TechTrends (2024) 68:195–199
https://doi.org/10.1007/s11528-023-00926-x

COLUMN: EDITORIAL

Innovation of Instructional Design and Assessment in the Age of Generative Artificial Intelligence

Charles B. Hodges[1] · **Paul A. Kirschner**[2,3]

Published online: 23 December 2023

The release of Generative Artificial Intelligence (GenAI) tools like ChatGPT has sparked vigorous discussion about its potential effects on education. Depending on what you read, GenAI will have an effect somewhere between solving all of education's problems and destroying learning and education completely. The claim that a new technology can and will solve all of education's problems is not new, though history tells us that these so-called technological silver bullets have not produced the predicted outcomes (see Thomas Edison's claims about the phonograph and moving pictures (1878, 1888) or Norman (n.d.)). Is GenAI different? Will it destroy education as we know it? GenAI is fundamentally different from other technologies of the last 20 + years due to its ability to generate original written work that is virtually indistinguishable from that of human authors. While this capability has disruptive implications for education, it is not likely to destroy it. It may, though, destroy the legitimacy of some long-held educational practices.

Take plagiarism. Unlike plagiarism, which can be readily detected with text-matching software, GenAI represents an AI-human hybrid authorship that current tools cannot reliably identify (Casal & Kessler, 2023) and for which current tech firms are racing/fumbling to develop. Hodges and Ocak (2023) note that past technological innovations such as affordable portable scientific calculators or Wikipedia caused instructors some distress with respect to the product (e.g., will students learn to calculate). GenAI's ability to produce human-like text without being detected as machine-generated, in contrast, presents both unique product and process challenges (e.g., will students learn to write and even think) that elude the policies of some instructors and institutions. While some embrace this new technology, others are seeking ways to ban their use. The European Union very recently (i.e., just the week before we wrote this) passed the EU AI Act[1] to protect people against the dangers of AI. The new rules establish obligations for providers and users depending on the level of risk from artificial intelligence.

Many major problems that historically have been observed regarding technology in education also exist in the context of GenAI. What are the ethics of integrating GenAI in education? What data do GenAI tools collect, and what do the companies that own the tools do with the collected data? Are these tools accessible to all learners? What issues of equity are created or perpetuated by these tools? The large-scale deployment and uptake of these tools have even brought attention to questions like, "What is the environmental impact of these technologies?" (Luccioni, 2023). While all these questions are important, we focus here on an instructional design problem related to GenAI.

Much so-called research being produced in the fields of instructional design and educational technology focuses on *things*; that is, the GenAI tools themselves. Reeves and Lin (2020) warned educational technology and instructional design scholars against conducting research on 'things'. And much of the research now being published on GenAI may not even qualify as research. At best, the work might be considered to be evaluation, but often it is more of a *gee whiz look at this* type of report focusing on the technology, the thing. Reeves and Lin recommend focusing on problems faced by teachers and students rather than on *things*.

✉ Charles B. Hodges
hodges.chuck@gmail.com

Paul A. Kirschner
paul.kirschner@ou.nl

[1] Georgia Southern University, Statesboro, GA, USA

[2] Open University, Heerlen, Netherlands

[3] Thomas More University of Applied Sciences, Mechelen, Belgium

[1] https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

New Problems

Along with the fundamental problems that teachers and instructional designers have always tried to tackle, namely "Did the **student** get the right answer?" (i.e., the product problem) and "How do we know that a student has learned?" (i.e., the process problem), there's a third problem now, namely, "What instructional strategies will support learning?". GenAI tools have made answering all three questions more complicated, especially in the context of asynchronous learning due to their overreliance on a particular set of assessment types and strategies – written responses, quizzes, and the like. GenAI tools can produce human-like responses to these types of assessments, allowing students to avoid [on purpose or accidentally] the hard work of learning. Even generating or interpreting graphics and diagrams is now within the reach of many GenAI tools. All of this makes generally productive learning activities, like the generative strategies highlighted by Fiorella and Mayer (2016), possible for GenAI tools to complete. The following figures are an outline, produced by ChatGPT (Fig. 1) and a summary (Fig. 2) of the first section of this editorial.

These problems, developing or utilizing instructional strategies and assessments that help students learn despite easy access to GenAI tools, require flexible thinking and innovation on the part of teachers and instructional designers. When instructors and students are together physically in the same room, it is possible to know what resources students use to complete tasks.[2] Therefore, GenAI presents less of a problem (no problem?) in that specific context. The remainder of this paper will consider the context where students are completing tasks away from their instructor, either as homework or in the context of a hybrid or fully online course.

One may be tempted to assume that GenAI tools, like ChatGPT, have negated the need for many types of knowledge. Asking for facts, procedures, or an analysis of facts is easily within the range of many GenAI tools now. However, Neelen and Kirschner (2020) respond to this type of thinking in detail in the context of learners and the Google search engine. They address the learning myth, "Google can replace human knowledge" by examining types of knowledge (e.g., propositional, tacit, etc.) and present well-documented arguments for such statements as:

> "Let's assume for a second that Google can replace our own knowledge. We'd still have to interpret the information that Google gives us to make it meaningful" (p. 122) and;

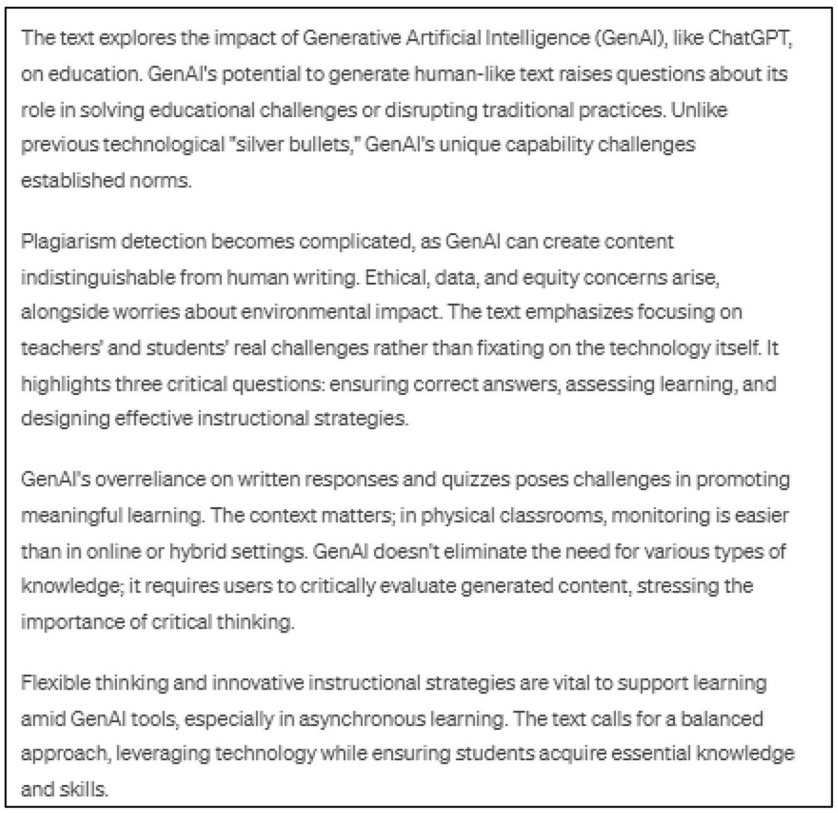
- GenAI in Education
  - Impact
    - *Solving Problems vs. Destroying Learning*
    - *Historical Technological Predictions*
  - Uniqueness
    - *Human-Like Text Generation*
    - *Disruption of Educational Practices*
  - Challenges
    - *Plagiarism Detection*
    - *Ethical, Data, and Equity Concerns*
    - *Environmental Impact*
  - Instructional Design
    - *Focus on Problems Over Tools*
    - *Complex Assessment Strategies*
    - *Adapting to Online Learning Context*
  - Role in Knowledge
    - *Types of Knowledge*
    - *Comparison with Google*
    - *Limitations and Critical Thinking*
    - *Instructional Design Strategies*

Fig. 1 Outline of the first section of this editorial

> "If we're trying to solve very complex problems, we run into several issues when relying on Google. The main problems are that we need to know what we're looking for and that we need to be able to judge the information we find based on the knowledge that's in our head" (p. 130).

The information retrieved from a Google search is not the same as what is produced by prompting a GenAI tool, but Neelen and Kirschner's arguments are still valid. GenAI tools are trained on massive data sets that may include inaccuracies and misconceptions. They do not think; they create human-like responses based on probabilities and, in doing so, also tend to make things up (i.e., hallucinate). It is not difficult to find, or generate, absurd results; for example, see Kirschner (2023). As Ian Bogost (2022) writes, GenAI "does not have the ability to truly comprehend the meaning behind those words. This means that any responses it generates are likely to be shallow and lacking in depth and insight." Just as Neelen and Kirschner observed for information obtained from a Google search, for the results of GenAI "we need to know what we're

[2] It does however limit what can be asked. You cannot ask a student to write a term paper in a 2-h exam sitting.


The text explores the impact of Generative Artificial Intelligence (GenAI), like ChatGPT, on education. GenAI's potential to generate human-like text raises questions about its role in solving educational challenges or disrupting traditional practices. Unlike previous technological "silver bullets," GenAI's unique capability challenges established norms.

Plagiarism detection becomes complicated, as GenAI can create content indistinguishable from human writing. Ethical, data, and equity concerns arise, alongside worries about environmental impact. The text emphasizes focusing on teachers' and students' real challenges rather than fixating on the technology itself. It highlights three critical questions: ensuring correct answers, assessing learning, and designing effective instructional strategies.

GenAI's overreliance on written responses and quizzes poses challenges in promoting meaningful learning. The context matters; in physical classrooms, monitoring is easier than in online or hybrid settings. GenAI doesn't eliminate the need for various types of knowledge; it requires users to critically evaluate generated content, stressing the importance of critical thinking.

Flexible thinking and innovative instructional strategies are vital to support learning amid GenAI tools, especially in asynchronous learning. The text calls for a balanced approach, leveraging technology while ensuring students acquire essential knowledge and skills.

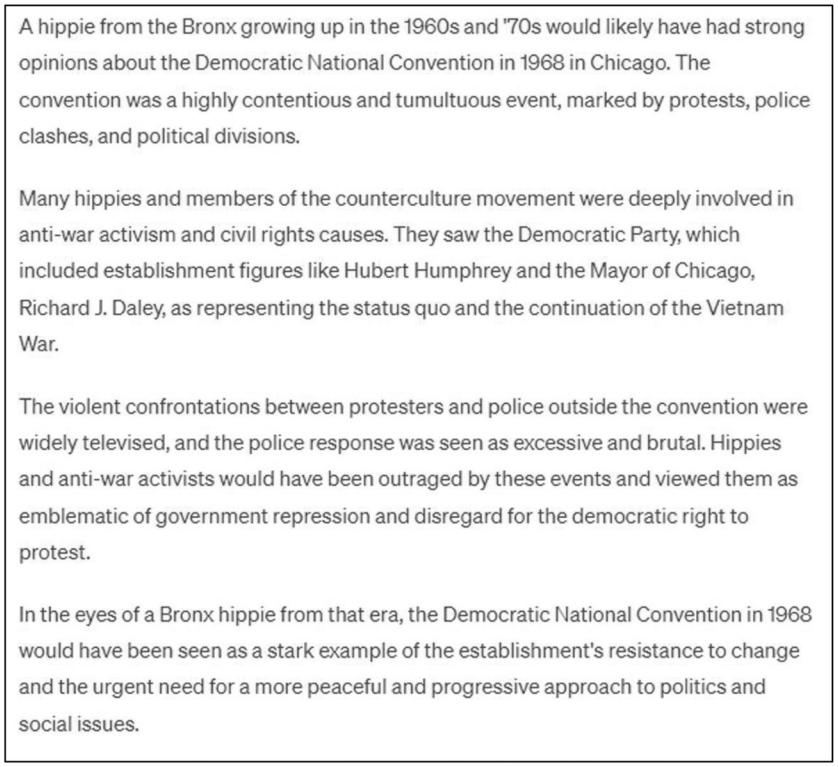
Fig. 2 Two-hundred-word summary of the first section of this editorial

looking for and that we need to be able to judge the information we find based on the knowledge that's in our head" (p. 130). So, as instructional designers, we must determine the learning strategies that will help students develop the knowledge they need, and the assessments for us to have confidence that it is there, regardless of whether the students do or do not have access to GenAI tools. So, what do we do?

Different Strategies

Instead of traditional learning strategies and assessments that can be completed by students with at least some degree of accuracy using GenAI, we must look to innovation in both the design of instruction and the assessment of learning. There are several strategies that educators can adopt to adapt to this new environment:

1. **Emphasize Process Over Product:** Shift the focus of assignments from the final product to the process of learning. This can include requiring students to submit drafts, outlines, or annotated bibliographies along with their final submission. By assessing the process, educators can better understand the student's learning journey. A snag here is that GenAI tools can also do this, as shown in the two figures.
2. **Incorporate Oral Assessments:** Oral exams or presentations can be effective in assessing a student's understanding. These assessments are more challenging to prepare for using AI and can give insights into the student's genuine grasp of the material. This, however, is virtually impossible in a lecture theater course like 'Psychology 101' with 600 students.
3. **Use AI Detection Tools:** There are emerging technologies designed to detect whether a piece of writing was generated by AI. Incorporating these tools may help educators identify work created by GenAI. However, the accuracy of these programs, both with respect to false negatives (i.e., GenAI was used but not detected) and false positives (i.e., GanAI was not used, but the student is accused of using it) is wanting. Educators who wish to incorporate AI monitoring tools should stay informed on the capabilities and limitations of these technologies in order to use them responsibly and effectively.
4. **Modify Assignment Design:** Tailor assignments to be more specific, personalized, or context-dependent. This can include prompts that relate directly to class discussions, current events, or unique scenarios that are less likely to be successfully addressed by AI.
5. **Encourage Critical Thinking and Analysis:** Assignments that require students to critique, analyze, or apply concepts in new contexts are less easily com-


A hippie from the Bronx growing up in the 1960s and '70s would likely have had strong opinions about the Democratic National Convention in 1968 in Chicago. The convention was a highly contentious and tumultuous event, marked by protests, police clashes, and political divisions.

Many hippies and members of the counterculture movement were deeply involved in anti-war activism and civil rights causes. They saw the Democratic Party, which included establishment figures like Hubert Humphrey and the Mayor of Chicago, Richard J. Daley, as representing the status quo and the continuation of the Vietnam War.

The violent confrontations between protesters and police outside the convention were widely televised, and the police response was seen as excessive and brutal. Hippies and anti-war activists would have been outraged by these events and viewed them as emblematic of government repression and disregard for the democratic right to protest.

In the eyes of a Bronx hippie from that era, the Democratic National Convention in 1968 would have been seen as a stark example of the establishment's resistance to change and the urgent need for a more peaceful and progressive approach to politics and social issues.

Fig. 3 ChatGPT response when prompted to reflect as indicated in item 12

pleted by AI. These tasks often require a depth of understanding that AI – at this moment—cannot mimic. ChatGPT 5.0, however, is expected to be able to do this.

6. **Focus on Application and Creativity:** Projects that require creative thinking, application of knowledge to new situations, or the solving of real-world problems can be more indicative of a student's own work and understanding. A recent article[3] in the Harvard Business Review, however, states "It [GenAI] can augment the creativity of employees and customers and help them generate and identify novel ideas".
7. **In-Class Assignments:** Conducting assessments during class time, either in person or via educator-monitored online platforms, can help ensure that the work submitted is the student's own. This also has the limitations mentioned earlier regarding class size.
8. **Peer Review and Collaborative Work:** Incorporating peer review processes and collaborative projects can encourage students to engage more deeply with the material and rely less on AI-generated content. A caveat here is that GenAI is capable of critiquing a text that is keyed in. For example, the prompt "Please critique the editorial by Hodges and Kirschner. What are its strong and weak points?" followed by copy and paste will return a 'peer assessment'.
9. **Develop Digital Literacy and Ethics Curriculum:** Educate students about the ethical use of AI, including discussions about academic integrity, the limitations of AI, and the importance of original work. The wall that this runs up against is human nature. In economics it's called the *homo economicus*; one who avoids unnecessary work by using rational judgment. In education this is called the *calculating learner*; a student who carries out the minimum of effort for the maximum benefit. Kirschner[4] calls this the *discipulus economicus*. In both of these, the key questions for the person are: What's in it for me? Do the costs of doing something weigh up against the benefits I receive if I do it? If the balance shifts to profits, we do it. If it shifts to costs, we don't.
10. **Personalized Learning Paths:** Customizing learning experiences and assignments for individual students can reduce the feasibility of using generic AI-generated content. This, again, is extremely difficult in large, theater lecture courses.
11. **Frequent, Low-Stakes Assessments:** Replace high-stakes testing with more frequent, low-stakes assess-

[3] https://hbr.org/2023/07/how-generative-ai-can-augment-human-creativity

[4] https://www.kirschnered.nl/2022/11/09/discipulus-economicus-the-calculating-learner/

ments that gauge student understanding over time. This, however, can be a problem if the low-stakes testing is online.

12. **Encourage Reflective Writing:** Assignments that ask students to reflect on their personal experiences or opinions are less amenable to AI generation. Here too, a 'smart' or 'savvy' student can work around this. With the prompt: 'Reflect on what the 1968 Democratic National Convention in Chicago would mean to a hippie from the Bronx in high school the 1960s and '70 s'. GenAI is able to answer in a way that the second author might, and the teacher has no way to know if this is the case (See Fig. 3).

By incorporating combinations of these strategies, educators can try to mitigate the challenges posed by AI and leverage the situation to foster deeper learning, critical thinking, and ethical understanding among students. However, as GenAI tools get 'smarter', the job gets tougher.

Conclusions

The GenAI tools will likely only get better (What does better mean? More powerful?) and more ubiquitous, so disabling them or banning them from schools and universities will neither be feasible nor productive. The focus should not be to try and design GenAI out of the learning experience, or necessarily to design it into the learning experience, but simply to design instruction so that students actually learn. The strategies suggested above, and others, may be productive paths to consider in this regard. One critical element that must be considered is whether university administrations will be able to support the changes in instructional design and assessment that are necessary to ensure integrity in the learning process, as some of the changes recommended will likely require instructors to spend considerably more time per student assessing students' work thus impacting workloads or personnel assignments.

Postscript: This editorial was written by humans without the help of GenAI tools (except for the two examples used in the figures).

References

Bogost, I. (2022, December 7) ChatGPT is dumber than you think: Treat it like a toy, not a tool. *The Atlantic*. https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/. Accessed 21 Dec 2023

Casal, J. E., & Kessler, M. (2023). Can linguists distinguish between ChatGPT/AI and human writing?: A study of research ethics and academic publishing. *Research Methods in Applied Linguistics, 2*(3). https://doi.org/10.1016/j.rmal.2023.100068

Edison, T. (1878). The phonograph and its future. *The North American Review*, *126*(*262*), 527–536. http://edison.rutgers.edu/yearofinno/Dec13/Edison_Phonograph%20and%20It%27s%20Future_North%20American%20Review_1878.pdf

Edison, T. (1888). The perfected phonograph. *The North American Review, 146*(379), 641–650.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*, 717–741. https://doi.org/10.1007/s10648-015-9348-9

Hodges, C.B., & Ocak, C. (2023, August 30). Integrating generative AI into higher education: Considerations. *EDUCAUSE Review*. https://er.educause.edu/articles/2023/8/integrating-generative-ai-into-higher-education-considerations. Accessed 21 Dec 2023

Kirschner, P. A. (2023, November 23). A critical note by @matbury about using ChatGPT for lesson plans and feedback [Tweet]. X. https://twitter.com/P_A_Kirschner/status/1727786231231041801

Luccioni, S. (2023). *AI is dangerous, but not for the reasons you think*. TED. https://www.ted.com/talks/sasha_luccioni_ai_is_dangerous_but_not_for_the_reasons_you_think?autoplay=true&muted=true. Accessed 21 Dec 2023

Neelen, M., & Kirschner, P. A. (2020). *Evidence informed learning design: Creating training to improve performance*. Kogan Page.

Norman, J. (n.d.). *Edison describes future uses for his phonograph.* History of Information. Accessed September 15, 2022, at https://www.historyofinformation.com/detail.php?entryid=617

Reeves, T. C., & Lin, L. (2020). The research we have is not the research we need. *Educational Technology Research and Development, 68*, 1991–2001. https://doi.org/10.1007/s11423-020-09811-3