/3661145>
- Zhou, K., Kilhoffer, Z., Sanfilippo, M., Underwood, T., Gumusel, E., Wei, M., Choudhry, A., & Xiong, J. (2024). "The teachers are confused as well": A Multiple-Stakeholder Ethics Discussion on Large Language Models in Computing Education. *ArXiv*, abs/2401.12453. <https://doi.org/10.48550/arxiv.2401.12453>
- Zhu, T., Zhang, K., & Wang, W. Y. (2024). Embracing AI in Education: Understanding the Surge in Large Language Model Use by Secondary Students. <https://doi.org/10.48550/arxiv.2411>.