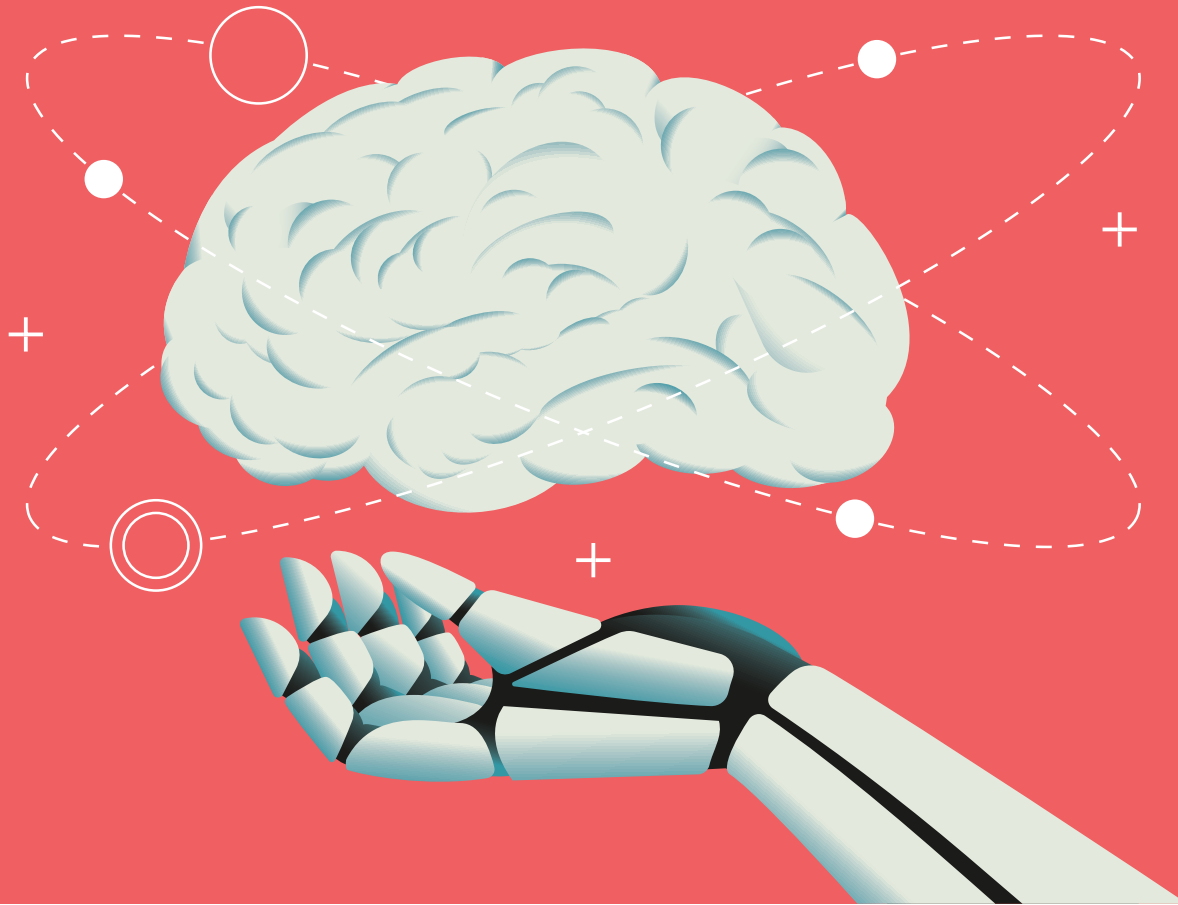


Leveraging AI to enhance learning



Engaging students in assessing and improving work generated by ChatGPT can promote higher-level creative and critical thinking that AI alone cannot achieve.

By Louis Volante, Christopher DeLuca, & Don A. Klinger

With 100 million active users in January 2023, just two months after its launch, ChatGPT is now considered the fastest-growing

consumer internet application. This popular artificial intelligence (AI) chatbot, developed by OpenAI, has brought a “tsunami effect” of changes to education (García-Peñalvo, 2023). Users as young as 13,

LOUIS VOLANTE (lvolante@brocku.ca; lvolante@maastrichtuniversity.nl; [@LouisVolante](https://twitter.com/LouisVolante)) is a professor at Brock University, St. Catherine’s, Ontario, Canada; a professorial fellow at the Maastricht Graduate School of Governance, United Nations University-MERIT, Maastricht, The Netherlands; and president of the Canadian Society for the Study of Education. **CHRISTOPHER DELUCA** (cdeluca@queensu.ca; [@ChrisDeLuca20](https://twitter.com/ChrisDeLuca20)) is a professor in the faculty of education and associate dean of the School of Graduate Studies and Postdoctoral Affairs at Queen’s University, Kingston, Ontario, Canada. **DON A. KLINGER** (dklinger@waikato.ac.nz) is pro vice-chancellor of the Division of Education at the University of Waikato, Hamilton, New Zealand.

with parent or guardian permission, can now use the application to generate complete essays in seconds. Unsurprisingly, other companies are developing similar AI language models, such as Google's Bart or Microsoft's Sydney, to rival ChatGPT.

The rapid emergence of AI applications leaves little doubt that the use of AI will continue to expand, presenting formidable challenges to educators, who must find ways to teach students to use such tools responsibly and effectively.

AI and academic integrity

Some have responded to the rise of AI chatbots by advocating that schools ban the tools altogether, citing cheating, plagiarism, and other forms of academic misconduct as serious issues. Indeed, the emergence of AI applications such as ChatGPT presents a new set of challenges to teachers, because the product ChatGPT generates is not a statement or report from a previously published piece of work that teachers can easily find using plagiarism detection software. Rather, the application constantly modifies its outputs to produce seemingly unique pieces of writing.

One study (Khalil & Err, 2023) found that popular detection tools were largely ineffective in identifying plagiarism in a sample of 50 essays generated by ChatGPT. AI detection tools can also generate false positives, meaning they identify text written by a human as AI-generated (Dalalah & Dalalah, 2023). Detection software companies, such as Turnitin, are promising to improve their products to garner more effective results (e.g., a 97% success rate for ChatGPT-authored content, Turnitin, 2023). But, even with these improvements, many students will uncover ways to make simple adjustments, such as inserting synonyms into AI-generated texts, to evade detection.

Recognizing the threat to academic integrity, many school districts across the U.S. have banned this AI tool (Rose, 2023). But how effective is this when students can use the tool on their own devices, without teachers knowing? Collectively, the education community must face the reality that it will be difficult, if not impossible, to identify all instances of AI-generated text with the current tools available. For this reason, instead of continuing to debate whether to allow these products, we must begin to consider how to integrate these applications in an ethically and educationally defensible manner. To do otherwise would be to ignore “the elephant in our schools” (Volante, DeLuca, & Klinger, 2023b). Our view is that AI can help spur long-overdue innovations in how students think about written work (Volante, DeLuca, & Klinger, 2023a).

AT A GLANCE



- **Artificial intelligence (AI) technologies, such as ChatGPT, present formidable challenges to educators concerned about cheating and plagiarism.**
- **AI also presents opportunities for reforming assessment practices to focus more on authentic tasks.**
- **Louis Volante, Christopher DeLuca, and Don A. Klinger suggest that teachers guide students to formatively assess and improve text generated by AI.**
- **They recommend a three-step pathway built on the ICE model: Students understand ideas in the text, make connections to their own contexts, and create extensions that demonstrate critical and creative thinking.**

AI in the classroom

The emergent literature on how to integrate AI into the teaching and learning process is markedly skewed toward college and university classrooms. This is understandable given that Open AI's (2023b) Terms of Use restrict use of the application to those 13 or older, or those who have a parent or guardian's permission.

At the higher education level, professors have had students use AI to complete assignments and then reflect on the experiment (Fyfe, 2022; Griffin, 2023). Overall, there appears to be a growing sentiment in higher education that AI will become ubiquitous in workplaces and appropriate ways to integrate these applications into higher education settings will be necessary.

Resources already exist to illustrate how teachers at the secondary level can use ChatGPT to improve their instruction (e.g., Herft, 2023). And it is likely that secondary students are turning to ChatGPT as a shortcut to complete assignments. A 2018 survey of 70,000 high school students indicated more than 60% of students plagiarized papers (Simmons, 2018). Despite school district bans, it is hard to imagine AI applications not contributing to this ongoing problem.

Recent surveys suggest that most teachers have not received guidance in the use of ChatGPT (Jimenez, 2023), even though its use has grown among teachers and students ages 12-17 (Kingston, 2023). There is an urgent need to offer pragmatic solutions to this pressing challenge so that teachers and students will use AI in ways that support learning.

A formative assessment pathway

Research suggests ChatGPT and other AI applications generate texts that look sophisticated at first glance but that are prone to factual errors and distortions (van Dis et al., 2023). This presents an opportunity for teachers to help students further develop research literacy skills. Specifically, we propose that teachers use the ICE model (Fostaty-Young & Wilson, 1995) as a pathway for formatively assessing and building on work generated by ChatGPT.

The ICE model shows how student learning develops, starting with understanding foundational *ideas*, to establishing *connections* between the ideas, to making novel and creative *extensions*. Figure 1 illustrates how ICE maps onto the six levels of cognition in Bloom's taxonomy.

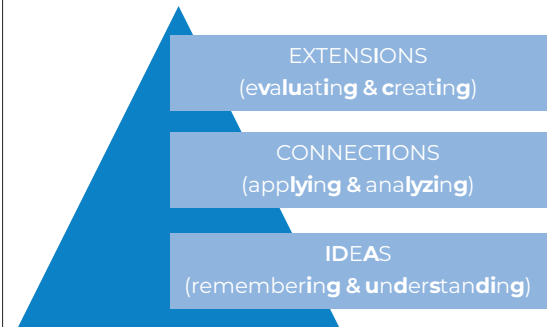
Teachers can use this framework to guide students through the steps of analyzing and improving AI-generated writing, using the same kinds of skills used to formatively assess student-generated work. As Sarah Beck and Sarah Levine (2023) explain, exercises like this enable students to practice steps of the writing process that are sometimes neglected in classrooms, where the focus tends to be on creating initial drafts. Engaging in this process can help students learn how to formatively assess their own and others' writing, using rubrics and peer feedback to improve the work at each step.

Step 1. Understanding ideas

The first step in using AI-generated text is for the teacher to engage students in fact-checking the key ideas in a sample output. Such fact-checking would involve the development of research skills where a student gathers information from multiple sources to confirm a "fact." This step ensures students understand and can work with the foundational ideas at play in the text and helps develop students' capabilities to understand facts and figures.

Self-assessment and peer assessment are widely recognized as important formative assessment strategies to help promote student learning (Double, McGrane, & Hopfenbeck, 2019), and these strategies could be incorporated into students' work with AI-generated text. After students complete an initial set of revisions to the writing sample ChatGPT generated, verifying the facts and ideas and correcting any errors, teachers could facilitate a peer-assessment feedback process using cooperative learning strategies, such as a jigsaw or a think-pair-share activity with an elbow partner. At this stage, the formative assessment focuses on ensuring the ideas are clear, while also serving as an additional opportunity for fact-checking.

Figure 1.
The ICE framework and Bloom's taxonomy



Note: Components of Bloom's taxonomy are in parentheses.

Step 2. Making connections

Two measures used to assess AI language models are perplexity and burstiness. To put it simply, perplexity relates to the complexity of the words in a sentence, while burstiness relates to the varying length of sentences. AI language models tend to generate sentences with low perplexity and burstiness, while humans tend to write in bursts with short and long sentences — similar to this article. For this reason, AI-generated content tends to be more uniform, and less interesting, than human-generated content. This is why AI-generated content is often described as "shallow."

Teachers could incorporate self- and peer-review exercises in which students generate their own perplexity and burstiness scores for AI-generated work and then revise the work with a goal of raising those scores. Using co-constructed rubrics and exemplars (Bacchus et al., 2019), students can learn to connect ideas within and between paragraphs in ways that make the work livelier and more engaging and less like text generated by a machine.

This step also invites students to connect ideas in the text to their personal life and local context, something no AI application can do. As students learn to actively "make a connection between the ideas and [their] own experience," they build their ability to analyze ideas and apply them to new situations (Fostaty-Young & Toop, 2022), which aligns with the connections level in the ICE model.

Step 3. Creating extensions

The final step in our pathway requires students to undertake a final round of revisions that extend the ideas in the text in ways that demonstrate critical, creative, and higher-order thinking. This might

involve evaluating the limits of the presented argument, suggesting alternatives, and presenting a novel way forward, one not suggested by ChatGPT. This step separates human work from AI-generated content, and it is where secondary teachers should increasingly focus their instruction. In many respects, AI makes the need for authentic assessment more evident than ever and can therefore push us to make education more human, not less (Cope, Kalantzis, & Sears, 2020).

Inviting students to relate extensions to their personal contexts (i.e., their classroom or community) is one way to encourage deep thinking. For example, having students outline actions they plan to take to support specific sustainable development goals would require independent thinking. Another extension method is to invite students to represent their learning through alternative formats or more authentic tasks, such as an oral presentation of the assignment, an artistic representation, or a community-based project. These assessment experiences naturally promote deeper learning by involving multiple modalities and requiring connections between ideas.

The extension requirement and assessment criteria should be available from the outset, so students know that generating and refining AI content is an insufficient demonstration of learning.

Turning a negative into a positive

The threat posed by AI language models has provoked strong reactions from academics, teachers, and even well-recognized public figures. For example, in a *New York Times* op-ed, Noam Chomsky (2023) wrote, “given the amorality, faux science, and linguistic incompetence of these systems, we can only laugh or cry at their popularity.” Certainly, there are legitimate reasons to be concerned that AI language models will impoverish our schools and society. But teachers always have found ways to use technological advances, good and bad, to help their students learn.

If students are going to use AI language models, and they undoubtedly will (and are), then we need to point out the shortcomings of these applications and leverage them, where possible, to promote deeper learning. A just-say-no approach will not suffice. We need to confront this new elephant in our schools and turn it into a teaching and learning opportunity.

AI tools have passed national high school examinations, with mean grades similar to those of human students (de Winter, 2023). Simulations using ChatGPT have obtained passing grades in college and university law and science courses (Choi

Instead of continuing to debate whether to allow these products, we must begin to consider how to integrate these applications in an ethically and educationally defensible manner.

et al., 2023; Gilson et al., 2023). And there’s reason to believe their abilities will improve. While we were writing this article in spring 2023, OpenAI unveiled ChatGPT-4. In a series of simulated tests, ChatGPT-4 demonstrated a substantial increase in power and accuracy over previous iterations. For example, it performed at the 90th percentile on a simulated Uniform Bar Exam, compared to the 10th percentile for ChatGPT-3.5, launched four months earlier in November 2022. ChatGPT-4 scored a 5 on nine out of 15 simulated Advanced Placement (AP) examinations, while ChatGPT-3.5 obtained a 5 on only three AP exams (OpenAI, 2023a). And the improvements went beyond test scores. Samantha Murphy Kelly reported for CNN (2023) that ChatGPT-4 could use visual information to produce a functioning website or view the contents of a refrigerator to plan a meal. It even demonstrated its ability to provide coding for game and app development (Marr, 2023).

As much as we may wish otherwise, the “incompetence” of these models is somewhat overstated.



“Nothing says ‘Teacher Appreciation’ like an administrator winning the raffle.”

LEVERAGING AI

If students are going to use AI language models, and they undoubtedly will (and are), then we need to point out the shortcomings of these applications and to leverage them, where possible, to promote deeper learning.

Their ability to do well on standardized assessments underscores an urgent need to significantly reform our assessment systems so that they are “AI proof.” The easiest way to do this is to emphasize qualities that AI language models are notoriously poor at generating and that make us uniquely human. In writing, these qualities include humor, sarcasm, subtle nuances, and connections to personal and local contexts. Humans can extend the work of AI in authentic ways that demonstrate critical, creative, and higher-order thinking. AI text is procedurally generated, word by word, based on weighted options in a sample. But humans, like their fingerprints, are unique. Their writing would never follow a strict

logic model. AI would be hard pressed to generate the perplexity or burstiness — let alone the connections to previous topics — of the last few sentences.

OpenAI continues to caution that ChatGPT still has challenges with certain tasks. Nevertheless, the expansive growth in just over four months demonstrates the current power and future potential of AI language models. With the addition of AI to applications like Microsoft Word and Outlook, language models will undoubtedly become even more mainstream (Hadero, 2023). As educators, we need to find ways to use these resources to support teaching and learning, and the formative assessment pathway provides an ideal doorway to explore what these tools can do, and what humans can do better. ■

Note: The authors of this article recommend teachers consult relevant terms of use and their district and school leadership personnel prior to using any AI applications within their classrooms.

PDK Members

Discuss this article with your colleagues using the discussion guide at https://members.pdkintl.org/PDK_Member_Discussion_Questions



References

- Bacchus, R., Colvin, E., Knight, E.B., & Ritter, L. (2018). When rubrics aren't enough: Exploring exemplars and student rubric co-construction. *Journal of Curriculum and Pedagogy*, 17, 48-61.
- Beck, S.W. & Levine, S.R. (2023). ChatGPT: A powerful technology tool for writing instruction. *Phi Delta Kappan*, 105 (1), 66.
- Choi, J.H., Hickman, K.E., Monahan, A., & Schwarcz, D.B. (2023). ChatGPT goes to law school. *Journal of Legal Education*.
- Chomsky, N. (2023, March 8). Noam Chomsky: The false promise of ChatGPT. *The New York Times*.
- Cope, B., Kalantzis, M., & Searsmith, D. (2020). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Education Philosophy and Theory*, 53, 1229-1245.
- Dalalah, D. & Dalalah, O.M.A. (2023). The false positives and false negatives of generative AI detection tools in education and academic research. *The International Journal of Management Education*, 21 (2).
- de Winter, J. (2023). Can ChatGPT pass high school exams on English language comprehension? *ResearchGate*.
- Double, K.S., McGrane, J.A., & Hopfenbeck, T. (2019). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review*, 32, 481-509.
- Fostaty-Young, S. & Toop, M. (2022). *Teaching, learning, and assessment across the disciplines: ICE stories*. Open Library.
- Fyfe, P. (2022). How to cheat on your final paper: Assigning AI for student writing. *AI & Society*.
- García-Peñalvo, F.J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: disruption or panic? *Ediciones Universidad de Salamanca*.
- Gilson, A., Safranek, C.W., Huang, T., Socrates, V., Chi, L., Taylor, R.A., & Chartash, D. (2023). How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*.
- Griffin, T.L. (2023, February 14). Why using AI tools like ChatGPT in my MBA innovation course is expected and not cheating. *The Conversation*.
- Hadero, H. (2023, March 16). Microsoft adds AI tools to office apps like Outlook, Word. *The Washington Post*.
- Herft, A. (2023). A teacher's prompt guide to ChatGPT aligned with "what works best." Department of Education, New South Wales, Sydney.
- Jimenez, K. (2023, March 1). ChatGPT in the classroom: Here's what teachers and students are saying. *USA Today*.
- Kelly, S.M. (2023, March 16). 5 jaw-dropping things GPT-4 can do that ChatGPT couldn't. *CNN Business*.
- Khalil, M. & Er, K. (2023). Will ChatGPT get you caught? Rethinking of plagiarism detection. *arXiv*.
- Kingston, J.A. (2023, March 7). Students and teachers warm up to ChatGPT. *Axios*.
- Marr, B. (2023, February 24). GPT-4 is coming: What we know so far. *Forbes*.
- OpenAI. (2023b). GPT-4. <https://openai.com/research/gpt-4>
- OpenAI. (2023a). Terms of use. <https://openai.com/policies/terms-of-use>
- Rose, K. (2023, February 3). How ChatGPT kicked off an A.I. arms race. *The New York Times*.
- Simmons, A. (2018, April 27). Why students cheat – and what to do about it. *Edutopia*.
- Turnitin. (2023, February 13). Turnitin announces AI writing detector and AI writing resource centre for educators [Press release].
- van Dis, E.A.M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C.L. (2023). ChatGPT: Five priorities for research. *Nature*, 614, 224-225.
- Volante, L., DeLuca, C., & Klinger, D.A. (2023a, May 15). Forward-thinking assessment in the era of artificial intelligence: Strategies to facilitate deep learning. *Education Canada*.
- Volante, L., DeLuca, C., & Klinger, D.A. (2023b). How can teachers integrate AI within schools? Five steps to follow. *Education Canada: The Facts on Education*.

Copyright of Phi Delta Kappan is the property of Sage Publications Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.