

# Perils and promise of online exams

Madeleine Schultz<sup>1</sup> and Damien L. Callahan<sup>1</sup>

As academic staff scrambled into emergency remote teaching during COVID-19 restrictions, we also had to move in-person exams online without compromising integrity. This disruption caused us to think carefully about how chemistry is assessed, because ‘business as usual’ was no longer possible — and at some institutions, there are no plans for in-person exams to return.

For many academic staff, a knee-jerk response to being forced to run assessments online was to question whether academic integrity could be maintained. It is known that learning-oriented environments result in less cheating than do goal- or grade-oriented environments, so a starting point is to build relationships with students that encourage them to value learning above exam performance.

At the same time, authenticating assessment is vital to maintaining the value of the qualification. For any task submitted online, the question “Who did the work?” arises. We have seen three distinct sources of information improperly used by students: collusion, facilitated by easy sharing of information online; search engines, which permit students to look up answers that they were expected to memorize or work out; and contract cheating, which allows students to pay for and obtain answers within a very short time frame.

A common approach to attempt to reduce collusion is to use pools of questions, whereby each student receives a random exam consisting of one question from each pool. We found that questions that appeared very similar to us were sometimes answered quite differently by students, meaning that students might inadvertently be given exams of different difficulties. It was shocking to find that when some versions of a question were answered much worse by students, those were the versions to which Google gave a wrong answer, which had been input by some students. This showed us that up to a third of our students used Google searches during exams. The obvious response that we should just write better questions is extremely difficult to enact for foundation chemistry, for which the bulk of the content is readily available online.

Contract cheating has emerged as a major challenge for higher education, and has expanded dramatically during COVID-19<sup>1</sup>. It is quick, cheap and easy for students to obtain written answers to any question, no matter how high level or how much reasoning is required. Students are most likely to make use of these services in time-limited, high-stakes assessment tasks such as exams<sup>2</sup>. Entering our own questions into Google was a disheartening experience, because we recognized some of our own wording, diagrams and fonts on contract-cheating sites.

Remote proctoring has been adopted by many institutions in an attempt to prevent students from cheating during online exams. This involves monitoring student keystrokes, cameras and/or microphones to detect behaviours that might indicate cheating, such as looking away from the screen, talking to others or changing between browser tabs. However, such surveillance invades student privacy and is likely to be particularly disadvantageous to students from marginalized groups<sup>3</sup>, and we decided not to use any of these services. Rather than viewing students as adversaries, who must be outmaneuvered to prevent them from cheating, we are attempting to address academic integrity by redesigning our assessment tasks.

In our high-enrolment units (1,000 students per semester), staff time and funding constraints mean that significant portions of exams must be automatically marked, either through the use of multiple choice or ‘fill in the blank’-style questions. Having students enter a single word or number themselves, which can be marked automatically, is our preferred approach because it avoids the cueing problems associated with multiple-choice

questions, whereby students can be prompted towards the correct answer based on the options provided. We have used ‘fill in the blank’-style questions in weekly online quizzes for many years, so our students had already encountered this question style prior to their first online exam in 2020. Typographical errors or unexpected punctuation are automarked incorrect in this format; consequently, we have learned to be very explicit when structuring the questions.

As part of the redevelopment of exams for online delivery, we reduced memorization requirements — a change that we plan to retain even if in-person exams return. Students are now encouraged to create their own notes to refer to during their exam, because the preparation of those notes may have a positive learning effect, and knowing that they will be available can reduce stress<sup>4</sup>. Well-designed, open exams can improve learning outcomes because students employ approaches focused on understanding the meaning of the content in preparing for assessment when they know that their higher-order skills will be tested<sup>5</sup>.

Regarding the online interface used to deliver the exam, we explored multiple options to allow students to input drawings such as chemical structures into an automarked platform, but we encountered technical difficulties with the integration of the answers into the learning management system. The easiest solution is for students to hand-draw structures and write out multi-step responses, then upload their answers for manual marking, which allows the awarding of partial marks. This approach enables us to ask higher-order questions, which are harder to find answers to online, for part of the exam. We provide a template for handwritten responses to clarify expectations and make marking easier, and we offer students opportunities to practise using both submission formats (short answer and handwritten).

Importantly, after each exam, we now review the difficulty of all versions of the questions within question pools based on student scores. We also input the questions into a search engine to see if there is a characteristic wrong answer that can be identified in student responses. In addition, we check whether any questions have been posted to contract-cheating sites. We follow question-writing guidelines<sup>6</sup>, peer review all questions within the teaching team, and



Thus, assessing chemical reasoning rather than the manipulation of formulae and memorizing facts can have a double benefit of making assessment more authentic and supporting diverse groups of students to achieve. There are no easy solutions, but we urge educators to think carefully about what we want our students to know and be able to do, and how we are teaching and testing them.

Madeleine Schultz  and Damien L. Callahan   
School of Life and Environmental Sciences,  
Deakin University, Victoria, Australia.  
✉e-mail: [madeleine.schultz@deakin.edu.au](mailto:madeleine.schultz@deakin.edu.au)

<https://doi.org/10.1038/s41570-022-00385-7>

1. Lancaster, T. & Cotarlan, C. Contract cheating by STEM students through a file sharing website: a COVID-19 pandemic perspective. *Int. J. Ed. Integ.* **17**, 3 (2021).
2. Bretag, T. et al. Contract cheating and assessment design: exploring the relationship. *Assess. Eval. High. Ed.* **44**, 676–691 (2019).
3. Silverman, S. et al. What happens when you close the door on remote proctoring? Moving toward authentic assessments with a people-centered approach. *To Improve the Academy* **39**, 115–131 (2021).
4. Piontkivska, H., Gassensmith, J. J. & Gallardo-Williams, M. T. Expanding inclusivity with learner-generated study aids in three different science courses. *J. Chem. Educ.* **98**, 3379–3383 (2021).
5. Scouler, K. The influence of assessment method on students' learning approaches: multiple choice question examination versus assignment essay. *High. Educ.* **35**, 453–472 (1998).
6. Haladyna, T. M., Downing, S. M. & Rodriguez, M. C. A review of multiple-choice item-writing guidelines for classroom assessment. *App. Meas. Ed.* **15**, 309–333 (2002).
7. McCabe, D. L., Trevino, L. K. & Butterfield, K. D. Cheating in academic institutions: a decade of research. *Ethics Behav.* **11**, 219–232 (2001).
8. Van Dusen, B., Nissen, J., Talbot, R. M., Huvard, H. & Shultz, M. A QuantCrit investigation of society's educational debts due to racism and sexism in chemistry student learning. *J. Chem. Educ.* **99**, 25–34 (2022).
9. Ralph, V. R., Scharlott, L. J., Schwarz, C. E., Becker, N. M. & Stowe, R. L. Beyond instructional practices: characterizing learning environments that support students in explaining chemical phenomena. *J. Res. Sc. Teach.* <https://doi.org/10.1002/tea.21746> (2022).

#### Acknowledgements

The authors thank M. Gallardo-Williams, V. Rosa Ralph, R. Sorensen-Unruh and all participants in the #CER100 reading challenge for helpful discussions.

#### Competing interests

The authors declare no competing interests.

make iterative improvements to our exams by analysing student outcomes.

A teaching style of 'this is how you pass the exam' contributes directly to a goal-oriented environment and thus an increased likelihood of cheating. Instead, we focus our teaching and assessment on critical thinking, and encourage students to take ownership of their learning. A further predictor of cheating is if students think their peers are getting away with it<sup>7</sup>,

“ Rather than viewing students as adversaries, who must be outmaneuvered to prevent them from cheating, we are attempting to address academic integrity by redesigning our assessment tasks ”

therefore we are explicit about our strict response to detected breaches of academic integrity, so that our students feel confident that others are not benefitting from cheating.

The problems with fairly assessing large classes are not new or unique to the online environment, but online assessment has exacerbated some issues and drawn attention to the weaknesses of traditional exams. The shift to online examination serves as an opportunity to rethink assessment and curriculum more broadly. We want to emphasize that we disagree with deficit framing, whereby students are blamed for academic difficulties resulting from racist and sexist power structures, and we believe that the injustice of current academic systems<sup>8</sup> is at the root of student behaviours. Recent findings show that teaching focused on reasoning can reduce gaps in performance between marginalized groups in STEM compared with their privileged peers<sup>9</sup>.