Using e-assessment to learn about students and learning

Sally Jordan sally.jordan@open.ac.uk

Department of Physical Sciences
The Open University
United Kingdom

Correspondence address The Open University (Cambridge Office)
Cintra House
12 Hills Road
Cambridge
CB2 1PF

Using e-assessment to learn about students and learning

Sally Jordan, The Open University

Abstract
Analysis of student responses to interactive computer-marked questions has provided insight into specific student misconceptions and also to the identification of characteristic patterns of engagement with assignments and the feedback provided. The paper summarises and updates previously reported analyses of student usage of interactive computer-marked assessment at the UK Open University. It then considers the engagement of two different student populations on assignments known to be of similar difficulty. One group was found to be more likely than the other to attempt summative questions just before the due date, less likely to use the feedback provided, and less likely to engage with a formative practice assignment, a factor associated with a lower score on the summative assignment. Reasons for the different engagement of the two groups are discussed, with the less engaged students likely to be those who are studying other modules concurrently and who are underprepared for study at this level. The paper concludes with a general consideration of the use of learning analytics and assessment analytics to find out more about student behaviour and learning.

Keywords
Interactive computer-marked assessment; student engagement; data analysis; learning analytics; assessment analytics.

Introduction
The focus of high quality e-assessment is, rightly, usually on improving the quality of assessment and learning from the student perspective. However Glaser (1981) recognised the importance of synthesising information from students' assessment performance in order to form an accurate picture of their misconceptions, whilst Erwin (1995) recognised that information gathered from assessment has a useful role to play in guiding broader decisions about teaching and the curriculum.

Online e-assessment systems can easily gather every response that is entered by every user, along with information about, for example, when a particular response was given, which variant of the question had been received and what feedback had been given previously. Many e-assessment systems make this information available to users (e.g. MasteringPhysics[1], Moodle Gradebook[2]). Analysing the data that have been collected presents more of a challenge, but such analysis can provide rich evidence about:

- Student mistakes. As discussed by Nicol (2008), this information can be used to improve the questions (e.g. Walet and Birch 2012) and for deeper understanding of misconceptions (e.g. Jordan 2007).

[1] http://www.masteringphysics.com/ (Accessed 16th November 2013)

[2] http://docs.moodle.org/24/en/Gradebook (Accessed 16th November 2013)

- The behaviour of different variants of questions (e.g. Jordan, Jordan and Jordan 2012) or of different questions extracted from a question bank (e.g. Dermo 2010).
- The performance of the assignment as a whole (e.g. Ding and Beichner 2009) and correlations between this and overall performance on a module (Pritchard and Warnakulasooriya 2005).
- Links between the performance of an assignment and its mode of use e.g. formative, summative, thresholded (e.g. Jordan 2010).

This paper summarises earlier findings from the UK Open University and extends the work to consider the impact of two different student populations on engagement with an e-assessment task.

Context

Module A and Module B

The context for most of the work is a 10-credit *Maths for Science* module that was studied by more than 12,000 students at the Open University (OU) between 2002 and 2012, and assessed by means of formative and summative interactive computer-marked assignments (iCMAs). In 2012, a new edition of the book on which *Maths for Science* was based was published (Jordan, Ross and Murphy 2013). This book is now being used in two modules (identified as Module A and Module B for the purposes of this paper) with different assessment strategies and different student populations, as described below. Module A and Module B each use a 41-question iCMA that is a slightly modified version of one from the original *Maths for Science.* These two iCMAs are different but they are known to be of equivalent overall difficulty.

On Module A (30-credits), *Maths for Science* forms one third of the whole module, and its iCMA comprises one third of the overall thresholded continuous assessment score. It was presented for the first time from October 2012 to June 2013. Module B (10-credits) comprises *Maths for Science* only and it was presented for the first time from October 2012 to March 2013; it is assessed entirely by its iCMA-based end-of-module assessment (iEMA).

Students on Module B have longer to study the content of *Maths for Science* (20 weeks for Module B compared with 10 weeks for Module A) with the Module A iCMA being available for 3 weeks whilst the Module B iEMA is available for 5 weeks before the cut-off date. Within these time-periods, students are allowed to spend as long as they want to on each assignment and they can access it as often as they wish, using a web browser from any computer.

The iCMAs run in the OpenMark[3] system and students are virtually always given three tries at each question (even in summative use), with increasing feedback after each try. A purely formative 47-question practice assignment, using the same technology and with similar questions, is available to students on both Module A and Module B for the duration of each module. Students are able to repeat practice assignment questions as often as they wish, with different variants of some questions providing extra opportunities for practice.

[3] http://www.open.ac.uk/openmarkexamples/ (Accessed 16th November 2013)

The student population
Open University students have a wide range of ages and the University's 'open' mission means that there are no formal entry requirements for most of the undergraduate curriculum, including Module A and Module B. Students have always come to OU study with a wide range of previous qualifications, from nothing to a higher degree. The OU is a UK-wide and increasingly a global university, but about 75% of students on Module A and Module B are studying in England, where changes to funding are resulting in a decreasing number of students with higher previous qualifications, whilst an increased access mission is leading to a higher proportion of students with lower qualifications than those required for entry to most other universities.

Open University students study at home and most are studying part-time alongside employment or caring responsibilities. Module A and Module B are both 'upper level 1' modules, being designed to be studied after the 60-credit module *Exploring Science*, but some students choose to take them at the same time as *Exploring Science*. This is a cause for concern, especially for the presentation of Module A under investigation (October 2012 to June 2013), on which 68% of the 457 students who started the module were completely new to the OU (so potentially unprepared for Module A) whilst at least 54% were both new and studying this module alongside at least one other (so potentially both unprepared and over-committed).

As shown in Table 1, students on the presentation of Module B under investigation (October 2012 to March 2013) tended to be older and to have higher previous educational qualifications than those on Module A. 457 students commenced study of Module A in October 2012 whilst 322 students commenced study of Module B at the same time.

Table1: A comparison of the student populations of Module A and Module B at the start of each module in October 2012.

	Module A	*Module B*
Percentage of students with lower than 'usual' university entry qualifications (< A-level).	*33%*	*22%*
Percentage of students who were completely new to Open University study.	*68%*	*7%*
Age $\leq$ 21 years	*19%*	*6%*
Age 22-29 years	*45%*	*21%*
Age 30-39 years	*20%*	*30%*
Age 40-49 years	*9%*	*25%*
Age 50-59 years	*5%*	*12%*
Age $\geq$ 60 years	*2%*	*6%*

Presentation of methodology and results
The paper is structured into sections that highlight the lessons that can be learnt from four different types of analysis of student responses to computer-marked questions, with sub-sections indicating different types of findings. The examples are intended to be illustrative, providing an indication of the power of data analysis of this sort.

In each case the analysis is based on data gathered from all the students who submitted either the question or the assignment under consideration. For the presentations that started in October 2012, 316 students submitted the Module A iCMA and 272 students submitted the Module B iEMA. Within each assignment, the number of students who attempted each question varied from question to question and the number of responses gathered to each question was always greater than the number of users because multiple tries were allowed. Where responses were repeated or blank, they were removed from analyses that were investigating specific misunderstandings, to prevent undue weight being placed on guessed responses. In the following sections, the results always indicate how many responses, usages or users were included in the analysis.

Analysis of responses to individual questions

Analysis for information about student misunderstandings

A previous analysis of student responses to *Maths for Science* questions identified misconceptions relating to units, powers notation, arithmetic fractions and the rules of precedence (Jordan 2007). The findings were more reliable for questions in summative use (so students were trying their best to get the right answer) and for constructed response rather than selected response question types (so the answer could not have been obtained by guesswork or with guidance from the options offered). When the equivalent mistake was seen in different variants of a question, the underlying misconception could be identified with more confidence.

The question by question analysis was repeated prior to the writing of the new edition of the *Maths for Science* book, with the aim of identifying areas where the teaching might be improved. Unsurprisingly, the common misconceptions were found to be unchanged, but a number of previously unidentified errors were spotted. For example, in the simple question shown in Figure 1, some students appeared to correctly find a common denominator, but then to add both the numerator and the denominator, leading to an answer half the size of the one expected. 1.9% of student responses (24 out of 1233) to five variants of this question included an error of this type. Inspection of the old version of *Maths for Science* showed that the teaching on this point was unclear, so the text was amended for the new edition. This question is used in the new Module B iEMA and in the first presentation only 2 responses out of 300 (0.7%) included an error of this type. This is promising, though insufficient data is available at the present time to show statistical significance**.**

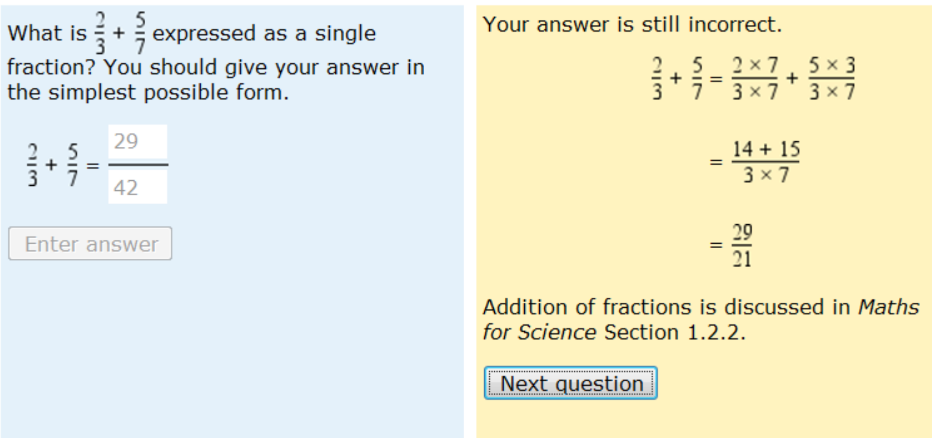

Figure1: A common error in a Maths for Science question, identified by data analysis.

Analysis for information about repeating of responses
On a small number of occasions, users have been observed to insert 'blank' responses or to repeat the same response on subsequent attempts at a question, despite having received feedback on their previous attempt. The extent of this behaviour varies from question to question and also depends on whether the iCMA is in summative, thresholded or purely formative use (Jordan and Butcher 2010).

The behaviour has been associated with:

- Lack of seriousness of engagement (so it is more common for questions in formative than summative use).
- Lack of understanding of what the question wants, or of the meaning of the feedback (an 'I haven't a clue' reaction to the question).
- Questions that are time consuming to complete, perhaps with multiple boxes for completion, or which require students to access a course component such as a video.

A small number of questions used in the assignments for Module A and Module B were essentially identical, but for some of these, the extent to which responses were left blank and repeated was nevertheless quite different for the two modules.

Figure 2 shows the different behaviour for a question assessing part of the penultimate chapter of *Maths for Science*, which required students to calculate the standard deviation of a data-set.

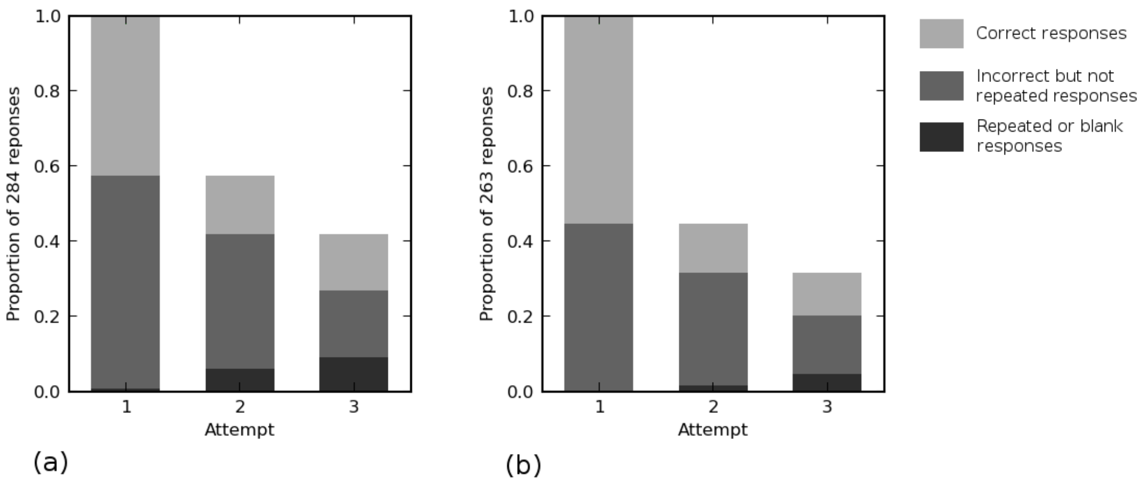

Figure2: The proportion of blank and repeated responses for a very similar question in (a) the Module A iCMA and (b) the Module B iEMA.

Responses shown in the palest tone (top) were correct, responses shown in the medium tone (middle) were incorrect but not blank or repeated and responses shown in the darkest tone were identical to the student's previously entered response (i.e. repeated) or were left blank.

The proportion of repeated and blank responses is greater for Module A (Figure 2a) than for Module B (Figure 2b). For example, of the 284 first-attempt responses for Module A, 17 (6%) were repeated at second attempt or left blank, whilst of the 263 first-attempt responses for Module B, only 4 (1.5%) were repeated or blank at second attempt.

Analysis for information about use of feedback
When a response is repeated after the delivery of feedback, as shown in Figure 2, the student has clearly not learned from that feedback. Figure 2 shows that the feedback was less well attended to for Module A than for Module B.

Figure 3 shows data for a question from a different module where the feedback provided to students, especially after a specific incorrect answer at first attempt, was improved between two presentations (identified as 09B and 10B). It is clear that the improved feedback has led to a higher proportion of responses being correct at the second attempt. This question has also become much more popular as a result; students were able to understand why the answer they gave was not correct.

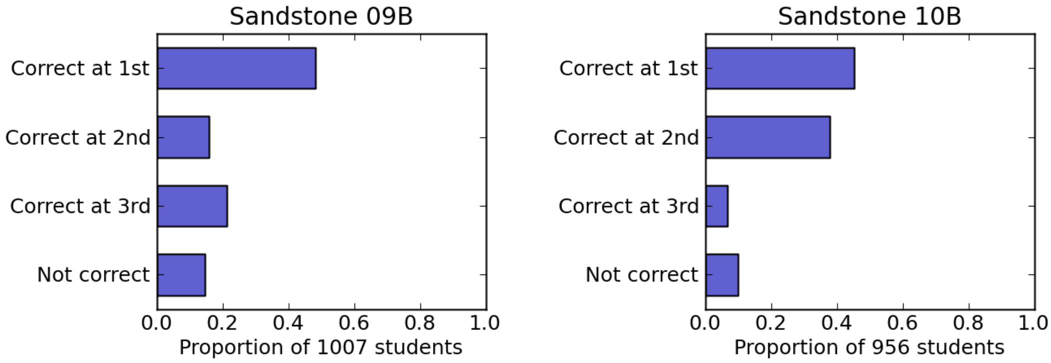

Figure3: The proportion of students who got the same question right at first, second, third attempt, or not at all, for the same question in the 09B and 10B presentation.

Analysis of responses to all questions in an assignment
The number of uses and number of users of each question
A simple plot of the number of individual users of each question in the *Maths for Science* Practice Assignment (Figure 4, dark blue bars) illustrates that, in common with most purely formative uses of e-assessment, usage drops off as the assignment progresses. The same effect is seen if an assignment is broken up into several shorter iCMAs. When students have the ‘carrot’ of summative assessment (even if lightly weighted) they usually attempt all the questions (Jordan and Butcher 2010).

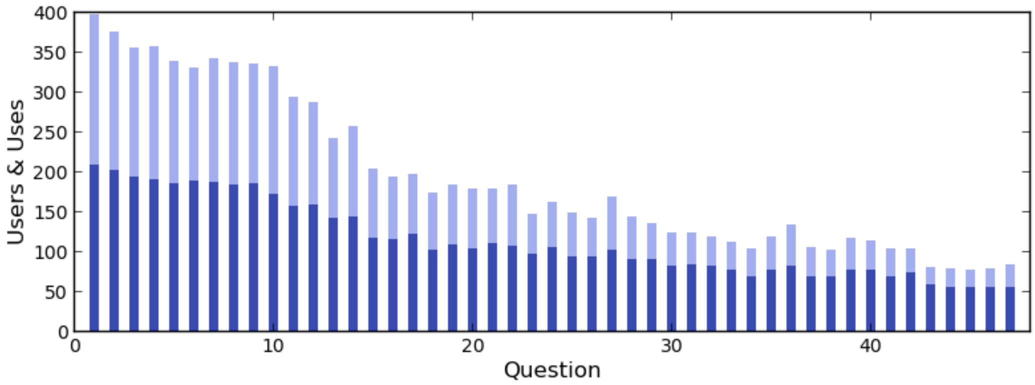

Figure4: The number of users (dark blue bars) and the number of separate usages (light blue) of each question on the Maths for Science Practice Assignment, for students studying Module B.

The light blue bars in Figure 4 indicate the number of separate usages of each question, giving an indication of the number of times each question is repeated; this also drops off as the assignment progresses. However information about the extent that individual students repeat questions cannot be extracted from Figure 4. Figure 5 makes it clear that whilst one or two students repeated certain questions many times, most students only attempted them once.

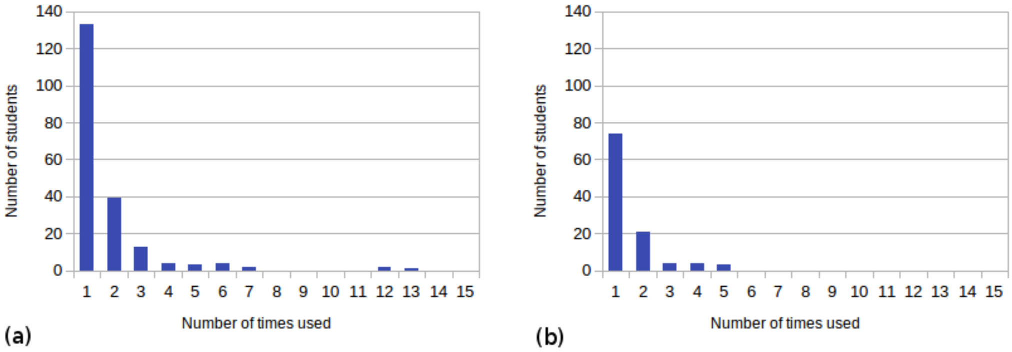

Figure5: Distributions showing the number of uses of (a) Question 2 and (b) Question 22, for Module B students on the Maths for Science Practice Assignment.

When do students attempt questions?

As reported in Jordan (2010), simply looking at all actions on an iCMA can provide an interesting insight into when students are engaging with an assignment. Figure 6 illustrates a case when students were using a formative iCMA for revision purposes, even when this use had not been suggested to them.

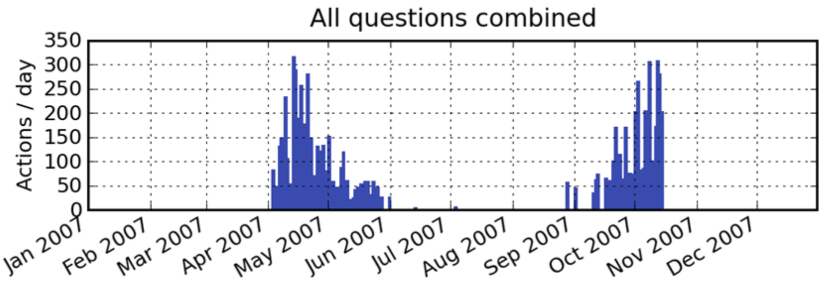

Figure6: All actions on an iCMA, illustrating activity around the time students are studying the relevant topic (April-May) and in the run up to an exam in October.

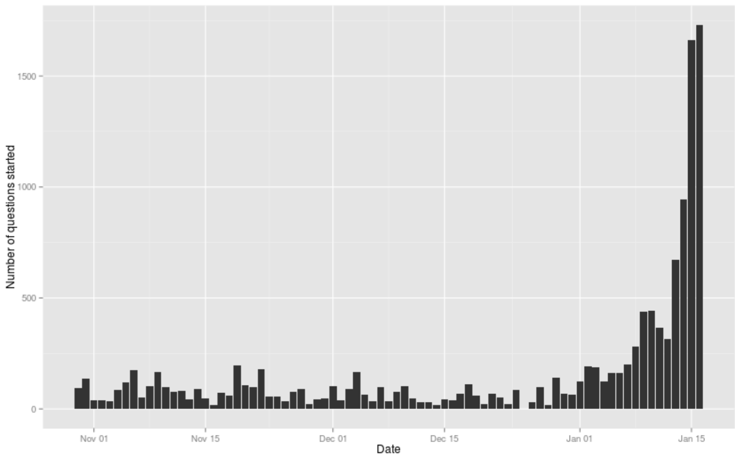

Figure7a: The total number of questions started by date for the Module A iCMA.

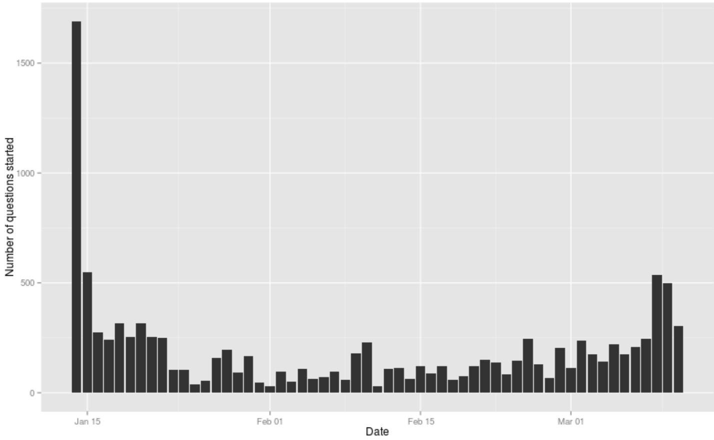

Figure7b: The total number of questions started by date for the Module B iEMA.

The number of separate questions started on each day provides a useful proxy for activity. Figure 7 compares the number of questions started on each day that the Module A iCMA and the Module B iEMA were live. Note that most Module B students appear to have been waiting for the iEMA to open, whilst most Module A students appear to have been rushing to complete the iCMA before it closed.

Correlations within an assignment

Impact of time spent on score

Maths for Science assignments are all available to students for several weeks and within that time, there is no limit to the amount of time that students can spend online in answering questions. Students can leave the iCMAs at any time and return to the same point in the assignment at a later date. Overall elapsed time is not a useful measure of student engagement, because some students look at the questions soon after the assignment opens but do not enter any responses until very much later.

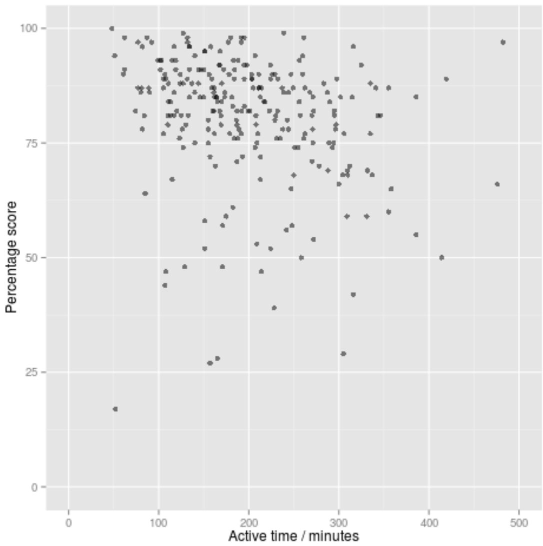

Figure8: Scatter plot showing percentage score achieved against the active time spent on a Maths for Science EMA in 2003. n = 250. Spearman rank correlation coefficient = −0.30.

An attempt was made in the early days of *Maths for Science* to estimate 'active time' (assuming for example, that students had 'gone away' if there were no actions for 30 minutes and so removing all time intervals of 30 minutes or more with no activity from the summation of active time). Figure 8 shows that there was a slight negative correlation between estimated active time on the assignment and overall score. Perhaps this is to be expected – mathematically more confident students answered the questions more quickly – but it is pleasing that many of the students who spent a long time answering the questions also did well.

Impact of date submitted on score

Figures 9a and 9b illustrate quite starkly the different behaviour of students on the very similar assignments for Module A and Module B.

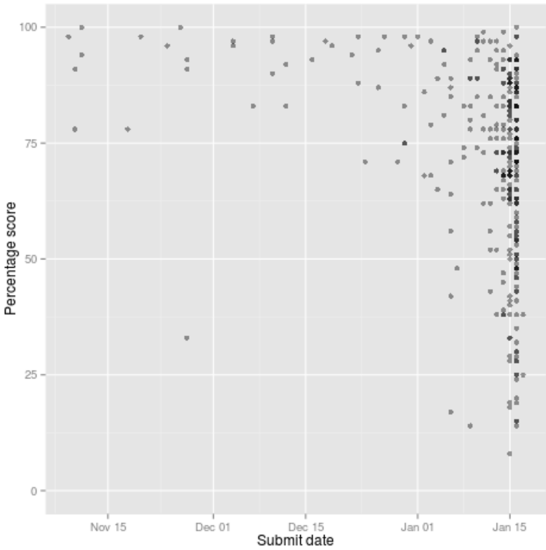

Figure9a: Scatter plot showing score against date submitted for the Module A iCMA. n = 316. Spearman rank correlation coefficient = −0.41.

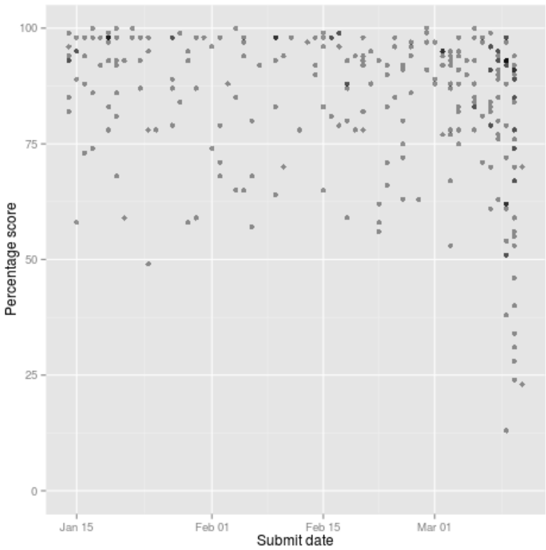

Figure9b: Scatter plot showing score against date submitted for the Module B iEMA. n = 272. Spearman rank correlation coefficient = −0.32.

On Module A, the small number of students who completed the iCMA more than a few days before the cut-off date all did well, but there are a large number of students who appear to have been rushing to complete the iCMA by the due date and who did not do well. In contrast, for Module B, there is a spread of marks, but all above

50%, for the duration of the assignment. The nine students who submitted the iEMA right at the end were probably encouraged to do so by an email reminding students of the due date, and they are unlikely to have been surprised to learn that they had failed the module. Most of them also started the iEMA late, one just 10 minutes before it closed!

Correlations between assignments

Actions by date on practice assignment and summative iCMA

When student engagement with e-assessment is investigated on a student-by-student level, several characteristic patterns appear. In the plots shown in Figure 10, it can be seen that the student worked through the Practice Assignment steadily, on a chapter by chapter basis, though she or he only attempted each question on one day (with probably means that they only attempted each question once). In contrast, they completed most of the questions in the summative assignment on one day, after completing the questions on the practice assignment.

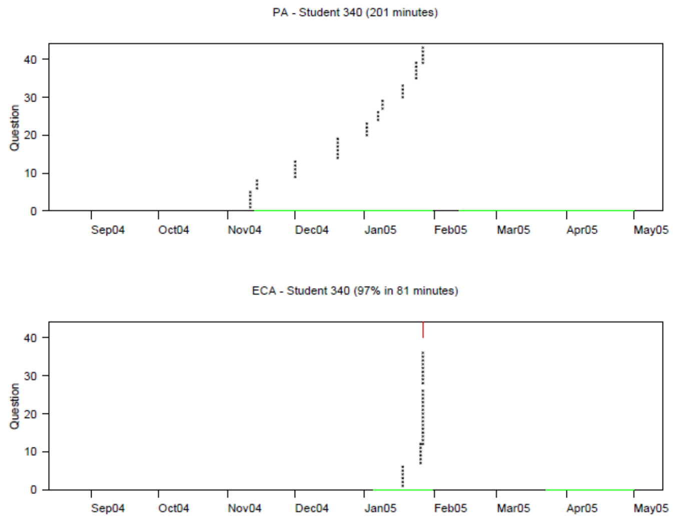

Figure10: Questions attempted by date by one student on the Maths for Science Practice Assignment (PA) and the summative assignment (referred to here as the ECA).

Impact of engagement with Practice Assignment on iCMA score

One of the other large differences between student behaviour on Module A and Module B was found to be the proportion of students who engaged with the *Maths for Science* Practice Assignment. Of the 316 students who submitted the Module A iCMA, just 121 (36%) had attempted the practice assignment. In contrast, of the 272 students who submitted the Module B iEMA, 214 (80%) had attempted the practice assignment. This was despite the fact that the practice assignment had been advertised to both groups of students, if anything slightly more to Module A students than to those studying Module B. However, it is well known that busy students do not engage with aspects of a module that are not assessed, especially if they fail to appreciate the benefit to themselves of doing so.

The Module A and Module B students who attempted the practice assignment did so in similar manner (so a plot like Figure 4 would be similar for Module A) and in both cases engagement with the practice assignment was positively correlated with score on the iCMA or the iEMA. The median iCMA score for Module A students who had attempted the practice assessment was 97 out of 120 (81%) whilst the median iCMA score for those who had not attempted the practice assessment was 86 out of 120 (72%). For Module B, the median iEMA score for students who had attempted the practice assessment was 109 out of 120 (91%) whilst the median iEMA score for those who had not attempted the practice assessment was 96 out of 120 (80%). The link between use of the practice assignment and final score is unlikely to be directly causal – it is more likely that students who were well motivated and had plenty of time were likely to both engage with the practice assignment and to do better overall.

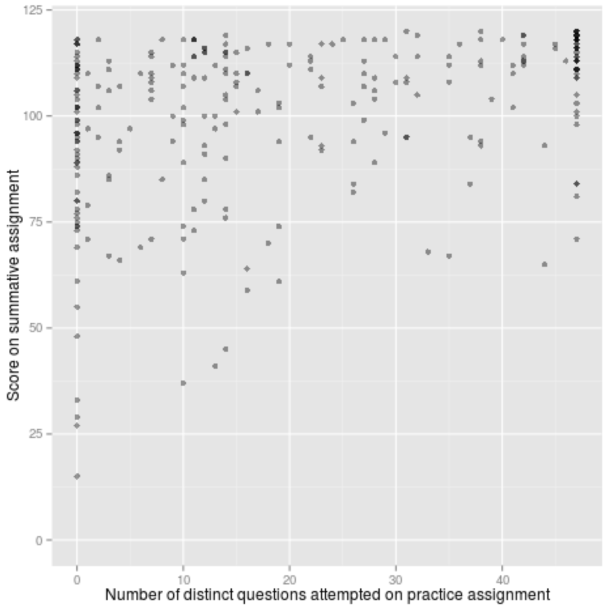

Figure11: Scatter plot showing score on Module B iEMA against the number of distinct questions attempted on the Maths for Science Practice Assignment. n = 272. Spearman rank correlation coefficient = 0.39.

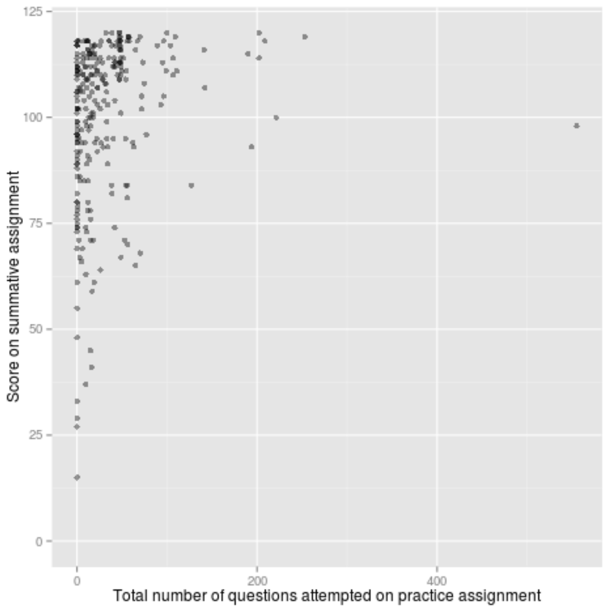

Figure12: Scatter plot showing score on Module B iEMA against the total number of questions attempted on the Maths for Science Practice Assignment. n = 272. Spearman rank correlation coefficient = 0.39.

In addition to the correlation with iEMA score of whether or not students had attempted the *Maths for Science* Practice Assignment at all, there was a correlation between iCMA/iEMA score with both the number of the 47 questions in the practice assignment that had been attempted (Figure 11) and the total number of practice assignment questions that had been attempted, taking account of repeats (Figure 12). Similar graphs and correlations were obtained for Module A and Module B students.

Discussion and conclusions

Analysis of student responses to interactive computed-marked questions has led to powerful evidence about student misunderstandings, student engagement with interactive computer-marked questions, and contrasting patterns of engagement for assignments in summative and purely formative use.

For a similar assignment used by two different student populations, contrasting patterns of use have been observed. Students on Module A were less likely to alter their responses after receiving feedback, and more likely to submit the iCMA just before the due date than were students on Module B. Students on Module A were also considerably less likely to attempt the formative *Maths for Science* Practice Assignment, a factor associated with a lower score on the final assignment.

All of these factors point towards good student engagement on Module B, but to many students on Module A who were lacking in the time and motivation to engage fully, and whose success was compromised as a result. This is entirely consistent with the less well prepared and overcommitted student population on Module A, as described in the ‘context’ section of this paper. Indeed, student behaviour on the Module A iCMA provided an early warning of later study difficulties for these students. Given that Module A was designed for study after *Exploring Science*, it is perhaps not surprising that there was a 33 percentage point difference in Module A completion rate between those who studied the module after successfully completing *Exploring Science* and those who were new to OU study[4].

Those students who were known to be over-committed (the 54% who were new and studying Module A alongside at least one other module) fared even worse: there was a 35 percentage point difference in completion rate between these students and those who had completed *Exploring Science* previously. Thankfully, for the presentation of Module A that started in October 2013, a considerably smaller proportion of the students (25% for 2013 compared with 68% for 2012) were new to the University and more (37% for 2013 compared with 24% for 2012) had completed *Exploring Science* previously. It is expected that a correspondingly different pattern of engagement with the Module A iCMA and the *Maths for Science* Practice Assignment will be observed.

More generally, lessons from the analysis of student responses to e-assessment tasks cannot be ignored. There is some subjectivity in the interpretation of the data e.g. in saying that students gave a particular answer to the question shown in Figure 1 as a result of a particular misconception, or that Module A students attempted questions close to the due date because they were time-poor . However the actual

[4] The data given in this paragraph were not available at the time the paper was presented at the CAA conference in July 2013.

data are unequivocal and provide vital evidence concerning the factors underpinning student engagement. In addition to discovering more about student misunderstandings, lessons can be learned which are highly relevant for assessment and curriculum design in the future.

Learning analytics (which can be defined as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs” (Ferguson 2012, p.305)) can inform teachers about the learning of a cohort of students as well as about individual learners. Ellis (2013) calls for ‘assessment analytics’ (the analysis of assessment data), pointing out that assessment is ubiquitous in higher education whilst student interactions in other online learning environments are not.

Redecker et al. (2012) suggest that we should “move beyond the testing paradigm” and start employing learning analytics in assessment itself. Data collected from student interaction in an online environment offers the possibility to assess students on their actual interactions rather than adding assessment as a separate event. The analysis of student engagement with e-assessment is thus useful in discovering more about our students and, perhaps, directly in assessing them.

There are two lessons for ‘models for learning’ in this work. The first is that, in order to benefit from learning opportunities, students need to be sufficiently prepared and to have sufficient time to study. More generally, we should not assume that we understand our students’ learning, even at the most basic level. The analysis of student responses to e-assessment tasks provides a powerful technique for learning more and it should not be under-rated.

References

Dermo, J. 2010. In search of OSIRIS : random item selection, fairness and defensibility in high-stakes e-assessment. In *Proceedings of the International Computer Assisted Assessment (CAA) Conference, 20th-21st July 2010.* Southampton: University of Southampton.

Ding, L., and Beichner, R. 2009. Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics : Physics Education Research* 5: 020103.

Ellis, C. 2013. Broadening the scope and increasing the usefulness of learning analytics: The case for assessment analytics. *British Journal of Educational Technology* 44(4): 662-664.

Erwin, T.D. 1995. Assessment and learning: contradictory or complementary? In *Assessment for Learning in Higher Education* ed. P. Knight, 49-59. London: Kogan Page in association with SEDA.

Ferguson, R. 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning* 4(5/6): 304-317.

Glaser, R. 1981. *The future of testing. American Psychologist* 36: 923-936.

Jordan, S. 2007. The mathematical misconceptions of adult distance-learning science students. In *Proceedings of the CETL-MSOR Conference, Loughborough University, 11th-12th September 2006*. 87-92. Birmingham: Maths, Stats & OR Network.

Jordan,S. 2010. Student engagement with e-assesment. Poster presented at the International Computer Assisted Assessment (CAA) Conference, Southampton, 20th-21st July 2010.

Jordan, S. and Butcher, P. 2010. Using e-assessment to support distance learners of science. In *Physics Community and Cooperation: Selected Contributions from the GIREP-EPEC and PHEC 2009 International Conference*, 202-216. Leicester: Lula/The Centre for Interdisciplinary Science.

Jordan, S., Jordan, H and Jordan, R. 2012. Same but different, but is it fair? An analysis of the use of variants of interactive computer-marked questions. *International Journal of e-Assessment* 2(1).

Jordan, S., Ross,S., and Murphy,P. 2013. *Maths for Science*. Oxford: Oxford University Press in association with the Open University.

Nicol, D. 2008. *Technology-supported assessment: A review of research.* Unpublished manuscript available at http://www.reap.ac.uk/Resources.aspx (accessed 17th November 2013).

Pritchard, D. and Warnakulasooriya, R. 2005. Data from a web-based homework tutor can predict student's final exam score. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* Vol. 2005, No. 1: 2523-2529.

Redecker, C., Punie, Y. and Ferrari, A. 2012. eAssessment for 21st Century Learning and Skills. In *21st Century Learning for 21st Century Skills, Proceedings of the 7th European Conference of Technology Enhanced Learning, Saarbrücken, Germany*.

Walet, N. and Birch, M. (2012). Using online assessment to provide instant feedback. In *Proceedings of the Higher Education Academy STEM Conference, Imperial College, London, 13th-14th April 2012*. York: Higher Education Academy.