Implications of Generative AI on Learning and Assessment in Higher Education and Design Research Practice

Roger Whitham, Glynn Stockton, Daniel Richards, Joseph Lindley, Naomi Jacobs, Paul Coulton
{r.whitham, g.stockton, d.richards, j.lindley, naomi.jacobs, p.coulton}
@lancaster.ac.uk

ImaginationLancaster
School of Design
Lancaster University
United Kingdom

ABSTRACT

In this paper, we explore the impact of generative AI (GenAI) on assessing student work in further education contexts. As GenAI becomes widely adopted, it engenders potential risks to assessment, such as false evidence of learning, student vulnerability to academic integrity injustice, and implications for independent learning and creativity. We report on the interim findings of a working group seeking to understand these risks and propose possible mitigations. Alongside desk research and discussion, a 'wargaming' activity was used by the team to attempt assignments with the help of AI tools. Based on our research, we propose possible mitigations to the assessment challenges, including adapting assessment methods, minimising reliance on automated enforcement tools, and reimagining course structures to integrate GenAI into them. In conclusion, we discuss critical issues and potential mitigations surrounding the intersection of GenAI, student assessment and Design Research practice and reflect upon how these considerations may inform and reshape the future of Design Research practice in general.

CCS CONCEPTS

•Applied computing~Education~Computer-assisted instruction
•Human-centered computing~Human computer interaction (HCI)

KEYWORDS

Generative AI; education; design research.

Background

We present this workshop paper as interim findings in part of a process intended to help our institution pair our teaching to robust and fair assessment methods. Alongside cross-cutting impact on many facets of creativity and criticality [2], the rapid adoption of a wide range of GenAI technologies in 2022 and 2023 is changing the nature of the work submitted by students. Our current modes of assessment—written assignments, but also visually-oriented design work and project-based work are intended to assess the students' abilities to demonstrate their learning. But they do not account for how GenAI may assist student learning or assist students in producing the submissions that are assessed.

In the following, we briefly explain some challenges we face as educators. Then, we describe a series of 'wargames' we conducted, where our team members aimed to better understand these challenges by attempting to complete assignments solely by using GenAI tools. Next, we use our background research and experience gained during the wargaming to suggest—so far untested—strategies to mitigate the risks posed by the previously identified challenges. We conclude by discussing how these issues relating to teaching and assessment have significance for parallel and related discussions in the context of Design Research practice.

We note that the work presented here does not represent the view of our institution and is offered as a provocation to help illuminate critical issues that GenAI is raising in the academy.

Discussion of Challenges

The challenges discussed here relate to teaching and assessment of that teaching. To offer a very general example, in a course on HCI, students may be given a lecture on Speculative Design and then asked to go and do their own independent reading/research, after which they should create a sketch and essay to demonstrate their learning about what Speculative Design is and how it works. Historically the sketch and the essay would have been sufficient evidence to assess whether the students had met the learning objectives. GenAI, however, may provide ways to produce materials that make it appear as if the student has learned when they have not. The following list is not exhaustive but illustrates some challenges.

False Evidence of Learning

Referring to the example above (assessing whether a student understands speculative design) a student could instruct a large language model (e.g., ChatGPT) to generate a list of speculative design concepts. Using an image generator, the student could produce a corresponding design sketch. Based on describing the sketch and further prompting, a student could then produce an essay explaining the design, why speculative design was a useful approach in this context, and even what was learned through the process.

Student Vulnerability to Academic Integrity Injustices

By 'academic integrity injustices', we refer to institutional mechanisms that have historically been focused on ensuring that students do not fool assessment systems through plagiarism (copying). The chance of creating a carbon copy of somebody else's text by chance is virtually zero. Hence, automated systems have successfully identified students who chose to reproduce the work of others. Services such as TurnItIn are commonplace in this academic integrity process and considering the widescale adoption of GenAI text generators, have implemented systems to try and automatically flag text that may have been written by a machine rather than a human. At the current time, it is not clear how such systems function or what their susceptibility to false negatives (not realising text is AI generated) or false positives (accidentally citing human generated text as AI generated) is. Moreover, proving the situation one way or the other seems virtually impossible. The upshot is a risk that students could become vulnerable to being accused of using an AI in their assessments when they did not.

Implications for Independent Learning and Creativity

In the example above (assessing whether a student understands speculative design), we suggested that after a lecture a student should then do their own additional reading and research before using what they have learned to propose a project, create a sketch, and describe the whole process in an essay to demonstrate their learning. A significant portion of this involves independent learning and creativity. GenAI technologies can be useful in enhancing these activities but can also be detrimental to the learning outcomes that should be associated with these activities. If GenAI is used to compare two texts and generate a critical analysis of the contrasts between them, this avoids having to use one's own critical faculties. Similarly, suppose GenAI is used to create a list of ideas based on a context (e.g., a prompt such as "Generate 5 speculative design concepts relating to creativity support tools that use big data"). In that case, plausible ideas may be generated without an underlying creative process which is crucial to developing students' skills having taken place.

Wargaming of Assignments

These challenges are simple explanations of complex and interlinked issues. To better understand how some of these challenges may manifest for students facing actual learning briefs and for the teaching staff responsible for assessing their responses, we adopted an approach we termed 'wargaming'. This took the form of three team members choosing to complete specific assignments relating to modules offered as part of our courses at the School of Design at Lancaster University. We played two 'rounds' of our wargame. Two of those playing games are teaching staff, and one is a researcher. One of the teaching staff is an expert in academic integrity. In this section, the 'we' refers to the three team members who were playing the wargames. All the wargaming activity (including planning, execution, and discussion) took place on the same day, i.e., it was a fast and dynamic approach to understanding the challenges better.

Game 1

We limited ourselves to 20 minutes of 'game time' in the first round. During those 20 minutes, each player had to produce their assignment. We limited ourselves to basic prompt engineering and 'one shot' generations (i.e., taking what was produced by the GenAI verbatim without manual editing).

Game 2

We allowed ourselves around an hour of 'game time' in the second round. In this longer game, we expected to refine our prompts, tinker with the AI output, and explore using 'pipelines' of multiple tools alongside some basic research to enhance the quality of our outputs.

What did we find?

Taking the abstract notion of 'how will GenAI impact our students' submissions' and concretising the question by attempting to create those submissions ourselves was illuminating—we encourage *practical* engagement with the technologies that are causing concern in this context to understand the concern better. This was not a comparative study, i.e., each player had a different assignment and used different tools.

Given such a short time, in Game 1, we simulated somebody putting next-to-zero effort into the assignment. We adopted strategies such as using the course brief as a prompt, with the addition of "Write me an essay on this". While impressive, the results were obviously not proper essays. To improve the output

within the 20-minute timeframe, we adopted strategies such as asking for an essay plan rather than the full text. Similarly, generating individual sections provided better results quickly.

Game 2 simulates somebody who is prepared to put a significant amount of effort into avoiding doing the actual learning. We adopted a wider range of approaches here. Strategies included creating a 'mega prompt' relating to the assignment. This included describing the style of writing required, the background readings and references to use, and the topics/structure that the essay should have. Other strategies related to incrementally developing and creating assignment content start by generating ideas, then creating a structure and then generating individual parts. A final strategy involved 'pipelining' several different technologies to enrich the content created by the GenAI and disguise that a human did not create it.

The team concluded that some of the content created during Game 1 could achieve a passing mark for a student submitting it. Some of the content created in Game 2 would likely have achieved a good mark if it were submitted. By using simple strategies it was easy to 'fool' automated detection tools into assuming that GenAIs did not write our assignments.

The team also concluded that the challenge of harnessing the GenAIs to create high-quality content is something of an art form or craft. As such, it seems *plausible* that creating prompts on a given topic—i.e., to use GenAI as a tool to respond to a given brief—can help teach how to use GenAI but in parallel, may also help teach about the topic in question. However, we also noted that, if 'playing the game' (i.e., trying to produce a suitable output with the minimum viable effort), an intelligent player could avoid this kind of incidental learning. We realized that in cases where the GenAI's initial response diverged from the thinking behind the prompt, it seemed easier to 'capitulate' and let the GenAI go in its own direction. This mechanism reduces the agentic capacity of the student to influence the overall outcome, but it means that a quality outcome can be produced with minimal effort.

Mitigating the Challenges

These proposed mitigations explore ways to think about teaching and assessment to promote optimal learning outcomes for students and embrace the adoption of GenAI. The following is not an exhaustive list but represents some points of view that may be used generatively to explore how to respond to the challenges GenAI poses in education. We have also chosen specific examples that will relate to our concluding section where we draw parallels to Design Research practice.

Do not rely on automated GenAI detection

We managed to beat current versions of some leading detection services with minimal effort. Reliably trusting what such services say (even if their outcome is a probability, e.g., '60% chance of being AI') is likely ill advised, especially if cognitive biases are considered. The risk of false positives and integrity injustice is likely to outweigh the reliability of identifying true fraud. Reframing *why* it is that one wants to detect something created by a GenAI is a more useful strategy.

Update assessment to be robust against AI-spoofed learning

Many higher education assessments take the form of written assignments. The learning that this writing is supposed to demonstrate can easily be spoofed using GenAI. Types of assessment that include more verbal presentations, controlled examinations (or other tasks completed in controlled environments) and practical/portfolio work may be more robust in terms of protecting against AI-spoofed learning.

Reward explanation, rationale, and contextualization

Whatever format an assessment takes it may be useful to weight merit towards detailed explanations or rationale and context. Doing so may reduce the attractiveness of relying on ideas created by GenAIs. Moreover, even if an idea was ultimately AI generated, if there is a requirement to meaningfully offer explanation, rationale and context then the student's learning and the insights represented in the work are likely to be maintained, rather than diminished by the GenAI.

Require AI Appendices

There are many reasons to encourage the use of GenAI tools but given how disruptive they can be to long-held notions of ownership, authenticity, and creativity, it is important that their use—particularly in an education or research context—is transparent. We suggest that any assignment should have an AI appendix to explain how and why GenAI was used to create the work. This would function both to support and maintain academic integrity, as well as helping to contribute towards a common understanding of what GenAI tools can do, and how to achieve it.

Generative AI and Design Research Practice

This short paper is content concerns teaching and assessing students' work in a higher education context and the impact that GenAI may have. This section draws parallels between that context and the context of Design Research practice, which is increasingly using and impacted by AI [1]. While most of our discussion relates to written work, our team is situated in a design school, and many of the findings resonate with aspects of Design Research. For example, much of what we assess with our students involves understanding and critically responding to complex contexts and theory, being creative to imagine and explore possible responses, and then articulating the value of that somehow. This kind of structure is reproduced in many Design Research projects. So, we ask, what risks are there for future Design Research practice? And what opportunities does GenAI pose?

We suggest that popular Research through Design practices such as Speculative Design, Design Fiction, and Critical Design stand to gain much from GenAI and also have much to lose. On the positive side, these practices (in particular, Design Fiction) don't necessitate any physical product, functional prototype, or tangible outcome to be perceived as valuable in a research context. Using GenAIs to iterate rapidly, producing high-quality digital artefacts, and using the process of imagining, refining, and reflecting on these artefacts as part of the design process could be extremely valuable. Established formats such as imaginary abstracts and fictional papers are analogous to this process; those forms take entirely invented abstracts and papers as proxies for possible futures. Although the worlds the abstracts and papers describe aren't real, by describing how they *could* be real and unpacking their textures and details this is considered a legitimate research practice. Similarly, as opposed to designing prototypes, using GenAI to help refine concepts, visualize them, and even articulate what was learned in the process seems as though it could be a similarly legitimate research practice. In the same way that describing fictional events that never happened is a helpful shortcut for researchers, having an AI help with the heavy lifting of ideation or sketching doesn't seem inherently flawed. So long as the insights that come out at the end are accurate reflections of the creative and critical process, arguably, there's little distinction between 'normal' Design Research and Gen-AI-infused Design Research.

Conversely, as demonstrated with our wargaming exercise, any Design Researcher who was minded to, could easily simulate a process that resulted in prototypes, processes, and reflections that are almost indistinguishable from those predominantly informed by human-generated insight, even if none is present. A clear departure from our discussion relating to education is, in that context, the onus is on the student to demonstrate a specific level of learning and the institution to assess the quality of that quite concretely (e.g., grade A, B, or C). In the context of Design Research, the outcomes are much softer and more challenging to define. This begs the question: if nobody can tell whether a Design Research project was entirely imagined, constructed, and analyzed by an AI—does it really matter? Whilst we do not give an answer to this question here, we offer some reflections on related concepts based on our work in an education and assessment context.

We noted how careful attention to prompt crafting could likely become useful for independent learning and creativity. It seems likely that, in a Design Research context, using GenAI as a tool to hone and refine the practitioner's perspective may be generative and productive in a similar way. In other words, if used diligently, GenAI may enhance Design Researchers' understanding of the problem spaces they are exploring.

In a practical sense, textual GenAI's can be useful tools for translating complex and hard-to-articulate ideas into simpler, more transferable language and insights. Considering the type of 'intermediate knowledge' [3] that often arises from Design Research can be hard to articulate, GenAI may improve Design Researchers' abilities to communicate the value of their work.

As in education, it seems absurd to imagine that things created with GenAI are not already pervading many examples of Design Research. As such, and in the same spirit of transparency and contribution towards a common understanding of what GenAI tools can do and how they achieve it, we suggest that the adoption of AI appendices for Design Research would be valuable. Given that some scholars will likely believe it does matter whether aspects of Design Research projects were imagined, created and analysed by GenAI—and some may hold the counter view—rapidly adopting transparent practices regarding communicating the role of GenAI in a particular Design Research may be useful to reduce the risk that AI-infused work is stigmatized and help develop consensus relating to acceptable practices.

Additional Facets of AI in Teaching

In the process of producing this paper several other related concerns arose, but our commentary on them is not sufficiently rationalised to incorporate into the core argument. The climate emergency demands careful attention is paid to energy and resource consumption, the scale of these matters in relation to broad adoption of GenAI is not well understood. Developing frameworks to contextualise the environmental impact is crucial for using GenAI in either teaching or Design Research. Cost as a barrier to access was also identified as a potential issue. Whilst the widely adopted models are currently within reach of most students in UK Universities (e.g., £20/month) more expensive or exclusive models may give wealthier individuals an unfair advantage. We also noted data protection concerns, the training of many models is opaque, and the processing of models may occur in jurisdictions with incompatible data protection legislations from the place of use. We suggest each of these areas of concern demand further attention to explore their implications in teaching, Design Research, and beyond.

Conclusion

In this short paper we describe a process intended to respond to the impact of GenAI on higher education teaching and assessment. We described some high-level challenges, our exploration of those challenges through a practical wargaming activity, and some potential mitigations. We concluded by discussing parallels and shared logics between the impact of GenAI on education and assessment and in the context of Design Research. We suggest that in each context there are some common risks that may have common solutions, for example the potential to include AI appendices to explain what role AI played in each project or assignment. GenAI is being adopted and the full scale of what its impact will be is hard to predict. We hope that these interim findings may contribute to useful discussions about what a preferable future, where GenAI is more integrated into the mainstream of both teaching and research, will look like.

ACKNOWLEDGMENTS

Thanks to Dr Claire Coulton for assisting with this article. This work was supported by Design Research Works (https://designresearch.works) under UK Research and Innovation (UKRI) grant reference MR/T019220/1.

AI APPENDIX

A version of OpenAI's GPT-4 was used to formulate this article's abstract based on a bullet-point list of topics. The original version of the article content was written manually but with interventions from Microsoft Word's inbuilt grammar checker and with the support of the Grammar.ly style, spelling, and grammar agent. During the wargaming session a broad range of AI tools were used by the players.

REFERENCES

1. Jesse Josua Benjamin, Heidi Biggs, Arne Berger, Julija Rukanskaitė, Michael B. Heidt, Nick Merrill, James Pierce, and Joseph Lindley. 2023. The Entoptic Field Camera as Metaphor-Driven Research-through-Design with AI Technologies. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–19. https://doi.org/10.1145/3544548.3581175
2. Ziv Epstein, Aaron Hertzmann, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R. Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex "Sandy" Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith. 2023. Art and the science of generative AI. *Science* 380, 6650: 1110–1111. https://doi.org/10.1126/science.adh4451
3. Kristina Höök and Jonas Löwgren. 2012. Strong concepts. *ACM Transactions on Computer-Human Interaction* 19, 3: 1–18. https://doi.org/10.1145/2362364.2362371