

Designing for Learner Engagement with e-Assessment Practices: The LE_e-AP framework

Richard Walker and Zoe Handley

E-Learning Development Team, University of York; richard.walker@york.ac.uk

Department of Education, University of York; zoe.handley@york.ac.uk

Abstract

This paper explores the issues influencing student engagement with high stakes computer-based exams, with our research drawing on feedback from a participant-informed assessment design for a master's level module in research methods. Computer-based testing was introduced to replace paper-based assessment methods for this module - supporting the swift return of marks to a large cohort of over 150 postgraduate students. The assessment design was based on a combination of short answer and multiple-choice questions to test the full range of cognitive processes including the higher order processes of analysis and evaluation.

First introduced in 2013-2014, the examination has undergone an iterative design process, drawing on feedback from key stakeholders which has informed subsequent changes to the balance of multiple choice and short answer questions in the question-set, as well as enhancements to the way that students are prepared to take the examination. This paper reports on the results of the second phase of evaluation (2014-15) in which participant feedback was solicited after the mock examination and again after the final examination in order to factor out the influence of students' perceptions of their performance on the final examination.

The combined feedback has helped us to identify a range of issues to consider in the delivery of high stakes computer-based assessments, which we present in the concluding part of this paper as an engagement framework for assessment coordinators to consider when designing computer-based assessments. The socialisation and preparation of students for computer-based assessment emerge as key responsibilities for instructors to address, with students requesting increased opportunities for practice and training in the skills necessary to succeed in the examination, enabling them to adapt test-taking strategies and develop their 'online exam craft'. Our findings suggest that students' effective engagement with computer-based testing has less to do with digital literacy and is related far more to their IT proficiency for assessment and their ability to adapt organisational and cognitive strategies to the assessment environment successfully.

Keywords: computer-based testing; summative assessment; participant-informed design

1 Introduction

High stakes computer-based testing is growing in importance within UK higher education and has developed from computer-based formative assessment practices which are already well established across the sector (Jenkins, Walker and Voce, 2014). The most recent UCISA Technology Enhanced Learning Survey (Walker et al., 2014) reported that 71% of the responding higher education institutions are now using summative e-assessment tools in their course delivery, with nine institutions having implemented defined response (multiple choice) summative tests for over 50% of their courses. This activity has been driven by a number of institutional factors, ranging from strategies to cope with large cohorts and reduce marking workloads to learner-centred issues such as the opportunity to offer students appropriate and rapid feedback on their test performance and to deliver objective and equitable testing (Jisc, 2007).

We observe that whilst there has been a strong institutional uptake in high stakes e-assessment, there has been little corresponding research conducted on students' attitudes towards and experiences of computer-based testing. A review of the literature reveals that most research to date has focused on investigating the differential impact of computer-based versus pen-and paper assessments on student achievement (Kingston, 2008; Leeson, 2006; Mead and Drasgow, 1993), and the relationship between individual differences including gender, race, socio-economic status and digital literacy and computer anxiety and performance on computer-based assessments (Leeson, 2006). Only a smaller number of studies have investigated students' attitudes towards computer-based assessment (e.g. Dermo, 2009; Deutsch et al. 2012; Hillier, 2014; Walker, Topping and Rodrigues, 2008), and these studies have tended to focus on undergraduate students and report on a specific testing intervention, rather than present a longitudinal picture of student attitudes to high stakes e-assessment over a series of assessments.

Research studies comparing performance on computer-based and pen-and-paper examinations have on the whole found that differences are generally small and not of practical significance - in particular for objective and multiple-choice assessments (Kingston, 2008; Leeson, 2006; Mead and Drasgow, 1993). Results of studies investigating the relationship between learner variables and performance on computer-based assessments are mixed and where effects of learner variables are observed these tend to be small (Leeson, 2006). While these studies are important and serve to allay concerns about equity that have been raised in explorations of students' attitudes towards and experiences of e-assessment (e.g. Hillier et al. 2014), the few post-examination surveys and focus groups that have been conducted have identified a wider range of concerns which merit further exploration. Concerns frequently cited in these studies include computer anxiety and perceptions of computer-self-efficacy and concerns relating to test security, the potential for cheating, and technical difficulties (Deutsch et al. 2012; Hillier et al. 2014; Ozden et al. 2004). Of these issues it is noteworthy that concerns about technical difficulties persisted in Deutsch et al.'s (2012) investigation of the impact of socialisation on attitudes towards computer-based assessment through the introduction of computer-based formative assessment.

This paper attempts to fill the gap in the literature on the postgraduate student experience with computer-based assessment. Our study investigates student engagement issues for a computer-based examination of research methods module at master's level, with our study conducted over an extended period of time, drawing on research from two cohorts of international MA students from the 2013-14 and 2014-15 academic years. Combining pre- and post-examination surveys and follow-

up focus group interviews, this paper provides in-depth insights into students' concerns and examines the impact of socialisation on those concerns. We discuss the consequences of these findings for instructors, presenting an engagement framework in the concluding part of this paper as a way of conceptualising engagement issues and the instructional responsibilities to be addressed in preparing students for computer-based testing

2. Designing computer-based assessment: a participative design approach

The design of effective computer-based assessments is a challenging task in terms of the preparation and review of defined or selected response questions which may serve as valid and reliable tests of student knowledge (Crisp and Palmer, 2007). Developing multiple-choice questions to tap criticality and higher order thinking skills such as how to 'analyse' and 'evaluate' (Krathwohl, 2002) is complex - see Williams (2006) for a full discussion on this theme – and this is particularly true for assertion / reason questions items and not all assessment systems support these items effectively. Criticality is however a key criterion for achieving passes standard on postgraduate master's programmes and is therefore a key outcome which must be addressed in the overall test design. One way of achieving is through the combination of short answer and multiple-choice style questions, although there is limited evidence of how best to combine open and closed questions and construct a question-set suitable for postgraduate assessment which may be delivered through computer-based testing. This was the challenge facing us in designing a computer-based exam for a research methods module for international students enrolled on the MA Language Learning & Teaching programme.

MA Research Methods in Language Learning & Teaching: the drivers for computer-based assessment

A cohort of around 150 international students is registered each year on the research methods module for this programme, with Chinese female students representing the typical profile of learner that we teach. Prior to 2013 the module was assessed through a two-hour pen-and-paper examination, made up of a combination of multiple-choice style questions (accounting for two-thirds of the question-set) and longer answers questions (one-third) which students sat at the beginning of their second term on the MA programme. Multiple-choice questions were used to assess knowledge and understanding of a wide range of methods for carrying out research in language learning and teaching. Longer answer questions were employed to assess students' ability to critically examine research reports in terms of the suitability of the methods used and the implications for the substantive claims that are made. The module team found that the marking effort was challenging to complete in the required 6-week period, when staff were busy teaching other modules and were keen to explore another approach.

The rationale for the switch to an e-assessment approach was to continue to assess the learning outcomes as before, but at the same time improve the efficiency of the delivery of the examination and consistency of marking for the multiple-choice style questions. Multiple-choice items again accounted for the main component (70%) of the examination, with the open questions employed to address the assessment of critical thinking and higher order skills. It was anticipated that the adoption of computer-based testing methods would reduce the marking load and ensure that teaching staff could continue to teach and complete their marking duties over the busy spring term.

Participant-informed design: an exploratory approach

The decision was taken by the module convenor and tutors in 2012 to approach the University's E-Learning Development Team to investigate the possibility of designing and delivering a computer-based assessment for this module. This request was timely and well received, as the University's IT and e-learning support services had been actively experimenting with e-assessment software solutions in the preceding period (2009-2012) and were keen to explore computer-based design and delivery issues with module teams with a view to establishing a centrally supported service in this domain.

Given the dearth of research evidence on effective e-assessment design and student reception of these methods, combined with the limited institutional experience of delivering high stakes computer-based tests at this time, it was felt that an evolutionary development model would be the most suitable way of approaching the planning, development and evaluation of the proposed computer-based assessment provision for this module. Our approach followed principles similar to those described by Zakrzewinski and Steven (2003) in their implementation model for computer-based assessment. Initial piloting was conducted in a 'low-stakes' context through iterative cycles of reviews and evolutionary development, with the aim of establishing a stable system before wider roll-out and use in a 'high-stakes' setting, as illustrated in Figure 1 below.

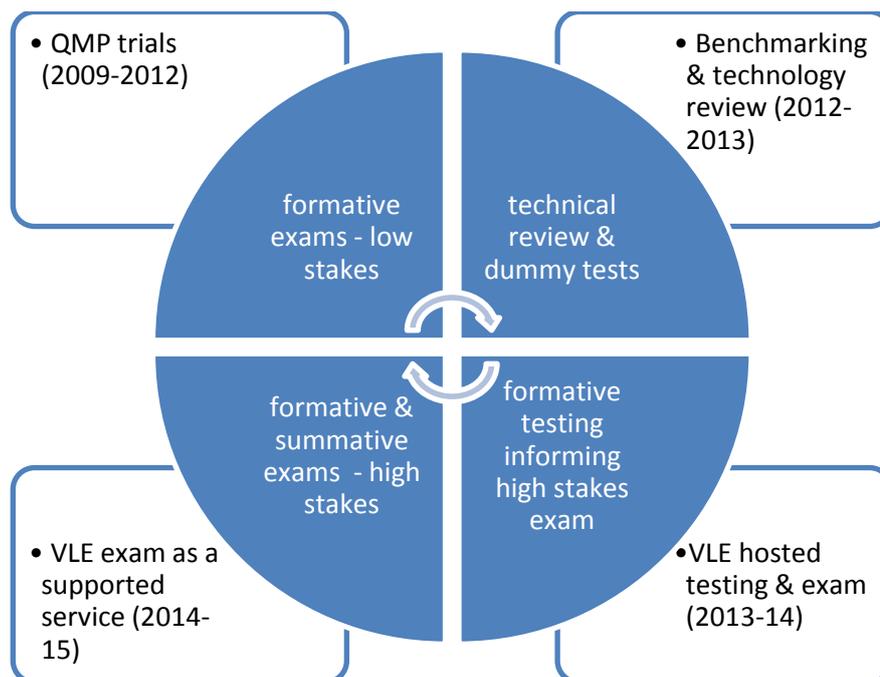


Figure 1: Evolutionary development of computer-based testing provision at the University of York (2009 – 2015)

Each cycle of development and review drew on specialist feedback from stakeholders who would be required to play a part in the successful delivery of the assessment, encompassing e-learning advisers, network administrators, examinations administrators and module tutors. Academic support staff and academics were also invited to join a test group of volunteers to trial the software in simulated examinations. We also consulted the University's Standing Committee on Assessment at regular intervals on assessment policy and procedure, keeping them updated on our progress.

Our initial trials focused on question design and testing through the use of *QuestionMark Perception* software (v.5) – involving trials with a locally installed version and the on-demand hosted service, with both solutions proving challenging to configure with our locked-down desktop service. Through a benchmarking exercise with comparator institutions and the undertaking of a full technical review into our network and supporting learning systems, we arrived at an alternative technical solution, based on the use of the assessment engine within our institutional *Blackboard Learn* (v. 9.1) virtual learning environment. A separate production instance of the institutional VLE was configured with its own dedicated servers, providing examinees with access only to the assessment area of the platform. The system was locked down so that it was available only on the university network and for the duration of the delivery of the examination. Further, during the delivery of the examination, the PCs in the computer labs in which the exam was delivered were also locked down, so that students could only access the examination and no other resources on internal networks or the internet. We also took the additional step of serving up the examination questions in a randomised order, with the intention of preventing cheating in labs which were not equipped with privacy screens between work stations. Multiple labs were prepared in this way for simultaneous testing across different venues; this was seen as the best way of managing the large cohort of examinees, given the limitations of the physical estate and absence of a dedicated test venue on campus which we could use to accommodate a cohort of this size.

Once confident with the technical performance of the VLE hosted solution and the examination support arrangements, we proceeded to formative and summative testing with students - conducted for the first cycle of VLE hosted testing for the 2013-14 cohort. Students in both the one-hour formative exam and two-hour summative exam tackled a combination of multiple-choice style question and open questions: the former were worth between 1 and 6 marks depending on the number of parts and the difficulty of the content assessed – the latter each worth 10 marks and requiring students to type around five sentences in their responses.

We recognise students as key stakeholders in the assessment process and their feedback was actively solicited on the assessment experience through questionnaires and focus groups activities. We also conducted difficulty and discrimination analyses of the combined question-set based on their formative and summative exam performances. Specifically tests were undertaken to identify whether any questions were too easy or too difficult and to determine whether the questions enabled us to differentiate between higher and lower proficiency students (Haladyna, 1999). As a result of the analyses and the feedback we received from students relating to their concerns over typing proficiency and noise levels in the completion of open response questions, adjustments were made to the number of open response items included in the summative examination. The number of multiple-choice style questions was also reduced from 32 to 19, with these changes being finalised for the delivery of the next cycle of VLE hosted testing for the 2014-15 cohort. The randomised order in which the questions were served up to students was retained though across both formative and summative examinations in order to prevent cheating and ensure the integrity of the testing process.

3. Research & evaluation methods

The framework for student engagement presented in this paper is based on students' experiences of the rollout of the VLE hosted examination, drawing on data derived from free-text responses to questionnaires and transcripts of focus group discussion as an evidence-base for our findings. The

questionnaires and focus group instruments were informed by Dermo's (2009) Student Perceptions of eAssessment Questionnaire (SPEAQ) and a review of previous instruments and research on students' attitudes towards and experiences of computer-based testing (Deutsch et al. 2012; Ferrao, 2014; Frein, 2011; Hillier, 2014; Williams and Wong, 2009). The research explored students' prior exposure to computer-based testing and their reflections on the preparation for this exam and their experiences of taking the exam and thoughts on the suitability of the method and recommendations for how it could be improved in the future.

For the 2013-14 cohort, the students were surveyed immediately after the summative examination; for the following year (2014-15 cohort), the evaluation was administered after the formative examination and again after the summative examination, in order to control for students' perceived performance in the final summative examination. Survey returns were received from 48 students from the 2013-14 cohort and 42 students from the 2014-15 cohort completed both the formative and summative questionnaires. Profile questions were asked in the 2014-15 post-mid-term exam and revealed that the vast majority of respondents were Chinese female postgraduates in the 20-24 years of age range, who were entirely new to our institution and from a wide range of undergraduate disciplines.

A combined total of 18 Chinese female students from the 2014-15 cohort volunteered to participate in three separate focus groups in January 2015, with five students from the 2013-14 cohort participating in February 2014. Transcripts from the focus groups were generated and a qualitative content analysis performed (Hsieh and Shannon, 2005). The unit of analysis was a line in the transcript, which in turn could address multiple units of meaning. Comments were categorised and then mapped against an evaluative framework based on Dermo's key themes, addressing affective variables, validity, practicality, reliability, security, and pedagogical issues related to the assessment method. The categorisation and mapping processes were then repeated for the open comments from the surveys, and the outputs from these analyses were then compared to form a rich picture of student experiences with computer-based testing, from which common themes in the student assessment experience were derived.

4. Findings and discussion

For the cohorts under review, feedback predominately came from Chinese female students encountering computer-based testing for the first time at the University of York. Given the timing of the mid-term (formative) and summative exams occurring early on in this one-year MA programme, they had received limited opportunities to familiarise themselves with institutional assessment practices and came with fresh perspectives on computer-based testing.

The combined findings highlight a range of issues which appear to have influenced their reception of computer-based testing methods, for which they had no control over the method (paper-based or computer-based) or timing of the assessment. These issues are presented below in a discussion of the key reception themes and recommended actions for instructors to consider in assessment design and delivery. The **Learner Engagement with e-Assessment Practices (LEe-AP)** framework derived from this feedback is presented in the Appendix to this paper.

4.1. Socialisation of learners

Student feedback indicated a need for the instructor to articulate the rationale and the complementary nature of the assessment methods to the discipline being assessed, with some students unconvinced by the need for computer-based testing methods.

“I think this exam could also be paper based because in research methods we always use computers to analyse some data, always computer but in this exam only a few questions is about how to analyse data. The others could also be shown on paper, so for me, I think, it’s not really necessary to be done with computer.” (Post-Test Focus Group 2014-15)

This chimes with findings from Hillier’s cross-disciplinary review of undergraduate students’ attitudes towards e-assessment and the perceived ‘readiness’ of disciplines to match a computerised assessment approach (Hillier, 2014). Disciplines and module area which have a less obvious fit with e-assessment methods will require greater support to students in the introductory phase in managing anxiety levels.

The rationale for e-assessment should also ideally address the benefits of this approach. Our findings confirm previous research (e.g. Nouandegani, 2012) in highlighting the perceived presentational benefits of typed open responses as opposed to poorly handwritten answers and the favourable impression that students believe this will have on markers:

“If we are bad at writing, yes especially in China, we lose marks because of that.” (Post-Test Focus Group 2014-15)

Students also welcomed the stricter enforcement of examination conditions in the UK compared with China, contributing to a fairer assessment experience:

“In China we sit close to each other and the teacher is not strict, so we can sometimes we can see other answers. In UK, I think the atmosphere is very strict.” (Post-Test Focus Group 2014-15)

Notwithstanding these benefits some students felt that there was a generational bias in favour of more technically literate students and an equity issue affecting mature students returning to education, who might take longer to adjust to computer-based exams.

“It’s kind of fair for most students because using a computer is almost a necessity for us and especially for our generation but it is not as much fair, as other generation.” (Post-Test Focus Group 2014-15)

“Some of our classmates after they have had some experience ... they return to school to get more experience in teaching. It could be some difficulty for them to use a computer in typing when they attend the examination, so it could take them longer time to get used to the system, so I think it could unfair for them.” (Post-Test Focus Group 2014-15)

This reflects a commonplace perspective on the existence of Net generation (Tapscott, 2008) or Digital Natives (Palfrey and Gasser, 2008; Prensky, 2001, 2001a) with distinctive skills and aptitudes for digital learning, which still persists and perhaps is more keenly felt by mature and returning students to full-time education, despite recent studies in the UK debunking this notion (Jones et al. 2010; Margaryan, Littlejohn and Vojt, 2011). Whilst research on computer-based testing (e.g. Frein,

2011; Leeson, 2009) has found no difference in test results between highly experience and less experienced users of computers, our findings align with previous studies of student perceptions of e-assessment (Fluck, 2013; Hillier, 2014) in underlining the need to address student concerns about fairness in the orientation phase and to offer reassurance from the outset. The orientation of learners to a new assessment method – particularly when this method departs from established assessment norms within a study programme – emerges as a key responsibility for instructors to address. This is consistent with previous studies (e.g. Deutsch, 2012; Zakrzewski and Steven, 2003) which have identified the need for instructors to integrate the assessment method at an early stage into the curriculum and in this way address and reduce student anxiety levels, which may stem from a variety of factors touching on computer self-efficacy, concerns over security and cheating, to fears over computers crashing during an online examination (Mogey and Sarab, 2006).

4.2. Preparation of students for assessment

Computer aversion or anxiety as a barrier to the adoption of computer-based assessment is a well-researched concept (e.g. Durndell and Lightbody, 1994; Meir, 1985, 1988). Meir defines this as a negative affective and cognitive state which often occurs when individuals have low expectations about the rewards of using computers or confidence in their ability to use computers effectively. Adequate preparation appears key to counteracting this mind-set, but there is less agreement on the nature of the interventions that are required to address student anxiety levels. Student feedback in our study highlighted three areas where preparation may be needed, touching on digital skills, examination technique and revision strategy.

(a) Digital skills

Zakrzewski and Steven (2003) have highlighted the enhancement of students' IT skills as a prerequisite for student preparation, but it is far from clear what this actually entails. The literature study by Leeson (2006) has shown that there is no established relationship between examinees' level of computer familiarity and performance on computer-based tests. Similarly, we found no association between any of the indicators of digital literacy that we surveyed which included previous experience of computer-based testing and keyboarding skills and preferences for computer-based versus paper-based assessment. Keyboarding skills were however frequently raised by students, with some perceiving lack of adequate keyboarding skills as a potential source of equity issues:

'Typing skills can be very important which is unfair to those who are not good at typing.'
(2013-2014 Post-Test Questionnaire)

It is important though to acknowledge that preparation of students may need to go beyond equipping them with the requisite level of keyboarding skills. Students also raised issues associated with their familiarisation and use of university-provided hardware and software:

"I'm not used to using the keyboard because it's different from laptop keyboard." (Post-Test Focus Group 2014-2015)

"I feel in the real exam, I found there's no correction tools for you to correct." (Post-Test Focus Group 2014-2015)

“I find that when I type really faster and quickly, there are typing errors I won’t recognise because there are no right lines on the line of words that I typed wrong.” (Post-Test Focus Group 2014-2015)

Our research suggests that we need to draw a distinction between *digital proficiency* – reflected in the effective day-to-day use of technology for learning (e.g. from email to essay writing) – with the capability to use unfamiliar technology under time pressure in computer-based exams, which we may refer to as *IT proficiency for assessment*. In other words our research suggests we need to go beyond equipping students with basic IT skills and familiarising them with the assessment environment and providing them with opportunities to develop proficiency in typing under time pressure, including use of unfamiliar hardware and software under authentic examination conditions. This is best achieved through the creation of computer-based formative assessments which align with the format of the summative examination.

(b) Exam technique: test-taking strategies for online examinations

Another important determinant of performance in examinations which students identified was effective test-taking strategies. Like the students in Hong et al.’s (2006) investigation of the test-taking strategies of high school students in a paper-based mathematics examination and Walker et al.’s (2008) investigation of first and second year university students’ expectations and perceptions of a computer-based science examination, the students in our study attempted to deploy a range of *organisational* (e.g. time allocation and sequencing) and *cognitive* (e.g. checking, eliminating, and using memory aids) strategies. Transferring strategies developed for paper-based examinations to the computer-based examination was, however, not always straightforward for our students and the items related to time management and checking answers were among those that students displayed the most negative attitudes towards in both pre-test and post-test questionnaires. Difficulties with time management appeared to be associated with the fact that questions were presented one at a time and it was difficult to get an overview of the exam:

‘We can’t see the exam as a whole. Reading the exam as a whole by clicking may waste time.’ (2014-2015, Pre-Test Questionnaire)

“In terms of time management, I think we, when we are doing the handwriting exam, I know what questions to I have, but in e-exam I just didn’t know what I am currently facing and I don’t know what kind of questions, you know closed or open question or is coming next.” (2014-2015, Post-Test Focus Group)

With respect to sequencing some students noted a preference for answering the multiple-choice questions before open response questions due to their lack of confidence in their typing:

“The exams I used to have at the beginning I used the multiple choice but this time at the beginning was open questions, so I feel a little nervous when I should first type so I’m not confident on typing, so just makes me a little nervous, so I choose a multiple-choice first.” (2014-2015, Post-Test Focus Group)

Note that the questions in the 2014-2015 examination were organised according to weighting, with questions allocated the highest number of marks presented first in response to feedback on the organisation of the 2013-2014 examination:

'A good exam paper should put the easy part first, while the on-line exam, the order of the questions are not the same, the difficult question could appear first, which will give more pressure to the students. I experienced this in the Research Methods exam. I spent about 5-10 minutes on the very first question, it made me nervous in the following part, which made me feel awful. And the last question was to give the mean, mode, median, which was supposed to be very easy, but there was only one minute left, I added the numbers very quickly and added one more number. It was awful.' (2013-2014 Post-Test Questionnaire)

The possibility to annotate the examination and externalise their knowledge of topics were among the most frequently noted cognitive strategies:

'I do not like not having the ability to circle questions I am unsure about or make notes to myself about which questions to come back to. During written assessments, I often write all over my test questions with arrows, circles, and other brainstorming sketches and it is difficult to work through the online assessment without these techniques.' (2014-2015 Pre-Test Questionnaire)

Similarly some students felt that the lack of annotation facilities made it difficult for them to check their work:

'Easy-go-back and check approach. Don't just provide question numbers and the function of flagging as in some online exams. Provide key terms of the question for later reminding or shown only in checking process. Student could even have a chance to make notes for reminding'. (2014-2015 Pre-Test Questionnaire)

As the students themselves suggest, some of the above concerns about the appropriation of test-taking strategies might be addressed through modifications to the design of the test and the platform on which it is delivered. However, it is also clear from the feedback that one assessment design will not fit all students' test taking strategies. Students differed with respect to their preferences for the sequencing of questions. As recommended by Hillier et al. (2014) and Zakrzewski and Steven (2003) the best solution may be to provide students with opportunities to familiarise themselves with the computer-based assessment environment and adjust the test-taking strategies that they have developed in paper-based contexts accordingly to suit the computer-based environment:

"I have to say that we should probably arrange weekly or monthly test to get us familiar with the test, if we want to grade enough exam." (2014-2015 Post-Test Focus Group)

'More practice to help the students be familiar with the system as well as question type.' (2014-2015 Pre-Test Questionnaire)

(c) Revision strategy

Focus group feedback also shed light on how students approach their preparation for computer-based examinations. The novelty of the assessment method with the combination of multiple choice and short answer open questions challenged students in terms of their personal revision strategies, which had previously been based on memorising content for the production of essay responses:

“I will prepare more about how to deal with the multiple choices, but in China we will focus on the memory so we try to remember the long answers to these questions.....for the Chinese exam, I will remember all of the answers, long sentences, but I will not do this for this module.[Just] the keywords because in China, the teacher will ask us to recite all the principles, all aims, something like this, long sentences but I think it’s not necessary to recite some for the assessment in York.” (Post-Test 2014-2015 Focus Group)

This suggests that students may need clearer guidance on how to manage their self-study time, taking account of the assessment format so that they come to the exam prepared to apply their learning to problems rather than to merely reproduce knowledge. As one focus group participant acknowledged:

“The questions (in the) exams are practical use of what you’ve learnt not just the knowledge or something else.” (Post-Test 2014-2015 Focus Group)

4.3. Assessment design and interface

(a) Assessment design

Two main issues drove the design of the original summative examination that was delivered in 2013-2014, namely: (i) ensuring that the question set tapped higher order thinking skills and (ii) preventing cheating in PC labs which were not equipped with privacy screens. Our research - like much before it - suggests that appropriately pitching a multiple-choice examination requires careful planning and iterative cycles of question review (Crisp and Palmer, 2007; Haldyna, 1999). Analyses of students’ performance on the 2013-2014 examination suggested that the balance of open response to multiple-choice questions was not appropriate. Despite careful review of the 2012-2013 paper version of the examination including content and cognitive behaviour reviews (i.e. reviews of the coverage of the examination and the cognitive behaviours that each item taps) and subsequent revisions, students’ marks on the 2013-2014 were skewed towards the lower end of the university scale. The automatic difficulty and discrimination analyses generated by Blackboard Learn’s assessment engine suggested that students found the open-response questions particularly difficult. The balance between open-response and multiple-choice questions was therefore adjusted for 2014-2015. Careful planning and reviewing of questions before administration and annual reviews including difficulty and discrimination analyses are therefore recommended until an appropriate and stable pitch is achieved.

In relation to cheating, while some students were positive about the randomised order in which questions were served up to them, others expressed concerns about equity:

‘Random questions for each student don’t represent the level of difficulties, for some students could encounter long answer question at Q1 which gives little confidence of students to move on. More it could also waste time in trying to answer that question and therefore time is not enough.’ (2013-2014 Post-Test Questionnaire)

Further, randomisation caused distraction and appeared to raise some students' levels of anxiety – the noise associated with typing meant that students were aware of the progress their neighbours were making with the questions:

“Maybe the first three questions should all be typing, maybe, so all the people are typing maybe, so they were not just distracted .”(Post-Test Focus Group 2014-2015)

“... I will tidy so much words but other students they finish one by one and very very quickly because the person needs a multiple choice item and I was ‘oh my god’ maybe they have written number 7 or number question, but I still do the first.” (2014-2015 Post-Test Focus Group)

Ideally, all PC labs used for examination purposes should be equipped with privacy screens between work stations. This, however, may not be feasible in ‘greenfield’ e-assessment sites like our own where dedicated computer-based testing venues have not been established on campus. Where randomisation in the serving up of question items to work stations is necessary to prevent cheating taking place, consistency in the presentation of questions to students is essential to promote fairness, avoid distractions and avoid increasing students' levels of anxiety. Further, consideration of possible test-taking strategies ought to be taken into account. It is however acknowledged that students may have individual preferences with respect to test-taking strategies and the presentation of the examination. It is therefore unlikely to be possible to design the examination to accommodate all students' preferences. Further to students' comments on sequencing presented above, students also noted:

“It’s no sense to put an open question for ten points at the beginning, so because our brain doesn’t work at the beginning to write/type so much.” (Post-Test Focus Group 2014-2015)

It is therefore essential that the assessment platform is flexible and that students have ample opportunities for ‘practice’ through formative computer-based assessments in order to develop dedicated test taking strategies for computer-based assessments.

(b) Design of assessment interface

Our research also suggests that the design of the assessment interface requires careful consideration in order to accommodate students' varying test taking strategies. As previously discussed, students reported using a range of organisational and cognitive strategies. In relation to organisational strategies, in order to help students get an overview of the examination and allocate time accordingly, examinees wanted to be able to see all of the questions on one screen first -a protocol which has been reported on in previous studies (e.g. Frein, 2011):

“Maybe one page with all the questions on the first page and then one by one.” (Post-Test Focus Group 2014-2015)

There was no consensus on item presentation thereafter, although Leeson’s summary of the research literature suggests that multiple items on screen may have a facilitating effect in allowing examinees to skip, scan, and build off previous item information, and may counter the effects of single item presentation on screen which may encourage hurried responses and in this way increase errors.

Easy navigation of the examination is also essential given that it is unlikely to be possible to accommodate students' differing preferences for the sequencing of questions. In order to support navigation, students made a number of suggestions including tagging questions with keywords, and grouping them into categories:

'Easy-go-back and check approach. Don't just provide question numbers and the function of flagging as in some online exams. Provide key tests of the question for later reminding or shown only in checking process. Students could even have the change to make notes for reminding.' (2014-2015 Pre-Test Questionnaire)

'Yes, also, I don't think it's necessary to just complete one question one time because maybe two or three that's OK, because one and you have to click, click, click. How about one hundred questions ... Maybe here, here, here and to like categorise the options, which one appears here to go back and which one to out of that, and which to submit.' (2014-2015 Post-Test Focus Group)

(c) Preparation and management of assessment centres

Our research also has implications for the preparation of multi-purpose PC labs which are to be used as assessment centres. Start-up and preparation of the exam environment should be completed before students enter the exam room. Equity in exam conditions should be established across all assessment centres in use for an examination, touching on issues such as provision of exam papers and note paper to students. Noise management is a key concern in the delivery of computer-based exams and one that has been reported in previous studies (Fluck, 2013; Hillier, 2014), requiring thought over the size of assessment venues and the numbers of students grouped together, as well as the quality of keyboards which are provided for students to use.

5. Conclusions

In drawing conclusions from this study, we should acknowledge the limitations of the research evidence that has been generated, which is derived from full-time postgraduate taught students within a 'greenfield' institutional context where high stakes e-assessment practices had not previously been supported by the university. The cohort was predominantly comprised of Chinese female nationals with limited prior exposure to e-assessment practices in UK higher education and there had been little opportunity to normalise online assessment procedures prior to the delivery of the research methods module. We acknowledge the distinctive features of this study which may impact on the generalisability of the findings and recognise that the assessment context for other institutions and cohorts will differ greatly; notably for undergraduate students on multi-year programmes of study there are likely to be greater opportunities to embed and normalise e-assessment practices and their experience may be quite distinct from students enrolled on one-year taught postgraduate programmes.

Notwithstanding these limitations, the study offers an insight into the factors influencing postgraduate students' reception of computer-based testing when encountering these methods for the first time. Our collective findings are summarised in the engagement framework for student acceptance of computer-based exams, which is presented in the Appendix to this paper. Building on the existing literature (Dermo, 2009; Hillier, 2014) we have identified a range of issues influencing

students' acceptance of e-assessment methods, which we have grouped into three key categories focusing on: (i) the socialisation of learners to the assessment method; (ii) the guidance and preparation which should be offered to students for the computer-based assessment; and (iii) issues impacting on the assessment design and infrastructure that will be employed. In response to these issues we propose a framework of recommended actions for assessment coordinators to consider in managing students' transition to computer-based assessments.

We have observed that the socialisation of learners is a key step in this process, requiring an upfront investment of time to explain to students the rationale and value of computer-based assessment methods, particularly when this is a new departure from established assessment practices. Students will also need adequate preparation to negotiate the transition from formative to summative computer based assessments. Our findings show that the provision of opportunities for students to practise the exam format under authentic conditions (i.e. timed conditions within the test environment) is as important as orientation to the question types that they will encounter in the exam and outweighs issues of digital literacy and keyboarding skills. Attention should instead be directed to fostering students' IT proficiency for assessment – i.e. the exam craft that students will need to display in the performance of the online exam. This will require students to adapt rather than directly transfer paper-based exam techniques to the online context, addressing both organisational and cognitive strategies in their approach.

References

- Crisp, G. & Palmer, E. (2007). Engaging academics with a simplified analysis of their multiple-choice question (MCQ) assessment results. *Journal of University Teaching & Learning Practice*, 4 (2), 88-106.
- Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, 40(2), 203-214
- Deutsch, T., Herrmann, K., Frese, T. & Sandholzer, H. (2012). Implementing computer-based assessment – A web-based mock examination changes attitudes. *Computers & Education*, 58(4), 1068-1075
- Durndell, A. & Lightbody, P. (1993). Gender and computing: Change over time? *Computers in Education*, 21, 331-336.
- Escudier, M. P., Newton, T. J., Cox, M.J., Reynolds, P. A., & Odell, E. W. (2011). University students' attainment and perceptions of computer delivered assessment; a comparison between computer-based and traditional tests in a 'high-stakes' examination. *Journal of Computer Assisted Learning*, 27, 440-447
- Ferrao (2010). E-assessment within the Bologna paradigm: Evidence from Portugal. *Assessment & Evaluation in Higher Education*, 35(7), 819-830.
- Fluck, A. (2013). Implementation of on-campus digital examination practices. Presented at Transforming Assessment in a Digital Era, Melbourne, Australia. Retrieved 30th July 2015 from: <http://www.eexams.org/stories/Documents/TA%20in%20a%20Digital%20Era%20-%20Fluck.pdf>
- Frein (2011). Comparing in-class and out-of-class computer-based tests to traditional paper-and-pencil tests in introductory psychology courses. *Teaching of Psychology*, 38, 282
- Haladyna, T. (1999). *Developing and validating multiple-choice test items*. Lawrence Erlbaum: New Jersey.
- Hillier, M. (2014). The very idea of e-Exams: Student (pre) conceptions. . In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin 2014* (pp. 77-88).
- Hong, E., Sas, M. & Sas, J. C. (2006). Test-taking strategies of high and low mathematics achievers. *The Journal of Educational Research*, 99(3), 144-155.
- Hsieh, H-F., Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15 (9), (pp.1277-1288).
- Jenkins, M., Walker, R., & Voce, J. (2014). Achieving flexibility? The rhetoric and reality of the role of learning technologies in UK higher education. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin 2014* (pp. 544-548).
- Jisc (2007). Effective Practice with e-Assessment, HEFCE. Retrieved 31st July 2015 from: http://www.jisc.ac.uk/publications/documents/pub_eassesspracticeguide.aspx
- Jones, C., Ramanau, R., Cross, S. & Healing, G. (2010). Net generation or Digital Natives: Is there a distinct new generation entering university? *Computers & Education*, 54 (3), pp. 722–732. Retrieved 31st July 2015 from: <http://oro.open.ac.uk/19890/>
- Kingston, N. M. (2008). Comparability of computer- and paper- administered multiple-choice tests for K-12 populations: A synthesis. *Applied Measurement in Education*, 22(1), 22-37
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212-218
- Leeson, H. V. (2006). The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing*, 6(1), 1-24
- MacCann, R. (2006). The equivalence of online and traditional testing for different subpopulations and item types. *British Journal of Educational Technology*, 37(1), 79-91

- Margaryan, A., Littlejohn, A., & Vojt, G. (2011). Are digital natives a myth or reality? University students' use of digital technologies. *Computers & Education*, 56 (2011): 429-440. Retrieved 7th August 2015 from: http://www.unil.ch/files/live/sites/magellan/files/shared/Are_digital_natives_a_myth_or_reality.pdf
- Marsden, E. (2011). Questionnaire feedback for the Question Mark Perception Mock Exam. Unpublished report, University of York.
- Mead, A. D. & Drasgow, F. (1993). Equivalence of computerized and paper and pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114, 449-458.
- Meier, S. (1985). Computer Aversion. *Computers in Human Behavior*, 1: 171-179.
- Meier, S.T. (1988). Predicting individual differences in performance on computer-administered tests and tasks: development of the computer aversion scale. *Computers in Human Behavior* 4: 175-187.
- Mogey, N. & Sarab, G. (2006). Exams, essays and tablet computers – trying to make the pill more palatable. In: Danson, M. (ed.). 10th CAA *International Computer Assisted Assessment Conference: Proceedings of the Conference on 4th and 5th July 2006 at Loughborough University*, 313-320. Retrieved 7th August from: https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/4414/1/Mogey_N_Sarab_G_h2.pdf
- Noubandegani, P. (2012). Students' Perceptions of Computerized TOEFL Test. *Language Testing in Asia*, 2 (2), 73. Retrieved 8th August 2015 from: <http://www.languagetestingasia.com/content/pdf/2229-0443-2-2-73.pdf>
- Ozden, M. Y., Erturk, I., & Sanli, R. (2004). Students' perceptions of online assessment: A case study. *Journal of Distance Education*, 19(2), 77-92.
- Palfrey, J. and Gasser, U. (2008) *Born Digital: Understanding the First Generation of Digital Natives*. New York: Basic Books
- Prensky, M. (2001.) Digital natives, digital immigrants. *On the Horizon*. NCB University Press, Vol 9(5).
- Prensky, M (2001a). Digital Natives, Digital Immigrants Part II: Do they really think differently? *On the Horizon*. NCB University Press, Vol. 9 (6).
- Tapscott, D. (2008) *Grown up digital: How the Net generation is changing your world*. New York: McGraw-Hill.
- Venables, H. & Fernandez, K. (2011). Evaluation of Question Mark Perception: An online assessment system. Unpublished Report, University of York.
- Walker, D. J., Topping, K., & Rodrigues, S. (2008). Student reflections on formative e-assessment: Expectations and perceptions. *Learning, Media and Technology*, 33(3), 221-234.
- Walker, R., Voce, J., Nicholls, J, Swift, E., Ahmed, J., Horrigan, S., & Vincent, P. (2014). 2014 Survey of Technology Enhanced Learning for higher education in the UK. Universities and Colleges Information Systems Association, Oxford, UK. Retrieved from: <http://www.ucisa.ac.uk/tel>
- Williams, J. B. (2006). Assertion-reason multiple-choice testing as a tool for deep learning: A qualitative analysis. *Assessment & Evaluation in Higher Education*, 31(3), 287-301
- Williams & Wong (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, 40(2), 227-236.
- Zakrzewski, S. & Steven, C. (2003). Computer-based assessment: quality assurance issues, the hub of the wheel. *Assessment & Evaluation in Higher Education*, 28 (6), 609-623.

APPENDIX

Learner Engagement with e-Assessment Practices (LEe-AP) framework

Theme	Issues	Recommended actions
<p>1. Socialisation of learners <i>Orientation of learners to assessment methods</i></p>	<ul style="list-style-type: none"> (i) Prior exposure to e-assessment / familiarisation (ii) Timing within study programme (iii) Rationale for e-assessment: perceived fit with curriculum & teaching methods and perceived benefits (iv) Perceived fairness and equity in assessment methods 	<p>Offer comprehensive induction procedure addressing:</p> <ul style="list-style-type: none"> - rationale for e-assessment - suitability of the methods; their fit within the study programme and complementarity to discipline and teaching methods. <p>Reassurance to students on exam procedure and equity of methods</p>
<p>2. Preparation of students for assessment <i>Guidance and preparation to students for e-assessment</i></p> <p>(a) Digital skills</p>	<ul style="list-style-type: none"> (i) Keyboard proficiency under exam conditions (ii) Reading from screen (iii) Familiarisation with exam environment & controls for navigation / question selection / writing 	<p>Prior exposure to exam hardware or BYOD approach, based on students using their own devices (locked down).</p> <p>Authentic practice opportunities under timed conditions: frequent opportunities for students new to e-assessment</p>

<p>(b) Exam technique: test-taking strategies for online exams</p>	<p>(i) Familiarisation with exam format and question types</p> <p>(ii) 'Online exam craft': addressing organisational (question selection; time management) and cognitive strategies (reviewing & editing answers)</p>	<p>Video tutorials and class-based guidance and feedback on formative tests.</p> <p>Provision of multiple practice opportunities, enabling students to review and adjust paper-based test taking strategies</p>
<p>(c) Revision strategy</p>	<p>(i) Alignment of formative and summative test formats</p> <p>(ii) Study skills support</p>	<p>Provision of authentic practice opportunities – replicating question format for summative test (question type / level of difficulty)</p> <p>Guidance on revision strategies</p>
<p>3. Assessment design & interface <i>Organisation and presentation of question-set, preparation of user interface and assessment venue</i></p> <p>(a) Assessment design</p>	<p>(i) Balance (open/closed items) and volume of questions per allotted time</p> <p>(ii) Sequencing and order of questions</p>	<p>Perform iterative question review (difficulty & discrimination analyses)</p> <p>Ensure consistency in ordering of questions (question value) to individual work stations when randomised question selection methods are employed.</p>
<p>(b) Design of assessment interface</p>	<p>(i) Item presentation</p> <p>(ii) Intuitive controls (e.g. flagging & navigation)</p>	<p>Comprehensive induction and guidance on exam environment and controls</p>
<p>(c) Preparation and management of assessment centre(s)</p>	<p>(i) Consistency in presentation of assessment venues (set-up and invigilation procedures)</p> <p>(ii) Noise management: keyboards</p> <p>(iii) Specialist support & contingency measures (technical failures)</p>	<p>Exam management ensuring consistency in presentation and management of exam venues.</p> <p>Review of supported hardware</p> <p>Contingency policy & procedures in place</p>