The Yin/Yang of Innovative Technology Enhanced Assessment for Promoting Student Learning

Maggie Hutchings, Anne Quinney, Kate Galvin, Vince Clark
Bournemouth University, Bournemouth, UK
mhutchings@bournem
aquinney@bournem
kgalvin@bournem
vclark@bournem

Abstract: While more sophisticated and constructively aligned assessment is encouraged to promote higher level learning, it is easier to assess knowledge and comprehension than critical thinking and making judgements (Bryan & Clegg 2006). Managing the logistics and resources required for assessing large numbers of students challenges the ethos of placing students at the heart of the learning process and helping them take responsibility for their own learning. The introduction of innovative technology enhanced assessment strategies contests our understanding of the purposes of assessment and affords opportunities for more integrated and personalised approaches to learning and assessment across disciplines.

This paper will examine the design, implementation and impacts of innovative assessment strategies forming an integral part of a collaborative lifeworld-led transprofessional curriculum delivered to cohorts of 600 students in health and social work using technology to connect learners to wide-ranging, humanising perspectives on evidence for guiding practice. Innovative assessment technologies included group blogs, multiple choice electronic or computer assisted assessment (CAA), and an audience response system (ARS) affording combinations of assessment for learning and assessment of learning.

We will explore, through analyses of student assessment experiences and student and staff evaluations, how these innovative assessment approaches contribute to effective and efficient blended education enabling students to enhance their practice through promoting and developing critical thinking and reflection for judgement-based practice (Polkinghorne 2004). Secondly, we will debate the yin and yang of contrasting and connecting values associated with the controlled, systematic measurement and objectivity of multiple choice assessments, compared with the formative, iterative and subjective nature of reflective blogging. We will consider relationships between teaching and learning strategies and experiences, breadth and depth of knowledge, passive and active approaches to learning, efficiency and effectiveness, individual and group, multiple choice and discursive assessments, face-to-face and online, on-campus and off-campus learning and assessment experiences.

Keywords
Innovative assessment; computer assisted assessment; technology enhanced learning; group blogs; audience response systems

1. Introduction: innovative technology-enhanced assessment

For assessment to be considered innovative it needs to do much more than introduce new technologies into the assessment diet for students. Innovative assessment strategies need to encourage assessment for learning over assessment of learning (Bryan & Clegg 2006) argue that innovative assessment should enable student learning and judgements, rather than acting as instruments of justification, measurement and limitation. But introducing innovative assessment can be a high-stake investment fraught with risk for students and higher education institutions (HEIs). The significance of assessment for students cannot be underestimated. Assessment frames learning by creating learning activities which orientate learning behaviour (Gibbs in Bryan & Clegg 2006) and while students can escape bad teaching, they cannot avoid bad assessment (Boud 1995). Boud and Falchikov (2007) point out that assessment is also a major concern and burden for those teaching students and suggest HEIs are afraid to change assessment systems because of the risks and major
effort entailed, leading to slow incremental change, compromise, and inertia. The aspiration for innovative assessment approaches to provide ways of redefining assessment “as an instrument of liberation” (Bryan & Clegg 2006, p.1) for both students and HEIs may prove elusive within the realities, challenges, and constraints of organisations.

While the value of different forms of technology-based or e-assessment and their potential for offering learning benefits to students, through more frequent and immediate feedback, has been recognised, the design and implementation of e-assessments brings major challenges for HEIs (Whitelock 2009). Computer assisted assessment (CAA), encompassing computer-based assessment (CBA) and web-based assessment (WBA), offers immediate scoring and feedback to students and the potential to reduce the marking workload of educators. Initial start-up costs are associated with question authoring and review, and development of question banks (Ricketts & Wilks 2002; Deutsch et al 2012). Securing and maintaining support of University services including computing, estates, administration and quality assurance (Bull 1999) also needs to be factored in to development and running costs. But once the technologies and systems are in place, benefits can be realised for students, with online scoring enabling faster and more regular feedback, and for organisations, with automated marking and feedback offering human resource efficiencies, reducing marking workloads and ensuring greater accuracy at the point of marking.

CAA is associated with provision of multiple choice questions (MCQs) (Bull & Danson 2004). Once a databank of questions has been developed, the technology enables items to be selected for online delivery, automated scoring and feedback, and report generation. This type of assessment has been described as ‘objective’ but Biggs and Tang (2007, p.203) argue that it is ‘not more scientific’ or any less prone to error. The potential for error has simply moved from the point of marking to the point of authoring questions and response options; this is not necessarily any less subjective than assessing an essay or student presentation. Further the value of MCQs to test anything beyond basic knowledge has been challenged. Biggs and Tang (2007) suggest that while well-designed MCQs can assess higher order learning they are rarely developed in this way. Breadth of learning required to answer lots of questions challenges the value of MCQs for depth of understanding and as a means of assessment for learning. Biggs (2007, p. 238) articulates the dangers of CAA where it is used to reinforce the idea that ‘knowing more’ is equated with good learning. The pedagogical foundations for a constructively aligned curriculum where assessment strategies, learning activities and learning outcomes are clearly aligned are at risk if CAA adopters cannot rise to the challenge of designing and delivering more sophisticated questions that can assess higher order learning. The yin/yang of CAA is highlighted here with its main advantages apparently lying in more efficient assessment procedures bringing organisational and logistical benefits for assessing large student numbers through technology but calling into question its pedagogical value and effectiveness.

Audience response systems (ARS) have been employed in a range of subjects and described as a viable and flexible tool for engaging students in more active learning in classroom settings (Medina et al 2008), by providing rapid feedback about understandings, misunderstandings and clarification (Miller & Hartung 2012, Cain & Robinson 2008), and for testing and evaluation (Mareno et al 2010). Caldwell (2007, p. 2) advises that usage in ‘summative high-stakes testing’ is rare. ARS, or electronic voting systems (EVS) enable participants to respond to MCQs displayed to the whole class (Draper 2009). The aggregated answers, which can be presented using a variety of formats, for example, bar charts or pie charts, are displayed on screen for all to see how responses are distributed across options. But as with CAA, the pedagogical value of ARS is also questionable, relying as it does on using MCQs. Draper (2009, p.286-287) describes how the MCQ format has been associated with game shows like “Who wants to be a millionaire?” where questions are based on ‘the lowest kind of learning’ of disconnected facts but goes on to identify learning designs that can transcend this apparent disadvantage.

Various claims have been made about online methods of developing skills of critical reflection and critical thinking through discussion boards and blogs. In dental education Wetmore et al (2010) found that whilst blogs did not appear to enhance grades they improved levels of reflection and Hanson (2011) indicated that blogging encouraged group engagement. Goldman et al (2008) evaluated public health students’ use of assessed weekly seminar blogs, highlighting opportunities for increased interaction, participation and learning. Students positively evaluated “exchanges with other students, hearing different perspectives, flexibility of time of participation, having other students see and comment on their postings and opinions, helping them stay on top of class work” (Goldman et al
The ability for technology to provide interactivity and asynchronicity complements the opportunity it also brings for peer-to-peer collaborative learning. Fischer et al (2011) compared traditional written assignments to a group blog in medical education, with levels of reflection being comparable in both groups, going some way to demonstrate that deeper learning can be facilitated and evidenced through online shared spaces as well as the more traditional private assignment writing.

The objective of this study is to examine the design, implementation and impacts of innovative assessment strategies, with CAA, ARS and blogging, using analyses of student assessments and student and staff evaluations to identify the effects on student learning and the implications for educators. This will lead to a discussion of the yin and yang of innovative assessment, highlighting contrasts and connections between the controlled, systematic measurement and objectivity associated with multiple choice assessments, compared with the formative, iterative and subjective nature of reflective blogging and differences between individual and group assessment.

2. Context and method: collaborative lifeworld-led learning strategies

The introduction of innovative assessment strategies form an integral part of a collaborative lifeworld-led transprofessional curriculum delivered to cohorts of 600 students in health and social work where technology has been harnessed to provide a multi-layered blended learning experience, connecting learners to humanising perspectives for guiding their practice. Assessment technologies included use of group blogs facilitated through Blackboard™, CAA using Questionmark Perception™ and ARS using TurningPoint™. The innovative assessment strategies are situated within an undergraduate Year 2 unit, Exploring Evidence to Guide Practice (EE2GP). The humanising philosophy underpinning the unit encouraged students to integrate understandings about different kinds of knowledge for practice; conventional evidence, understandings about the person’s experience, and the student’s personal insights that come from imagining ‘what it is like’ for the person experiencing human services (Galvin & Todres 2011). This was facilitated through a series of 17 web-based case studies providing topic-related resources for learners to consider experiences of specific illnesses and conditions, such as dementia, social isolation, and substance misuse through narratives and poems, topic-specific qualitative and quantitative research, and policy and practice issues (Pulman et al 2012). These resources were supplemented with face-to-face (f2f) lectures and group work.

The intended learning outcomes were constructively aligned with the case study resources and teaching and learning strategies (Biggs & Tang 2007). The students were directed through weekly student managed guided learning (SMGL) activities over a four week period (Figure 1) using a detailed guide with tasks and questions to structure and scaffold their learning involving reading, listening, and viewing in preparation for critical reflection and blogging.

![Figure 1: Weekly student experience](image-url)
Firstly students were asked to explore what a health or social care condition or situation might be like for people experiencing it by reading and viewing stories, poems, and videos as evidence drawn from the arts and humanities. Secondly, students were asked to examine published research embedded within practice issues relevant to their field through reading and comparing a number of case study, topic-specific, research papers and listening to research active staff, talk about research through short podcasts. Thirdly, students were asked to consider how these different kinds of evidence could usefully guide their practice by comparing, reflecting and demonstrating their knowledge, facilitated by group work, which took place face-to-face and through the group blogs.

Students, working in groups of six to nine, were allocated case studies relevant to their field of study. Opportunities for considering transprofessional issues of what is required in humanly sensitive care, and associated tensions, risks and dilemmas were facilitated through inter-group discussions. Students, working in their groups, compiled and submitted individual blogs as formative coursework based on the weekly SMGL activities and received online formative feedback from educators. The purpose of the weekly activities was to enable students to build their knowledge progressively and collaboratively towards the summatively assessed group coursework blog (50%) and online MCQ examination (50%) at the end of the unit.

Building on projects supported by the Higher Education Academy (HEA) (Hutchings et al. 2011) and JISC/SEDA (Hutchings et al 2010b), this study examines the contributions of different forms of assessment to student learning through analysis of student assessment, feedback and evaluation. Data was collected over a two year period (2011 and 2012) using ARS voting pads to collect regular in-class feedback on the experience of undertaking the unit and an online end of unit evaluation, tailored to the specific features of the unit. The unit was delivered in two blocks each year with approximately 300 students per block. Data will be identified by cohort year 2011 and 2012 and block numbers 1 and 2, ie. 2011.1. The end of unit evaluation included, firstly, a set of item statements using a 5-point Likert scale from Strongly Agree to Strongly Disagree; secondly, a series of open response questions asking what "enhanced learning", what "challenged learning", what they "enjoyed most", what they "enjoyed least" and "recommendations for change". This was deployed immediately following the online examination and engendered very high response rates; 2011.1 98% (n=301), 2011.2 94% (n=243), 2012.1 86% (n=188) and 2012.2 94% (n=283). Staff evaluation was conducted by means of a staff focus group in 2011 and a questionnaire in 2012. Ethical processes and procedures were followed.

3. Findings: Adoption of innovative assessment strategies

The design, implementation and impacts of the innovative assessment strategies adopted will be examined in preparation for discussion of the challenges and implications for student learning, pedagogy, and organisation. The technologies deployed included CAA, ARS, and group blogs, affording combinations of assessment for learning and assessment of learning. We will examine the contributions of each of these in turn and relate these to our findings.

3.1 Computer assisted assessment (CAA)

The online MCQ examination consisted of 30 questions, 25 generic research questions and 5 case study specific questions, randomised by question and by options. Questions were presented one at a time and students were able to navigate between questions by using buttons to review their responses before submission. Students were given one hour to complete the exam; there was an on-screen timer and the exams were invigilated. Students were provided with copies of the journal articles they had read for their case studies.

The technology interface can impact on student acceptance and student performance in CAA. Ricketts and Wilks (2002) found the student assessment interface has a major impact on how acceptable CAA is to students with question-by-question delivery improving student performance compared with paper tests marked with OMR, where they identified a small improvement, and online scrolling, where they identified a large difference. Did the nature of the CAA assessment strategy impact on the student experience in this study? Student experiences of the online MCQ examination are identified in Table 1, which lists assessment-related item responses collated from the End of unit evaluation Likert-scale statements. The majority of students were positive about logging on to the MCQ examination (item1) and the user interface with questions identified as easy to read and answer on screen (item 2).
Table 1: End of unit evaluation statements

<table>
<thead>
<tr>
<th>Item</th>
<th>End of unit evaluation statements</th>
<th>2011.1</th>
<th>2011.2</th>
<th>2012.1</th>
<th>2012.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Logging on to the multiple choice exam was easy and straightforward</td>
<td>Disagree/strongly disagree</td>
<td>61%</td>
<td>4%</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agree/strongly agree</td>
<td>11%</td>
<td>86%</td>
<td>61%</td>
</tr>
<tr>
<td>2.</td>
<td>The questions were easy to read and answer on the computer screen</td>
<td>Disagree/strongly disagree</td>
<td>48%</td>
<td>12%</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agree/strongly agree</td>
<td>10%</td>
<td>70%</td>
<td>66%</td>
</tr>
<tr>
<td>3.</td>
<td>Using computerised tests would be appropriate for other summative assessment in my course</td>
<td>Disagree/strongly disagree</td>
<td>27%</td>
<td>9%</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agree/strongly agree</td>
<td>35%</td>
<td>58%</td>
<td>58%</td>
</tr>
<tr>
<td>4.</td>
<td>I would prefer the assessment for this unit to be a 3,000 word essay on applying research evidence to practice</td>
<td>Disagree/ Strongly disagree</td>
<td>81%*</td>
<td>71%*</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agree/strongly agree</td>
<td>19%*</td>
<td>18%*</td>
<td>9%</td>
</tr>
<tr>
<td>5.</td>
<td>I preferred submitting the group coursework assessment on the computer rather than on paper</td>
<td>Disagree/strongly disagree</td>
<td>16%</td>
<td>9%</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agree/strongly agree</td>
<td>56%</td>
<td>62%</td>
<td>59%</td>
</tr>
<tr>
<td>6.</td>
<td>The group blog has been helpful for learning collaboratively with my group</td>
<td>Disagree/strongly disagree</td>
<td>**</td>
<td>**</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agree/strongly agree</td>
<td>**</td>
<td>**</td>
<td>48%</td>
</tr>
</tbody>
</table>

* Statement 4 was included in the ARS for 2011.1 but not the end of unit evaluation.

** Statement 6 was not included in the end of unit evaluation for 2011.

But our findings also reveal that technical issues impacted adversely on student perceptions of CAA. A comparison of results for 2011.1 (items 1-3) compared with the other cohorts highlights the effects of technical issues experienced with logging on to the exam using a lockdown browser which delayed exam start times and added to exam anxiety for the 2011.1 block. Deutsch et al (2012) found that a positive CAA experience effects positive attitudinal changes towards the role of CBA, perceived ease of use and perceived objectivity. Our findings agree in showing that those students (2011.1) who experienced difficulties with the technology when accessing the online examination were less likely to agree that using computerised tests would be appropriate for other summative assessment (item 3).

Building student familiarity with innovative types of assessment is an important factor with students concerned:

_Having to revise for a test, which is something I am not used to as I am used to preparing presentations, essays etc._ (2012.1)

Students were given opportunities weekly to practice MCQ questions and receive feedback in lectures using the ARS. This enabled learning through shared feedback contributing to assessment for learning. Additionally a formative online mock assessment, consisting of 20 questions, was provided for students to familiarise themselves with the CAA interface in preparation for the summative online examination. But the challenges of the online MCQ examination were relative with the majority of
students in all four cohorts identifying a preference for the innovative assessment strategies adopted in this unit over the more traditional and familiar 3,000 word essay (item 4).

### 3.2 Audience Response System (ARS)

The ARS was used in the lecture theatre, with groups of approximately 300 students with the dual purpose of gathering opinion and gauging knowledge; through ongoing responses on the experience of the unit and formative self-assessment in exam preparation. Whilst it enabled students to experience typical MC exam questions and to gain rapid feedback and clarification it did not provide the facility for multiple responses as would be included in the exam. The overall trend, identified in Caldwell’s (2007) literature review, is that, on the whole, students and educators like the ARS, but effectiveness is dependent on the pedagogic strategies informing their use (Cain & Robinson 2008; Draper 2009). In our evaluation student responses ranged from those who valued the learning opportunities provided to those who preferred less use of the ARS. The 2012.2 cohort were the most positive in their responses, recommending more usage of the ARS. The ability of the ARS to provide interactivity in large groups, to be enjoyable and provide feedback in order to check learning in preparation for the exam themes were noted by students, with the ARS providing a bridge between the process of learning, and the outcomes of that learning. The number of questions each week using the ARS was reduced in the 2012 cohort, taking into account response times needed to register the votes for such a large group.

_The voting pads in lectures made my learning more interactive and engaged my attention so I thought this was useful._ (2011.1)

_The voting pads were fun, and really made me think about what I did and didn't know._ (2012.1)

### 3.3 Group blogs

The group blog facility was part of the standard provision in the University’s VLE, enabling the large cohort to be allocated to subgroups of 6–9, with a requirement for weekly individual formative blog contributions with the facility for intergroup comment and discussion, culminating in a final summative group blog. It was complemented by weekly face-to-face group work, with an opportunity to share learning with another group. The feedback from students on the helpfulness of the blogs for learning collaboratively indicated that whilst positive comments outweighed negative comments (Table 1, item 6) there was no strong trend in either direction. Themes in the responses included the opportunity to gather other opinions and learn from and with peers, acknowledging the opportunity for collaborative learning, whilst some expressed preferences for an individualised approach with expectations of individual feedback and a preference for an individual rather than group mark.

_Learning more about how others interpret the same material and in some ways it improved my understanding if there was something I was unsure of._ (2012.2)

_The group blog and posting individual blogs every week helped me to reflect on what I had learnt. The smgl were useful in informing what you had to complete each week._ (2011.1)

Some spoke of tensions within the subgroups in managing both the process and the product of the blog, with others recognising that learning to work effectively in groups represents the reality of health or social care practice. Some appeared to embrace the use of blogs whilst others commented on the unfamiliarity with blogs as an educational tool.

Staff feedback indicated that the group blog was:

…better for assessing ‘integration’ and ‘reflective ability’, academic and professional skills that are important for the course’s central topic: ‘exploring evidence for practice’.

### 4. Discussion and Conclusions

Student statements in the end of unit evaluations reveal significant ying/yang connections and interrelationships at work in the student experiences of innovative technology-enhanced assessment. The dynamic between breadth and depth of knowledge required and relationships between teaching and learning were clearly demonstrated. Students were challenged both by the depth of knowledge
suggested it was “complicated to understand” (2011.1) and “extremely difficult” (2012.1) and by its breadth:

The exam has challenged me; it was a lot of information to take in in 5 weeks as it is such a broad subject (2011.1)

There was so much information I needed to know, I felt very overwhelmed! (2011.1)

Gibbs (2006, p.18) argues where CAA is included as a component of assessment, “students tend to adopt a surface approach to their studies (attempting to reproduce) rather than a deep approach (trying to make sense)”, pointing to potential adverse impacts of CAA assessment strategies on student learning. However it is significant that student comments relating to this assessment strategy were associated as much with it being an exam as being facilitated through technology, reflected in a student response to “what I enjoyed least”:

The exam, just because it was an exam! (2011.2)

Notwithstanding exam anxiety as a factor, the MCQ assessment, consisting as it does of many questions on multiple topics, challenges the relationship between breadth and depth and highlights the importance of designing questions for deep learning and balancing breadth (numbers of questions) with depth (levels of difficulty). The yin yang features and benefits of CAA were identified by a member of staff:

It was efficient in that it tested both breadth and depth, covering a range of learning that was wider than I believe an essay would do. I also believe that it was efficient because it reduces the assessment burden on students (compared to a longer assignment), thus freeing their time for learning.

Further, the assessed discursive group blogs provided a dynamic counterbalance and complementarity to the controlled, time-limited and ‘objective’ exam. Informed by the unit philosophy of active, collaborative and reflective learning, in line with the shift from teaching to learning, it helps to address the concern that MCQ examinations do not facilitate deep learning. The ARS acted as a focal point for monitoring research knowledge and understanding and facilitating interactivity, with potential to be used more to facilitate collaborative and constructivist learning through students working together to debate and choose answers.

Our findings also highlight the dynamics of relationships between teaching and learning and the impact of technology, where students recommended reverting to familiar face-to-face lecture and seminar strategies to improve their exam preparedness:

I would have preferred to have structured seminars to prepare us better for the examination. (2012.2)

Some students appeared to have unrealistic expectations of the technology or concerns about its reliability and some emphasised technical problems, potential and real, rather than the capacity to learn or be assessed differently. Technology was afforded considerable potential yet subjected to scrutiny and criticism with user acceptance fluctuating according to perceived and actual experiences. Readiness for technology-enhanced assessment strategies appeared to be linked to the predominant learning culture with a lag between what was technically feasible but risky and what was accepted. Benefits can be longer term, as Wyllie (2011) discovered that students who engaged in online learning and assessment became more independent learners, taking more responsibility for their own and their peers’ learning.

The multi-layered blended learning strategy underpinning the unit sought to balance efficiencies with effectiveness, face-to-face on-campus and online off-campus teaching and learning, group and individual learning, discursive and MCQ assessment, employing a range of technology enhanced modes of assessment. Student feedback drew attention to passive and active approaches to learning, particularly in relation to group participation, face-to-face and online. In developing such a complex approach it is important for HEIs “not to be dazzled or seduced by what the technology can do but to adapt and apply” it to what we want students to be able to do (Quinney 2005, p. 449) in a more radical shift from just replacing current teaching and learning strategies to transforming the process of
learning, in line with the constructivist pedagogy and learning theory at the root of the EE2GP unit. Hutchings et al (2010a, p.201), in an earlier study, drew attention to experiences of educators who may be uncertain or unconvinced of the efficacy of disruptive technologies in teaching and learning concluding that "the challenge is to achieve ‘optimum disruption’, where transformation is seen as achievable and realistic, rather than being experienced as too uncomfortable". These challenges and associated risks must be addressed through strategic engagement and co-partnering, at institution, school, programme and unit levels, with management, staff and students, in order to change cultures and practices.

Acknowledgements

Development and evaluation of the EE2GP unit was supported by the HEA Discipline-focused Learning Technology Enhancement Academy (Hutchings et al 2011) and project funding from JISC/SEDA Embedding Work-with-IT (Hutchings et al. 2010b).

References


