An Empirical Analysis Exploring the Impact of Traditional Exams and Multi-Stage Assignments on Academic Workload in a Final Year Engineering Context

Practitioner Research
In Higher Education

Vol 14(1) pages 16-27

Tom O'Mahony
Cork Institute of Technology, Republic of Ireland

Abstract
This article documents the time taken to assess two different types of final year engineering courses – a more traditional course in which the purpose of assessment is mostly to measure performance, and a second, that uses a multi-stage assessment process to implement an assessment for learning perspective. The research question is to determine what impact these different assessment models have on instructor workload (as measured by time). A quantitative research methodology was adopted. The time taken to assess both courses was carefully recorded during one semester. The main finding was that the multi-stage assignment with feedback originating from the instructor, increased the instructor's assessment workload by 23% and also resulted in a temporal shift in that workload. The data collected was confined to two courses and this may limit the generalisability of these findings. However, the data is consistent with and supports recent calls to reconceptualise feedback in higher education.

Keywords
Assessment practicability; assessment *for* learning; multi-stage assignments; assessment workload; quantitative research.

Introduction

Assessment is increasingly understood as having a significant, albeit frequently unrealised, influence on learning (Wiliam, 2011). To actualise its potential the assessment discourse argues for a number of changes. One such recommendation is that learners act on the feedback that is generated to enhance learning (Handley et al., 2008; O'Donovan et al., 2016). Two-stage or multi-stage assignments are often advocated as a means of addressing this recommendation, as feedback from initial stages can be directly applied to final submissions (Carless et al., 2011; Boud and Molloy, 2013; Winstone et al., 2016). Notwithstanding other voices (Sadler, 2013; Boud and Molloy, 2013; Carless et al., 2018), in adopting two-stage assignments, instructors frequently assume responsibility for providing feedback (Scaife and Wellington, 2010; Cranny, 2017). The workload that ensues has generated some debate, with some sources claiming that workload is increased (Carless et al., 2011) while others suggest that it is unaffected (Prowse et al., 2007). However, little empirical evidence exists to support either opinion.

This article aims to contribute to this issue by quantifying the time taken to assess two final year engineering courses with different perspectives on assessment – a measurement perspective and a *for learning* perspective. The main contribution, then, is to investigate the practicability of instructor generated feedback in multi-stage assignments. Specifically, this research documents how much more instructor time is required to assess these types of assignments when compared with the more traditional unseen terminal examination. By explicitly documenting this time-commitment, it is hoped that this article would encourage instructors to make more informed decisions when adopting multi-

Citation
O'Mahony, T. (2021) 'An Empirical Analysis Exploring the Impact of Traditional Exams and Multi-Stage Assignments on Academic Workload in a Final Year Engineering Context', *Practitioner Research in Higher Education Journal*, 13(2), pp. 16-27.

stage assignments, and in particular, nudge instructors to explore alternative ways of generating formative feedback.

Throughout this contribution, the verb *to assess* (and its variants) is used quite generally and it can take on quite different meanings. In the measurement context, *to assess* refers to the process of reviewing assignments, making a judgement, assigning a grade and providing brief summative comments which mostly serve to justify the allocated grade. In the assessment for learning context, the same verb mostly refers to the process of reviewing assignments, making a judgement and generating formative feedback i.e. feedback which serves to improve and accelerate learning (Sadler, 1989). At the end of such a course, *to assess* may revert to the measurement perspective.

A Review of Some Relevant Literature

Perspectives on Assessment

Few of the main players are satisfied with assessment in higher education. Students report that too often feedback is late and does not facilitate learning (Carless, 2006; Orsmond, 2011). Staff claim that students fail to act on the feedback provided (Handley et al., 2008; Weaver, 2006). Advocates for assessment claim that current practice in higher education it is "not fit for purpose" (Carless et al., 2011), failing to fulfil its potential (Knight and Yorke, 2003), and that it is in "disarray" (Knight, 2002). As a consequence, over the past ~15 years there has been a drive to enhance assessment as it is enacted in higher education (Nicol and Macfarlane-dick, 2006; Carless, 2007; Gibbs, 2010; Sambell, 2011; Y1Feedback, 2016; Chase et al., 2017). There is evidence of this sustained interest having an impact, but that evidence also indicates that students are still less satisfied with assessment than any other aspect of their higher education experience (HEFCE, 2014).

The resulting discourse has argued for a shift in how assessment is conceptualised. Traditionally, assessment is seen as something that mostly relates to a quality assurance and grading function. Academic staff are now encouraged to consider assessment as enhancing learning (Evans, 2013; Wiliam, 2011). This requires practitioners to understand that assessment is a process that learners engage with, as opposed to a product that is received (Carless and Boud, 2018; Orsmond et al., 2013; Winstone et al., 2016). Similarly, feedback needs to be understood as information that learners interact with to produce an identifiable impact on learning (Carless and Boud, 2018; Wiliam, 2011). Models and principles exist to enable this new view to be enacted (Nicol and Macfarlane-dick, 2006; Gibbs, 2010; Orsmond et al., 2013; O'Donovan et al., 2016). In general, this research advocates for authentic assessment tasks, a shared understanding of assessment criteria and standards, quality feedback and opportunities to act on that feedback. Key enablers are student engagement (Carless and Boud, 2018; Sambell, 2011; O'Donovan et al., 2016;), feedback that is a dialogic process (Nicol, 2010) and learner's ability to judge, evaluate and regulate their own learning (Nicol, 2014; Carless et al., 2018).

Acting on Feedback

While improving assessment in higher education is a multi-faceted problem, one of the issues identified in the literature is the fact that learners are unable to do anything much with the feedback that is provided. Within a singular course, Duncan (2007) reports on how students lamented the absence of opportunities to implement changes (and improve grades) in response to feedback information. Similar frustrations were identified by, for example, Pokorny and Pickford (2010), Orsmond and Merry (2011), Vardi (2013) and Hepplestone and Chikwa (2016). The cause, in part, is the trend in higher education towards modularised structures which reduce "opportunities for feedback as tasks get squeezed into fewer and fewer weeks" (Boud and Molloy, 2013, p. 699).While it can be argued that information provided could be utilised in subsequent courses, students in the Hepplestone and Chikwa (2016) study made it explicit that much feedback was difficult to use in future assignments because "it's a very unique piece of work, and it won't be able to be used for anything

else" (*ibid*. p.8). Likewise, Prowse et al. (2007) suggest that deferring the "implementation of such advice to subsequent modules reduced its relevance and weakened its impact" (*ibid*. p.438). The net consequence then, is that the effort learners make to remember, and understand, the feedback they receive is often wasted.

Gibbs and Simpson (2004) offer a number of ways by which learners can be encouraged to attend to and act on the feedback that they receive – one of which is the use of two-stage or multi-stage assignments. A multi-stage assignment consists of two or more related phases or tasks in which feedback from the initial phase(s) can be used to enhance the quality of work (and assumedly learning) of subsequent phases. Specifically, Scaife and Wellington (2010) provide this operational definition for two-stage assignments:

- Phase 1: the student's assignment is assessed formatively by the tutor. The student receives detailed guidance but no grade.
- The student uses the tutor's feedback to improve the assignment.
- Phase 2: the student submits a revised assignment, in which the revisions are clearly shown. This is assessed summatively by the tutor. The student receives a grade with summary feedback.

However, the exact implementation can vary. For example, in the implementation discussed by Prowse et al. (2007), the formative and summative elements are mixed. The initial phase or draft is graded and feedback provided. This feedback can be used to revise the submission in phase 2, but a limit is placed on the maximum additional marks that can be earned in this phase. The authors suggest that this format encourages learners to "avoid submitting speculative first attempts" and then relying on feedback to develop the work. Regardless of the format, an advantage of multi-stage assignments is that likely problems are dealt with pre-emptively rather than summatively. This affords learners real opportunities to learn from their mistakes, not just through being advised of the mistake, but by being offered the opportunity to revisit and revise (O'Mahony, 2017). The obvious relationship between feedback and subsequent tasks is a particular advantage as it motivates learners to engage with the assessment process, and in particular to act on the feedback. There is evidence that learners do act on this feedback. For example, in the Prowse et al. (2007) study, learners averaged a grade increase of 2.4 points (out of a maximum of three) in the revised or phase 2 submission. Finally, the two-stage or multi-stage assessment process encourages timely feedback, and evidences how learners act on the feedback received. In summary, multi-stage assignments offer real potential to address a key issue associated with assessment, namely, that too much "feedback in higher education comes too late for it to be useful to students" (Carless, 2007) and consequently the multi-stage assessment process is frequently recommended in the assessment discourse (Handley et al., 2008; Carless et al., 2011; Boud and Molloy, 2013; Winstone et al., 2016).

Practicability of Multi-Stage Assignments

While there are clear advantages, the instructor workload associated with multi-stage assignments is a concern. The workload is particularly acute when the instructor assumes responsibility for generating individual feedback. This would appear to be common practice (Carless, 2007; Scaife and Wellington, 2010; Cranny, 2017; O'Mahony, 2017). The practitioner literature on multi-stage assessments mostly reports on implementations with classes of less than 50 students (Prowse et al., 2007; Carless, 2007; Harney, 2017; Cranny 2017; O'Mahony 2017). This would raise questions as to the scalability of the approach. Indeed, Boud and Molloy (2013) criticise the assessment *for* learning framework in general, by observing that the "practical dilemma of higher education is that the amount and type of feedback that can realistically be given is severely limited by resource constraints" (*ibid.*, p.703). More specifically, Carless et al. report that one of their participants "acknowledged the heavy marking workload generated by the two submissions of the portfolio" (Carless et al., 2011, p.401). In

direct contrast, a participant in Gibbs' study reported that "commenting on the 1,000-word drafts did not take long as many of the students made similar mistakes" (Gibbs, 2010, p.37). In the case study presented by Prowse et al. (2007), the authors noted a general concern about increases to staff workload but claim that this "was addressed by the mechanism of revisiting only highlighted passages in the amended work", noting that this "worked very well" (*ibid*, p.442). It is evident that this commentary is conflicting and poorly supported by data. This article aims to inform this issue by generating some empirical data to evidence the workload associated with the multi-stage assessment process, as it is commonly operationalised, relative to more traditional end-of-semester examinations.

Method

Methodology

This small-scale research adopted a quantitative research method derived from the post-positivist research epistemology (Cohen et al., 2011). A post-positivist perspective acknowledges the existence of an external reality and the appropriateness of empirical methods in coming to know that reality. However, it also recognises that all observations are fallible and susceptible to bias. A quantitative approach was selected as it was the most appropriate fit for the research question – *what impact does assessing multi-stage assignments have on instructor workload, as measured by time, when compared with a more conventional exam-based assessment process?* In this case, the multi-stage assignment predominantly relied on instructor feedback – and it is acknowledged that other forms, that focus on self or peer feedback, may have different outcomes.

Participants

The two courses that comprise this study were delivered by the author during Semester 2 (January – May) of the 2017-18 academic year. Both courses are worth 5 European Credit Transfer System (ECTS) credits, are final (4^{th}) year engineering courses in the same sub-discipline (control engineering) and have the same number (4) of contact hours per week. The course employing the multi-stage assessment process was designed by the author. 11 students participated in this course during the 2017-18 academic year. On this course, learners worked in teams to complete a design problem over the course of the semester. The assessment process required learners to maintain a website to document and share contributions to the team process (worth 20%), to complete a series of short answer or multiple choice questions each week (worth 10%) and to write an individual conference-style paper (IEEE, 2018) that was worth 70%. The objective of the paper was to document the design and implementation process and summarise outcomes. Two drafts of the paper were submitted, one at week 6, the second at week 11 and the final version was due on week 15. Individual instructor feedback was provided on both drafts via short annotations on an assessment rubric and detailed audio feedback recorded as an MP3 file. Exemplars, representing past work, were provided via the institutes Virtual Learning Environment (VLE). Selected exemplars were discussed and graded in class to support the development of a shared understanding of the assessment rubric and therefore, to facilitate the development of self-assessment. These various elements, the assessment rubric, exemplars, weekly questions, in-class assessment workshops and instructor feedback all represent formative assessment elements that are designed to feedforward to the final summative paper.

The author was assigned to teach the second course during the 2017-18 academic year but was not involved in the design of the course. It adopts a more traditional lecture-based approach to teaching and an exams-oriented approach to assessment. During the 2017-18 academic year, 27 learners participated in this course. In this case, the assessment process was principally concerned with measuring and grading students. The laboratory component was assessed via three laboratory reports, which combined were worth 15%. The other assessment components consisted of a mid-term closed-book written assessment (15%) and a terminal closed-book written assessment which was worth 70%. The laboratory reports were marked based on a grading sheet and short evaluative written comments were included in the grading sheet which were returned to the students. Short written

comments were also included on assessment scripts associated with both the mid-term and end-of-term assessments. Scripts from the mid-term assessment were returned to learners and an hour of contact-time devoted to providing feedback by discussing common errors and addressing individual questions. Of the assessment components, this is the only one that the author would consider to have a "formative" function because the feedback provided has the potential to be acted upon in the end-of-semester examination. In contrast, each of the three laboratory reports focus on a different learning outcome and therefore the feedback, which is mostly technical and corrective, does not easily transfer. The author would also argue that because of the different contexts, the laboratories being more practical and problem-oriented and the written exams being more theoretical and routine, the feedback may not easily translate from one context to the other. Hence while feedback information is provided, it might be considered more as "dangling data" (Sadler, 1989) rather than something than can be easily acted on. Table 1. provides a summary of these different contexts.

Table 1. Key features of the courses that comprise this study.

	Course A	Course B
Programme	BEng. (Hons) in Electronic Eng.	BEng (Hons) in Mechanical Eng.
Credits	5 ECTS	5 ECTS
No of students	11	27
Assessment process	Multi-stage assignment	Exams plus laboratory reports
Assessment elements	Team Website (20%) Weekly SAQ/MCQ (10%) Individual paper (70%)	Laboratory reports (15%) Mid-term assessment (15%) End-of-term assessment (70%)
Formative assessment	Website: feedback on work-in-progress on week 2 & week 8 Paper: feedback on drafts at week 6 & week 11	Feedback on mid-term assessment

Data Collection

The data that was collected consisted of the actual time spent assessing student work. For Course A (see Table 1), all of the assessment material was electronic in format and submitted through the institute's VLE. Individual papers (including draft) were downloaded and the time from when an individual assignment was opened to when it was closed was recorded. This time would include reading the assignment, grading the assignment via the assessment rubric, recording oral feedback and processing that as an MP3 file. If this process was interrupted the time duration of this interruption was also carefully recorded. A similar process applied to the websites, but in this case the time required to assess each website (which typically represented a group of four students) was recorded. Again this time included reviewing the websites, grading the websites and providing feedback on the websites.

The main assessment element for Course B (see Table 1.) was hand-written examination scripts resulting from the mid-term and end-of-term assessments. The scripts were graded question by question or for the end-of-term assessment, one question component e.g. Q1 part (a) at a time. In order to minimise the bias that might arise from fatigue, the author would typically take a short break every hour when grading examination scripts. Based on past experience, it would usually require one hour to assess a single question from 10 scripts. Hence, the process adopted was to select 10 scripts, record the time, review and grade the particular question on one script and repeat this process for

the 10 scripts. After the tenth script, the time was again recorded. This time included brief annotations on the actual script e.g. tick marks or comments to the effect "No, this is incorrect. See model answer" but did not include administration time e.g. formally recording grades on the institute's system or communicating results back to students, etc. The laboratory reports associated with Course B were submitted electronically and the time required to assess each individual lab report was recorded using the same process as was used to record the times for the individual papers in Course A. All times were recorded electronically in a word-processing document.

Analysis

To minimise numerical errors, the start/end times were copied from the word-processing document into a spreadsheet and the duration calculated for each assessment component. Summing these provided the *total time* required to assess each component and subsequently the arithmetic mean was determined. This gives the *average time taken per student* to assess each of the assessment components. The *total time per student* was then calculated by summing the times for each of the individual assessment components.

Results

Tables 2 and 3 present the empirical results that this study generated. Both tables record the total time spent assessing each student, along with the time spent assessing each component per student and the contribution of each assessment component to the overall student grade. A direct comparison reveals that the multi-stage assignment with instructor feedback utilised in Course A (table 2) requires a 23% increase in time-investment per student. Examining the breakdown, it is evident that the time spent assessing is dominated by the second draft and final versions of the individual paper (45.3 and 38.4 minutes per student respectively) for Course A and by the terminal exam for Course B (44.3 minutes per student). Significantly, the second draft is submitted during the teaching semester and the assessment process for this component needs to be completed in a timely fashion such that learners have a reasonable opportunity to act on the feedback. This implies that the assessment process for this component must also happen during the teaching semester and this limits the viability and scalability of the approach. To illustrate this problem, Figure 1. graphs the assessment times for both courses chronologically over the course of the semester. For comparison purposes, it was assumed that 35 students take each course and figure 1 is generated based on this assumption using the data from Table 2. and Table 3. The timing in Figure 1., is representative of when the submissions are typically assessed (by the author), but varies with workload demands in any particular year. While both models incur a significant workload in week 15/16, this workload happens at a time when there is no face-to-face teaching which makes the workload viable. The multi-stage assignment however, also requires ~25 hours of assessment time (assuming a class size of 35 students) during week 12 when classes are ongoing. Creating this time in the midst of a busy teaching week is always a challenge and, arguably larger class sizes could only be accommodated by allocating additional resources.

Table 2. Time taken to assess each component for Course A (see Table 1.).

Assessment Component	Average Time taken per Student (Minutes)	Contribution of Assessment Component to Grade
Draft No 1	16	0%
Draft No 2	45.3	0%
Final Paper	38.4	70%
Website (Draft)	9.5	0%
Website (Final)	4.5	20%
Total	113.7	100%

Table 3. Time taken to assess each component for Course B (see Table 1).

Assessment Component	Average Time taken per Student (Minutes)	Contribution of Assessment Component to Grade
Terminal Exam	44.3	70%
Mid-term Assessment	13.4	15%
Lab Report 3	8.4	5%
Lab Report 2	15.2	5%
Lab Report 1	11.2	5%
Total	92.5	100%

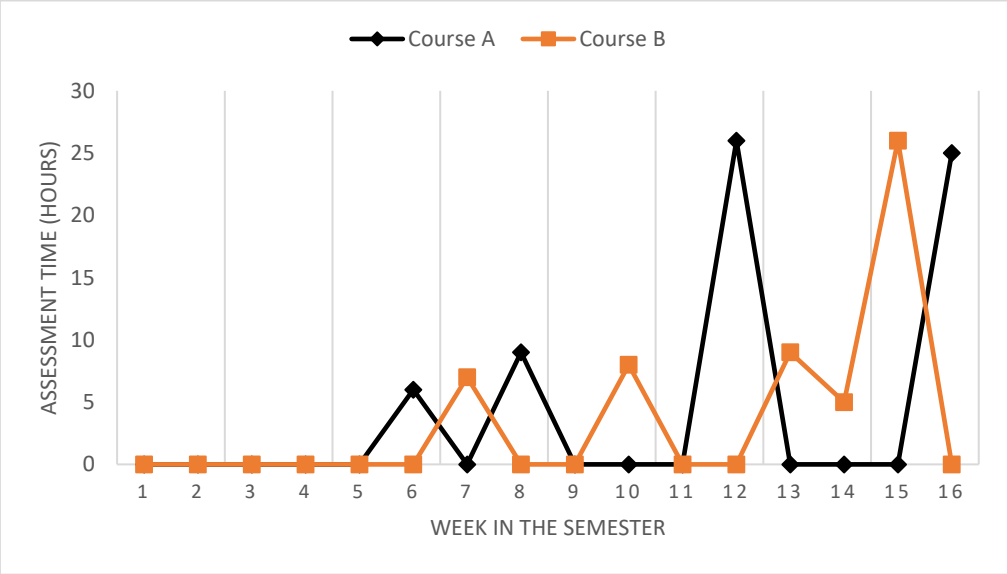

Figure 1. A chronological representation of assessment workload for both courses over the duration of one semester *assuming* an equal class size of 35 students.

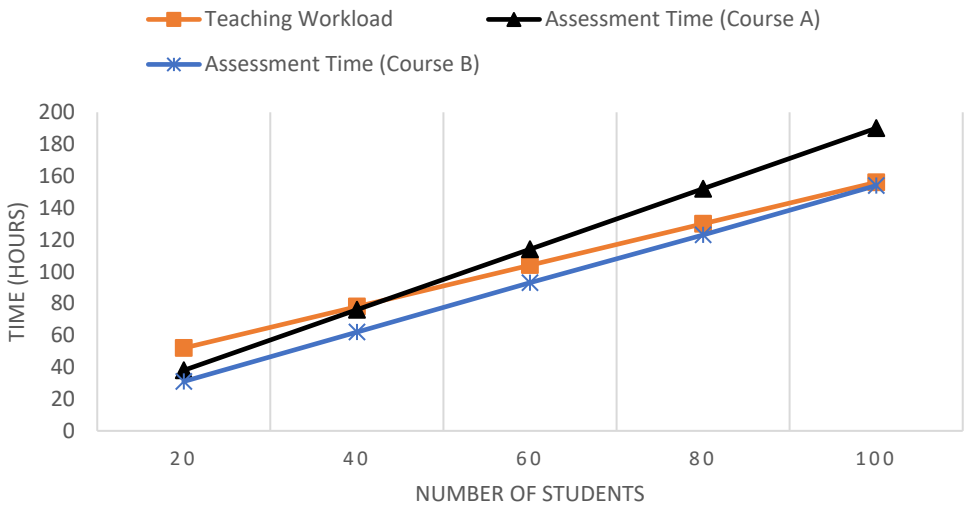

Figure 2. Face-to-face teaching and assessment workload as a function of N, the number of students taking the course.

The scalability issue can also be explored by using Table 2. & 3. to predict workload as a function of the number of students (N) in the class. Figure 2. graphs this trend in comparison to the face-to-face teaching workload for the same number of students. In calculating the face-to-face teaching workload, it was assumed that the course was structured as a two hour lecture and a two hour laboratory component, which would be typical for engineering courses at the author's institution. At the author's institution, the norm is that the maximum number of students in a laboratory is 20, and therefore the teaching workload scales as a multiple of 20. Figure 2 illustrates that, while the assessment workload for Course B (exam-based course) approximately tracks the teaching workload as N varies from 20 to 100, the same cannot be said for Course A. In the case of Course A (multi-stage assignment), once the cohort size increases beyond 40 students the assessment workload becomes greater than the face-to-face teaching workload. Furthermore, examining the trend, it is clear that as N increases the assessment workload for Course A continues to diverge from that of Course B and the face-to-face teaching workload.

Discussion

Multi-stage assignments can both enhance student learning generally (Handley et al., 2008) and directly impact grades (Prowse et al., 2007). Previous research by this author confirms both the general learning impact (O'Mahony, 2018), and a specific increase in grades (O'Mahony, 2013). This prior research was conducted in a context similar to the Course A considered in this paper, and while these benefits were not evaluated here it is believed that they apply to this course. The focus therefore of this article has been on the cost, which has not been explored and empirically documented. In this article, the multi-stage assessment process outlined for Course A has been shown to increase the assessment workload by 23% relative to Course B (a more traditional exams-based course). In Course B, 38% of the assessment time was spent grading and providing feedback on laboratory reports which were only worth 15% of the final module grade, and efficiencies here would make the multi-stage assignment even less attractive. Given the already heavy workload of many academic staff in higher education, there is a need to question whether this increase can be justified? Figure 2. suggests that multi-stage assignments – where the instructor generates the feedback - are only practicable with small class sizes. This finding is consistent with the practitioner literature on multi-stage assessments which mostly reports on implementations with classes of less than 50 students (Prowse et al*.,* 2007; Carless, 2007; Harney, 2017; Cranny 2017; O'Mahony, 2017).

These findings are not intended to suggest that multi-stage assessments have limited value in higher education or that they are only viable with small class sizes. Rather these findings question how multi-stage assessments are typically operationalised – i.e. the instructor assumes responsibility for generating feedback. This common conceptualisation of feedback places strict bounds on the practicability of the approach. Hence, the empirical data and findings presented in this article support the position of Sadler (2013), Boud and Molly (2013), Nichol (2014) and Carless and Boud (2018) who argue for a re-conceptualisation of the assessment process. These authors consistently recommend that the burden of feedback shift from the instructor onto learners and that practitioner's support learner's capacity to generate quality feedback though self-assessment and peer review. In order to achieve this, the use of exemplars, dialogue and peer review is strongly recommended (Nicol, 2014; Carless & Boud 2018). This literature argues that engagement with different representations of quality, interpreting those representations in the context of the assessment criteria and standard and generating critical commentary all help to develop evaluative judgement and the capacity to self-assess. Developing this capacity to self-assess and supporting a peer review process represents an obvious solution to the assessment workload issue documented here. In the next iteration of Course A, the intention is to develop and evaluate a peer review process as an alternative to instructor feedback. A key issue to be unravelled in this model is trust (Carless, 2009) as learners will only act on

the basis of trustworthy information (Handley et al., 2008) but trust is "hard won" (Boud and Molloy, 2013). Empowering individuals and developing a sense of mutual trust, is a key challenge that will need to be addressed to realise the full potential of peer-reviews.

While the research presented in this article evidences an increase in assessment workload of 23%, the generalisability of this result can be questioned. This is the main limitation of the study. The discipline (engineering) and nature of learning (systematic and logical), the assessment requirements and instructor characteristics are all likely to influence the data that was recorded and, consequently, the main result. It is encouraging to see that the data presented in this paper is broadly consistent with other reports in the literature (Voelkel and Mello, 2014; Posa-Lujan et al., 2016). For example, Posa-Lujan et al. (2016) report spending 15 hours correcting a terminal exam and 67.5 hours teaching. This equates to spending 18% of the total course time on assessment. If Course B was to be assessed only by a terminal exam, 23% of the total course time would have been spent on assessment. Posa-Lujan et al. (2016) also present numbers for a modified course which includes continuous-evaluations, consisting of regular exercises. Conceptually, this is not unlike Course B. For their modified course, Posa-Lujan et al. (2016) spent 40% of the total course time on assessment, while the comparable number for Course B is 39%. There are also issues with the reliability of the main result. Given the scale of the study (two courses, over one semester with a small number of students on both) it is not possible to argue or claim that the result is likely to be stable. A larger scale study, involving a greater diversity of courses, and academic staff would be required to establish both the reliability and generalisability of the findings. So while the actual result established in this paper (an increase of 23%) should be interpreted with some caution, the data and findings do support the argument that conceptualisations of feedback which position the instructor as the main provider of feedback are unsustainable (Boud and Molloy, 2013).

Conclusion

Much of the existing discourse on the practicability of multi-stage assignments (where feedback predominantly originates from the instructor) is anecdotal and contentious. For example, Carless et al. (2011) note the heavy marking workload associated with multi-stage assignments while Prowse et al. (2007) indicate that the workload is quite manageable. This article intends to ground this debate by generating some empirical data. The data compares the assessment effort, as measured by time, for two courses which have adopted different assessment philosophies. Course A aligns well with the assessment *for* learning literature and utilises a multi-stage assignment as the main assessment instrument. Course B would be more consistent with assessment *of* learning and predominantly uses unseen written exams to measure learning. The main findings are that there is a significant cost implication associated with the adoption of a multi-stage assignment – a 23% increase in the assessment workload, relative to Course B. This increase in instructor workload might be feasible, if it happened at the end of the teaching semester, and resulted in a measurable impact on learning. However, to provide learners with the opportunity to act on the feedback, a significant portion of the effort needs to happen during the active teaching semester and this limits the practicability of the approach. This perhaps accounts for why many of the existing practitioner reports on multi-stage assignments document results for classes that have less than 50 students (Prowse et al., 2007; Carless, 2007; O'Mahony, 2013; Harney, 2017; Cranny, 2017; O'Mahony, 2017).

This finding should not be interpreted to suggest that multi-stage assignments be abandoned. The existing discourse presents strong evidence that learners value, engage with and learn from this assessment process (Handley et al., 2008; O'Mahony, 2018). The heart of the issue is the assumption that feedback is (pre-dominantly) a uni-directional transmission from the instructor to the learner. Under these conditions, this article would support the assertion made by Boud and Molloy (2013) that practitioners need to carefully consider the types of contexts that would warrant the additional workload that this type of assessment demands. More generally, the results would support their call

to reconceptualise feedback as originating from multiple sources and travelling in multiple directions (Nicol, 2010; Orsmond et al., 2013; Sadler, 2013; Carless and Boud, 2018). This is the main implication that arises from the findings presented in this paper. If multi-stage assignments are to become more commonly used, the potential of peer-review needs to be harnessed. While learners may initially struggle with peer-reviews (Handley et al., 2008; Bloxham and Campbell, 2010; Carless et al., 2011), the assessment discourse evidences the role that peer-review can play in promoting evaluative judgement, self-regulation and independent learners (Nicol, 2014; Carless and Boud, 2018). By developing these competences practitioners can reduce learner's dependence on a single other and offer real opportunities for them to engage with an academic community of practice. The literature contains recommendations on how to do this well (Thomas et al., 2011; Sadler, 2013; Kearney and Perkins, 2014; Nicol, 2014) and also evidences that when it is done well learners recognise the benefits of peer-reviews (Cartney, 2010). Future work will explore how these ideas can be used to transform the author's practice.

References

Bloxham, S. and Campbell, L. (2010) 'Generating dialogue in assessment feedback: exploring the use of interactive cover sheets', *Assessment & Evaluation in Higher Education*, *35*(3), pp.291–300. https://doi.org/10.1080/02602931003650045

Boud, D., and Molloy, E. (2013) 'Rethinking models of feedback for learning: the challenge of design', *Assessment & Evaluation in Higher Education*, *38*(6), pp.698–712. https://doi.org/10.1080/02602938.2012.691462

Carless, D. (2007) 'Learning-oriented assessment: Conceptual bases and practical implications', *Innovations in Education and Teaching International*, *44*(1), pp.57–66. https://doi.org/10.1080/14703290601081332

Carless, D. (2009) 'Trust, distrust and their impact on assessment reform', *Assessment and Evaluation in Higher Education*, *34*(1), pp.79–89. https://doi.org/10.1080/02602930801895786

Carless, D. and Boud, D. (2018) 'The development of student feedback literacy: enabling uptake of feedback', *Assessment and Evaluation in Higher Education*, *2938* (May), pp.1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Carless, D., Kam, K., Chan, H., To, J., Lo, M. and Barrett, E. (2018) 'Developing students' capacities for evaluative judgment through analysing exemplars', in Boud, D., Ajjawi, R., Dawson, P. and Tai, J. (eds) *Developing Evaluative Judgement in Higher Education: Assessment for knowing and producing quality work*. Abingdon: Taylor & Francis.

Carless, D., Salter, D., Yang, M. and Lam, J. (2011) 'Developing sustainable feedback practices', *Studies in Higher Education*, *36*(4), pp.395–407. https://doi.org/10.1080/03075071003642449

Cartney, P. (2010) 'Exploring the use of peer-assessment as a vehicle for closing the gap between feedback given and feedback used', *Assessment & Evaluation in Higher Education*, *35*(5), pp.551–564.

Cohen, L., Manion, L. and Morrison, K. (2011) Research Methods in Education. Availble at: https://gtu.ge/Agro-Lib/RESEARCH%20METHOD%20COHEN%20ok.pdf (Accessed: 15 January 2021.

Cranny, D. (2017) Using Screencasting to Close the Feedback Loop. *Y1Feedback.Com*, 1–5. http://y1feedback.ie/wp-content/uploads/2017/01/DKIT-6-DCranny-2.pdf

Duncan, N. (2007) '"Feed-forward": improving students' use of tutors' comments', *Assessment & Evaluation in Higher Education*, *32*(3), pp.271–283. https://doi.org/10.1080/02602930600896498

Evans, C. (2013) 'Making sense of assessment feedback in higher education', *Review of Educational Research*, *83*(1), pp.70–120. https://doi.org/10.3102/0034654312474350

Gibbs, G. (2010) *Using assessment to support student learning*. http://repository-intralibrary.leedsmet.ac.uk/

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning.

Learning and Teaching in Higher Education, *1*, 3–31.

Handley, K., Price, M. and Millar, J. (2008) *Engaging students with assessment feedback* (Issue June). Final Report for FDTL5 Project 144/03, Oxford Brookes University. https://doi.org/FDTL5 Project 144/03.

Harney, S. (2017) *Visual Audio Screencasts to Enrich Feedback and Learner Engagement*. 1–5.

HEFCE. (2014) *UK review of the provision of information about higher education: National Student Survey results and trend analysis 2005- 2013* (Issue July).

Hepplestone, S. and Chikwa, G. (2016) 'Exploring the processes used by students to apply feedback', *Student Engagement and Experience Journal*, *5*(1), pp.1–15. https://doi.org/10.7190/seej.v5i1.104

IEEE (2018) *Institute of Electrical and Electronics Engineers*. https://www.ieee.org

Knight, P. (2002) 'Summative Assessment in Higher Education: practices in disarray', *Studies in Higher Education*, *27*(3), pp.275–286. https://doi.org/10.1080/0307507022000

Knight, P. and Yorke, M. (2003) *Assessment, learning and employability*. Maidenhead: Society for Research into Higher Education and the Open University Press.

Nicol, D. (2010) 'From monologue to dialogue: improving written feedback processes in mass higher education', *Assessment & Evaluation in Higher Education*, *35*(5), pp.501–517. https://doi.org/10.1080/02602931003786559

Nicol, D. (2014) 'Guiding Principles for Peer Review: Unlocking Learners' Evaluative Skills', in Kreber, C., Anderson, C., Entwhistle, N. and McArthur, J. (eds) *Advances and Innovations in University Assessment and Feedback* (Issue March, pp. 197–224). Edinburgh: Edinburgh University Press. https://doi.org/10.3366/edinburgh/9780748694549.003.0011

Nicol, D. J. and Macfarlane-dick, D. (2006) 'Formative assessment and self-regulated learning: A model and seven principles of good feedback practice', *Studies in Higher Education*, *31*(2), pp.199–218. https://doi.org/10.1080/03075070600572090

O'Donovan, B., Rust, C. and Price, M. (2016) 'A scholarly approach to solving the feedback dilemma in practice', *Assessment & Evaluation in Higher Education*, *41*(6), pp.938–949. https://doi.org/10.1080/02602938.2015.1052774

O'Mahony, T. (2017) 'The Impact of a Constructivist Approach to Assessment and Feedback on Student Satisfaction and Learning: A case-study', *AISHE*, *9*(2), pp.2871–28719.

O'Mahony, T. (2013) Assessing engineering learning: Formative, summative and dialogical. *2013 IEEE Global Engineering Education Conference (EDUCON)*, 990–995. https://doi.org/10.1109/EduCon.2013.6530228

Orsmond, P., Maw, S. J., Park, J. R., Gomez, S. and Crook, A. C. (2013) 'Moving feedback forward: Theory to practice', *Assessment and Evaluation in Higher Education*, *38*(2), pp.240–252. https://doi.org/10.1080/02602938.2011.625472

Orsmond, P. and Merry, S. (2011) 'Feedback alignment: effective and ineffective links between tutors' and students' understanding of coursework feedback', *Assessment & Evaluation in Higher Education*, *36*(2), pp.125–136. https://doi.org/10.1080/02602930903201651

Pokorny, H. and Pickford, P. (2010) 'Complexity, cues and relationships: Student perceptions of feedback', *Active Learning in Higher Education*, *11*(1), pp.21–30. https://doi.org/10.1177/1469787409355872

Prowse, S., Duncan, N., Hughes, J. and Burke, D. (2007) '"...do that and I'll raise your grade". Innovative module design and recursive feedback', *Teaching in Higher Education*, *12*(4), pp.437–445. https://doi.org/10.1080/13562510701415359

Sadler, D. Royce. (2013)'Opening up feedback: Teaching learners to see', in Merry, S., Price, M.P., Carless, D. and Taras M. (eds) *Reconceptualising Feedback in Higher Education: Developing Dialogue with Students* (pp. 54–63). Routledge. https://doi.org/10.4324/9780203522813

Sadler, D.R. (1989) 'Formative assessment and the design of instructional systems', *Instructional Science*, *18*, pp.119–144.

Sambell, K. (2011) Rethinking feedback in higher education: an assessment for learning perspective.

York: The Higher Education Academy. Available at: https://www.plymouth.ac.uk/uploads/production/document/path/2/2729/RethinkingFeedbackInHigherEducation.pdf (Accessed: 15 January 2021).

Scaife, J. and Wellington, J. (2010) 'Varying perspectives and practices in formative and diagnostic assessment: a case study', *Journal of Education for Teaching*, *36*(2), pp.137–151. https://doi.org/10.1080/02607471003651656

Vardi, I. (2013) 'Effectively feeding forward from one written assessment task to the next', *Assessment & Evaluation in Higher Education*, *38*(5), pp.599–610. https://doi.org/10.1080/02602938.2012.670197

Voelkel, S. and Mello, L. V. (2014) 'Audio feedback - Better feedback?', *Bioscience Education*, *22*(July), pp.16–30. https://doi.org/10.11120/beej.2014.00022

Weaver, M. R. (2006) 'Do students value feedback? Student perceptions of tutors' written responses', *Assessment & Evaluation in Higher Education*, *31*(3), pp.379–394. https://doi.org/10.1080/02602930500353061

Wiliam, D. (2011) 'What is assessment for learning?', *Studies in Educational Evaluation*, *37*(1), pp.3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

Winstone, N. E., Nash, R. A., Parker, M. and Rowntree, J. (2016) 'Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes', *Educational Psychologist*, *1520*(September), pp.1–21. https://doi.org/10.1080/00461520.2016.1207538

Y1Feedback (2016) *Feedback in First Year: A Landscape Snapshot Across Four Irish Higher Education Institutions*. Available at: http://y1feedback.ie/wp-content/uploads/2016/04/ALandscapeSnapshot2016.pdf (Accessed: 15 January 2021.