



INSIGHT 14

AI IN ASSESSMENTS AND EVALUATIONS

AI is disrupting established assessment and evaluation processes, such as job screenings, grant application evaluations, peer review, and grading. Screening processes could see a significant increase in applicants, and new forms of evaluation could emerge that focus less on written work as a measure of competence.

TODAY

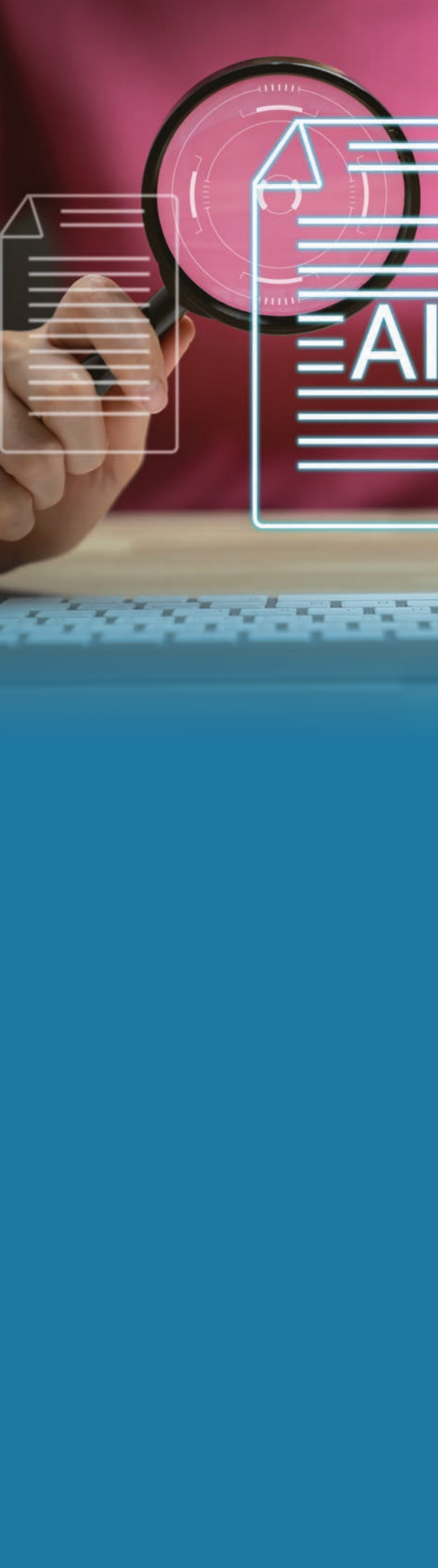
AI-generated applications are flooding institutions, who are, in turn, turning to AI to conduct assessments and evaluations. More than half of UK undergraduates reported using AI to help with essays.¹ Likewise, nearly half of all job seekers today are using AI tools to improve their resumes.² A survey of scientists found that nearly 30% had used generative AI to write their scientific papers.³ One in six researchers reported using generative AI to help write their grant applications.⁴ As employers and granting agencies are being inundated with high volumes of AI-generated applications, it has become increasingly common for them to use AI tools to screen, recruit and manage employees and potential grantees, despite concerns raised by workers, unions, and employee rights groups.^{5, 6, 7, 8, 9, 10}





The widespread use of generative AI to write applications is disrupting assessment and evaluation processes across various domains. The use of AI by applicants is straining existing evaluation processes rooted in the assumption that written communications are an accurate representation of an individual's competence. AI-generated writing in student work is difficult to detect, leading to questions about how to grade and assess student learning. Academic publishers have expressed concerns about how AI-written articles submitted for peer review are undermining scientific integrity.¹¹ Employers are questioning the extent to which a job application can continue to be used as an indicator of an applicant's competence or fit for a given role. Across a variety of domains, many of the modern methods and tools used to evaluate people and their abilities, such as take-home writing tests, may no longer be useful.

Institutions are divided on how to handle this disruption. Some organizations, such as the journal Science, have banned the submission of AI-generated content, as well as the use of AI to evaluate submissions.^{12, 13} Other organizations have dismissed bans as impractical and ultimately unenforceable, in large part because current AI-detection technology is not effective enough to be useful.^{14, 15, 16} While granting bodies have issued warnings about the use of AI in grant applications, researchers are seeing the benefits in using AI to assist in writing proposals.^{17, 18} Some funders are asking if AI could help address inequities related to the noted "snowball effect," where grant winners tend to have an advantage in winning future funding.¹⁹ Universities have largely left the problem of what student assessment should look like in the AI era to individual professors. The lack of a coordinated response has prompted some professors to call for a one-year pause in student assessments to determine a viable path forward.²⁰ Some universities have moved back to old-fashioned testing methods – where supervised written exams constitute a majority of the grade. This approach is being questioned by professors who note that those methods were abandoned because they failed to evaluate skills important to modern social and work contexts, such as collaboration, teamwork and communication.²¹



FUTURES

Evaluations may evolve to address the problems posed by people passing off AI-generated content as their own. The importance of written work as a valid object of assessment could diminish, relative to other factors such as in-person character assessments, group work, supervised technical exams, or professional references. Future application processes may become more rigorous, holistic, in-person, or in real-time. Assessments could also begin to test a person's ability to use AI effectively and appropriately to support the role in question.

The use of AI for screening, sorting and decision-making may also increase in response to high volumes of job and grant applications. Human decision-makers could play less of a role in assessment, grading, and funding decisions, or play a more specific role at certain points in the evaluation process.²² AI could be used strategically at certain points in assessment or evaluation processes to mitigate human bias, for example by quickly surfacing job applicants who possess relevant skills, but may lack formal certifications.^{23, 24} These kinds of uses could also perpetuate or exacerbate existing human biases if they are programmed into AI systems, either inadvertently or by design.

Questions may continue to be raised about AI's purported objectivity and neutrality in screening processes. Increased calls for monitoring and transparency coupled with new forms of oversight for AI-mediated decisions could become widespread, if AI begins to serve a more central role as a gatekeeper in screening processes.²⁵

If AI truly breaks assessment, it could prompt a search for a new path forward. While ranking and ordering processes are central to the functioning of modern institutions, they also tend to perpetuate harmful forms of exclusion.²⁶ While the past decades have seen efforts to adjust assessments to account for human bias – blind assessments, affirmative action, or the turn toward diversity and inclusion, for example – discrimination remains a real problem in hiring. In Canada, resumes with English-sounding names are still 35 percent more likely to receive callbacks than those with Indian or Chinese names.²⁷ If more candidates using AI makes established forms of assessment untenable, new ways of evaluating an individual's present competence or potential for future success may begin to emerge.



IMPLICATIONS

- Applicants could use AI to **create manufactured evidence of competence or expertise** such as false websites, academic articles, or testimonials.
- Large volumes of applications completed with the assistance of generative AI could **undermine employment screening processes** that rely heavily upon standardized questions that generative models are good at answering.
- Overwhelmed with large volumes of applications that are hard to differentiate, assessors may turn more to **transferable skills, tests, in-depth character assessments, in-person interviews, strong personal references and personal networks to identify top applicants.**
 - › This could exacerbate already existing forms of **nepotism and homogeneity** across workplaces and funding streams. This could continue to disadvantage some individuals, such as those that are neurodiverse or who don't have strong social or professional networks.
- Those who are skilled at using AI to generate work while having a knowledge base strong enough to identify errors **will be most competitive.**
 - › **Those who rely heavily on AI could be disadvantaged** by changes to evaluation or assessment processes, such as effective AI detection tools or requirements that written content is produced by individuals.²⁸
- › If more applicants use AI and produce more generic-sounding applications that resemble one another, it could become challenging to assess them.
- AI screening tools could be specifically programed to surface more diverse applicants or characteristics of previously overlooked talent, contributing to more **inclusive and diverse academic institutions and workplaces.**
 - › This may present new opportunities for people with disabilities, neurodivergent people, and English-as-a-second-language speakers.²⁹
- Increased adoption of AI systems to evaluate students or assess applicants could create **legal and ethical challenges** that put increased pressure on institutions to demonstrate how AI systems are fair or and transparent.
- Without the application of a consistent approach to the use of generative AI by students, it may become difficult or **impossible to compare assessments or grades between institutions or across jurisdictions**, challenging university admission processes, amongst other things.
 - › If university-level grades cease to be an accurate representation of student ability, the true value of a degree may be challenging to determine.

Endnotes

- 1 Adams, Richard. "More than Half of UK Undergraduates Say They Use AI to Help with Essays." *The Guardian*, February 1, 2024, sec. Technology. <https://www.theguardian.com/technology/2024/feb/01/more-than-half-uk-undergraduates-ai-essays-artificial-intelligence>.
- 2 Westfall, Chris. "Study Says Hiring Managers Expect (and Prefer) AI-Enhanced Resumes." *Forbes*. Accessed August 29, 2024. <https://www.forbes.com/sites/chriswestfall/2024/01/26/study-says-hiring-managers-expect-and-prefer-ai-enhanced-resumes/>.
- 3 Singh Chawla, Dalmeet. "Is ChatGPT Corrupting Peer Review? Telltale Words Hint at AI Use." *Nature* 628, no. 8008 (April 10, 2024): 483–84. <https://doi.org/10.1038/d41586-024-01051-2>.
- 4 Moller-Neilsen, Thomas. "European Research Council Issues Warning on AI's Use in Grant Applications." *ScienceBusiness*, December 19, 2023. <https://sciencebusiness.net/news/ai/european-research-council-issues-warning-ais-use-grant-applications>.
- 5 Sheen, Robert. "What The Coming AI Transparency Wave Means For HR." *Forbes*, July 3, 2024. <https://www.forbes.com/councils/forbeshumanresourcescouncil/2024/07/03/what-the-coming-ai-transparency-wave-means-for-hr/>.
- 6 Mearian, Lucas. "Legislation to Rein in AI's Use in Hiring Grows." *Computerworld*. Accessed October 1, 2024. <https://www.computerworld.com/article/1621393/legislation-to-rein-in-ais-use-in-hiring-grows.html>.
- 7 Adarlo, Sharon. "Recruiters Are Getting Bombarded With Crappy, AI-Generated CVs." *Futurism*, August 16, 2024. <https://futurism.com/the-byte/recruiters-ai-generated-job-cvs>.
- 8 Hager, Leah. "AI and the Future of Grant Writing." *Medium*, May 7, 2024. <https://medium.com/@lhager0505/ai-and-the-future-of-grant-writing-dbf083b5a7f7>.
- 9 Gelles-Watnick, Lee Rainie, Monica Anderson, Colleen McClain, Emily A. Vogels and Risa. "AI in Hiring and Evaluating Workers: What Americans Think." *Pew Research Center* (blog), April 20, 2023. <https://www.pewresearch.org/internet/2023/04/20/ai-in-hiring-and-evaluating-workers-what-americans-think/>.
- 10 Tangermann, Victor. "Unemployed Man Uses AI to Apply for 5,000 Jobs, Gets 20 Interviews." *Futurism*, November 7, 2023. <https://futurism.com/the-byte/unemployed-man-ai-apply-5000-jobs-gets-20-interviews>.
- 11 Coldewey, Kyle Wiggers, Devin. "This Week in AI: Generative AI Is Spamming up Academic Journals." *TechCrunch* (blog), June 19, 2024. <https://techcrunch.com/2024/06/19/this-week-in-ai-generative-ai-is-spamming-up-academic-journals/>.
- 12 Thorp, H. Holden. "ChatGPT Is Fun, but Not an Author." *Science (New York, N.Y.)* 379, no. 6630 (January 27, 2023): 313. <https://doi.org/10.1126/science.adg7879>.
- 13 Soeteman, Krijn. "Use of Generative AI Still Not Allowed by Dutch National Funders." *Research Professional News* (blog), May 10, 2024. <https://www.researchprofessionalnews.com/rr-news-europe-netherlands-2024-5-use-of-generative-ai-still-not-allowed-by-dutch-national-funders/>.
- 14 *The Economist*. "At Least 10% of Research May Already Be Co-Authored by AI." Accessed October 16, 2024. <https://www.economist.com/science-and-technology/2024/06/26/at-least-10-of-research-may-already-be-co-authored-by-ai>.
- 15 Science and Innovation Canada. "Advice from the Ad Hoc Generative AI Panel of External Experts." *Innovation, Science and Economic Development Canada*, January 12, 2024. <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/use-generative-artificial-intelligence-development-and-review-research-proposals/advice-ad-hoc-generative-ai-panel-external-experts>.
- 16 Cook, Jodie. "AI Content Detectors Don't Work (The Biggest Mistakes They Have Made)." *Forbes*, July 4, 2024. <https://www.forbes.com/sites/jodiecook/2024/07/04/ai-content-detectors-dont-work-the-biggest-mistakes-they-have-made/>.

- 17 Moller-Neilsen, Thomas. "European Research Council Issues Warning on AI's Use in Grant Applications." SciencelBusiness, December 19, 2023. <https://sciencebusiness.net/news/ai/european-research-council-issues-warning-ais-use-grant-applications>.
- 18 Seckel, Elizabeth, Brandi Y. Stephens, and Fatima Rodriguez. "Ten Simple Rules to Leverage Large Language Models for Getting Grants." *PLOS Computational Biology* 20, no. 3 (March 1, 2024): e1011863. <https://doi.org/10.1371/journal.pcbi.1011863>.
- 19 Baker, Simon. "AI Might Help Science Break out of Narrow Funding Focus." Nature Index, March 5, 2024. <https://www.nature.com/nature-index/news/ai-might-help-science-break-out-narrow-funding-focus>.
- 20 Janos, Zach Justus, Nik. "Assessment of Student Learning Is Broken." Inside Higher Ed. Accessed October 1, 2024. <https://www.insidehighered.com/opinion/views/2024/03/28/assessment-student-learning-broken-opinion>.
- 21 Sims, Alex. "ChatGPT and the Future of University Assessment." Times Higher Education, March 9, 2023. <https://www.timeshighereducation.com/campus/chatgpt-and-future-university-assessment>.
- 22 Naujokaitytė, Goda. "Commission Gears up to Confront the Risks Generative Artificial Intelligence Poses to Science." SciencelBusiness. Accessed October 16, 2024. <https://sciencebusiness.net/news/ai/commission-gears-confront-risks-generative-artificial-intelligence-poses-science>.
- 23 Fergus, J. "A Bookshelf in Your Job Screening Video Makes You More Hirable to AI." Input, February 18, 2021. <https://www.inverse.com/culture/a-bookshelf-in-your-job-screening-video-makes-you-more-hirable-to-ai>.
- 24 Reed, Jon. "Can Generative AI Make HR More Productive - and Less Biased? A Closer Look at Oracle's HCM AI Announcements." Diginomica, June 30, 2024. <https://diginomica.com/can-generative-ai-make-hr-more-productive-and-less-biased-closer-look-oracles-hcm-ai-announcements>.
- 25 Wahid, Ishraat Saira. "From Recruitment to Performance: AI and ML Transforming HR Operations | The Business Standard." Accessed October 16, 2024. <https://www.tbsnews.net/thoughts/recruitment-performance-ai-and-ml-transforming-hr-operations-915836>.
- 26 Foucault, Michel. "Discipline and Punish." A. Sheridan, Tr., Paris, FR, Gallimard, 1975. https://www.joycerain.com/uploads/2/3/2/0/23207256/foucault_discipline_punish.pdf.
- 27 Oreopoulos, Philip, and Diane Dechief. "Why Do Some Employers Prefer to Interview Matthew, but Not Samir? New Evidence from Toronto, Montreal, and Vancouver." SSRN Scholarly Paper. Rochester, NY, February 1, 2012. <https://doi.org/10.2139/ssrn.2018047>.
- 28 Ortiz, Sabrina. "Using ChatGPT for Job Applications Could Backfire on You, According to This Study." ZDNET. Accessed October 16, 2024. <https://www.zdnet.com/article/using-chatgpt-for-job-applications-could-backfire-on-you-according-to-this-study/>.
- 29 Lindzon, Jared. "Recruiters Advise Candidates to Use ChatGPT When Applying for Jobs, but with Some Caveats." The Globe and Mail, February 13, 2023. <https://www.theglobeandmail.com/business/careers/article-recruiters-advise-candidates-to-use-chatgpt-when-applying-for-jobs-but/>.

© His Majesty the King in Right of Canada, 2025

For information regarding reproduction rights: <https://horizons.gc.ca/en/contact-us/>

PDF: PH4-224/2025E-PDF

ISBN: 978-0-660-76921-9

Aussi disponible en français sous le titre : L'IA et les évaluations.

DISCLAIMER

Policy Horizons Canada (Policy Horizons) is the Government of Canada's centre of excellence in foresight. Our mandate is to empower the Government of Canada with a future-oriented mindset and outlook to strengthen decision making. The content of this document does not necessarily represent the views of the Government of Canada, or participating departments and agencies.