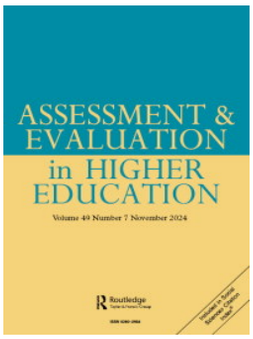

Assessment & Evaluation in Higher Education

ISSN: 0260-2938 (Print) 1469-297X (Online) Journal homepage: www.tandfonline.com/journals/caeh20

Validity matters more than cheating

Phillip Dawson, Margaret Bearman, Mollie Dollinger & David Boud

To cite this article: Phillip Dawson, Margaret Bearman, Mollie Dollinger & David Boud (2024) Validity matters more than cheating, Assessment & Evaluation in Higher Education, 49:7, 1005-1016, DOI: 10.1080/02602938.2024.2386662

To link to this article: https://doi.org/10.1080/02602938.2024.2386662

Published online: 08 Aug 2024.

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
2024, VOL. 49, NO. 7, 1005–1016
https://doi.org/10.1080/02602938.2024.2386662

Validity matters more than cheating

Phillip Dawson[a], Margaret Bearman[a], Mollie Dollinger[a,b] and David Boud[a,c,d]

[a]Centre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Melbourne, Australia; [b]Faculty of Health Sciences, Curtin University, Perth, Australia; [c]Faculty of Arts and Social Sciences, University of Technology Sydney, Sydney, Australia; [d]Work and Learning Research Centre, Middlesex University, London, UK

ABSTRACT

Cheating attracts a significant amount of attention in conversations about assessment, and with good reason: if students cheat, we cannot be sure they have met the learning outcomes of their course. In this conceptual article we question the attention given to cheating as a concept and argue that the broader concept of validity is a more important concern. We begin by questioning what cheating is, why it is wrong, and how justifiable the approaches used to address cheating are. We then propose a reframing of cheating as subsumed by assessment validity. In this view, cheating is addressed without moralising, as part of the broader positive mission of assurance of learning. This perspective highlights how attempts to improve validity by addressing cheating can sometimes make validity worse, for example when an anti-cheating technology reduces cheating but creates problems for inclusion. In shifting focus from cheating to validity, we hope to draw renewed attention to what matters most in assessment: that we know our graduates are capable of what we say they are.

KEYWORDS

Academic integrity; cheating; assessment; validity; assessment security; artificial intelligence

There is a large body of literature on student cheating and academic integrity in higher education, including dedicated journals, which sit alongside multiple international conferences and hundred-chapter handbooks. Major topics of study include students' motivations for cheating (e.g. Daumiller et al. 2023), frameworks and strategies to reduce cheating (e.g. Ellis and Murdoch 2024) as well as calls for a reframing of cheating to better recognise historical and cultural roots (e.g. Eaton 2024). Underpinning the broad field of cheating is a range of assumptions, including that we agree on what cheating is, that cheating is wrong, and that cheating is a very important problem. But perhaps the biggest shared assumption underlying the field is its dominant framing: *students* are doing wrong things, such as plagiarising, breaching academic integrity or inappropriately using artificial intelligence. In this paper, we explore the term 'cheating' and ask if this frame is truly the most productive way to think about the phenomenon, and if our expensive strategies to combat cheating are indeed fit-for-purpose. We further explore what moving away from the idea of cheating might afford the sector, and how the underlying phenomenon might be more productively conceptualised.

We come to the concept of cheating through our background as assessment researchers. To us, assessment is very important for a range of reasons, one of which is that it enables universities to certify to society that a student has achieved particular outcomes (Boud 2000).

CONTACT Phillip Dawson p.dawson@deakin.edu.au Centre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Melbourne, Australia.

Cheating is therefore a problem because it makes it difficult to certify what a particular student has achieved. In other words, cheating poses a threat to the validity of the assessment. In its broadest sense, valid university assessments identify students who have met the standards of a course to an agreed level of performance and distinguish these from those who have not met the requirements. From a more technical perspective, as we expand in the later parts of this paper, we can think of validity as a term with a plurality of meanings (St-Onge et al. 2017), which focus on the inferences and implications that can be drawn from an act of assessment. Thus cheating and validity must interplay. As Cizek and Wollack (2017) note:

> There may be disagreement about the ethical dimensions of cheating, but it is uncontestable that cheating represents a threat to the valid interpretation of a test score. When cheating takes place—whether in the form of copying from another test taker, collusion, prior access to secure test materials, inappropriate manipulation of answer documents, or any other form—the resulting test scores are not likely to be an accurate measurement of an examinee's true level of knowledge, skill, or ability. In short, a concern about cheating can be viewed as a psychometric concern about the validity or "interpretive accuracy" of test scores. (Cizek and Wollack 2017, p. 8)

In this paper we go beyond drawing a connection between 'validity' and 'cheating' to propose that validity is a more useful conceptualisation than cheating for all but the moralistic elements. We avoid engaging in what Makridis and Englander (2021) would term as "revisionism" – an attempt to redefine cheating to be morally acceptable – and instead argue for a different conversation entirely that is focused on validity. But before we propose our alternative perspective, we mount a case for what may be limiting with the current concept of cheating. We commence with an exploration of how cheating is currently defined.

What is cheating?

There is often an assumption that we all know what cheating means with respect to university assessment, however, this may not be the case. While the term has formal definitions in the academic literature, these do not necessarily represent how 'cheating' is viewed in day-to-day teaching and learning practice. Everyday educational usage of the term involves an act of deception (OED 2023), resulting in unfair advantage or personal gain. While validity is treated as a powerful 'god term' in assessment (St-Onge et al. 2017) – a word that is evaluated as inherently good and one that all other language must be subservient to – the inverse label, a 'devil term' (Chandler and Munday 2011) could easily be applied to cheating. Cheating is a serious, accusatory word that brings an assumption of intent and moral wrongness. It is negative and potent.

In the literature, cheating is often defined in a concrete way, such as through a list of prohibited behaviours or an instrument asking students if they have committed particular acts (Barnhardt 2016). This approach is often used in institutional policy, procedure and processes around academic integrity, where lists of proscribed behaviours and accompanying penalties are presented to students. A similar approach is taken in some legislation, such as in Australia where commercial third parties are prohibited from doing assessed work for students (Parliament of Australia 2020). As these lists have come to define what cheating means, their contents matter – and there is little consensus on what belongs in them. For example, reading a condensed study guide rather than a set text is regarded as cheating by some scholars (McKibban and Burdsal 2013) but is absent from most concrete definitions of cheating.

The dominance of list-based definitions allows for exhaustive checklists of prohibited behaviour. But when a tick against even the most benign item can lead to claims that in excess of 90% of students cheat (Barnhardt 2016), it may be that such lists are doing more harm than good. And as the technology and cheating landscapes are changing rapidly in recent years, list-based definitions have struggled to keep up. For example, with the introduction of ChatGPT in 2022 and its ability to satisfactorily complete many assessed tasks (Gilson et al. 2023; Nikolic et al. 2023),

some uses of these technologies are now regarded by some as cheating (Cotton, Cotton, and Shipway 2023). List-based definitions of cheating require constant attention as the set of prohibitions expands.

In contrast, abstract definitions of cheating avoid specifying prohibited acts, instead attempting to describe cheating as a concept (Barnhardt 2016). These definitions often hold underlying assumptions about education, for example "gaining an unfair advantage" (Barnhardt 2016, 330) could be interpreted as suggesting that assessment is viewed as a competition. Some definitions require a moral element, suggesting that intentionality is necessary for an act of cheating to have occurred (Barnhardt 2016). Some view education as a game, which necessarily has rules, and cheating is what happens when they are broken (Makridis and Englander 2021). Some are tautological, taking the form that cheating is what happens when the rules against cheating are broken (Bouville 2009).

As with other aspects of educational practice, understandings of cheating are deeply connected to place, discipline, and culture. Cheating is socially constructed; it is not some objective natural phenomenon with universal rights and wrongs (Dawson 2021). Understandings of cheating change over time, such as the emphasis on citation practices in the wake of rampant Internet-enabled copy-pasting in the late 1990s, and the current questions about when the use of artificial intelligence crosses the line into being cheating. Just as students who move across continents can find themselves engaged in practices not acceptable in their new setting (Amsberry 2009), older scholars may wish to reflect on how their undergraduate work would fare if put through a modern text-matching tool, or through the anti-cheating technologies that are likely to arise in the future (Dawson 2021).

In a digital world, nuanced concepts such as cheating often get implemented into algorithms. In practice, the assessment security industry – vendors of products such as remote proctoring, text-matching, and AI writing detection tools – are influential in deciding what cheating is (Dawson 2021). While vendors of these products are careful not to position themselves as the cheating police, their algorithms often trigger the human processes that investigate if someone has cheated or not (Introna 2015; Penketh and Beaumont 2014). Thus, unwittingly, the academy may be ceding some of the responsibility for defining cheating to for-profit companies who have a vested interest in the appearance of high rates of cheating. With no clear consensus or stability of understanding on what cheating is, or who gets to decide, there is significant room for nuanced work in defining cheating. But more fundamentally, there also needs to be an examination of the wrongness of cheating.

Why is cheating bad?

The substantial attention that cheating gets in the literature and in practice is predicated on the wrongness of cheating. But why is cheating bad? A range of justifications are possible, and each shapes the ways cheating is viewed and the responses that are deemed appropriate. Bouville (2009) discusses the two most prominent justifications: that *cheating is unfair*, and that *cheating hurts student learning*, and argues that neither holds up to scrutiny. The unfairness argument proposes that cheating is wrong because it provides those students who cheat with an advantage. Bouville (2009) counters by pointing out that this requires a view of education as a race of all against all, something that he thinks is not compatible with current understandings of education; indeed, in theory in a standards-based assessment system, the unearned high grades of one student should not affect another student's grades. The problem may be that in a standards-based system, when students cheat and graduate without the required capabilities, they de-value their peers' degrees. The argument that cheating is wrong because it hurts learning is also described by Bouville (2009) as inadequate as there are also many other activities students do that could similarly rob them of learning opportunities that we do not punish. If we were to view cheating as wrong because students missed out on learning opportunities, why

would we not also view part-time work or engagement in hobbies as wrong if those activities led to less time on task?

Taking a purely utilitarian perspective on cheating, it can be argued that cheating is wrong because it poses a threat to validity, as we suggest using Cizek and Wollack (2017)'s argument at the start of this paper. That is, students pass, or are recorded as reaching a particular level of achievement, when they have not truly demonstrated their learning. But if this is the case, is the problem that needs to be addressed one of cheating in particular, or one of validity in general? If validity is the primary reason for cheating's wrongness, it is just one of many validity threats facing assessment, and the argument is yet to be successfully made that it is the most important one to address. For example, assessment that is not inclusive poses a threat to validity by making the assessment of some students less able to represent what they are truly capable of Dawson (2022). While validity as an overarching frame forms the main thrust of this paper, which we will come back to, it is by no mean the only view on why cheating is bad.

The field of academic integrity presents its own varied justifications for the wrongness of cheating, but the dominant one is that it is a breach of a set of values that students should have. The fundamental values of academic integrity are honesty, trust, fairness, respect, responsibility and courage (Fishman 2014). Cheating is wrong in this view because it is a demonstration that students do not live these values. However, the absence of cheating is not necessarily evidence of academic integrity. If students do not cheat in an examination out of fear of getting caught, this is not a demonstration of the values; it may just be a rational choice. Academic integrity therefore may be a justification for the wrongness of cheating, but it is no justification for the surveillance and policing of students to stop cheating; that is the purview of assessment security not academic integrity (Dawson 2021). As a result, academic integrity only justifies part of the cheating conversation, and not the part that is most contentious.

If cheating is taken to be wrong, is it necessarily wrong all of the time? The qualitative literature on the experiences of those students who cheat contains heartbreaking narratives (e.g. Pitt, Dullaghan, and Sutherland-Smith 2020). In the hypothetical situation of being a student who has to choose between completing an assignment without cheating, or earning money to feed their family, is it possible to place the decision to cheat in the same category as 'stealing a loaf of bread' (Hammer, Norton, and Star 2004)? Are there exceptions to the wrongness of cheating when to not cheat would result in disaster for the individual or people they care for? More broadly, there are arguments that cheating is an act of resistance or a reasonable response to a broken system of assessment and education (e.g. Högberg 2011; Pulfrey & Butera 2013; Redding 2017): busywork, credentialism, massification, students as consumers, the overemphasis of grades, or other societal evils. Such an act of resistance is simultaneously courageous and at odds with the fundamental academic integrity value of courage.

This trend for questioning the wrongness of cheating, and even arguing it to be morally acceptable is termed "transvaluation" by Makridis and Englander (2021, p. 2); transvaluation being in this instance the flipping of the negative moral value generally attached to cheating, into a positive. They argue that the work has not been successfully done to support a case that cheating is anything but morally wrong. While there have been many attempts, Makridis and Englander argue they contain logical fallacies at their core, and that the burden of proof lies not on those arguing for the wrongness of cheating, but on those who are seeking transvaluation.

If cheating is wrong, whose wrong is it? By default, students are regarded as the ones who have wronged. Kohn (2007) argues that this is a fundamental attribution error, in that we ascribe the problem of cheating to individual choice, whereas it is likely more the cause of the context the individual finds themselves within. In this view, cheating may be wrong, but it is the system's wrong. Anti-cheating legislation in many countries explicitly targets only purveyors of cheating resources, regarding them as the criminal wrongdoers, rather than the students. Pabian (2015, 818) argues that if an education system focuses on the transmission and replication of knowledge, then students engaging in copying is "integral (rather than deviant) to the educational

process". Who is really wrong in cheating – the student, the educational system, assessment designers who set inappropriate tasks, commercialised cheating providers, or a society that places so much emphasis on grades and credentials? There is a need for much more work in the space of the wrongness of cheating, both in the form of new arguments and perspectives, and critiques of existing arguments.

What are we doing about cheating?

Significant resources are put into addressing cheating. Individual educators spend substantial time in designing assessment, conducting exams, investigating suspected cheating, and attending hearings. There has been a growth of specialist roles in this area (Vogt and Eaton 2022), and entire teams now exist at some institutions that are solely concerned with cheating. Anti-cheating technology (such as text-matching and remote proctoring tools) is a multi-billion-dollar industry, largely funded by universities. Is this money, time and effort well spent?

In evaluating if the time and money devoted to the problem of cheating is well spent, it is useful to consider the concepts of importance, neglectedness and tractability. These concepts are used in other contexts to determine the best use of limited resources to address complex social problems (MacAskill 2015), and they have been applied to the prioritisation of research and practice problems in other disciplines, such as psychology (Gainsburg et al. 2022). Through this lens, a problem is considered a good use of resourcing if it is very important, highly neglected, and if there are effective interventions that can demonstrate value for money. How does cheating rate against these criteria? The field of academic integrity has worked to establish the importance of the problem of cheating over several decades, and this is largely a mixture of the moral arguments presented earlier and a plethora of studies into the prevalence of cheating; we would add that the validity threats raised by cheating further the importance of the problem of cheating. The extent of the neglectedness of cheating as a problem for research or practice is more challenging to establish. Entire journals, conferences, teams of professional staff, policies, legislation and courses exist to address cheating. However, there are always calls for more resourcing for the problem on the basis that it is neglected. The tractability of cheating – the extent to which investment in it can meaningfully reduce the scale or severity of the problem – is not very well established.

To address the tractability problem there is a need for cost-benefit studies that compare the efficacy of different approaches to addressing cheating against what they cost, not just in terms of financial cost, but also how they distort educational practices, and their broader impacts on staff, students and society. Efficacy is difficult to measure as there are currently no accepted metrics (Dawson 2021). It is unclear what efficacy would look like in terms of cheating: is it reduced rates of cheating; reduced rates of undetected cheating; improvements to the detection of cheating; or some sort of delayed behavioural indicator such as professionalism in graduate employment?

Given the lack of a shared understanding of efficacy in addressing the problem of cheating, it is unsurprising that there is a lack of evidence in support of many of the dominant anti-cheating practices. There is no literature of which we are aware that demonstrates that the academic integrity modules in widespread use have led to reduced rates of cheating – despite such modules having been in use for decades. There is strong evidence that 'honor codes' which require students to state they will not cheat, and they will act with integrity are associated with reduced rates of cheating – but those reductions are very small (McCabe, Treviño, and Butterfield 2002). There is evidence that remote proctored exams are associated with poorer performance compared to unsupervised online exams, which is inferred to mean that proctoring deters cheating (Hylton, Levy, and Dringus 2016), but there is no peer reviewed evidence we are aware of that remote proctoring effectively detects cheating, and, in fact, emerging evidence that it does not (Bergmans et al. 2021; Burgess et al. 2022). Even common assumptions that face-to-face exams

are more cheating-proof than take-home assignments are questionable (Buckley 2023), especially given evidence of more third-party cheating in exams than in assignments (Harper, Bretag, and Rundle 2021). These activities have unintended consequences, not just to educational institutions, but to students and their learning.

Anti-cheating approaches and technologies can stifle collaboration, peer learning and feedback seeking. When students are told they may not discuss their work with others, it robs them of the opportunity to engage in informal peer learning. Collaboration with others is widely seen as a desirable learning outcome, but how are students to distinguish between collaboration and collusion when messages are mixed? This sets up an individualistic approach to work that is highly inauthentic (Boud and Bearman 2024); in most workplaces collaboration is encouraged. It also runs counter to efforts to encourage students' feedback seeking behaviour. In workplace settings it is highly desirable for employees to seek feedback (Joughin et al. 2020), but when students are told doing so is cheating, they miss out on opportunities to develop their capabilities in seeking feedback. If the importance of the problem of cheating is due to its harms for learning, then these harms to learning from anti-cheating need to be weighed carefully. Most students do not engage in the most egregious forms of cheating such as contract cheating (Curtis et al. 2021) or exam cheating (Barnhardt 2016), but anti-cheating approaches are applied to them regardless. Are the harms to their feedback seeking and collaboration capabilities worth the (as yet unquantified) benefits?

Anti-cheating approaches attempt to enforce ethical behaviour. But are anti-cheating approaches themselves ethical? Significant critiques have been made around the ethics of anti-cheating approaches and especially anti-cheating software. For example, remote proctoring has been labelled racist and ableist (Logan 2020); and text-matching software has been critiqued for its capturing and monetisation of student data (Morris and Stommel 2017). These approaches fall into the category of 'cop shit', which Moro (2020) defines as "any pedagogical technique or technology that presumes an adversarial relationship between students and teachers". The metaphor of policing can extend beyond the detection of cheating and onto its punishments, which can be severe and affect students differently. The immediate punishments of cheating can include failure and exclusion from an institution, but the consequences of failure affect some students more than others. For example, an international student can lose their visa if they are found to have cheated repeatedly, but a domestic student does not face this sanction. Are these harms to individuals proportional to the wrongs they have done?

Compared with other problems in teaching and learning, cheating gets a lot of resources. We argue that in focusing so much on cheating, we have lost sight of some of these more fundamental problems of higher education, which don't get as much attention. In the next section, we argue that validity could be viewed as a more important concern, and one that can subsume the notion of cheating.

Reframing cheating as a problem of assessment validity

If the major damage done by cheating is in the harms it poses to assessment validity and the flow-on effects from that (such as incompetent and unsafe graduates), then perhaps that validity threat should be the key focus. The moral element could be downplayed, if not discarded, or reserved for egregious harmful behaviours such as blackmail or taking people's work without their permission. The remainder of this section makes an argument as to why validity should be the core focus of our efforts and cheating should be on the periphery.

A validity perspective makes the claim: a students' assessment submission is valid if it represents their actual capability. Validity subsumes the concept of cheating within a broader, positive mission: one of assurance of learning. Seeking to ensure that assessments are valid makes no moral judgements about students who cheat but does imply that such students should not graduate if their cheating compromises their demonstration of the learning outcomes. It does

not necessarily preclude resitting an exam or re-submitting an assignment – it treats cheating as a form of failure to demonstrate capability. Such an approach is only possible within a standards-based system of assessment (Boud 2017), which is what is current in most major higher education jurisdictions now.

A standards-based system is one where the only students who can legitimately graduate are those who have demonstrated all the learning outcomes specified for the program of study. This is not radical; it is a vision for assessment that represents the minimum required in some contexts, such as Australia, where legislation dictates that institutions may only graduate those students the institution is confident have met the declared learning outcomes (Australian Government 2015). In a ranking or norm-based system, validity is difficult to grapple with, as by its nature the assessments focus on relative performance, not real capability. The adoption of standards-based assurance of learning requires assessment that maps student performances to learning outcomes, determining if the learning outcomes have been demonstrated sufficiently (and optionally, to what degree of sophistication). If behavioural or ethical characteristics are desired of graduates, such as those typically associated with academic integrity, these can be assessed and specified as intended outcomes of the qualification. In this way, thus we reconceptualise behaviours such as ethical conduct, collaboration and so on as intended outcomes rather than describing exhaustive lists of prohibited behaviours.

Assuring learning positions cheating as a *subset* of assessment validity. A validity perspective requires assessments that all capable students can complete within the specified conditions, just as assessment for inclusion seeks to redesign assessment so it suits all students without the need for adjustments (Tai et al. 2022). Judgements about what students are capable of under given conditions are invalidated if students do not work within those conditions. This becomes a technical question about whether the conditions can be met, rather than a moral question about whether students will do what we ask them to. Such a reframing avoids the fundamental attribution error of regarding breaching these rules as the sole fault of the individual student (Kohn 2007). Rather, it suggests that institutions should not set rules that will likely be broken. Doing so only hurts validity, as it results in judgements being made about student outcomes based on the incorrect assumption the rules were adhered to. This calls into question all manner of restrictions imposed on students in their assessments, such as the use of closed-book exams (Dawson, Nicola-Richmond, and Partridge 2024).

So what are the implications for cheating of a validity approach? It is still important for students to know the rules, and when they break the rules, and they are caught, there is often some sort of sanction involved. However, from a validity perspective, when a student is found to have completed a task in a way that makes their assessor's judgement invalid, awarding the student no credit for the relevant parts of the task is simply assessment, not punishment. The student's capacity to do the task cannot be assured and therefore they can gain no credit for their effort. This is entirely compatible with, and indeed necessary for, a validity view.

This does not preclude sanctions. There can be patterns of behaviour in a student which suggest they may be engaged in persistent or highly problematic behaviour that threatens validity, such as outsourcing their work, which may require a different response. From a validity perspective, punishment is not inherently required; instead, we seek approaches that will improve the validity for assessing the student in the future. Approaches that support students accused of breaches to understand and address their problematic behaviour, such as frank discussions with them or 'courageous conversations' (Murdoch and House 2023) may meet this goal without the need for traditional punishment approaches. Students may need to retake a unit of study or even a whole year of their course. It may also be that an institution will judge that, due to patterns of behaviour, that they will not be able to assure learning for a student, therefore that student will not be able to continue. While the consequence may feel the same as punishment, the deficit is not located in the character of the student.

There has been much handwringing about artificial intelligence and cheating since the rise of generative AI tools such as ChatGPT. Much of this discourse has centred around when the use of AI is cheating. Through our reframing of cheating, we argue that the use of artificial intelligence becomes unacceptable when it threatens validity. Therefore, assessments that depend on students not using artificial intelligence but are incapable of preventing students from doing so, are not particularly useful for high-stakes assessment of learning. This again locates the problem of inappropriate use of artificial intelligence as a property of the assessment, rather than as a property of the student. The use of artificial intelligence in assessment is not 'cheating', it is a condition to be attended to alongside other validity matters. A validity lens also enables us to consider both the learning benefits of preparing students for a world where artificial intelligence will be ubiquitous, alongside the need to sometimes assess students without artificial intelligence (Lodge et al. 2023).

Making a claim: validity as more than the property of the assessment

So far in this paper, we have used a basic, everyday understanding of validity as simply assessing what we intend to assess. In practice, this usually sits somewhere between face validity (does it look like the assessment assesses what it purports to?) and construct validity (what evidence is there that it assesses the underlying capability it is supposed to?). We have done this partly because validity is often seen as a somewhat arcane term in higher education, something that is the domain of the psychometrician or assessment specialist rather than the everyday assessor. However, the term validity has a range of different meanings in assessment (St-Onge et al. 2017) and some of these are more relevant to an assurance of learning perspective than others. In particular, we suggest moving beyond the traditional psychometric conception that validity is a property of the assessment tool but is rather a property of the assessment occasion.

Why is this useful with respect to cheating? Firstly, it moves us from a focus on just the instruments of assessment to also consider the implications of the assessment on what students do, which has been termed consequential validity (Boud 1995; Sambell, McDowell, and Brown 1997). Cook and Lineberry (2016) consider these consequences to be: "impact, beneficial or harmful, of the assessment itself and the decisions and actions that result…" (p. 786). Thus a consequential validity perspective on cheating can view student misbehaviour as a function of the assessment design. Just as pre-exam cramming can be viewed as part of the consequential validity of high-stakes testing, the use of assignment sharing sites can be viewed as part of the consequential validity of single-right-answer assignments. This view of validity not as an immutable property of an assessment, but as a social imperative (St-Onge et al. 2017), causes us to move beyond concerns of accuracy to also consider the positive and negative consequences of assessment for students, teachers and society more broadly.

An associated conceptualisation is that validity cannot be proven, but it can only be argued for. In short, validity is a claim not an absolute. Thus, validity can also be building of chains of evidence in support of judgements about student capability (St-Onge et al. 2017). This means that no single act of assessment can address all potential validity problems, however multiple types of assessment, at multiple time points, with multiple assessors can better address the challenges found in a single assessment (van der Vleuten et al. 2012). This is compatible both with the notion of programmatic assessment, and with the Swiss Cheese model (Reason 2000) that has been adapted to the context of integrity and cheating by Rundle, Curtis, and Clare (2020). They argue that any one approach to address cheating will have 'holes' but that many layers of different types of approaches will better address the deficiencies of any one layer.

Any assessment design is ultimately about trade-offs, and this perspective of validity as a claim reinforces that notion. There are many potential ways to improve one aspect of the evidentiary chain that address 'cheating' concerns, which also threaten validity in other ways. For example, the benefits to validity achieved through using a remote proctored exam might be ultimately

not worth the harms to validity created by that technology's exclusionary nature (Dawson 2023). Using artificial intelligence detection software might improve validity by identifying probable cases where a requirement to not use AI was breached, but this might hurt validity by giving assessors false confidence that other students did not use artificial intelligence. A take-home authentic assignment might be vulnerable to all manner of unintended student behaviours, but it may have better face validity than a high-security multiple-choice exam. In a task that privileges the learning purposes of assessment, the construct validity benefits of strict rules against collusion might not be worth the consequential validity effects of discouraging peer learning and feedback seeking.

Viewing validity as separate to cheating means that these tensions cannot be considered holistically. It also means that new advances such as generative artificial intelligence have needed to be cast as either cheating or not cheating, when the real challenge they pose in this context is not one of student morality but one of assessment design. Challenging the concept of 'cheating' and recentering validity in conversations about assessment enables us to better weigh up these pros and cons, and address new challenges as they arise.

Conclusion

For higher education assessment to function, it has seemed reasonable to set some ground rules for students, for example: cite your sources, don't pay someone else to do your work for you, and check with us before you use artificial intelligence. However, the overwhelming focus on adherence to rules has led the sector to focus on cheating as if adherence to the rules were an end in itself. In this paper we have drawn attention to the thing the rules are supposed to support: assessment validity. We hope that a renewed focus on validity will help shift focus from moralising about students towards assuring what matters most in assessment: that we are confident students who graduate from our programs of study are capable of doing what we say they are.

Acknowledgments

The authors thank the attendees at the 2022 CRADLE Symposium *Challenging Cheating*, and our colleagues Rola Ajjawi and Joanna Tai, for their feedback on earlier versions of these ideas.

Disclosure statement

Phillip Dawson has received funding from educational technology companies. The two that are relevant to this paper are Turnitin and Inspera. He has also received funding from Australia's higher education regulator, the Tertiary Education Quality and Standards Agency (TEQSA) for work addressing cheating and validity, including a project with Margaret Bearman where she also received funding. No funding was provided by any organisation for this article.

ORCID

Phillip Dawson http://orcid.org/0000-0002-4513-8287
Margaret Bearman http://orcid.org/0000-0002-6862-9871
Mollie Dollinger http://orcid.org/0000-0003-1105-9051
David Boud http://orcid.org/0000-0002-6883-2722

References

Amsberry, D. 2009. "Deconstructing Plagiarism: International Students and Textual Borrowing Practices." *The Reference Librarian* 51 (1): 31–44. doi:10.1080/02763870903362183.

Australian Government. 2015. *Higher Education Standards Framework (Threshold Standards)*. Canberra: Federal Reigster of Legislation.

Barnhardt, B. 2016. "The "Epidemic" of Cheating Depends on Its Definition: A Critique of Inferring the Moral Quality of "Cheating in Any Form." *Ethics & Behavior* 26 (4): 330–343. doi:10.1080/10508422.2015.1026595.

Bergmans, L., N. Bouali, M. Luttikhuis, and A. Rensink. 2021. "On the Efficacy of Online Proctoring Using Proctorio." In Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021) – Vol. *1*, pp. 279–290. SCITEPRESS – Science and Technology Publications, Lda.

Boud, D. 1995. "Assessment and Learning: Contradictory or Complementary." In *Assessment for Learning in Higher Education*, edited by P. Knight, 35–48. London: Kogan Page.

Boud, D. 2000. "Sustainable Assessment: Rethinking Assessment for the Learning Society." *Studies in Continuing Education* 22 (2): 151–167. doi:10.1080/713695728.

Boud, D. 2017. "Standards-Based Assessment for an Era of Increasing Transparency." In *Scaling up Assessment for Learning in Higher Education*, edited by D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski, 19–31. Singapore: Springer Singapore.

Boud, D., and M. Bearman. 2024. "The Assessment Challenge of Social and Collaborative Learning in Higher Education." *Educational Philosophy and Theory* 56 (5): 459–468. doi:10.1080/00131857.2022.2114346.

Bouville, M. 2009. "Why is Cheating Wrong?" *Studies in Philosophy and Education* 29 (1): 67–76. doi:10.1007/s11217-009-9148-0.

Buckley, A. 2023. "Are we Answering the Question That Has Been Set? Exploring the Gap between Research and Practice around Examinations in Higher Education." *Studies in Higher Education*. doi:10.1080/03075079.2023.2283784.

Burgess, B., A. Ginsberg, E. W. Felten, and S. Cohney. 2022. "Watching the Watchers: Bias and Vulnerability in Remote Proctoring Software." In 31st USENIX Security Symposium (USENIX Security 22), 571–588. Boston, MA: USENIX Association.

Chandler, D., & Munday, R. (Eds.) 2011. *A Dictionary of Media and Communication*. Oxford: Oxford University Press.

Cizek, G. J., and J. A. Wollack. 2017. "Exploring Cheating on Tests: The Context, the Concern, and the Challenges." In *Handbook of Quantitative Methods for Detecting Cheating on Tests*, edited by G. J. Cizek & J. A. Wollack, 3–19. New York: Routledge.

Cook, D. A., and M. Lineberry. 2016. "Consequences Validity Evidence: Evaluating the Impact of Educational Assessments." *Academic Medicine* 91 (6): 785–795. doi:10.1097/ACM.0000000000001114.

Cotton, D. R. E., P. A. Cotton, and J. R. Shipway. 2023. "Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT." *Innovations in Education and Teaching International* 61 (2): 228–239. doi:10.1080/14703297.2023.2190148.

Curtis, G. J., M. McNeill, C. Slade, K. Tremayne, R. Harper, K. Rundle, and R. Greenaway. 2021. "Moving beyond Self-Reports to Estimate the Prevalence of Commercial Contract Cheating: An Australian Study." *Studies in Higher Education* 47 (9): 1844–1856. doi:10.1080/03075079.2021.1972093.

Daumiller, M., T. Fritz, H. González Cruz, S. C. Rudert, and S. Janke. 2023. "Cheating as a Prosocial Act? Helping Others with Academic Cheating is Related to Social Goals and Cooperative Norms." *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2023.2290979.

Dawson, P. 2021. *Defending Assessment Security in a Digital World: Preventing e-Cheating and Supporting Academic Integrity in Higher Education*. Abingdon, Oxon: Routledge.

Dawson, P. 2022. "Inclusion, Cheating, and Academic Integrity: Validity as a Goal and a Mediating Concept." In *Assessment for Inclusion in Higher Education: Promoting Equity and Social Justice in Assessment*, edited by J. T. Rola Ajjawi, David Boud, Trina Jorre de St Jorre, 110–119. London: Routledge.

Dawson, P. 2023. "Remote Proctoring: Understanding the Debate." In *Handbook of Academic Integrity*, edited by S. E. Eaton. Singapore: Springer Nature Singapore.

Dawson, P., K. Nicola-Richmond, and H. Partridge. 2024. "Beyond Open Book versus Closed Book: A Taxonomy of Restrictions in Online Examinations." *Assessment & Evaluation in Higher Education* 49 (2): 262–274. doi:10.1080/02602938.2023.2209298.

Eaton, S. E. 2024. "Decolonizing Academic Integrity: Knowledge Caretaking as Ethical Practice." *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2024.2312918.

Ellis, C., and K. Murdoch. 2024. "The Educational Integrity Enforcement Pyramid: A New Framework for Challenging and Responding to Student Cheating." *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2024.2329167.

Fishman, T. 2014. *The Fundamental Values of Academic Integrity*. 2nd ed. Clemson: International Center for Academic Integrity, Clemson University.

Gainsburg, I., S. Pauer, N. Abboub, E. T. Aloyo, J.-C. Mourrat, and A. Cristia. 2022. "How Effective Altruism Can Help Psychologists Maximize Their Impact." *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 18 (1): 239–253. doi:10.1177/17456916221079596.

Gilson, A., C. W. Safranek, T. Huang, V. Socrates, L. Chi, R. A. Taylor, and D. Chartash. 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* 9 (1): E45312. doi:10.2196/45312.

Hammer, S., P. Norton, and C. Star. 2004. "Cheating to Cope in the Hothouse." *The Australian,* 16 June 2004.

Harper, R., T. Bretag, and K. Rundle. 2021. "Detecting Contract Cheating: Examining the Role of Assessment Type." *Higher Education Research & Development* 40 (2): 263–278. doi:10.1080/07294360.2020.1724899.

Högberg, R. 2011. "Cheating as Subversive and Strategic Resistance: Vocational Students' Resistance and Conformity towards Academic Subjects in a Swedish Upper Secondary School." *Ethnography and Education* 6 (3): 341–355. doi:10.1080/17457823.2011.610584.

Hylton, K., Y. Levy, and L. P. Dringus. 2016. "Utilizing Webcam-Based Proctoring to Deter Misconduct in Online Exams." *Computers & Education* 92–93: 53–63. doi:10.1016/j.compedu.2015.10.002.

Introna, L. D. 2015. "Algorithms, Governance, and Governmentality: On Governing Academic Writing." *Science, Technology, & Human Values* 41 (1): 17–49. doi:10.1177/0162243915587360.

Joughin, G., D. Boud, P. Dawson, and J. Tai. 2020. "What Can Higher Education Learn from Feedback Seeking Behaviour in Organisations? Implications for Feedback Literacy." *Assessment & Evaluation in Higher Education* 46 (1): 80–91. doi:10.1080/02602938.2020.1733491.

Kohn, A. 2007. "Who's Cheating Whom?" *Phi Delta Kappan* 89 (2): 89–97.

Lodge, J. M., S. Howard, M. Bearman, and P. Dawson. 2023. *Assessment Reform for the Age of Artificial Intelligence*. Canberra: Tertiary Education Quality and Standards Agency (TEQSA).

Logan, C. 2020. Refusal, Partnership, and Countering Educational Technology's Harms. *Hybrid Pedagogy*. https://hybridpedagogy.org/refusal-partnership-countering-harms/

MacAskill, W. 2015. *Doing Good Better: Effective Altruism and a Radical New Way to Make a Difference*. London: Guardian Faber Publishing.

Makridis, O., and F. Englander. 2021. "Normative Revisionism about Student Cheating." *Journal of Academic Ethics* 19 (1): 1–23. doi:10.1007/s10805-020-09384-z.

McCabe, D. L., L. K. Treviño, and K. D. Butterfield. 2002. "Honor Codes and Other Contextual Influences on Academic Integrity: A Replication and Extension to Modified Honor Code Settings." *Research in Higher Education* 43 (3): 357–378. doi:10.1023/A:1014893102151.

McKibban, A. R., and C. A. Burdsal. 2013. "Academic Dishonesty: An In-Depth Investigation of Assessing Measurable Constructs and a Call for Consistency in Scholarship." *Journal of Academic Ethics* 11 (3): 185–197. doi:10.1007/s10805-013-9187-6.

Moro, J. 2020. "Against Cop Shit." https://jeffreymoro.com/blog/2020-02-13-against-cop-shit/

Morris, S. M., and J. Stommel. 2017. "A Guide for Resisting Edtech: The Case against Turnitin." *Hybrid Pedagogy*. https://hybridpedagogy.org/resisting-edtech/

Murdoch, K., and D. House. 2023. "Courageous Conversations: Approaching Amnesty Through Honesty as Reparations to a Learning Community." In *Handbook of Academic Integrity*, edited by S. E. Eaton, 631–643. Cham: Springer International Publishing.

Nikolic, S., S. Daniel, R. Haque, M. Belkina, G. M. Hassan, S. Grundy, S. Lyden, P. Neal, and Caz Sandison. 2023. "ChatGPT versus Engineering Education Assessment: A Multidisciplinary and Multi-Institutional Benchmarking and Analysis of This Generative Artificial Intelligence Tool to Investigate Assessment Integrity." *European Journal of Engineering Education* 48 (4): 559–614. doi:10.1080/03043797.2023.2213169.

OED. 2023. *Oxford English Dictionary*. Oxford: Oxford University Press.

Pabian, P. 2015. "Why 'Cheating' Research is Wrong: New Departures for the Study of Student Copying in Higher Education." *Higher Education* 69 (5): 809–821. doi:10.1007/s10734-014-9806-1.

Parliament of Australia. 2020. *Tertiary Education Quality and Standards Agency Amendment (Prohibiting Academic Cheating Services) Bill 2019, Commonwealth of Australia*. Canberra: Office of Parliamentary Counsel.

Penketh, C., and C. Beaumont. 2014. "Turnitin Said It Wasn't Happy': Can the Regulatory Discourse of Plagiarism Detection Operate as a Change Artefact for Writing Development?" *Innovations in Education and Teaching International* 51 (1): 95–104. doi:10.1080/14703297.2013.796721.

Pitt, P., K. Dullaghan, and W. Sutherland-Smith. 2020. "Mess, Stress and Trauma': Students' Experiences of Formal Contract Cheating Processes." *Assessment & Evaluation in Higher Education* 46 (4): 659–672. doi:10.1080/02602938.2020.1787332.

Pulfrey, C., and F. Butera. 2013. "Why Neoliberal Values of Self-Enhancement Lead to Cheating in Higher Education: A Motivational Account." *Psychological Science* 24 (11): 2153–2162. doi:10.1177/0956797613487221.

Reason, J. 2000. "Human Error: Models and Management." *BMJ* 320 (7237): 768–770. doi:10.1136/bmj.320.7237.768.

Redding, A. B. 2017. "Fighting Back Against Achievement Culture: Cheating as an Act of Rebellion in a High-Pressure Secondary School." *Ethics & Behavior* 27 (2): 155–172. doi:10.1080/10508422.2016.1145058.

Rundle, K., G. Curtis, and J. Clare. 2020. "Why Students Choose Not to Cheat." In *A Research Agenda for Academic Integrity*, edited by T. Bretag. Cheltenham: Edward Elgar Publishing.

Sambell, K., L. McDowell, and S. Brown. 1997. "But is It Fair?": An Exploratory Study of Student Perceptions of the Consequential Validity of Assessment." *Studies in Educational Evaluation* 23 (4): 349–371. doi:10.1016/S0191-491X(97)86215-3.

St-Onge, C., M. Young, K. W. Eva, and B. Hodges. 2017. "Validity: One Word with a Plurality of Meanings." *Advances in Health Sciences Education* 22 (4): 853–867. doi:10.1007/s10459-016-9716-3.

Tai, J., R. Ajjawi, M. Bearman, D. Boud, P. Dawson, and T. Jorre de St Jorre. 2022. "Assessment for Inclusion: Rethinking Contemporary Strategies in Assessment Design." *Higher Education Research & Development* 42 (2): 483–497. doi:10.1080/07294360.2022.2057451.

van der Vleuten, C. P. M., L. W. T. Schuwirth, E. W. Driessen, J. Dijkstra, D. Tigelaar, L. K. J. Baartman, and J. van Tartwijk. 2012. "A Model for Programmatic Assessment Fit for Purpose." *Medical Teacher* 34 (3): 205–214. doi:10.3109/0142159X.2012.652239.

Vogt, L., and S. E. Eaton. 2022. "Make It Someone's Job: Documenting the Growth of the Academic Integrity Profession through a Collection of Position Postings." *Canadian Perspectives on Academic Integrity* 5 (1): 21–27. doi:10.55016/ojs/cpai.v5i1.74195.