

A review of a FACULTY-WIDE change in Assessment Practice for Open and Distance Learners of Science

Sally Jordan, The Open University, United Kingdom

Summary for the Book of Abstracts

Previous practice in the Open University Science Faculty has been for all modules to be assessed by a combination of summative continuous assessment, with extensive feedback comments, and an end-of-module task (an examination or an extended assignment). This practice, although well established and apparently well received, has led to concerns, as reported elsewhere, that staff and students have a different understanding of the purpose of continuous assessment: staff see its purpose as primarily formative whilst students are primarily concerned with obtaining high marks.

The revised practice still requires students to meet a threshold for their overall continuous assessment score, but the final grade is determined by the end-of-module assessment alone.

The evaluation of the change in practice has been split into small practitioner-led sub-projects, comparing impact across different modules and levels, with the aim of identifying factors that lead to improved engagement. Sub-projects are both quantitative, e.g. comparing assignment completion rates before and after the change, and qualitative e.g. investigating student and tutor perceptions and opinion.

The change to formative thresholded assessment has sometimes led to a reduction in submissions for the final piece of continuous assessment and to an increase in the number of partial submissions. However other factors, in particular a cut-off date for an assignment (whether summative or formative thresholded) close to an examination, had a considerably larger effect and the change in practice does not appear to have had any effect on overall completion and success rates.

Many students and associate lecturers have a poor understanding of our assessment strategies, including conventional summative continuous assessment. This is in line with a frequently found result that students have poor understanding of the nature and function of assessment, perhaps because it has not been properly explained. It is important that assessment strategies are clear and consistent across qualifications, and that they are made clear to students.

Whilst student motivation cannot be implied, it is possible to see evidence that supports a notion of two contrasting groups of students who are in borderline continuous assessment categories: those who do well on the minimum number of assignments but chose not to submit others and still do well on the final examination, and those who have a more modest performance on continuous assessment (perhaps just omitting one assignment) and fail the module as a result of their poor examination performance. Some students are probably best advised to spend their limited time on revision rather than attempting all components of continuous assessment.

Where a skill is best assessed by a project or experimental report, future policy will encourage a two-part summative component, including the project/experimental report as well as an examination. Other assignments will remain formative but thresholded, with a clear purpose of preparing students for the examinable components.

Background

On most of their modules of study, students at the UK Open University (OU) encounter both continuous assessment (comprising tutor-marked assignments (TMAs) and sometimes interactive computer-marked assignments (iCMAs)) and an end-of-module project or examination. Both the continuous assessment and the end-of-module assessment are summative in that they are graded and the outcomes of each contribute in some way to students' overall module results. However detailed feedback is provided on TMAs, and in terms of learning design, continuous assessment is largely formative⁽¹⁾ in the sense that it is "for learning"⁽²⁾.

It has been pointed out that in such a "double-duty" model (3), assessment's formative and summative functions can be difficult to balance (4, 5) with the formative function sometimes "getting lost" (6). Staff and students have also been observed to have a different understanding of the purpose of continuous assessment: Staff may see its purpose as primarily formative, but students are primarily concerned with obtaining high marks. There was a wish to free students from anxiety over the minutiae of grading of TMAs and iCMAs, placing greater focus on feedback and dialogue between students and their tutors.

There was also a wish to establish more honest assessment strategies. In practice, the summative nature of the continuous assessment was somewhat illusory for the vast majority of the students under consideration because their course grade was actually determined by their exam mark, which was generally significantly lower than their continuous assessment mark (7).

Finally, if assessment items could be re-used, more effort could be put into optimising the questions, the tutor notes and feedback, with improvements made in the light of experience. This would remove one of the barriers to more frequent presentation of the modules and to flexible study rates (since new assignments would not need to be produced for each presentation) and encourage the use of "little and often" assessment, for pacing (8).

Formative thresholded continuous assessment

In formative thresholded assessment, students are required to demonstrate engagement by meeting a modest threshold of some sort in the module's overall continuous assessment score (OCAS), but the student's overall grade is then determined on the strength overall examinable component score (OES) alone. The Science Faculty was given permission to move to formative thresholded assessment for all undergraduate modules at OU level 1¹ and for level 2 and 3 modules with an examination rather than an end-of-module assessment (EMA).

Two basic models of formative thresholded assessment are currently in use:

- Model A. Tutor-marked assignments (TMAs) and interactive computer-marked assignments (iCMAs) are weighted, and students are required to reach a threshold (usually 40%) overall.
- Model B. Students are required to demonstrate engagement by reaching a threshold (usually 30%) in, say, 5 out of 7 assignments.

Several of the modules included in the study use minor variants of the above, and there have been changes during the period of the investigation. Whilst, in general terms, such variation is to be discouraged because of the confusion caused to students, it has enabled useful comparison of some points of detail.

¹ OU levels 1, 2 and 3 are equivalent to Framework for Higher Educations Qualifications levels 4, 5 and 6 and Scottish Credit and Qualifications Framework levels 7, 8, 9 and 10 respectively.

Formative thresholded assessment Model B was first introduced into two existing and two new level 3 physical science modules from 2010; *Analytical science* (S240) adopted a similar approach from 2012. Also from 2012, *Exploring science* (S104), *Introducing health sciences* (SDK125) and *Human biology* (SK277) moved to formative thresholded Model A, alongside the new module *Investigative and mathematical skills in science* (S141), and other modules followed from 2013. Level 2 and level 3 modules with end-of-module assessments rather than examinations have retained summative OCAS, as have modules nearing the end of their lifetimes.

Evaluation methodology

Student opinion was canvassed prior to the introduction of formative thresholded assessment (9) and previous work had evaluated the introduction of purely formative iCMAs into a level 3 module (10) and investigated student engagement with iCMAs in a range of modules, in formative, thresholded and summative use (11).

This paper summarises some of the outcomes of a major project which ran from 2012-14 with the aim of evaluating the Faculty-wide change in practice. In more detail, the research questions whose results are reported in this paper were:

- Has the move to formative thresholded assessment led to a change in TMA and iCMA submission rates?
- Has the move to formative thresholded assessment led to a change in TMA performance?.
- Has the move to formative thresholded assessment led to a change in module completion and pass rates?
- What do students and associate lecturers know about and think of our assessment strategies?
- What are the relative merits of Model A and Model B formative thresholded assessment?

Other aspects of the evaluation, in response to the following research questions, are included in the final project report (12) and will be disseminated elsewhere:

- Has the move to formative thresholded assessment led to a change in the extent or type of plagiarism cases that are detected?
- How do the levels of student engagement on iCMAs compare for summative, formative thresholded, and purely formative use?
- What is the impact of other assessment-related factors e.g. additional thresholds?
- What is the impact of other student-related factors e.g. do new students and continuing students perform differently? If students are studying two modules concurrently, what is the impact on TMA submission?

The evaluation of the change in practice was split into small practitioner-led sub-projects, comparing impact across different modules and levels, with the aim of identifying factors that lead to improved engagement. Sub-projects were both quantitative, e.g. comparing assignment completion rates before and after the change, and qualitative e.g. investigating student and tutor perceptions and opinion.

Summary of results

The full results are given in full in the final project report (12).

TMA submissions, grades and overall completion and pass rates

Year-on-year variation was difficult to interpret because of the many other changes that had place on the same time-scale as the study, in particular the changing student population as a result of HE funding changes in England. In order to allow for the impact of these factors, some results were verified by considering TMA submission only for those students who attempted the examination. Figure 1 compares the TMA submission rate for the October 2011 start (11J) and October 12 start (12J) presentation of *Introducing health sciences* (SDK125), on which formative thresholded assessment had been used for the first time in 12J. The figure shows that submissions for the final TMA were slightly but not substantially reduced. On some other modules the drop in final TMA submission rate was more substantial, but the effect was not always present and the impact of other factors, in particular a cut-off date for a TMA (whether summative or formative thresholded) close to an examination date, had a considerably larger effect.

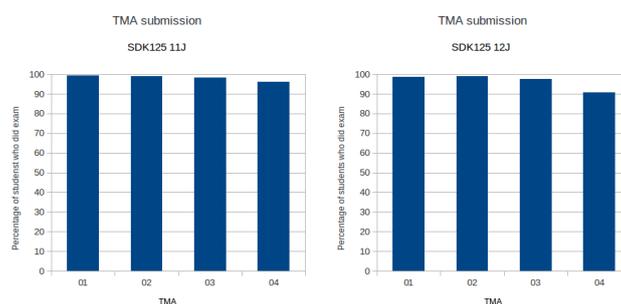


Figure 1 TMA submission rates for the 11J and 12J presentations of SDK125, for students who attempted the examination only (n = 630 for 11J and n = 574 for 12J)

Students were observed to be more likely to submit partial TMAs (measured by a proxy of scoring zero for one or more questions) in the final assignments of a module following the move to formative thresholded assessment (Table 1).

However, the most notable finding of this part of the evaluation was that there was no significant change in completion or success rates for any of the modules in the study as a result of the introduction of formative thresholded assessment. Other factors, for example, changing student populations, had considerably larger impact.

Correlations between TMA/iCMA submission and overall success

Correlations between assignment submission and overall success were investigated in detail for seven modules and, unsurprisingly, strong correlations were found. For example, for students on the presentation of *Exploring science* (S104) that started in October 2012, those who submitted all TMAs had a mean exam score was 53.4 whilst those who did not submit all TMAs had a mean exam score of 44.4%. However correlation should not be taken to imply causality. Whilst, it is possible that the TMAs helped students to prepare for the final assessment, but it may simply have been that the students who were more able or who had more time for study were more likely both to submit all TMAs and to be more successful in the examination.

Table 1 Percentages of S104 students who scored zero for one or more questions in a TMA. Note that TMA07 was discontinued from the October 2012 presentation and an examination replaced the previous EMA.

Oct 201 presentation. n =2109 at module start		Oct 2012 presentation. n = 2357 at module start	
TMA01	1.6%	TMA01	0.8%

TMA02	3.4%	TMA02	2.5%
TMA03	2.8%	TMA03	2.4%
TMA04	5.8%	TMA04	4.3%
TMA05	3.9%	TMA05	6.0%
TMA06	2.8%	TMA06	7.9%
TMA07	6.5%		

The submission of the final TMA (TMA04) for *Infectious disease and public health* (SK320) fell from 74% of registered students in 2012 (with summative OCAS) to 61% of registered students in 2013 (with formative thresholded OCAS, Model A); the mean score for this TMA for students who submitted it also fell from 73 to 69 between the two years, suggesting that some students may have put less effort into TMA04 following the change in assessment strategy. TMA04 was designed to prepare SK320 students for Part C of the final examination so this drop in engagement was a cause for concern. However, there was found to be at best a weak association between performance on TMA04 by students who submitted it in 2013 and their performance on Part C of the examination, and no association between non-submission of TMA04 and Part C score. It was therefore concluded that some students may have made a sensible decision in choosing to concentrate on revision rather than on submitting or gaining a high score in the final TMA.

Student and staff perception of assessment strategies

It is pleasing that some students have welcomed the change in assessment strategy, with one student on Human biology (SK277) commenting:

Looking back on my previous courses I think I stressed over my TMAs far too much, perhaps to the detriment of revising throughout the year in preparation for the exam. I always wanted each TMA to be perfect... The reality for me has always been that my exam grade is the one which has determined by final grade. This year I am still working hard on my TMAs but I am not stressing about every tiny detail like I used to. I think I can actually say I am enjoying my TMAs.

However, email dialogue with students and associate lecturers on several modules revealed a widespread and worrying misunderstanding as to the nature and purpose of assessment strategies, both formative thresholded and summative, with one experienced associate lecturer commenting:

I don't think they do understand this - and I don't either. My experience on S104 was that in the previous model of summative continuous assessment the final mark was made up of both OCAS (75%) and OES marks (25%) – or at least that's what I thought... maybe I was wrong all those years, oops!

Finally, there were pleas for more consistent practice, for example:

I do work on other modules with a variety of different assessment strategies...I do wish there was consistency in assessment strategy style across the faculty, for me and for students.

Potential impact of a change from Model A to Model B or vice versa

For three level 2 modules *Analytical science* (S240), *The physical world* (S207) and *Human biology* (SK277), the impact of a change in threshold or a change from Model B to Model A

or vice versa was investigated. Some students appear to have been strategic in their decision to omit assignments, for example one student only submitted one of SK277's TMAs, but scored 95 on it and thus passed the overall OCAS threshold. This student would not have passed either a 40% overall OCAS threshold or a "30% on 2 out of 3 TMAs" threshold, and obtained 54% in the exam, leading to a Grade 4 pass. Had the student been required to meet a more challenging OCAS threshold, they would presumably have submitted more TMAs, but it is not possible to tell whether this would have affected their final outcome. Another SK277 student scored 90% and 26% on two SK277 TMAs. They would not have passed either a 40% overall OCAs threshold or a "30% on 2 out of 3 TMAs" threshold, but they obtained 91% in the exam, leading to a distinction.

Other students were observed to score poor marks on assignments, and a fail or resit result whether or not they reached the OCAS threshold. For example, one student scored 40% on S207 CMA41, and 58%, 30%, 23%, 24%, 34% and 32% respectively on TMAs 01-06 respectively, leading to an OCAS of 37% which did not meet the threshold. This student only obtained 19% in the exam. Different thresholding may have led to a different OCAS result, but is unlikely to have affected the overall outcome.

Discussion, conclusions and suggestions for future work

The change to formative thresholded assessment does not appear to have had any detrimental effect. However, many students and associate lecturers have a poor understanding of our assessment strategies, including conventional summative continuous assessment. This is in line with a frequently found result that students have poor understanding of the nature and function of assessment (13, 14), perhaps because it has not been made clear (15). It is important that assessment strategies are clear and consistent across qualifications, and that they are explained carefully to students.

Whilst student motivation cannot be implied, it is possibly to see evidence that supports a notion of two contrasting groups of students who are in borderline OCAS categories: those who do well on the minimum number of assignments but chose not to submit others and still do well on OES, and those who have a more modest performance on OCAS (perhaps just omitting one assignment) and fail the module as a result of their poor OES performance. Some students are probably best advised to spend their limited time on revision. Further qualitative work is required to ascertain students' reasons for behaving in the way that they have been observed to do.

References

1. WILIAM, D. & BLACK, P. (1996). *Meanings and consequences: A basis for distinguishing formative and summative functions of assessment?* British Educational Research Journal, 22(5), 537-548.
2. BLACK, P. & WILIAM, D. (1998). *Assessment and classroom learning*. Assessment in Education, 5(1), 7-74.
3. BOUD, D. (2000). *Sustainable assessment: rethinking assessment for the learning society*. Studies in Continuing Education, 22(2), 151-167.
4. GIORKA, O. (2008). *Teacher or assessor? Balancing the tensions between formative and summative assessment in science teaching*. In A. Havnes, A. & L. McDowell, Balancing dilemmas in assessment and learning in contemporary educations. New York: Routledge (pp. 145-156).

5. PRICE, M.; CARROLL, J.; O'DONOVAN, B.; et.al. (2011). *If I was going there I wouldn't start from here: A critical commentary on current assessment practice*. *Assessment & Evaluation in Higher Education*, 36(4), 479-492.
6. BREARLEY, F.Q. & CULLEN, W.R. (2012). *Providing Students with Formative Audio Feedback*. *Bioscience Education*, 20, 22-36.
7. FREAKE, S. (2008). *New assessment models for Physics and Astronomy... "the times they are a-changing."* *OpenCETL Bulletin*, 3, 4-6.
8. GIBBS, G. & SIMPSON, C. (2004-5). *Conditions under which assessment supports students' learning*. *Learning and Teaching in Higher Education*, 1, 3-31.
9. MANNERS, J.; FREAKE, S.; BOLTON, J.; et.al. (2008). *Formative assessment and presentation patterns for Level 3 physics and astronomy courses: A survey of student views*. Reprinted in S. Jordan (Ed.) (2010), *e-Assessment: Compilation of final reports on Open University Physics Innovations CETL projects*. Milton Keynes: piCETL (pp. 109-131).
10. BOLTON, J. (2010). *Using interactive computer-marked assignments in Level 3 quantum mechanics*. Reprinted in S. Jordan (Ed.) (2010), *e-Assessment: Compilation of final reports on Open University Physics Innovations CETL projects*. Milton Keynes: piCETL (pp. 82-87).
11. JORDAN, S. (2011). *Using interactive computer-based assessment to support beginning distance learners of science*. *Open Learning*, 26(2), 147-164.
12. JORDAN, S.; BOLTON, J.; COOK, L.; et.al. (2014). *Thresholded assessment: Does it work? Report on an eSTEE M project*. Milton Keynes: eSTEE M.
13. CARLESS, D. (2006). *Differing perceptions in the feedback process*. *Studies in Higher Education*, 31(2), 219-233.
14. ORSMOND, P. & MERRY, S. (2011). *Feedback alignment: Effective and ineffective links between tutors' and students' understanding of coursework feedback*. *Assessment & Evaluation in Higher Education*, 36(2), 125-136.
15. SURGENOR, P.W.G. (2013). *Measuring up: comparing first year students' and tutors' expectations of assessment*. *Assessment & Evaluation in Higher Education*, 38(3), 288-302.