

Assessment as learning: A rethink of assessment in engineering education

Justine Lawson, Beata Francis, Rob Jarman and Roger Hadgraft
University of Technology Sydney
Corresponding Author Email: Justine.lawson@uts.edu.au

CONTEXT

As universities embrace 21st century pedagogies of flipped, hybrid, blended, online and student-led practices, assessment seems to lag. It remains separated from teaching and learning, and consumes a great deal of academic time and energy. Despite innovations in assessment in other sectors of education, engineering education has only relatively recently engaged with authentic and alternative assessment practices and has perhaps confined these to project-based or capstone experiences. End-point, high stakes exams, grades and marks continue to dominate assessment practices with resultant high academic workload and student behaviour skewed towards attainment of marks rather than learning.

The prominent place assessment plays in academic work prompted us to wonder: to what extent have we got it right? What could we be doing better?

PURPOSE

Our small scale study explored what a group of academics who were committed to and recognised as leaders in quality teaching, thought about assessment and how they might be provoked by new possibilities.

APPROACH

The study began with a workshop based on three provocations derived from key readings in the assessment space:

1. What if our assessment practices develop and sustain poor student performance? (Wilson & Scalise, 2006)
2. What if our assumptions about assessment are flawed? (Bloxham, 2009)
3. What if analytic (rubric) marking is indeterminate? (Sadler, 2009)

At this workshop, academics discussed their responses to these provocations and articulated their key challenges. From here, key participants took part in semi-structured interviews which were analysed for themes.

RESULTS

Outcomes from the workshop revealed a list of concerns regarding assessment including student motivation, academic workload, plagiarism and students vying for marks rather than focusing on learning. There was also a clearly articulated need for a mindshift from conceptualising assessment as activity separate and distinct from learning and towards an understanding that assessment and learning can be seen as interchangeable terms. Interview data revealed academics grapple with some of the systemic features of assessment which hamper and sometimes punishes student learning. The place of formative assessment in supporting learning was identified as key. Academics articulated the need for much more transformative assessment practices that align more strongly with progressive teaching.

CONCLUSIONS

Assessment is often at the forefront of academic work and despite innovations in teaching, remains a largely conservative practice in engineering education. Persistent and enduring challenges of student motivation, plagiarism and focus on marks can distract us from alternatives that might mitigate some of these. One key way in which transformative assessment practice might be ameliorated is to re-vision it as learning.

KEYWORDS

Assessment, learning.

Introduction

There are a lot of things to 'get right' in teaching engineering. The 'what' (content, skills, dispositions) and the 'how' (delivery modes, studios, projects) are frequently addressed and debated. Advances in these areas are well documented and indeed celebrated at forums like AAEE. Advances in assessment remain more problematic. There is a great variety of the ways in which students are assessed (reports, presentations, quizzes, exams, portfolios), by whom they are assessed (peers, self, academic) and for what purpose (formative, summative). This is overlaid with questions of what grades actually mean. Do all students with a 'pass' know and do the same things? What is the difference between students with a 65 and 67?

Some of these issues are addressed in proposed new models of assessment such as specifications grading (Nilson, 2014). However, academics continue to grapple with issues of time, student motivation and confidence in assuring learning outcomes are being met. There remain questions of whether assessment practices serve either students or academics well. When students graduate, can we, with confidence, assure ourselves that students have achieved course intended learning outcomes? Does a final grade even communicate this?

Context

At University of Technology Sydney (UTS), there is an explicit agenda to innovate in teaching and to align practices with 21st century learner needs and experiences. Academics are impelled to adopt pedagogical practices that are active, collaborative and make use of open and online resources. They are also encouraged to ensure at least one assessment task in every subject is authentic (practice-based) and that students receive regular formative feedback.

These directions in assessment, whilst promising, do not address more fundamental questions or dominant practices of whether assessment is doing what it needs to do. That is, systemic (institutional level) features of assessment (such as grading and weightings, GPA calculations) and how it is routinely undertaken (e.g. centrally conducted written exams) remain firmly in place and unquestioned. Like many other improvement agendas, small iterations or incremental changes are made in isolation. We need a rethink of assessment and perhaps more disruptive and transformative practices.

There are several disconnects pertinent to this discussion. First is the separation of learning from assessment. There are institutional policies that separate learning from assessment, and students and academics alike will distinguish learning and assessment activity. Second, feedback and assessment as practised at university is different to that which will be experienced in the workplace. That is, despite assessment tasks approximating professional activity, such as a design report, the grading of this work, or proof of individual effort in collaborative tasks, is not mirrored in professional practice. Finally, there is also perhaps a fundamental mismatch between outcomes based education and graded assessment. Students achieve or not yet achieve outcomes; they don't 'very' achieve them as implied when awarded a high distinction. If they do exceed outcomes, they are likely to be achieving different or more advanced outcomes altogether.

Many academics within our faculty are interested in grappling with these big questions around assessment. The following reports on findings from a small-scale study with these academics. We held a workshop to address some fundamental questions of assessment and then followed up with more in-depth interviews.

Provocations

We began our exploration with a workshop framed around three provocations. These were derived from three papers each designed to challenge commonly held beliefs about

dominant assessment practices. Participants were asked to read one or all of the papers before attending the workshop.

Provocation 1: What if our assessment practices develop and sustain poor student performance?

Wilson and Scalise (2006) posit that when a student underperforms in a subject, academics (and sometimes the student) will attribute this to lack of ability or application of the student. An alternative explanation is that enshrined practices such as not providing timely formative feedback not only allows for poor student performance, but sustains it. That is, a student who receives no feedback on a mistake in learning will go on to practice that mistake over and over, only to discover too late (i.e. in a final exam) that they don't know or can't do something. This has less to do with the student and more to do with a system that lets them underperform unchecked.

At UTS, all subject coordinators must specify when students receive formative feedback and at least one of these must occur before census date. (This, of course, assumes that early evaluation of performance should inform a student's decision to proceed with the subject rather than inform the academic that changes might be necessary.) However, this is likely to be only part of the picture. Wilson and Scalise (2006, p. 637, emphasis added) add:

Students tasks needed to be aligned, or on target, with learning goals, and students receive meaningful and timely feedback on their performance **as well as targeted follow up work.**

Often, academics have planned the term or semester's work and "after receiving grades, students typically move on to a new topic and work for another set of grades" (Bransford, in Wilson and Scalise p. 643). This limits the extent to which targeted follow up work can take place.

Sometimes feedback is blurred when performance is conflated with effort. We hear anecdotally that these early tasks are set to motivate students, impel attendance and engagement rather than to enter a dialogue about learning.

Wilson and Scalise go on to propose an entire system for assessment reform. The principle of matching instruction and assessment is of greatest interest to us here. This is consistent with, but goes beyond Bigg's (1996) notion of constructive alignment. It suggests that the frameworks for assessment and instruction must be one and the same (Wilson & Scalise, 2006, p. 652). Further, if done well, where opportunities for formative feedback are built into teaching activity, 'the richness and vibrancy of curriculum development (are integrated) into assessment) and the discipline and hard-headedness of assessment data (are integrated) into the design of instruction' (p. 652).

Provocation 2: What if our assumptions about assessment are flawed?

As suggested in the introduction and context sections, we often iterate for improvement at the periphery of assessment regimes and leave unquestioned the substantive practices of marking and grading. Bloxham (2009, p. 209) suggests that the hours academics invest in marking and moderation practices do little to enhance either reliability or accuracy.

Bloxham presents and challenges four commonly held assumptions (myths) about assessment practices in higher education. The context is the UK, with strong applicability to the Australian university sector, and are worth restating here:

- a. We can accurately and reliably give a mark to most students' work.
- b. Even if individuals' marking may sometimes be inaccurate, internal moderation ensures fair and appropriate standards in marking.
- c. Even if internal moderation does not reflect expected standards, external moderation ensures students are assessed against consistent standards across the UK university sector.

- d. Students' final award (degree classification) reflects their achievement in a consistent way within and, to a certain extent, across universities.

(Bloxham, 2009, p.209)

Assessment here is seen as a resource intensive practice which is largely disconnected from learning. Bloxham argues that the assumptions above go largely unquestioned but should be debated because they are, at best, unverified, and yet to underpin marking and moderation practices.

Provocation 3: What if analytic (rubric) marking is indeterminate?

Consistent with the above challenge to marking practices, we asked participants to question if rubrics deliver on their promise of fairer and more transparent marking. Royce Sadler (2009) reminds readers of the high stakes of assessment for students and points to the failure of grading schemes (such as rubrics, scoring guides, criteria sheets) to 'meet the conditions for sound assessments of complex student works, and that this deficiency is inherent in the method' (p. 159).

As argued elsewhere (Bearman et al., 2016) there is often a mismatch between intention and practice in assessment. Sadler offers a series of observations attesting to this suggesting that assessment in practice is inherently more complex than students might expect. He presents several behaviours of academics undertaking marking using analytic grading, such as marking holistically and pragmatically adjusting criteria to fit the overall impression. He also highlights the common phenomena of markers agreeing on an overall grade but not at the level of individual criteria. Students might reasonably expect that their work has been considered in relation to separate criterion rather than what is the reality – a combination of global and analytic marking.

Findings

Workshop

As expected, the above articles sparked lively debate within our workshop as 18 academics discussed how the articles resonated or jarred with their experiences or practices. Academics were asked to write their most front-of-mind issues with assessment. One idea per sticky note was recorded. These concerns were then clustered. They fell broadly into four areas. See Table 1.

Participants were also keen to engage in discussion about possibilities of 'front-ending' assessment, increasing student self-assessment and co-construction of assessment tasks with students. In this group, there was a willingness to be creative and innovative in assessment practices, but also a recognition of the constraints or barriers.

Participants theorised student motivation. Many identify students who seek the minimum requirements for a passing grade, students who privilege mark acquisition over learning, surface over deep approaches. There was also speculation that students who vie for marks (alone) are more likely to engage in academic misconduct. These mark-oriented students contrasted with those with more positive attitudes to learning. An observation was shared that students motivated by learning can be negatively affected by grades. In both cases, grading possibly disengages learning.

Assessment purposes	Resources and marking
<ul style="list-style-type: none"> • Is it to assess or assure learning? • Assess process (reflect a growth mindset) • Single grade marks tell us nothing about what is learnt 	<ul style="list-style-type: none"> • Chasing external (casual) markers • Exam scheduling – adds to workload • Spend A LOT of time on marking • Pressure to return marks in timely manner

<ul style="list-style-type: none"> • What do students need to know and do? • Assessment is sometimes punitive • We work in isolation in assessment design • How can we create assessment to meet graduate outcomes? 	<ul style="list-style-type: none"> • Large classes limit timely feedback • Couldn't we find better IT systems to aid assessment?
<p>Feedback</p> <ul style="list-style-type: none"> • What is the quality of feedback we give to students? • Challenge of providing feedback in a timely manner before subsequent or final tasks are due • How to give students regular and useful feedback? • To what extent do students value or engage with feedback provided 	<p>Students</p> <ul style="list-style-type: none"> • Motivation of students – minimal time on tasks • Attitudes: learning versus marks • Misconduct/plagiarism • Cheating • Negative effects of grades

Table 1: Sample of assessment issues from sticky notes

Interviews

Following the workshop, participants were invited to take part in short semi-structured interviews. Questions included:

- What is a fundamental thing that matters to you in your subject?
- How do you assess that?
- Where do you set the bar?
- How do you know that all students have passed it?
- What do you see as the issues with assessment?
- If you could reimagine assessment, what would you do?
- What are some of the issues that might need to be addressed?

There were four main themes to emerge from the interviews. First, the tension between formative and summative assessment was raised, and relatedly the need for reconceptualising assessment as learning. All participants noted the value of formative assessment, and one explained how feedback for students' online questions was a regular feature of tutorials. Similarly, for a 'hurdle' task, understanding and drawing free body diagrams, students were given multiple attempts to reach the 75 per cent threshold pass mark. Teaching and discussions were tightly focused on achieving outcomes. Where formative assessment features like this in teaching, it is more than 'constructive alignment'. It is a blurring of assessment and learning. Assessment becomes the learning, or as one academic put it, "it's the conversation."

For another there was no separation between the activities:

"It's all about lifelong education...if we're going to help them launch into where they want to be, I think we've got to help them make decisions and then if they can make that decision, they can actually start to concentrate on things so they can do more work in that area."

Second, systemic constraints on innovative practices in assessment were identified. Consistent with ideas raised in the workshop, interviewees readily articulated those broader issues (accreditation, awarding grades, and limited resources) that curtailed more large-scale or fundamental improvements to assessment. One academic also pointed to the constraints of having predetermined outcomes for all students. He argued that "students should be able to have different outcomes and be judged against those different outcomes." Instead, we

have prescribed order, content and assessment. Where there is negotiation or co-construction of curriculum between academics and students, it is isolated and almost never changes the set outcomes for a subject or course.

Third, the interviewees noted the problematic nature of marking. One academic worried that academics readily 'flick pass' marking to say PhD students, and therefore see marking as a mechanical exercise and not one which is a process through which the academic can learn. This academic went on to say that it is foolish to think that assessment is anything other than a judgement but that:

"we need to bring students along in that as well. If we're trying to develop them as professionals, being a professional is about exercising your professional judgement. So opportunities for students to assess themselves, to assess each other."

This sentiment was echoed by another: "People have to make a holistic judgement." Two academics suggested that when we mark and assess, we fragment learning into bits (either a series of things required by Engineers Australia, or discrete criteria in rubrics) and obscure the larger, more important picture of development of the whole engineer.

Fourth, the social dimension of assessment and the need to motivate students rather than punish them emerged. As suggested above, assessment cannot be conceptualised simply as a mechanical exercise, and the importance of academic judgement was identified by those interviewed. This suggests assessment is a social practice where complex relationships around assessment are constructed. Several academics worried about students and the impact of grading:

"I think redoing things or allowing people to learn without failing, or without having failure recorded, would be a good thing."

This academic suggested that when you accumulate (sometimes low) marks from early in the semester, even if you go on to excel in the final task, you are 'punished' by your earlier work by having all marks count. Likewise, fail grades are recorded on transcripts even when you eventually succeed in being awarded a degree. These practices were seen as punitive, punishing and ultimately harmful to students and their learning. Arguably, academic transcripts are another systemic constraint. They are a record of enrolment (what subjects were undertaken and when), but also include marks and grades that may inadequately convey learning outcomes.

This notion of developing safe learning environments was reiterated by another:

"The most important thing we do with the students is give them feedback on the work that they've done. So opportunities to have conversations about the work that students have done...well multiple opportunities, so the students get a chance to practice without the wheels falling off...It gives people who learn at different rates a chance to learn it...I mean if they don't get it in week 2, why should they be punished then? That's why they've got multiple opportunities."

The interviewed academics more readily problematised assessment practices as they impacted on students than some participants in the workshop who in some cases, reverted to deficit explanations for students underperforming.

Discussion

This study was enlightening in several ways. It demonstrated that there is a preparedness to debate some commonly held truths about assessment and challenge some fundamental practices. Some academics feel forced to work with assessment in ways that do not resonate with their deeper philosophical beliefs about learning and work creatively to soften what they see as some of the harsh realities of assessment. Some for example, allow for multiple

attempts on a task and don't record early failure. Particularly for those interviewed, there was a deep worrying about the impact assessment has on students and the need to privilege formative feedback over summative assessment.

Further, where formative feedback is emphasised, there is a clear repositioning of assessment as learning, where teaching and learning activity is the assessment activity or vice versa. Re-envisioning assessment as learning is aligned with Wilson and Scalise's (2006) notion of embedded assessment where 'activities are embedded or become part of, class learning activities' (p. 645). Interviewed academics had no trouble at all in seeing assessment in this way. In fact, they found it hard not to.

Finally, the subjective and social nature of assessment was a clear theme and was not seen as something to be challenged, but rather accepted. This is consistent with observations elsewhere (Bloxham, 2009; Lawson et al., 2015). This means that assessment as a process needs closer examination.

Conclusion

Back to our opening questions, with regard to assessment: to what extent have we got it right? What could we be doing better? It would seem that the intentions of the academics we worked with are certainly aligned with instigating better practice. There is a promising willingness to tackle and grapple with difficult questions about our current practices. The lived experience for students and academics, however, suggests that we have some way to go to get assessment 'right'. This will involve confronting and reviewing long-standing and quite fundamental practices. The starting points perhaps are a re-visioning of assessment as learning, and importantly, consistent with a key theme of our findings, involving students in the journey to change.

References

- Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E., Boud, D., & Joughin, G. (2016). How university teachers design assessments: a cross-disciplinary study. *Higher Education*. DOI: 10.1007/s10734-016-0027-7. Published online, July 2016.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.
- Bloxham, S. (2009). Marking and moderation in the UK: false assumptions and wasted resources. *Assessment and Evaluation in Higher Education*, 34 (2), 209-220.
- Lawson, J., Rasul, M., Howard, P., Martin, F., Hadgraft, R & Jarman, R. (2015). Getting it Right: The Case for Supervisors Assessing Process in Capstone Projects. *International Journal of Engineering Education* Vol. 31, No. 6(B), pp. 1810–1818, 2015
- Nilson, L. (2014) *Specifications grading: Restoring rigor, motivating students and saving faulty time*. Stylus
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment and Evaluation in Higher Education*, 34 (2), 159-179.
- Wilson, M. & Scalise, K. (2006). Assessment to improve learning in higher education: The BEAR assessment system. *Higher Education*, 52, 635-663