Hacking assessment

Phillip Dawson
Deakin University, Victoria

Associate Professor Phillip Dawson is Associate Director of the Centre for Research in Assessment and Digital Learning, Deakin University. His recent paper 'Five ways to hack and cheat with bring-your-own-device electronic examinations' *(British Journal of Educational Technology*, 2015) is amongst the first published research on assessment hacking. Phill's most recent completed project explored how university teachers make decisions when designing assessment. Phill has a decade of teaching experience in higher education, for which he has received multiple vice-chancellors' awards and a citation from the Australian Learning and Teaching Council.

Abstract

Hackers exploit weaknesses in a system to achieve their own goals. In this paper I argue that hacking presents a significant threat to the growing world of online assessment. This threat needs to be addressed through a variety of means; technological anti-hacking approaches will not be sufficient. The most effective ways to prevent hacking may be changes to the assessment tasks themselves to make hacking less tempting; these approaches also have a range of positive side effects in terms of authenticity, transparency of criteria, and ensuring tasks involve work beyond the exam. I conclude with a brief exploration of the ways that teachers may also hack assessment systems.

The promise of online assessment

Vast bodies of research indicate that when used appropriately, educational technology can improve learning outcomes for students (Means, Toyama, Murphy, Bakia & Jones, 2010; Tamim, Bernard, Borokhovski, Abrami & Schmid, 2011). Benefits from educational technology are greatest when we adapt curriculum, instruction and assessment to take advantage of the affordances of technology.

Assessment can be adapted to use technology in a variety of ways. Student learning and performance can be improved through automatic feedback on an exam, or allowing typing instead of writing (Butler & Roediger, 2008; Charman, 2014; Mogey, Cowan, Paterson & Purcell, 2012; Mogey & Hartley, 2013). Student judgement can be improved through formative self- or peer-assessment procedures, which are made more efficient thorough online systems (Li et al., 2015). Examinations can be made more authentic by incorporating rich computer-based tasks (Hillier & Fluck, 2013). Technology even enables a vast array of new assessment types, ranging from social media tasks to high-fidelity simulations.

Threats to online assessment

In addition to providing additional affordances for learning, technology-supported assessments also provide potential affordances for cheating. Existing research suggests that an unsettlingly high proportion of students have engaged in copy-paste plagiarism, with one 2008 study finding almost three in five students copy-pasting without citing (Selwyn, 2008). In response an arms race has developed around anti-plagiarism 'text matching' software such as Turnitin, which compares student work against a database of sources. Cheating students have adapted their practices, and now employ a range of clever strategies like running their copy-pasted sections through translation engines like Google Translate or Babelfish (Jones & Sheridan, 2014). In addition to assisting do-it-yourself plagiarists, educational technology has also supported the logistics of pay-for plagiarism, with essays available made to order.

Although online plagiarism has received substantial attention, the online underbelly of assessment hacking has received little mainstream scrutiny. Unfortunately this lack of awareness hides real threats to assessment integrity. In another paper (Dawson, 2015) I document several 'proof of concept' hacks on a particular type of electronic assessment system:

Bring-your-own-device electronic examinations (BYOD e-exams) are a relatively new type of assessment where students sit an in-person exam under invigilated conditions with their own laptop. Special software restricts student access to prohibited computer functions and files, and provides access to any resources or software the examiner approves. In this study, the decades-old computer security principle that 'software security depends on hardware security' is applied to a range of BYOD e-exam tools. Five potential hacks are examined, four of which are confirmed to work against at least one BYOD e-exam tool. The consequences of these hacks are significant, ranging from removal of the exam paper from the venue through to receiving live assistance from an outside expert. Potential mitigation strategies are proposed; however, these are unlikely to completely protect the integrity of BYOD e-exams. Educational institutions are urged to balance the additional affordances of BYOD e-exams for examiners against the potential affordances for cheaters.

That paper has a troubling finding: even with in-person invigilation it is possible to circumvent all of the security features of some assessment software. Any assessment conducted on student-owned hardware is in theory vulnerable to similar sorts of hacks.

How can we deal with assessment hacking?

One possible approach to this problem is to do nothing, in the hopes that hacking remains a niche or hidden issue. However several of the attacks I present in that paper could be easily packaged up by one crafty student and shared or sold to others. In the parallel world of computer game hacking, this is the approach taken by gamers who want an unfair advantage.

Another approach to dealing with hacking is to invest heavily in clever security measures to counter the threat posed by hackers. This is the approach taken in the computer game hacking world, where intrusive software is installed alongside games to monitor for cheating and instantly ban offenders. Despite ever-increasing anti-cheating measures, hackers still identify new exploits on a regular basis, which sell for substantial sums online. In the online gaming world it appears that fighting hackers through technical means is still only partially successful.

An alternative solution to this problem may lie in educational rather than technological changes. If we start from the position that all of our assessment is vulnerable to hacking, what can we do to design tasks that still mostly achieve their purposes — even when hacked?

One of the threats posed by assessment hacking is that it may transform an examination from 'closed-book' to 'open-book', or even 'open-book, open-web' (Williams

& Wong, 2009). Open-book, open-web environments are often argued to be more 'authentic': in many cases, the actual practice of what is being assessed is usually conducted without restricted access to information. Changing the assessment to foil hackers may create a more real-world task.

Hacking also threatens to reveal the marking logic that sits behind electronic assessment, which ranges from answers to multiple-choice questions, to intelligent scoring of written responses. Educational workarounds to this sort of threat may require us to move away from some task types entirely. They may also force us to make our marking criteria more transparent for automatically marked tasks.

Hackers can also make identities of those involved in assessment more difficult to verify, through impersonation or unauthorised collusion. In my own work I have been able to hack around secure systems and allow a Skype call or instant messaging chat to run in the background. These hacks challenge assessment designers to consider what they can ask of students that is uniquely theirs. So the threat of hacking may lead to more tasks that incorporate evidence of students' work across a variety of verifiable situations over time.

Can hacking improve assessment?

Some of the adaptations required to combat hacking may result in assessment that is more authentic, transparent and sustained. But beyond changes to combat hacking, we can also think of hacking as a metaphor that can be applied to the process of assessment improvement.

In a recently completed Office for Learning and Teaching project (Dawson et al., 2014) we interviewed 33 university teachers about how they make changes to their assessment tasks. Several spoke about creatively interpreting the rules that surround assessment processes. Taking hacking as a metaphor, there is tentative evidence in our data that these teachers 'hacked' around bureaucracy and complexity, in order to implement changes to their assessment.

Hacking is thus a powerful force in assessment, and one that will be very difficult to eliminate. However through creative educational design, hacking may be the catalyst for improvements to assessment.

References

Butler, A. & Roediger, H. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36*(3), 604–616. doi: 10.3758/mc.36.3.604

Charman, M. (2014). Linguistic analysis of extended examination answers: Differences between on-screen and paper-based, high- and low-scoring answers. *British Journal of Educational Technology*, *45*(5), 834–843. doi: 10.1111/bjet.12100

Dawson, P. (2015). Five ways to hack and cheat with bring-your-own-device electronic examinations. *British Journal of Educational Technology*. doi: 10.1111/bjet.12246

Dawson, P., Bearman, M., Molloy, E., Boud, D., Joughin, G. & Bennett, S. (2014). *Improving assessment: Understanding educational decision-making in practice* (Final report). Sydney: Office for Learning and Teaching.

Hillier, M. & Fluck, A. (2013). Arguing again for e-exams in high stakes examinations. In H. Carter, M. Gosper & J. Hedberg (Eds.), *Electric dreams: Proceedings of the 30th ascilite Conference* (pp. 385–396). Sydney: ascilite.

Jones, M. & Sheridan, L. (2014). Back translation: an emerging sophisticated cyber strategy to subvert advances in 'digital age' plagiarism detection and prevention. *Assessment & Evaluation in Higher Education*, 1–13. doi: 10.1080/02602938.2014.950553

Li, H., Xiong, Y., Zang, X., Kornhaber, M., Lyu, Y., Chung, K. S. & Suen, H. (2015). Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education*, 1–20. doi: 10.1080/02602938.2014.999746

Means, B., Toyama, Y., Murphy, R., Bakia, M. & Jones, K. (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. Washington, DC: US Department of Education.

Mogey, N., Cowan, J., Paterson, J. & Purcell, M. (2012). Students' choices between typing and handwriting in examinations. *Active Learning in Higher Education*, 13(2), 117–128. doi: 10.1177/1469787412441297

Mogey, N. & Hartley, J. (2013). To write or to type? The effects of handwriting and word-processing on the written style of examination essays. *Innovations in Education and Teaching International*, 50(1), 85–93. doi: 10.1080/14703297.2012.748334

Selwyn, N. (2008). 'Not necessarily a bad thing …': A study of online plagiarism amongst undergraduate students. *Assessment & Evaluation in Higher Education*, 33(5), 465–479. doi: 10.1080/02602930701563104

Tamim, R.M., Bernard, R.M., Borokhovski, E., Abrami, P.C. & Schmid, R.F. (2011). What forty years of research says about the impact of technology on learning. *Review of Educational Research*, *81*(1), 4–28. doi: 10.3102/0034654310393361

Williams, J.B. & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, *40*(2), 227–236. doi: 10.1111/j.1467-8535.2008.00929.x