Abilities of design professors to distinguish design assignments generated by students and AI

Yuan Yin [id],✉, Boheng Wang [id], Haoyu Zuo [id], Ruocong Liu [id], Shafina Iqbal Vohra [id],
Saul Haydon-Rowe [id] and Peter R. N. Childs [id]

Imperial College London, United Kingdom

✉ y.yin19@ic.ac.uk

ABSTRACT: This study aims to detect the ability of professors to distinguish design assignments generated by students with and without using AI. Ten students were recruited to undertake a conceptual design task twice, one with and one without the help of AI. 105 higher-education associate, assistant and full professors from industrial and product design programmes were recruited to assess the generated designs using a 7-point Likert Scale with nine indexes. The results indicate that assessors have moderate ability to distinguish between design assignments of students using AI and those where students did not use AI. Three cues to suggest the risk of the design assignment is made with AI instead of students who did not use AI were identified. By considering the three cues, lecturers distinguish design assignments generated by students with or without AI.

KEYWORDS: design education, artificial intelligence, higher education, conceptual design, justification

1. Introduction and literature review

The development of artificial intelligence (AI) has brought new opportunities to higher education (Zawacki-Richter et al., 2019). Applied AI in education has become a popular trend and provided potential pedagogical opportunities to support higher education students (Baidoo-Anu & Ansah, 2023; Jafari & Keykha, 2024; Zawacki-Richter et al., 2019). Some research has indicated that AI can provide creative ideas to students and complete rendering or video production for students (Fitria, 2021; Zhao et al., 2024). Some studies in higher education indicate that students can use AI in design processes and that AI can help students generate more creative ideas and improve students' design ability (Zhang et al., 2024). However, AI cannot replace the role of students in design processes as students have their own design thinking which differs their design from each other (Razzouk & Shute, 2012). In addition, students may not adequately develop design skills in their higher education if they rely too much on AI to get high scores (Rudolph et al., 2023). This could potentially hinder the growth of their professional design abilities. Therefore, although professors may encourage students to use AI in their design processes, it does not mean students should overly rely on AI to undertake and finish their assignments (Vázquez-Cano et al., 2023). Further, some students can have a tendency to overly depend on technologies which negatively affect independent thinking abilities (Cladis, 2020; Cingillioglu, 2023).

The use of AI in assignments is also related to academic integrity (Fowler, 2023). Although some research has indicated positive effects of AI on education, concerns have also been raised about academic integrity such as academic dishonesty (Yusuf et al., 2024), leading to plagiarism, impeding critical thinking, suppressing creativity, and eroding originality (Khatri & Karki, 2023). Vázquez-Cano et al. (2023) recruited 30 professors to evaluate the abstract created by students and ChatGPT based on the context and styles. The results suggested that ChatGPT provided better performance in context and

style of the abstract than that of students without the aid of AI. This suggests that the use of AI may increase unfairness among the students and lead to inequities in the assessment process (Cotton et al., 2024).

Some approaches have been tried to alleviate these academic integrity and ethics concerns. For example, some universities have allowed students to use AI to help them in assignments but also require students to provide an AI usage statement in their reports (Gonsalves, 2024). Students can also submit a draft of their work to review before the final submission as a sign of their writing abilities. Plagiarism detection tools have been developed to detect whether the work was finished by AI, such as the AI plagiarism detection function in Turnitin (Baron, 2024). However, AI detection tools may produce uncertain results. To compound the issues, these tools may give different detection results (Simon, 2023). Some university staff have also tried to solve the problems by changing the assignment submission from a report or profile to a presentation or group discussion (Cotton et al., 2024).

Considering the negative effects of using AI on students' design thinking and potential academic integrity and ethics concerns behind using AI in assignments, it has become important for professors abd markers of students' assignments, to have the ability to distinguish whether the students' design assignment was generated by with or without the aid of AI (Metersky et al., 2024). Zhang et al. (2024) asked 168 industrial design professors to finish a questionnaire related to their attitudes toward students' use of AI, areas in which students use AI, assessing students' use of AI, and ethical standards for students' use of AI. The results revealed that 73% of professors reported that they could independently distinguish the work completed by AI from those completed by students as the work completed by AI lacked emotion and was more professional. 27% of professors pointed out that if students modify the work from AI, it was hard to distinguish whether the work was completed by AI or by students (Cotton et al., 2024). Although this study revealed some results, this study was based on a questionnaire. The ability to make a distinction was reported by professors subjectively and not verified practically. In addition, the conditions for the writing assignments were different. Fleckenstein et al. (2024) asked young and professional educators to assess the writing generated by students and ChatGPT. The results revealed that AI (ChatGPT) can help students finish their writing assignment in a way that professors cannot distinguish and help students get a higher score in report writing.

Therefore, this study aims to detect the abilities of design professors to distinguish whether the higher-education students' design assignment is generated by students with or without the aid of AI practically. The research question of this study thus is whether design professors can distinguish whether a higher-education students' design assignment is generated by a student with or without the aid of AI. For this study the term "professors" is used broadly to refer to higher-education educators (teaching assistants, associate professors, assistant professors and full professors) who have experience in assessing design assignment of students.

2. Methodology

To answer the research questions, the design work generated by ten students using or not using AI was collected and assessed by 105 professors. The protocol of this study is shown in Figure 1. Design assignments completed by the students with and without use of AI were collected. Then, professors were recruited to mark the designs.

2.1. Phase I: design generation

2.1.1. Participants of phase I

Ten undergraduate and postgraduate students (5 females, 5 males, average age = 23.5) were recruited to finish a conceptual design task. All of the students have an industrial design or product design background and have experience in finishing conceptual design assignments. Also, they all self-reported that have knowledge in how to use Midjourney and were able to use Midjourney to generate conceptual design images.
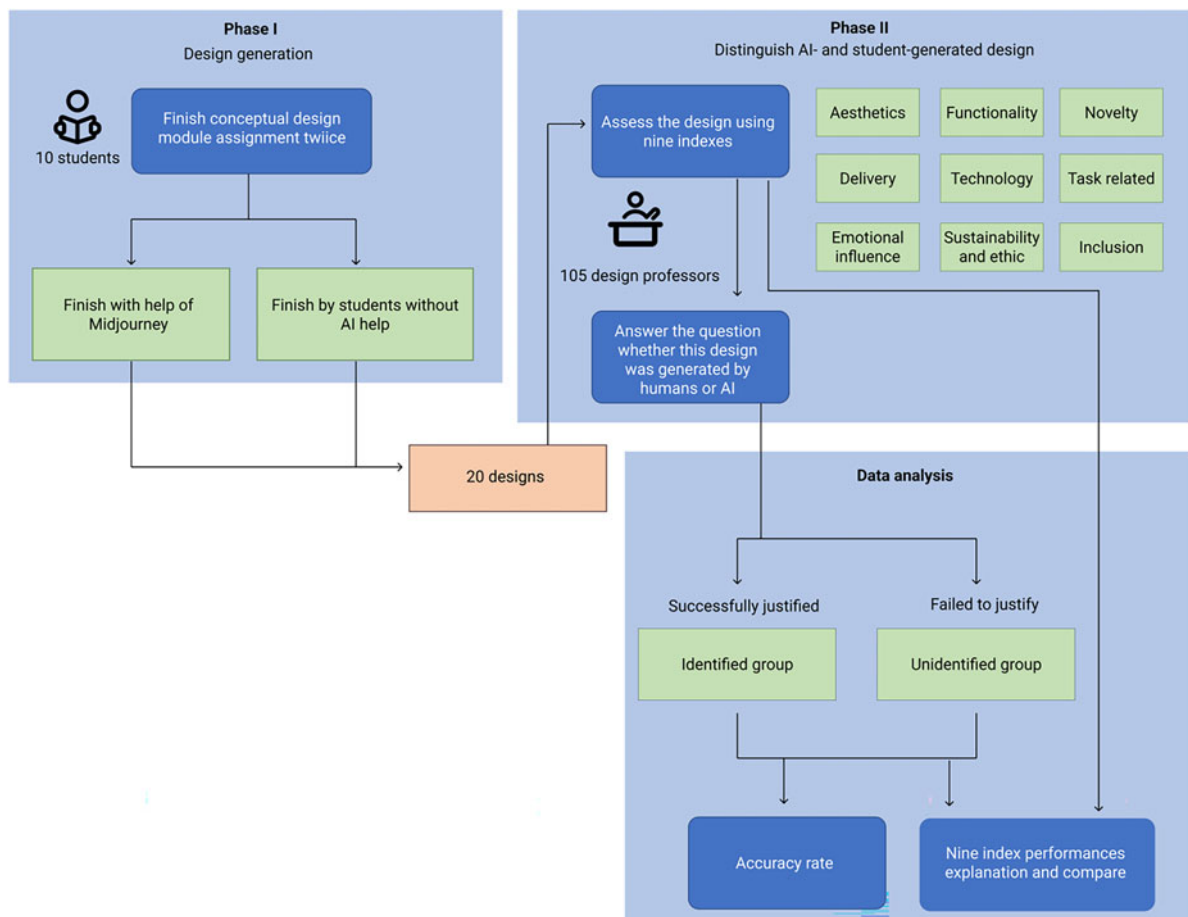
Figure 1. Study protocol

2.1.2. Protocol of phase I

Ten students were recruited to finish the phase I study. The participants information sheet was first sent to students through email. Students were provided an opportunity to ask any questions they had. If they did not have any questions, they were asked to sign the consent form. Then, each student needed to complete the design assignment twice. The design assignment was for a conceptual design module. Students were asked to "Design a conceptual product under one of the following topics: Smart devices, Furniture, and Accessible devices". This conceptual product should be submitted in a digital form. The reason why there were three alternative topics in this assignment was out of the consideration that students may have different interest areas in design. Three different topics gave students a flexible option to focus on what they were interested in (Pretorius et al., 2017). Among the ten students, four of them selected to design a smart device, three of them designed furniture, and three selected to design an accessible device.

To be specific, in one trial, students need to finish the design assignment by themselves without the help of AI (self-finished round). The design can be generated using any digital software, such as Rhino, Solidworks, Photoshop, and Illustrator. In the other trial, students need to finish the design assignment with the help of AI (AI-finished round). Midjourney (version 6) was selected for this study as the AI tool to finish the design assignment for the AI-finished round. In other words, instead of completing the assignment by themselves without the help of AI, in this trial, students need to finish the assignment only using Midjourney. Students need to use Midjourney to help them generate a conceptual design that they think can satisfy the assignment requirement as the final submitted design. The reason why Midjourney was selected as a representative AI tool was that Midjourney has been commercialized and widely tested by existing researchers (Arslan & Ghazal, 2024) and commonly used in design. Midjourney is a relatively mature text-to-image Generative AI and widely used in design communities for diverse tasks (Naseh et al., 2024). The AI engine behind Midjourney is deep learning. However, as Midjourney runs on closed-source and proprietary code, we cannot access to how the algorithms work in detail.

For both rounds, there were no time limitations for students to finish the task but it was suggested to finish the design within one hour. This time constraint may reduce the reliability of the study. In the assignment, although students were given the submission deadline, this deadline leaves enough time for students to finish the assignment. To mitigate this limitation, the assignment in this study was also simplified to only submit a digital design, instead of a profile of the design. All students finished the tasks within one hour. The outputs of both rounds were digital images. Students were voluntarily involved in this study.

In the second round task, students finished the same task as in the first round. As they already have an overview and considerations for their own design, this brought the second round task an order effect from the first round when students were thinking how to finish the second round task. To mitigate this limitation, the study randomly allocated five students to first use AI to finish the assignment and then finish the assignment by themselves, while the other five students were asked to first finish the assignment by themselves and then use AI to finish the assignment.

2.2. Phase II: distinguish AI- and student-generated design

2.2.1. Participants of phase II

105 design professors were recruited as markers (53 males, 52 females, average age = 29.59) to assess the product generated by students without use of AI and with AI (Midjourney). All have industrial design or product design backgrounds and have experience in assessing students' product design in higher institutions. Also, they all have knowledge of how to use Midjourney and were able to use Midjourney to generate conceptual design images.

2.2.2. Methodology of phase II - assignment criteria

To identify the ability of design professors in distinguishing student- or AI-generated assignments, the following nine indices were used as the assessment criteria – aesthetics (Christensen & Ball, 2016; Jansson-Boyd, 2011), functionality (Christensen & Ball, 2016), novelty (Christensen & Ball, 2016), delivery (Macmillan et al., 2001; Wessels & Roos, 2009), technology (Grunwald, 2009), task related (Mislevy et al, 2002; Wang et al., 2002), emotional influence (Ho & Siu, 2012; Lottridge et al., 2011), sustainability and ethic (Ceschin & Gaziulusoy, 2016; Hjalsted et al., 2021), and inclusion (Heylighen & Bianchin, 2013; Persad et al., 2007). These indices were selected because they have been widely used in conceptual design assessment and previous studies.

Aesthetics refers to whether the design has elements of visual attraction. Functionality refers to whether the design is easy and practical to use. Novelty refers to the creativity of the design. Delivery refers to whether the design clearly expressed the expectations and ideas of students and as percieved by markers. Technology refers to whether the design included technology details. Task related refers to whether the design satisfied the assignment requirement. Emotional influence refers to whether the design triggered the emotion of markers appropriately. Sustainability and ethics refers to whether the design considered the sustainability and ethical principles. Inclusion refers to whether the design considered and delivered inclusive design principles.

2.2.3. Protocol of phase II

105 markers were recruited to finish the phase II study. Participants information sheets were first sent to markers through email. Markers can ask any questions they have. If markers did not have any questions, they were asked to sign the consent form. Then, each marker was asked to evaluate 20 designs generated from phase I (10 students * (1 design generated by students + 1 design generated by AI)).

Markers were told that the assignment was for a conceptual design module. Students were asked to design a conceptual product under one of the following topics: Smart devices, Furniture, and Accessible devices. Markers were not told whether the design was generated by students without or with the aid of AI.

For each design assignment, markers were asked to first evaluate the design based on nine indices (aesthetics, functionality, novelty, delivery, technology, task related, emotional influence, sustainability and ethic, and inclusion) with a 7-point Likert scale (Score 1 means the design perform poorest on this

index, Score 7 means the design performed excellently on this index). To ensure markers understand the meaning of each index, an explanation of each index was also provided. The explanation of each index is included in Section Methodology of phase II - Assignment criteria. Then, markers need to answer the questions on whether they think this design assignment was completed by the student with or without the aid of AI. The 20 design outcomes were displayed in a random order to reduce the effect of display orders on the design in the assessment. The evaluation was displayed in the online questionnaire (Qualtrics). There was no time limitation on how long the markers should spend on the marking. Nevertheless, all the markers finished the assignment within two hours. The markers involved in this study were all voluntary.

3. Results

To answer the research questions, the results first explored whether design professors can distinguish between the students' generated assignments with an without the aid of AI. Then, the performance of AI- and student-generated assignments was compared. Examples of designs that students completed in both self-finished and AI-finished rounds are shown in Figure 2.

Figure 2. Examples of designs that students completed without and with the aid of AI

Based on the results, the accuracy rate of markers to distinguish AI- and student-generate assignments is 67.35%. For each marker, the assignment results were grouped into two categories. One group (Identified group) was related to these design assignments which have been successfully distinguished between students designs generated without and with the use of AI. The other group (Unidentified group) was related to these design assignments which failed to be distinguish between student designs produced without an with the use of AI. The results of all markers were integrated. Each index score of AI-generated design (and student-generated design) was represented based on the average score of each index. One-way ANOVA was used to calculate the significant levels of the results.

The Identified group results are shown in Table 1. From Table 1, it can be found that design professors gave a higher score to student-only generated design than AI-generated design on aesthetics, functionality, delivery, technology, task related, emotional influence, and inclusion. The higher scores on functionality ($p=0.013 < 0.05$), technology details ($p=0.005 < 0.05$), emotional influence ($p= 0.019$), and inclusion ($p=0.014$) were statistically significant. The table also reveals that design professors gave a lower score of student-generated design than AI-generated design on novelty, and sustainability and ethics. Both of the lower scores were statistically significant.

The Unidentified group results are shown in Table 2. From Table 2, it can be identified that design professors gave a higher score to student-generated design than AI-generated design on functionality, task related, emotional influence, and inclusion. The higher scores on task related ($p=0.0332 < 0.05$) and emotional influence ($p=0.027 < 0.05$), were statistically significant. It can also be found that design

Table 1. Aesthetics, functionality, novelty, delivery, technology, task related, emotional influence, sustainability and ethic, and inclusion scores for the Identified group

Index	Student-generated Average score	SD	AI-generated Average score	SD	P-value
Aesthetics	4.84	1.713	4.80	1.770	0.766
Functionality	4.80	1.842	4.49	1.891	0.013
Novelty	4.61	1.932	4.86	1.843	0.043
Delivery	4.67	1.809	4.63	1.892	0.785
Technology	4.84	1.847	4.48	1.953	0.005
Task related	4.75	1.766	4.71	1.844	0.716
Emotional influence	4.89	1.783	4.56	1.932	0.008
Sustainability and ethic	4.49	1.922	4.79	1.847	0.019
Inclusion	4.72	1.834	4.40	1.958	0.014

Table 2. Aesthetics, functionality, novelty, delivery, technology, task related, emotional influence, sustainability and ethic, and inclusion score for the Unidentified group

Index	Student-generated Average score	SD	AI-generated Average score	SD	P-value
Aesthetics	4.77	1.824	4.93	1.853	0.011
Functionality	4.79	1.763	4.68	1.987	0.247
Novelty	4.61	1.927	4.83	1.784	0.001
Delivery	4.66	1.758	4.71	1.745	0.763
Technology	4.69	1.936	4.82	1.737	0.001
Task related	4.72	1.729	4.65	1.871	0.032
Emotional influence	4.74	1.892	4.65	1.719	0.027
Sustainability and ethic	4.56	1.716	4.76	1.873	0.003
Inclusion	4.71	1.752	4.62	1.922	0.394

professors gave a lower score to student-generated design than AI-generated design on aesthetics, novelty, delivery, and technology. The lower scores on aesthetics ($p=0.011<0.05$), novelty ($p=0.001<0.05$), technology ($p=0.001<0.05$), and sustainability and ethics ($p=0.003<0.05$) were statistically significant.

4. Discussion

The results of this study are considered first, followed by a discussion on the contributions and limitations of this study.

4.1. Explanation of the results

The results of this study revealed that the accuracy rate of markers to distinguish between AI-aided and student-only generated design assignments was 67.35%. This accuracy rate indicated that design professors have moderate ability to distinguish between design assignments generated with and without the aid of AI. The results of this study indicate that design professors may not be able to distinguish between AI-aided and student-only design assignments with high accuracy. The results of this study echoed existing research which have indicated that design professors have limited ability to distinguish between AI- and student-generated design assignments (Lottridge et al., 2011). However, this study further enlarged this limitation in ability from text-based to image-based assignment. In addition, this result differs from some studies which have indicated that design professors can distinguish the student- and AI-generated design (Zhang et al., 2024). The differences may be because Zhang et al. (2024) asked participants to finish a questionnaire to report their ability to distinguish between AI- and student-generate design assignments, while this study involved judgement of a practical task to report the ability. This study has indicated which kind of design performance can affect design professors to distinguish the student- and AI-generated design assignments. Based on the Identified group results, it is shown that

student-generated design assignment has better performance than AI-generated design assignment, especially in aesthetics, functionality, delivery, technology, task related, emotional influence, and inclusion, while AI-generated design assignment has a better performance than student-generated design assignment in novelty, and sustainability and ethics. It is understandable that AI-generated design assignments have a higher novelty score as AI generated ideas are based on big data which allowed AI to access more information that students have accessed (Liao et al., 2020). The higher score of sustainability and ethics in AI-generated design assignments may be because students have not paid enough attention and realized the importance of sustainability and ethics while the training data of AI has included the characteristics of sustainability and ethic (Larsson et al., 2019). The reason why humans have a higher score on aesthetics may be because aesthetics is a subjective area that has not been standardized. Currently it is a significant challenge for AI to achieve meaningful and effective aesthetics. As for the delivery, AI can only generate images based on prompts. This lower score on AI-generated design assignments may indicate that students cannot express and display their ideas as flexibly as they draw it by themselves (Kim et al., 2022).

However, the Unidentified group showed a different result. Design professors gave a higher score of student-generated design assignments than AI-generated design assignments on functionality, task related, sustainability and ethics, and inclusion, while a lower score on aesthetics, novelty, delivery, technology, and emotional influence. A histogram (Figure 3) is created to compare the results between Identified group and Unidentified group results.
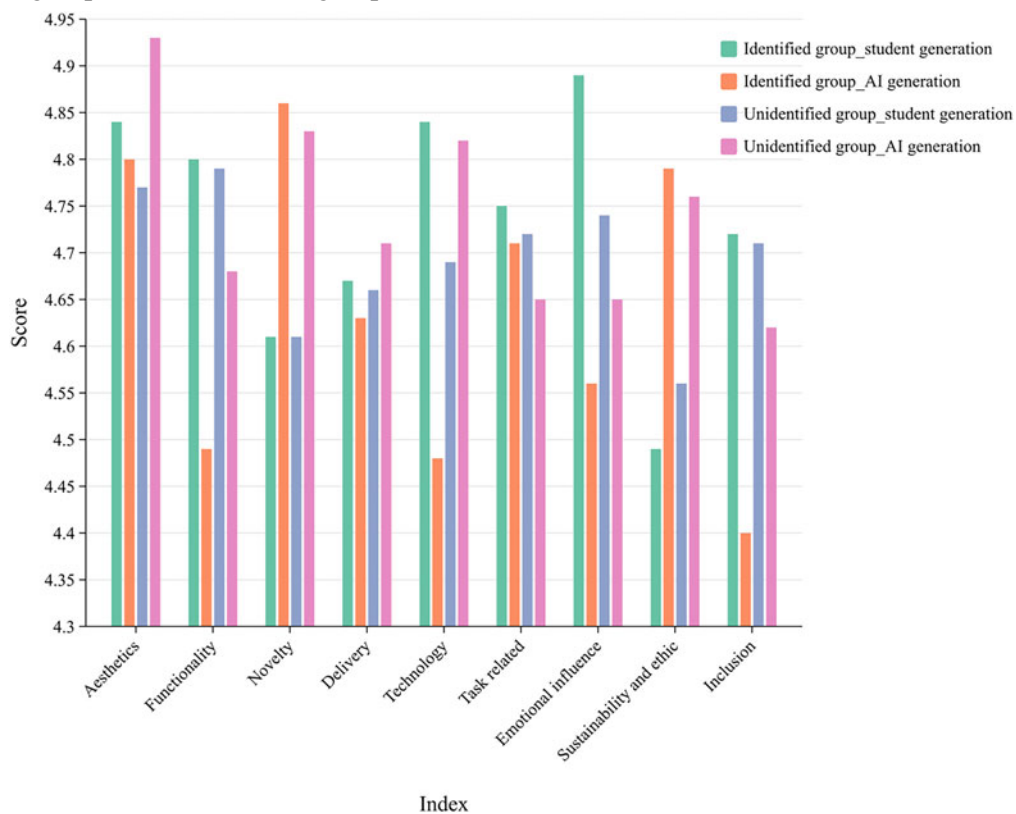
Figure 3. The Identified group and Unidentified group scores on the nine indices (aesthetics, functionality, novelty, delivery, technology, task related, emotional influence, sustainability and ethic, and inclusion) based on AI-generated results and student-generated results *(*This image is drawn in software Chiplot (https://www.chiplot.online/)*

Based on the comparison, the cue for design professors to justify whether the submitted design assignment was made by students or AI can be summarized. Firstly, the cue is based on the performance of functionality, technology, and inclusion. If the design assignment showed a low performance on these three indexes (functionality, technology, and inclusion), design professors may pay attention to whether the design assignment was completed by students or AI. Secondly, the cue is based on the performance of sustainability and ethics. If the sustainability and ethic performance is high in the submitted design tasks, design professors may realize that the design has a risk being finished by AI. Thirdly, design professors can use emotional influences as a cue. If the design has a low performance on emotional influences,

design professors may be aware that the design assignment has a higher change that was generated by AI. By synthetically considering the three cues, design professors can decide whether the design was generated by students or AI more clearly.

4.2. Contribution

This study revealed that design professors have moderate capacity to identify whether the conceptual design is finished by students with or without the aid of AI. This moderate accuracy level only provides indication that there is a need to help lectures distinguish between AI-aided and student-only generated design assignments. In addition, the results between the successfully identified and unidentified students (and AI) design were compared. The results identified three cues from design professors on how they justify the student-only and the AI-generated design assignments. These cues could be useful to aid design professors to further consider how to distinguish between AI-assisted and student-only generated design assignments. Also, the results can help developers to further consider how to produce tools to assist design professors to distinguish beween student-generated design assignments produce with and without the aid of AI.

4.3. Limitation and future work

Although the study brings new insight to researchers, educators, and developers, this study also included limitations. Firstly, the study did not fully simulate the design assignment in a form of a portfolio or report; Instead, the conceptual design tasks were used as the assignment task. This different submission format may bring a bias to the assignment performance. In the future, more studies need to be conducted to detect whether design professors have the ability to distinguish student design portfolios and design reports produced with and without the aid of AI. In addition, this study did not limit the professional levels of markers. No distinction was made between teaching assistants, associate professors, assistant professors, or professors. As long as they have experience in being a design professor and assessing conceptual design, they were considered as markers. Although some research studies have indicated that the professional levels of the markers does not affect assessment results (Yin et al., 2021), this was not based on distinguishing between student-generated design work produced with and without the aid of AI. Therefore, it is still worth exploring the effect of professional levels. Furthermore, although this study has indicated three cues for design professors to consider whether the design assignment was finished by AI or students alone, these cues need to be further validated and more studies are expected to detect how the interactive effect of these three cues can affect design professors' justification. Finally, only recruited 105 markers whose background was industrial design or product design were recruited for this study and the design assignment concerned only conceptual design image, while there are other forms of design assignment such as graphic or application design assignments. This limited participant and scope sample thus may be not be able to fully represent design professors who need to assess differing types of design assignment. In the future, more participants and types of design assignment could be considered.

5. Conclusions

This study concerned the ability of academic assessors to distinguish the design assignments generated by students with and without the use of AI. Ten students were recruited to finish a conceptual design task twice with one round completed by themselves and the other round completed with an AI tool, Midjourney. 105 associate, assistant and full professors from design programmes were recruited to assess the designs using nine indices (aesthetics, functionality, novelty, delivery, technology, task related, emotional influence, sustainability and ethics, and inclusion). The results of this study revealed that design professors have a moderate ability to distinguish between student-generated design assignments with and without the use of AI. Three cues that can be used for design professors to justify whether the submitted design assignment was made by students or AI were summarized. Design professors need to be aware that the design assignment has a risk of being completed by AI, (i) if the design showed a low performance of functionality, technology, and inclusion; (ii) if the sustainability and ethic performance of the design was high; (iii) if the design has a low performance on emotional influences. The results of this study can be used to support the need for additional assistance for design professors to distinguish between student-generated design assignments that have been produced with or without the use of AI. In addition, this study has potential to aid developers to produce assistance tools which can help design

professors distinguish student-generated design assignments that have been produced with and without the aid of AI.

References

Arslan, M., & Ghazal, A. (2024). Assessing the Role of AI in Generating Content for Digital Brand Communications: The Case of a Midjourney Generated Campaign. *Open Access Library Journal*, 11(11), 1–14.

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62.

Baron, P. (2024). Are AI detection and plagiarism similarity scores worthwhile in the age of ChatGPT and other Generative AI? *Scholarship of Teaching and Learning in the South*, 8(2), 151–179.

Ceschin, F., & Gaziulusoy, I. (2016). Evolution of design for sustainability: From product design to design for system innovations and transitions. *Design studies*, 47, 118–163.

Christensen, B. T., & Ball, L. J. (2016). Dimensions of creative evaluation: Distinct design and reasoning strategies for aesthetic, functional and originality judgments. *Design studies*, 45, 116–136.

Cingillioglu, I. (2023). Detecting AI-generated essays: the ChatGPT challenge. *The International Journal of Information and Learning Technology*, 40(3), 259–268.

Cladis, A. E. (2020). A shifting paradigm: An evaluation of the pervasive effects of digital technologies on language expression, creativity, critical thinking, political discourse, and interactive processes of human communications. *E-Learning and digital Media*, 17(5), 341–364.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in education and teaching international*, 61(2), 228–239.

Fitria, T. N. (2021). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. *Paper presented at the Prosiding Seminar Nasional & Call for Paper STIE AAS*.

Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence*, 6, 100209.

Fowler, D. S. (2023). AI in Higher Education: Academic Integrity, Harmony of Insights, and Recommendations. *Journal of Ethics in Higher Education*, 3, 127–143.

Gonsalves, C. (2024). Addressing student non-compliance in AI use declarations: implications for academic integrity and assessment in higher education. *Assessment and Evaluation in Higher Education*, 1–15.

Grunwald, A. (2009). Technology assessment: Concepts and methods. *In Philosophy of technology and engineering sciences* (pp. 1103–1146): Elsevier.

Heylighen, A., & Bianchin, M. (2013). How does inclusive design relate to good design? Designing as a deliberative enterprise. *Design studies*, 34(1), 93–110.

Hjalsted, A. W., Laurent, A., Andersen, M. M., Olsen, K. H., Ryberg, M., & Hauschild, M. (2021). Sharing the safe operating space: Exploring ethical allocation principles to operationalize the planetary boundaries and assess absolute sustainability at individual and industrial sector levels. *Journal of Industrial Ecology*, 25(1), 6–19.

Ho, A. G., & Siu, K. W. M. G. (2012). Emotion design, emotional design, emotionalize design: A review on their relationships from a new perspective. *The Design Journal*, 15(1), 9–32.

Jafari, F., & Keykha, A. (2024). Identifying the opportunities and challenges of artificial intelligence in higher education: a qualitative study. *Journal of Applied Research in Higher Education*, 16(4), 1228–1245.

Jansson-Boyd, C. V. (2011). Designing aesthetic concepts: Can it be done? *Psychology of Aesthetics, Creativity, and the Arts*, 5(3), 279.

Khatri, B. B., & Karki, P. D. (2023). Artificial Intelligence (AI) in Higher Education: Growing Academic Integrity and Ethical Concerns. *Nepalese Journal of Development and Rural Studies*, 20(01), 1–7.

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27(5), 6069–6104.

Larsson, S., Anneroth, M., Felländer, A., Felländer-Tsai, L., Heintz, F., & Ångström, R. C. (2019). Sustainable AI: An inventory of the state of knowledge of ethical, social, and legal challenges related to artificial intelligence.

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: informing design practices for explainable AI user experiences. *Paper presented at the Proceedings of the 2020 CHI conference on human factors in computing systems*.

Lottridge, D., Chignell, M., & Jovicic, A. (2011). Affective interaction: understanding, evaluating, and designing for human emotion. *Reviews of Human Factors and Ergonomics*, 7(1), 197–217.

Macmillan, S., Steele, J., Austin, S., Kirby, P., & Spence, R. (2001). Development and verification of a generic framework for conceptual design. *Design studies*, 22(2), 169–191.

Metersky, K., Chandrasekaran, K., Rahman, R., Haider, M., & Al-Hamad, A. (2024). AI-generated vs. student-crafted assignments and implications for evaluating student work in nursing: an exploratory reflection. *International Journal of Nursing Education Scholarship*, 21(1), 20230098.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2002). On the roles of task model variables in assessment design. *Item generation for test development*, 97–128.

Naseh, A., Thai, K., Iyyer, M., & Houmansadr, A. (2024). Iteratively Prompting Multimodal LLMs to Reproduce Natural and AI-Generated Images. *arXiv preprint arXiv:.13784*.

Persad, U., Langdon, P., & Clarkson, P. J. (2007). A framework for analytical inclusive design evaluation. *Paper presented at the DS 42: Proceedings of ICED 2007, the 16th International Conference on Engineering Design*, Paris, France, 28.-31.07. 2007.

Pretorius, L., van Mourik, G. P., & Barratt, C. (2017). Student choice and higher-order thinking: Using a novel flexible assessment regime combined with critical thinking activities to encourage the development of higher order thinking. *International Journal of Teaching and Learning in Higher Education*, 29(2), 389–401.

Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important?. *Review of educational research*, 82(3), 330–348.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of applied learning and teaching*, 6(1), 342–363.

Simon, W. (2023). Distinguishing between student and AI-generated writing: A critical reflection for teachers. *Metaphor*(3), 16–23.

Vázquez-Cano, E., Ramirez-Hurtado, J. M., Saez-Lopez, J. M., & Lopez-Meneses, E. (2023). ChatGPT: The brightest student in the class. *Thinking Skills and Creativity*, 49, 101380.

Wang, L., Shen, W., Xie, H., Neelamkavil, J., & Pardasani, A. (2002). Collaborative conceptual design—state of the art and future trends. *Computer-Aided Design*, 34(13), 981–996.

Wessels, P. L., & Roos, S. A. (2009). The development of a conceptual framework for the design, delivery, and assessment of a typical management accounting syllabus. *Accounting Perspectives*, 8(2), 147–164.

Yin, Y., Han, J., Huang, S., Zuo, H., & Childs, P. (2021). A Study on Student: Assessing Four Creativity Assessment Methods in Product Design. *Proceedings of the Design Society*, 1, 263–272.

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21(1), 21.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27.

Zhang, Y., Fan, Y., & Bohemia, E. (2024). Perceptions of lecturers of artificial intelligence on industrial design students. *Paper presented at the DS 131: Proceedings of the International Conference on Engineering and Product Design Education (E&PDE 2024)*.

Zhao, L., Hussam, E., Seong, J.-T., Elshenawy, A., Kamal, M., & Alshawarbeh, E. (2024). Revolutionizing art education: Integrating AI and multimedia for enhanced appreciation teaching. *Alexandria Engineering Journal*, 93, 33–43.