

Promoting Generative AI in Learning, Teaching, and Assessment: A Case Study in Human Anatomy Education

Chileka CHIYANIKA¹

¹ Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Abstract

This case study examined the impact of integrating Generative Artificial Intelligence (GenAI), specifically ChatGPT, into human anatomy education, comparing the academic performance of two cohorts from different academic years. Cohort 1 () studied using traditional methods, while all students in Cohort 2 () engaged with GenAI tools during collaborative learning sessions. Survey responses from Cohort 2 () indicated that 69.8% rated their GenAI experience positively, and 60.6% reported an enhanced understanding of anatomical concepts. Academic performance analysis revealed that Cohort 2, which utilised GenAI universally, achieved a higher mean score (78.73% vs 77.31%) and a greater proportion of top grades (27.6% vs 16.5%) compared to Cohort 1. While students valued GenAI for its ease of use and interactivity, they expressed concerns regarding information accuracy and highlighted the need for further improvements. The study underscores GenAI's potential as a valuable adjunct to traditional teaching, supporting digital literacy and pedagogical innovation in anatomy education.

Keywords: Generative Artificial Intelligence, Human Anatomy Education, ChatGPT, Collaborative learning, Digital literacy

Introduction

Generative Artificial Intelligence (GenAI) tools, such as ChatGPT, are reshaping higher education by granting students unprecedented access to information and enabling more interactive, personalised, and self-directed learning experiences. In response to this technological evolution, universities are increasingly embracing GenAI, prompting academic departments to rethink and update their approaches to Learning, Teaching, and Assessment (LTA).

Human anatomy is a cornerstone of medical and health sciences education, traditionally taught through cadaveric dissection, textbooks, and physical models. While effective, these methods are often limited by issues such as ac-

cessibility, cost, and the static nature of learning materials (Fahrni & Sabatasso, 2025). GenAI offers new opportunities to enhance anatomy education through interactive, adaptive, and scalable digital platforms, potentially overcoming many limitations of conventional approaches.

In line with our university's digital innovation strategy, ChatGPT was piloted to supplement traditional teaching and promote collaborative, critical engagement with technology. Given the limited evidence on GenAI's effectiveness in anatomy education, this study evaluated ChatGPT integration, assessed student experiences and perceptions, and compared academic performance between a traditional cohort and one with GenAI integration.

Methodology

Study Design and Setting

This study was conducted during the autumn semester of the 2024/25 academic year with ethics approval (HSEARS20250409003), targeting all 121 first-year undergraduate radiography students at our university.

Participants

Two cohorts were compared: Cohort 1 (2023/24) utilised traditional methods, while Cohort 2 (2024/25) participated in collaborative learning sessions enhanced with Generative AI (GenAI), specifically ChatGPT. All students were included in the academic performance analysis. Survey participation was voluntary and anonymous; responses were received from 43 students within Cohort 2.

Traditional Learning Methods

Cohort 1 followed traditional learning approaches, including structured lectures and seven collaborative sessions over 13 weeks, focused on revising key content. Activities such as real-life case scenarios, mini-presentations, and manikin practice reinforced theory and application. The remaining six weeks were reserved for optional individual consultations with the tutor, either in person or online.

Integrating ChatGPT into Collaborative Anatomy Learning

In Cohort 2, selected collaborative learning sessions (four out of seven in total) were supplemented with clinically relevant case scenarios designed by teaching staff to reflect real-world challenges. Students worked in small groups to analyse these scenarios, utilising ChatGPT to research and generate responses. Following AI-assisted exploration, students engaged in group discussions to critically evaluate the accuracy and relevance of AI-generated content, cross-referencing it with textbooks and trusted sources. Tutors facilitated these

sessions, providing expert feedback and clarifying misconceptions to reinforce theoretical understanding and promote digital literacy.

Survey Instrument

At the end of the semester, Cohort 2 students completed a structured questionnaire assessing their GenAI experience in anatomy sessions. The ten-item instrument used a 1–10 Likert scale to evaluate overall experience, ease of use, perceived enhancement of understanding, helpfulness, challenges, frequency of use, accuracy, comprehensiveness, likelihood of recommending GenAI, comparison to traditional methods, and suggestions for improvement.

Student Performance Assessment

Academic outcomes were compared between cohorts using final grades for the subject. Grade distributions and mean scores were extracted from institutional records, reflecting either traditional teaching (Cohort 1) or GenAI-supported learning (Cohort 2).

Data Management and Confidentiality

All data were managed in accordance with institutional ethics and data protection guidelines. Survey responses were anonymous, securely stored, and stripped of personal identifiers.

Data Analysis

Descriptive statistics summarised survey and academic data. Grade distributions and mean scores were compared using independent samples *t*-tests or Mann-Whitney tests, as appropriate. All tests were two-sided, with considered significant (SPSS version 28).

Results

A total of 43 students completed the survey. GenAI (ChatGPT) received positive feedback, with 69.8% rating it 7 or higher (the most frequent rating was 8 at 30.2%). Ease of navigation was also rated highly, with 69% giving a score of 7 or above. Regarding learning impact, 60.6% felt GenAI enhanced their understanding of human anatomy, and 60.5% found its features helpful. Few students reported significant challenges; only 30.2% noted minimal difficulties. Usage was moderate (67.5% rated it 5 or higher), and 48.8% rated its accuracy and comprehensiveness at 7 or higher. The likelihood of recommending GenAI was moderate (58.2% rated it 7 or above), and 55.8% considered it better than traditional methods. However, 78.7% expressed a desire for further improvements.

Baseline academic abilities were similar between cohorts, as shown by comparable university admission scores (Cohort 1: ; Cohort 2: ;). At the end of the semester, Cohort 2 (GenAI users) had a slightly higher mean anatomy score () than Cohort 1 (), though this difference was not statistically significant (). Notably, Cohort 2 achieved significantly more top grades (A+ and A: 27.6% vs 16.5%;). With similar baseline abilities, these results suggest that GenAI may have helped improve top-level performance in human anatomy. See Tables 1 and 2 for details.

Table 1. Survey results regarding student perceptions of Generative AI (ChatGPT) use in human anatomy education.

Survey Question	Likert scale responses																			
	1		2		3		4		5		6		7		8		9		10	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Overall Experience	0	0	0	0	1	2.3	0	0	7	16.3	5	11.6	11	25.6	13	30.2	4	9.3	2	4.7
Ease of Use	0	0	0	0	0	0	1	2.4	9	21.4	3	7.1	8	19.0	9	21.4	5	11.9	7	16.7
Enhanced Understanding	0	0	1	2.3	0	0	2	4.7	8	18.6	6	14.0	6	14.0	11	25.6	6	14.0	3	7.0
Helpful Features	0	0	0	0	2	4.7	3	7.0	6	14.0	6	14.0	8	18.6	10	23.3	5	11.6	3	7.0
Challenges Encountered	0	0	4	9.3	8	18.6	1	2.3	14	32.6	4	9.3	5	11.6	4	9.3	2	4.7	1	2.3
Frequency of Use	0	0	2	4.7	7	16.3	5	11.6	6	14.0	9	20.9	7	16.3	4	9.3	1	2.3	2	4.7
Accuracy/Comprehensiveness	0	0	0	0	4	9.3	2	4.7	8	18.6	8	18.6	10	23.3	9	20.9	1	2.3	1	2.3
Recommend to Others	1	2.3	0	0	3	7.0	2	4.7	7	16.3	5	11.6	7	16.3	10	23.3	3	7.0	5	11.6
Use of GenAI compared to Traditional Methods	0	0	0	0	2	4.7	2	4.7	9	20.9	6	14.0	8	18.6	8	18.6	5	11.6	3	7.0
Want Improvements	0	0	1	2.4	2	4.8	6	14.3	7	16.7	7	16.7	9	21.4	7	16.7	2	4.8	1	2.4

Note. for each item. Scores are based on a 1–10 Likert scale, where 1 indicates a strongly negative response and 10 indicates a strongly positive response.

Table 2. Comparison of human anatomy assessment results between student cohorts with and without Generative AI (GenAI) tool use in collaborative learning sessions.

Grade	Cohort 1 (without GenAI tools) n=121		Cohort 2 (with GenAI tools) n=123		P-Value
	<i>n</i>	%	<i>n</i>	%	
A+	7	5.8	18	14.6	0.023
A	13	10.7	16	13	0.586
A-	32	26.4	33	26.8	0.946
B+	27	22.3	18	14.6	0.123
B	24	19.8	15	12.2	0.104
B-	15	12.4	15	12.2	0.962
C+	2	1.7	6	4.9	0.158
C	1	0.8	1	0.8	0.991
C-	0	0	1	0.8	0.321
D+	0	0	0	0	-
D	0	0	0	0	-
F	0	0	0	0	-

Note. The Mann-Whitney test was used for statistical comparison. Grade classifications are defined as follows: A+ (90–100), A (85–89), A– (80–84), B+ (75–79), B (70–74), B– (65–69), C+ (60–64), C (55–59), C– (50–54), D+ (45–49), D (40–44), and F (0–39).

Discussion

This study highlights a broadly positive response to Generative AI (GenAI) tools, particularly ChatGPT, among undergraduate radiography students studying human anatomy. Students reported that GenAI enhanced their understanding of anatomical concepts, was easy to navigate, and was a tool they would recommend to their peers. These findings reflect the growing value placed on GenAI as a supplementary educational resource, aligning with trends in higher education where digital tools are increasingly integrated into learning, teaching, and assessment (LTA) (Khlaif et al., 2024).

GenAI usage was moderate, with 67.5% of students reporting usage levels between 5 and 10 on the Likert scale. This suggests that these tools are becoming a regular part of students' study routines. This shift indicates changing preferences, with students turning to digital support alongside traditional methods. The convenience and immediacy of GenAI—such as on-demand access to information and clarification of concepts outside scheduled teaching sessions—are clear advantages. This aligns with the broader move towards technology-enhanced learning across universities, where intuitive and accessible platforms drive adoption (Zenios, 2020).

Ease of use is a significant factor in GenAI uptake; 69% of students rated navigation highly. When educational technology is user-friendly, students are more likely to incorporate it into their daily study habits (Holmes et al., 2019). The perceived improvement in understanding, reported by 60.6% of students, further demonstrates GenAI's effectiveness in supporting learning. Human anatomy results reinforce this: the GenAI cohort performed better overall and achieved a higher proportion of top grades compared to the non-GenAI cohort. AI-powered platforms offer personalised explanations and instant feedback, which are particularly valuable in complex subjects like anatomy (Holmes et al., 2019). These features facilitate deeper learning and encourage active engagement (Lam, 2025).

Despite these advances, there remain areas for improvement. Many students expressed a desire for improved features and greater accuracy, reflecting ongoing concerns about current AI's ability to contextualise complex anatomical variations and ensure valid content (Davis et al., 2025). Only 48.8% rated GenAI's accuracy and comprehensiveness highly, highlighting the need for caution. Frequent reliance on these tools raises questions about their educational role and underscores the importance of educators guiding students in critically evaluating AI-generated content. GenAI should complement, not replace, expert instruction and validated resources. Issues such as group dynamics, fairness in

assessment, and academic integrity in AI-assisted assignments require attention to ensure effective learning and ethical standards.

This study has several limitations. The sample size was small, with a low survey response rate, introducing potential non-response and selection bias. The non-randomised, cross-sectional design compared two cohorts from different academic years without controlling for confounding factors such as demographic differences and changes in teaching staff. Furthermore, the lack of parallel survey data from Cohort 1 makes it difficult to determine whether the experiences reported by the GenAI-supported cohort reflect actual improvements or simply perceived differences, limiting the strength of comparative claims. Reliance on self-reported perceptions introduces response bias and may not reflect objective learning outcomes or long-term retention. Conducted at a single institution and focused on undergraduate radiography students, the findings may not generalise to other disciplines or settings. While Cohort 2 showed a modest improvement in mean scores and more top grades, these outcomes should be interpreted cautiously. Future research should use larger, more diverse samples, objective measures, and longitudinal or randomised controlled designs to better assess GenAI's educational impact.

Implications

Integrating GenAI tools like ChatGPT into anatomy education can enhance student engagement, understanding, and academic performance. However, concerns regarding information accuracy and the need for critical evaluation indicate that GenAI should supplement, not replace, expert instruction. Ethical considerations, including fairness and academic integrity, must also be addressed to ensure responsible use.

These findings suggest that GenAI could benefit diverse disciplines and contexts by promoting digital literacy and critical thinking. It can guide ethical integration and standards while helping students adapt to evolving technology, thereby supporting global educational priorities in a rapidly changing landscape.

Acknowledgments

This study was supported by a grant from The Hong Kong Polytechnic University on promoting the use of GenAI in LTA. The author gives special thanks to all the students for their participation.

Ethical Declaration Statement

The project has been approved by the Human Subjects Ethics Application Review Committee of The Hong Kong Polytechnic University (Reference No: HSEARS20250409003).

Generative AI Policy

The author used GPT-4.1 during the preparation of this work to correct grammatical errors and improve readability. After utilising the tool, the author thoroughly reviewed and edited the content as necessary and assumes full responsibility for the publication's content.

References

- Chan, K. S., & Zary, N. (2019). Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR Medical Education*, 5(1), e13930. <https://doi.org/10.2196/13930>
- Davis, P., Napole, A., Reddy, C., Palepu, S. K., Wu, A. Q., Mannam, S. S., Gajjar, A. A., Jabarkheel, R., Ajmera, S., & Ahmad, H. (2025). Systematic evaluation of AI-based text-to-image models for generating medical illustrations in neurosurgery: A multi-stage comparative study. *Clinical Neurology and Neurosurgery*, 109039. <https://doi.org/10.1016/j.clineuro.2025.109039>
- Fahrni, S., & Sabatasso, S. (2025). 3D modelling in anatomy teaching: State of the art and pilot investigations for its application. *Translational Research in Anatomy*, 100444. <https://doi.org/10.1016/j.tria.2025.100444>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign. <https://discovery.ucl.ac.uk/id/eprint/10139722/>
- Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A., Ayyoub, A., Hattab, M. K., & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, 14(10), 1090. <https://doi.org/10.3390/educsci14101090>
- Lam, N. (2025). Interactive learning meets AI: Elevating student experiences in fashion and culture education. *EDULEARN25 Proceedings*. <https://doi.org/10.21125/edulearn.2025.0609>
- Zenios, M. (2020). Educational theory in technology-enhanced learning revisited: A model for simulation-based learning in higher education. *Studies in Technology Enhanced Learning*, 1(1). <https://doi.org/10.21428/8c225f6e.1cf4dde8>