Embracing Large Language Models in Higher Education: A Case Study on Hong Kong's Students' Perceptions

Lok Him Lam, The Hong Kong University of Science and Technology, Hong Kong

The European Conference on Education 2025
Official Conference Proceedings

Abstract

A key feature of large language models (LLMs) is the ability to comprehend and generate human-like texts, and this feature undoubtedly endows LLMs with the capacity to revolutionize the education sector, particularly higher education. Although ChatGPT, the most well-known LLM, is still not available in Hong Kong to the public, universities in Hong Kong are no exception to this global race of embracing LLMs into teaching and learning. Although students' perceptions of LLMs are not typically an under-studied topic, such perceptions may constantly shift because of new developments in LLMs. In particular, LLMs have become more advanced, and there have been a growing number of LLMs. This study aims to understand the perceptions of Hong Kong university students on the adoption of LLMs in higher education. Thirty university students or recent graduates (within two years) constitute the sample of this study, who were asked to fill in a survey mostly consisting of qualitative questions that explore their perceptions of LLMs in regard to their usage, reliability/accuracy, negative impacts, and trade-offs. This study concludes that Hong Kong university students are generally aware of the known concerns of LLMs in education, and while most of them have used LLMs for educational purposes, the majority of those who have used them have not used them to directly write academic essays. At the same time, some students acknowledge that they are personally suffering from related adverse impacts from the use of LLMs, but most that acknowledge so would continue to use LLMs.

Keywords: large language models, higher education, students' perceptions, Hong Kong

iafor
The International Academic Forum
www.iafor.org

Introduction

The rise of large language models (LLMs) has brought both opportunities and concerns in higher education. A key and unique feature of LLMs, including ChatGPT or DeepSeek, is the ability to comprehend and generate human-like texts (Xing et al., 2025). It is undeniable that LLMs have the potential and capacity to revolutionise higher education. Some may even compare the birth of LLMs to the popularisation of computers in universities in the late 20th or early 21st century, when desktop computers were made available to students, and students no longer needed to hand-write their theses. LLMs are perceived with a mix of enthusiasm and criticism: some institutions recognise the transformative potential of these technologies and hence embrace them with support; others believe that the learning process of students would be compromised. This leads to the importance of understanding the perspectives of different stakeholders to seek the right balance between embracing such potential and addressing such concerns.

Universities in Hong Kong are early in the game in embracing LLMs into teaching and learning. In particular, the Hong Kong University of Science and Technology (HKUST) is the first to adopt LLMs into its pedagogy, providing students with access to ChatGPT (via cloud) and free token quotas. Other universities soon followed, and access was also expanded to other LLMs (notably, DeepSeek). Through qualitative surveying, this paper aims to gain in-depth insights into students' experiences, opinions, and concerns regarding the integration of LLMs into their education journey.

Methodology

As in most studies that attempt to gain in-depth perspectives from the relevant group of the population, this study is based on qualitative data collection.

2.1 Participants

This study sampled students pursuing a degree (bachelor's, master's, or doctorate) or recent graduates (within two years) who obtained a degree from a university in Hong Kong. The rationale for including recent graduates is to gather their perspectives at the later stages of their studies when LLMs emerged in education, as they might be able to draw comparisons from their earlier stages of studies. To potentially make this study comparable to other studies (both domestically or internationally), students who are enrolled in certificate programs, diplomas, higher diplomas, associate degrees, and postgraduate diplomas, as well as students from non-university self-financing post-secondary institutions and overseas institutions, who are not also enrolled in a university in Hong Kong as a degree-reading student are excluded from this study.

2.2 Sampling

A convenience sampling method was used. A total sample size of 30 was collected.

2.3 Instrument

A questionnaire is distributed to participants electronically (see Appendix A). It consists of multiple-choice questions to gather information from participants on their levels of study, home institutions, fields of study, and purposes for using LLMs; one Likert-type question for

them to rate the reliability of LLMS; open-ended questions to gather in-depth perspectives on participants' perceptions on the usage, accuracy, impact (on both learning process and academic integrity) on LLMs; and a text-box in the end for the participants to express any closing thoughts. All questions aside from the closing thought text box are compulsory.

2.4 Make-Up of Participants

Participants were required to answer background-related questions based on their latest educational status. For example, if a participant recently read for both a bachelor's and master's in Hong Kong, he/she should refer to his/her master's degree to answer such questions.

In total, 63.3% of respondents are reading for/ have received a bachelor's degree, 26.7% are reading for/ have received a master's degree, and 10% are reading for/ have received a doctorate (see Figure 1). 46.7% of participants chose HKU as their institution, 13.3% chose HKUST, 10% chose CityU, the remaining 30% chose other institutions, and no participants chose HKSYU, HSUHK, and HKSFU (see Figure 2). 66.7% of participants chose "humanities, social science, and law" as their field of study, 23.3% chose "science, engineering, and medicine," 10% chose "creative arts, music, and performing arts," and no participants chose "business and economics" (see Figure 3).

Figure 1

Levels of Study

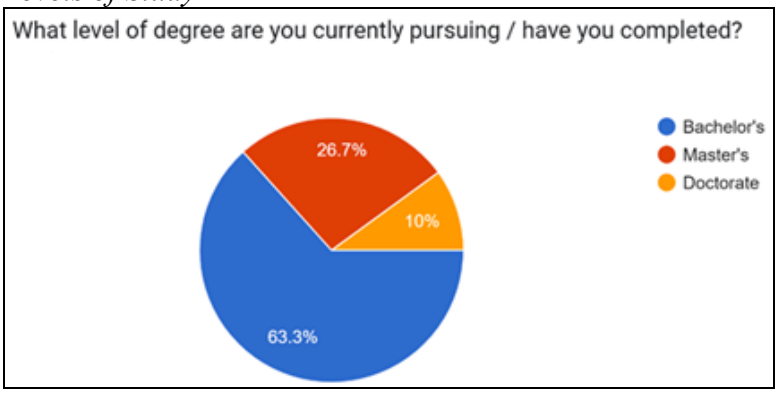


Figure 2

Instituional Make-Up

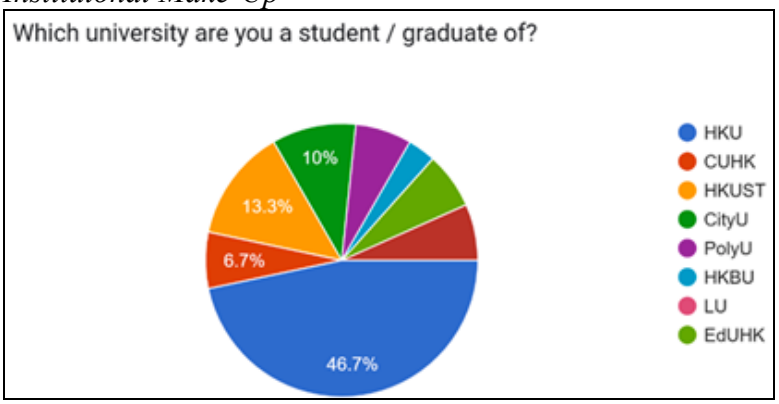
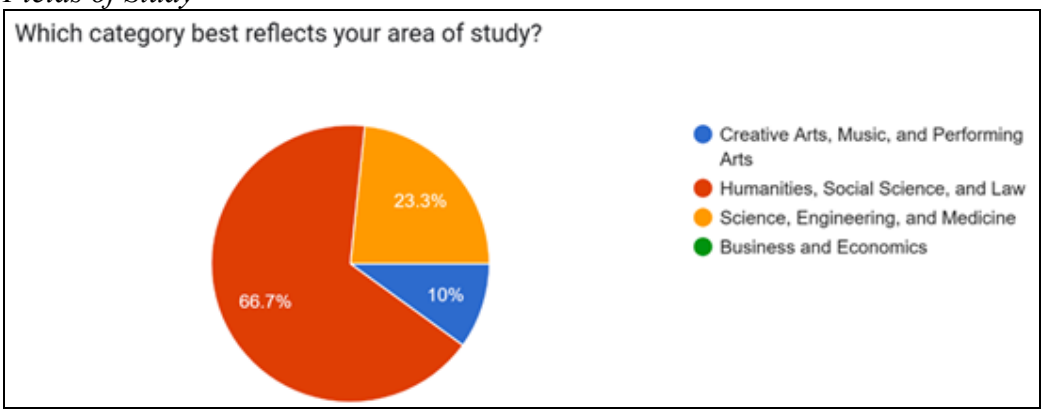


Figure 3

Fields of Study



Results and Discussion

3.1 Usage of LLMs

LLMs can be used to generate content, edit texts, generate ideas, search for information (although the database of some models is outdated), and seek solutions (e.g., generating codes for R or Python) [Bernabei et al., 2023]. Most participants (70%) claimed that they had used LLMs in their learning journey (see Figure 4). Figures from similar research are diverging, from ~30% to ~80% (see Paustian & Slinger, 2024; Renna, 2024; Zhu et al., 2024). When asked about their purpose for using LLMs, 53.3% of participants have used it for idea generation, 50% have used it for text editing, 36.7% have used it for information gathering/solution seeking, and only 23.3% have used it for content generation (see Figure 5).

Figure 4

Usage of LLMs

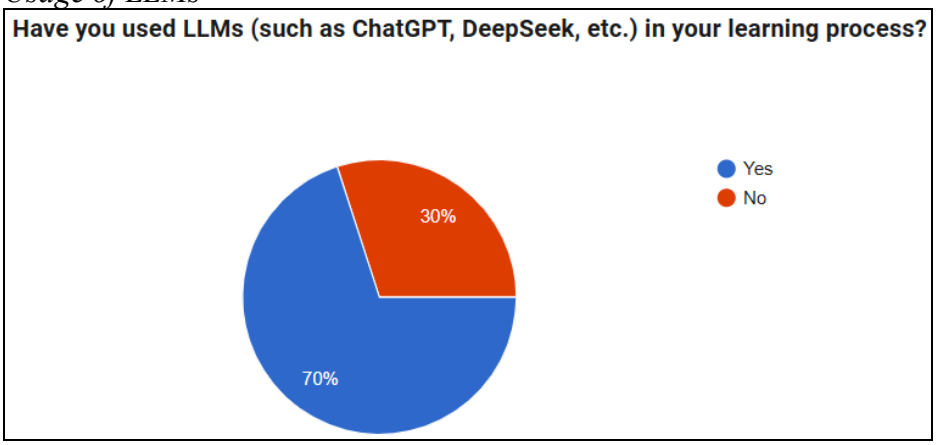
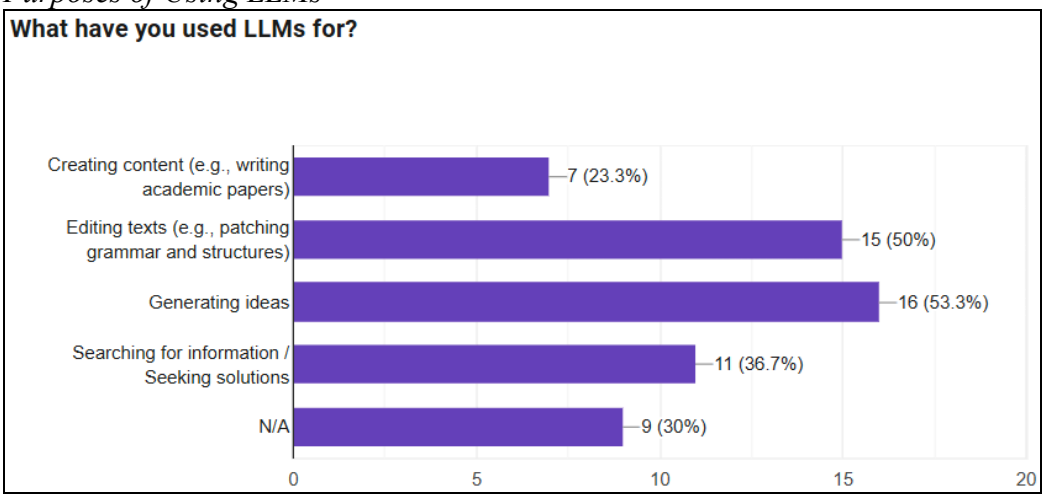


Figure 5

Purposes of Using LLMs



Many of the participants claimed that they are using LLMs for polishing their texts (including fixing their grammar), translating texts, and changing the tone of the writing:

"I used LLMs for grammar checks and transcribing historical texts in foreign language"

"Paraphrasing grammar because you no longer need effort in learning new vocab with LLM."

Others have used it to generate ideas for class, research ideas, and research methodologies:

"Help brainstorm some preliminary ideas for tutorial discussions and direction to approach essay questions."

"Using LLMs as a research assistant to generate research/methodological ideas, consolidate key points, and change writing style to scientific."

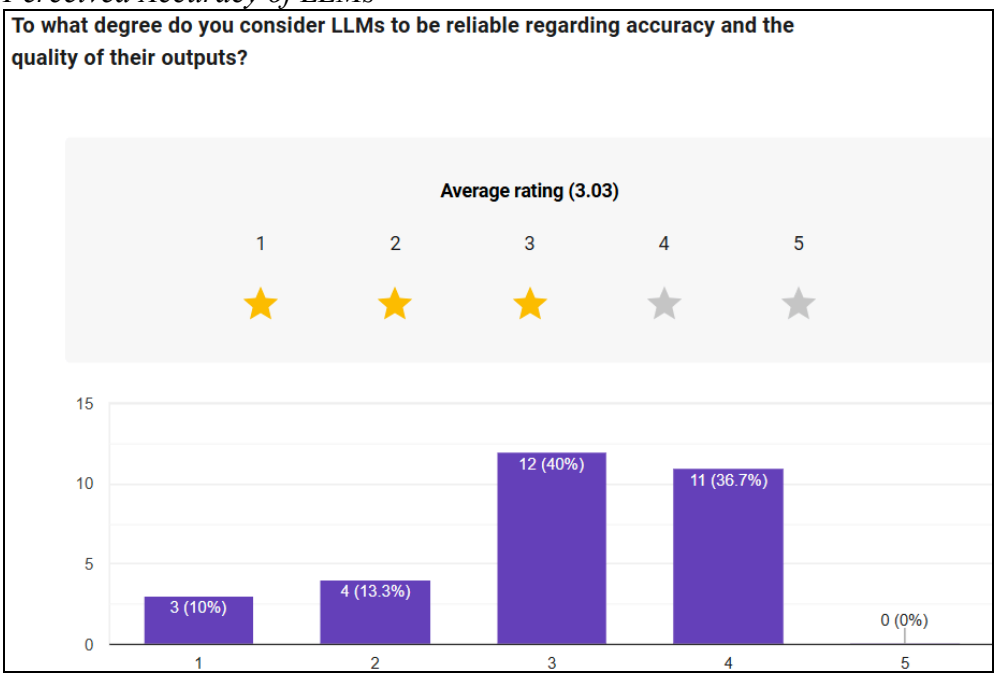
While most participants have used LLMs in their learning process, interpreting what the participants have mentioned, most treat LLMs as advanced translators or search engines. No participants explicitly mentioned that they had used LLMs to write academic papers. Most participants know that their assignments or other writing tasks are integral to their learning process. Among our participants, usage of LLMs is largely limited to "everyday" tasks in their learning or research journey that reduce their time costs, rather than using LLMs for a purpose that replaces them as agents of learning. Depending on the purpose of usage, LLMs can potentially be used as a tool to facilitate students' learning process and enhance productivity. However, as some of our participants showed, LLMs may discourage them from thinking originally or critically, but educators may still want to be aware of such a phenomenon (see 3.3 & 3.4).

3.2 Reliability and Accuracy of Output

Existing studies have established that LLMs can contain unverified or even, at times, factually incorrect content (Bélisle-Pipon, 2024; Qin, 2024; Schiller, 2024; Wang et al., 2024). Participants were asked to rate the reliability of LLMs, and an average score of 3.03 out of 5 was given. Most participants rated "3" or "4," with no participants rating "5."

Figure 6

Perceived Accuracy of LLMs



Most participants seem to be aware (and have experienced) the potential inaccuracies of LLMs:

"Sometimes, LLMs hallucinate information, and without own knowledge, it is hard to find out."

Some participants pointed out that LLMs are better at general tasks rather than specialized tasks in their field:

> When I am looking for something very general (common senses), I think the models are quite reliable. However, if I ask them to do some analysis, I will fact check by myself, since most responses are likely make[made] up, or not that accurate.
>
> It is convenient, but there are still many errors (e.g.: Grammatical mistakes and lack of content coherence). From my experience as a language learner at university, other languages (exclude English) still have a long way to improve its grammatical accuracy and content coherence (e.g. wrong sentence structure, wrong verb conjugation or wrong use of verbs etc.).

Some participants shared their personal experiences dealing with inaccurate output from LLMs with reference to their field of study. Such experience persists across all fields of study:

"For subjects on literary reviews that depends heavily on personal preference, the times we tried to seek inspiration from these models resulted in very vague and useless responses." – presumably a literature student

"Lack of details in case law." – presumably a law student

Contrary to some popular beliefs that the accuracy of LLMs might "perform" better in certain fields or can even "replace" students of humanities or social sciences, our participants largely disagree. Our participants find them, at times, unreliable in such fields. Other participants also recognize that content generated from LLMs can be inaccurate, unreliable, and even, at times, "*terrible.*" Contrary to some concerns of educators, students are largely aware of the potential problems of inaccuracies on LLMs.

3.3 Self-Perceived or Peer-Perceived Negative Impacts on Learning Process

Many participants believe that the potential inaccuracy from LLMs (also see 3.2) constitutes a negative impact on their peer's learning process:

"Some people don[']t check the LLM results."

"Some students may overuse LLMs and their works may not be original."

Some others expressed concerns about how LLMs may jeopardize the learning process:

> Most of the learning task in the curriculum is designed in a way that facilitate learning by doing. If one were to substitute conscious effort at thinking and writing with output generated from AI, the effectiveness of learning through doing this task is practically moot.

> At times, it makes me stop thinking originally on my own, even though I believe I have a solid foundation already. It's almost like taking a shortcut to get to faster and more outcomes, which from a utilitarian perspective is not a bad thing, but I'm genuinely afraid my brain might shrink because of this.

Many participants admit that using LLMs may discourage them from thinking critically and originally and admit that the effectiveness of learning might be compromised. These responses align with what some literature argues (Duenas & Ruiz, 2024; Jošt et al., 2024). Some others also believe that using LLMs has decreased their capability to write coherently, with one participant suggesting that his/her "*English become[s] worse.*" As discussed in 3.2, while students are aware of the potential inaccuracies of LLMs and potential adverse effects on their psychology of learning, many students may still opt to use LLMs in order to save time and effort. Educators must be aware of such a trend and potentially design learning tasks that encourage critical and original thinking.

3.4 Perceived Impact on Academic Integrity

The way LLMs obtain data and generate responses according without attributing to the original source may bring concerns regarding academic integrity (Pudasaini et al., 2024). Some popular sources even criticize LLMs as "plagiarism machines" (Keegin, 2023; Novak, 2023). Based on the responses that the participants have given, there are two dimensions to such concerns: 1) students committing plagiarism from the use of LLMs, and 2) teachers wrongly accusing students of using LLMs when they are forbidden.

One participant listed four major reasons how LLMs may compromise academic integrity: 1) accuracy, in terms of both the citation and the content being cited; 2) transparency, the process of response generation is not known to the user, such that the user cannot identify

potential bias or problematic information in a dataset, 3) credit, whether attributions are probably given, and 4) infringement of institutional policies, if the coursework requirements forbid the use of LLMs.

However, an under-investigated perspective by similar studies would be the other way around: how instructors may have misunderstood their students for using LLMs without permission when they have not:

“The teachers would often misunderstand students using LLMs when the so-called ‘ai detector’ detected 100% ai but they didn't use it.”

Despite the dependency of AI-generated detectors on some faculty members, it is well-known that they are largely unreliable and inconsistent (Elkhatat et al., 2023; Weber-Wulff et al., 2023). However, there has recently been a growing discussion among student communities that many are wrongly accused of AI-related plagiarism. The tendency of some institutions and faculty members to rely on AI-detection tools can penalize hardworking students and discourage grammatically accurate and well-written work. This form of misuse of AI-detection tools is, at best, an increased unnecessary burden on students and, at worst, an infringement of academic integrity on its own. More research on the perceptions of faculty members on LLMs should be conducted, especially on the ways of dealing with such risks.

Limitations

As convenience sampling was adopted, the sampling may not be very representative of the actual student population. This is reflected in the unbalanced institutional make-up of participants (see 2.4). The main scope of this study is to gain in-depth perspectives from participants to bridge future research, and present prelimnary findings, rather than being very conclusive at this stage. Moreover, as this is still a growing topic, future students’ perceptions of LLMs may shift drastically and significantly deviate from the findings of this paper. Moreover, as this is a growing field of research, preprints (e.g., arXiv articles) instead of other scholarly sources are often cited to capture the newest findings.

Conclusion

This study provides insights into the perceptions of Hong Kong students regarding integrating large language models (LLMs) into their educational journey. While most participants have used LLMs, they are also knowledgeable regarding the concerns regarding their learning process and academic integrity posed by LLMs. In particular, a number of participants pointed out that their language and critical thinking capabilities have deteriorated. Some may believe that the benefits of LLMs outweigh the potential adverse effects, seeing them as an effective tool for learning; others argue that the students' learning process is severely compromised. This study hopefully bridges further research on the perceptions of LLMs from a broader range of stakeholders, particularly educators. Looking forward, case studies on Hong Kong students on a larger scale can be conducted to reach a more conclusive result.

Acknowledgement

This paper is reproduced from a class paper (originally titled: *LLMs in Education: A case study on Hong Kong's students' Perceptions*) written individually by the author in partial fulfilment of the course requirements of *PPOL5380 Technology Disruptions and Public Policy* (Instructor: Masaru Yarime) at the Division of Public Policy, the Hong Kong University of Science and Technology. Feedback from the instructor and the teaching assistant, Meline Yesayan, is appreciated and partially adopted to improve this paper.

References

Bélisle-Pipon, J.-C. (2024). Why we need to be careful with LLMs in medicine. *Frontiers in Medicine*, 11, 1495582.

Bernabei, M., Colabianchi, S., Falegnami, A., & Costantino, F. (2023). Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence*, 5, 100172.

Duenas, T., & Ruiz, D. (2024). The Risks of Human Overreliance on Large Language Models for Critical Thinking [Preprint]. https://doi.org/10.13140/RG.2.2.26002.06082

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, *19*(1), 17–16.

Jošt, G., Taneski, V., & Karakatič, S. (2024). The Impact of Large Language Models on Programming Education and Student Learning Outcomes. *Applied Sciences*, *14*(10), 4115.

Keegin, J. M. (2023, May 23). ChatGPT Is a Plagiarism Machine. *The Chronicle of Higher Education*. https://www.chronicle.com/article/chatgpt-is-a-plagiarism-machine

Novak, M. (2023, May 30). Google's New AI-Powered Search Is A Beautiful Plagiarism Machine. *Forbes*. https://www.forbes.com/sites/mattnovak/2023/05/30/googles-new-ai-powered-search-is-a-beautiful-plagiarism-machine/

Paustian, T., & Slinger, B. (2024). Students are using large language models and AI detectors can often detect their use. *Frontiers in Education*, 9.

Pudasaini, S., Miralles-Pechuán, L., Lillis, D., & Llorens Salvador, M. (2024). Survey on AI-Generated Plagiarism Detection: The Impact of Large Language Models on Academic Integrity. *Journal of Academic Ethics*.

Renna, I. (2024). Is ChatGPT Massively Used by Students Nowadays? A Survey on the Use of Large Language Models such as ChatGPT in Educational Settings [Preprint]. *arXiv*. https://arxiv.org/html/2412.17486v1

Schiller, C. A. (2024). The Human Factor in Detecting Errors of Large Language Models: A Systematic Literature Review and Future Research Directions [Preprint]. arXiv. https://arxiv.org/html/2403.09743v1

Wang, Y., Wang, M., Manzoor, M. A., & Liu, F. (2024). Factuality of Large Language Models: A Survey [Preprint]. arXiv. https://arxiv.org/html/2402.02420v3

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, *19*(1), 26–39.

Xing, W., Nixon, N., Crossley, S., Denny, P., Lan, A., Stamper, J., & Yu, Z. (2025). The Use of Large Language Models in Education. *International Journal of Artificial Intelligence in Education, 35,* 439–443.

Zhu, T., Zhang, K., & Wang, W. Y. (2024). Embracing AI in Education: Understanding the Surge in Large Language Model Use by Secondary Students [Preprint]. *arXiv*. https://arxiv.org/html/2411.18708v1

Contact email: lhlamah@connect.ust.hk

Appendix

Survey Questions

Are you a student reading for a degree (bachelor's, master's, or doctorate) or a recent graduate (within two years) who read a degree at a university?
(Please refer to the "eligibility" session if you are not sure.)

Which university are you a student / graduate of?
(Please indicate your most recent affiliation from the list below if you have attended more than one of these institutions.)

What level of degree are you currently pursuing / have you completed?
(If you have completed multiple degrees or studied abroad, please indicate only your latest level of studies in Hong Kong.)

Which category best reflects your area of study?

Have you used LLMs (such as ChatGPT, DeepSeek, etc.) in your learning process?

What have you used LLMs for?

Please elaborate and provide the relevant details based on the above response.
(Fill in N/A if you have not used LLMs in your learning process.)

To what degree do you consider LLMs to be reliable regarding accuracy and the quality of their outputs?
(With 1 being completely unreliable and 5 being completely reliable.)

Please elaborate and provide the relevant details based on the above response.

What potential negative impacts do you think the use of LLMs could have on your learning experience or that of your peers?

What are your opinions on academic integrity concerning LLM usage? Do you feel that they create potential risks for plagiarism, like neglecting to credit the right sources?

How do you view the role of large language models (LLMs) in education? In your opinion, do the potential benefits of LLMs outweigh the challenges and negative impacts associated with their use?

Any closing thoughts?