11
Digitally supported assessment

Leo Havemann, Simon Katan,
Edward Anstead, Marco Gillies,
Joanna Stroud and Sarah Sherman

This chapter focuses on digital assessment and feedback practices in distance education. Providing evidence of learning through assessment is at the heart of students' experience of higher education (HE), whatever their mode of study. Open and distance education-focused institutions have justifiably been proud of their technical innovation, tending to move rapidly to harness available technologies (from post to broadcast media and, most recently, online media) in their mission to enable education for remote, distributed groups of learners. In recent years, distance education courses have, in the main, moved from paper and digital media delivered physically to wholly online delivery, except where the circumstances of target learners preclude reliance on a reliable and fast internet connection. In terms of content, discussion and collaboration, where distance education has forged ahead, campus-based, blended programmes have generally followed. However, in terms of assessment and feedback, distance education has remained somewhat conservative. While most assessment in distance education has taken place online along with content and communication, there has been a tendency to replicate fairly traditional assessment formats using digital tools.

The contexts and cases discussed in this chapter have been drawn from member institutions across the University of London (UoL) Federation. This chapter considers the future of assessment in distance education. It reviews some previous work conducted within a subset of the federation based in Bloomsbury, which aimed to provide a snapshot of current assessment practice within credit-bearing distance education

programmes (Weitz and Seddon, 2017). It then attempts to look beyond the current mainstream of practices by considering additional cases of emerging practice within the massive open online course (MOOC) space.

During their rise to prominence in the early 2010s, MOOCs were widely heralded by their proponents as a completely new and innovative educational format and, conversely, critiqued by detractors as simply a rebranding exercise for online distance education. Taking a view from somewhere between these points, this chapter goes on to discuss that while MOOCs certainly represent a form of distance education, they also possess features that distinguish them from 'traditional' distance education programmes.

A snapshot of mainstream distance education practice

The pre-COVID-19 practice in digitally supported assessment is explored through findings of a review of assessment activity in several programmes offered via the UoL network. The distance education programmes forming the UoL portfolio are run by a subset of UoL member institutions. The findings of this current practice review, which was conducted by the Bloomsbury Learning Exchange (BLE) as part of a larger focus on assessment throughout its member institutions (Havemann and Sherman, 2017), are categorised according to three broad approaches: functional, enhanced functional and innovative, defined as follows:

- A functional approach, in which compliance processes are met. Resources and results are generally available through technology. Summative assessment is more prevalent than formative assessment.
- An enhanced functional approach, where there is an increase in the use of formative assessment with tutor interaction and individual feedback enabled through technology.
- An innovative approach, with a strong collaborative pedagogical rationale and increased variety of learning activities.

Taken together, the examples discussed provide a 'snapshot' of typical practices, rather than representing an exhaustive documentation or evaluation. Furthermore, the examples highlighted do not discuss the tools employed in significant detail, instead focusing on assessment contexts, tasks and outcomes.

Functional approach

All departments, programmes and modules explored and referenced as part of the BLE review that can be grouped within the functional approach fulfil, at a minimum, the standards of academic quality expected by their respective institutions, inclusive of learning, teaching and assessment practices. Nevertheless, these courses may be termed 'legacy', having originally been delivered via correspondence with posted, hard-copy material and now updated with modern content delivery and communication tools rather than explicitly designed for an online context. The courses are now mostly paperless, with content and some aspects of assessment activity (such as provision of marks and feedback) available online or supported by technology. Summative assessment dominates the courses featured, although there are occasions when formative assessments are also present and delivered.

Several departments, including those at the School of Oriental and African Studies (SOAS), the London School of Hygiene & Tropical Medicine (LSHTM) and the Royal Veterinary College (RVC), indicated that their programmes operated summative assessment models primarily weighted towards examinations taken in traditional, invigilated face-to-face contexts, such as in local examination centres. These examinations take a variety of forms, including seen and unseen essay question papers and multiple-choice question (MCQ) tests. Weightings for these activities were typically up to 80 per cent of the grades and made little use of technology given the context within which they were completed. Nevertheless, programmes at LSHTM suggested that exam scripts completed in local centres were scanned and marked electronically.

Coursework activities within this approach were completed through varied assessment tasks appropriate to the discipline of study, such as reports, essays, audio-visual presentations, case studies, journals, logs and scientific or mathematical exercises. These were frequently weighted at approximately 20 per cent of the module mark and were more likely to be supported by technology, either through a pedagogic delivery mechanism inherent within the virtual learning environment (VLE) or by means of an upload through a VLE-based online submission facility (Birkbeck, LSHTM, RVC, SOAS and UCL). Marks and feedback for coursework submitted online were in most cases also made available via the VLE assignment facility or proprietary tools such as Turnitin, with the UCL Institute of Education noting that feedback was provided for drafts in advance of the final summative submission.

Enhanced functional approach

Courses following the enhanced functional approach were common among reviewed institutions, building on the functional approach to demonstrate greater recognition and use of formative assessment and feedback. These activities were typically enabled or enhanced by technology and would often feature an increased focus on tutor or peer interaction with individual students.

As with summative assessments, formative tasks take a variety of forms across member institutions, including essays, mock examinations, quizzes and portfolios, each with a strong focus on feedback. Several of these offer opportunities for students to engage with the actions and processes of self-regulated learning (Zimmerman, 1990) that have a defined relationship to academic achievement in both face-to-face and learning contexts. For instance, both LSHTM and the RVC offer quizzes that promote self-efficacy and assessment in relation to the individual's understanding of specific topics, while UoL has further developed its own custom self-assessment VLE plugin for essay questions, giving students the ability to mark and evaluate their responses against a series of model answers.

Programmes from LSHTM, RVC and UCL Institute of Education each identify discussion fora as environments in which peer-to-peer and peer-to-tutor interactions are used to generate formative feedback.

Synchronous tools such as Skype and Blackboard Collaborate are further identified by Birkbeck, LSHTM, RVC, SOAS and UCL as mechanisms for individual and group-based interactive tasks. These routinely take the form of tutorials, being an opportunity to review material but, as described by SOAS, additionally promote student voice and engagement in personal goal setting.

The UCL Institute of Education additionally reported tutors' use of audio to deliver feedback, offering greater flexibility in both the generation of the feedback and the potential to make feedback feel more personal, thereby building more meaningful connections with students studying at a distance.

Innovative approach

The key feature of this innovative approach is the collaborative, pedagogic rationale taken by departments, programmes and modules, which facilitate much more student interaction.

In such examples there typically exist an increased diversity of learning activities, which are often drawn from cultural or socially driven

learning theories such as social constructivism. While this rationale broadly informs course learning design, it is also evident within the assessment methods employed, with formative and peer-supported tasks being prevalent.

Multiple programmes considered as part of the BLE review featured assessment activity that could be grouped under the innovative approach and, notably, through formatively assessed tasks. Alongside traditional forms of assessment such as presentations and essays, UCL's MSc in paediatric dentistry further requires that students complete a logbook using the iPad minis with which they are supplied. The logbook is populated with treatment approaches and requires some peer interaction in relation to the rationale for the selections made.

Students enrolled on Birkbeck's MSc in geochemistry are provided with high-quality digital learning resources to analyse through Xerte tutorials. The initial analyses are used as the basis for collaborative discussions between peers and tutors, which subsequently inform the assessed portions of the course.

The SOAS MSc in financial sector management operates predominantly within the functional and enhanced functional approaches. However, one module uses a strategic simulation model in which students are placed in teams of five to conduct research and build a strategic case study. Within this case study there are specific assessed activities, including development of a business plan, a risk analysis and scenario planning, while long-term collaboration is established as a significant factor in the awarding of marks. Formative feedback is given at key milestones within the simulation and students can benchmark their progress against both their own plans and that of other teams.

The programme featured as part of the review that most fully embraces collaboration between students and tutors in its activities and assessments is the SOAS MA in global diplomacy. The course demonstrates constructive alignment in its learning and assessment design and course activities are explicitly mapped to assessments and learning outcomes (Biggs, 2003), with each module featuring five written online assessments or 'e-tivities' (based on the work of Salmon, 2002) comprising 30 per cent of the module mark and a longer-form essay for the remaining 70 per cent. The concept of e-tivities is drawn from the framework for participatory online learning in which learners are supported through five stages of progressive participation in an online learning community: access and motivation, online socialisation, information exchange, knowledge construction and development (Rofe, 2011; Salmon, 2002).

Emerging good practices in assessment at scale

Blurred boundaries: traditional and non-credit-bearing online assessments

MOOCs are generally delivered to large numbers of learners across diverse geographical and cultural contexts and are open in the sense of being at least initially free of fees or formal entry requirements. The bulk of such courses are made available as a result of HE or specialist providers' formal relationship with a privately operated platform provider, such as Coursera, FutureLearn or edX. Course content is typically designed using pedagogic patterns familiar within the distance education context (Bali, 2014; Daradoumis et al., 2013; Glance et al., 2013), such as text, video, audio and downloadable files of supplementary or longer-form material. However, the fundamental differences in delivery have seen implications for the administration and learning design of such courses, leading to a departure from the types of assessment activity most frequently employed as part of more traditional distance education. Tutor-led review, marking and feedback is a labour-intensive process and, on that basis, cannot be replicated at scale, meaning that MOOCs and their delivery platforms have instead adopted pedagogic strategies more appropriate for large cohorts. These strategies focus on peer-to-peer interaction, opportunities for self-regulated and self-evaluated learning activities and automated assessments.

The tools or pedagogic approaches in place to support assessment activity as part of MOOCs are ultimately dependent on the platform being used. However, those that are offered as standard across platforms include:

- In-content pause points, such as those delivered within video material. These are typically lightweight, ungraded MCQs that assess understanding and provide immediate contextual feedback.
- Discussion prompts, whereby learners are encouraged to engage with one another in social spaces by answering questions posited by tutors within the body of content. Although there may be occasional tutor interaction, learners are often asked to read and respond to the comments of others.
- Quizzes containing a range of automated question types, such as multiple choice, multiple answer and numeric response. These can be used in both a formative and summative context with pass/fail functions, but automatic grading and delivery of pre-generated

feedback is paramount to both accommodate large numbers of learners and deliver uniform results across the cohort.

- Peer-supported assignments, in which open-ended assessment tasks can be delivered and learners grade one another's work using a rubric or criteria provided within the course.
- Programming assignments that require learners to submit computer code, then a platform technology reviews and grades the script.

A light touch, automated tutor facilitation and lack of summative assessment has led these tools to exist solely in a non-credit-bearing context. This means that they exist outside of the assessment regulations and accreditation frameworks that govern delivery of more traditional distance education modules and programmes.

However, while the at-scale nature of MOOCs has seen them occupy a distinct space within the provision of education internationally, there are lessons to be learned or pedagogies that might be borrowed from the delivery format. This can happen through the application of at-scale pedagogies and formative assessment tasks to award-bearing courses and is happening in the sense that full, credit-bearing programmes are now being delivered via MOOC platforms. Practical examples of both are explored in the subsequent illustrations of innovative practice.

Formative peer review

In one assignment, learners are asked to select any single issue that negatively affects the lives of the residents of an African city or neighbourhood of their choice. The task assignment is an essay between 500 and 1,000 words that cannot be delivered at scale with the tutor support available. On that basis it is delivered as a peer review activity, with significant scaffolding to support its completion.

Learners are encouraged to draw upon and combine their prior personal, professional and educational experiences with their learning around issues and concepts as part of the course in their submitted response. A number of key and open-ended questions serve to implicitly scaffold the learner's response, asking them what the issue is, who is most affected by it and why, what issues underpin it, what factors and processes have contributed to it and how and which actors could contribute to its solution. Some basic examples of issues are provided but learners are encouraged to select their own.

The second stage of the assessment is to review the submission of another learner. It is only possible to progress and complete this stage

once an individual has submitted their own work. Learners do not provide a mark as part of their review but are asked to actively engage with the content of another learner's assignment and reflect upon it in a positive and constructive way. Assignments are reviewed using the same generic criteria applied to more traditional assignments, such as quality of critical reflection and originality, use of evidence, use of concepts and materials to illustrate the issue and build an argument, coherence and clarity and relevance and focus. Additionally, they are asked to write a short reflection guided by open questions relating to the subject matter, such as their thoughts on the assignment's most interesting points, what it made them think about, the connections they had made between the subject location of the assignment and a different place or situation and whether anything was missing that could have helped to better elucidate the situation or problem.

Autograding

Autograding is the automatic grading of student work by a computer program rather than a human tutor. Students submit work in digital form to a VLE. The work is then marked algorithmically by a computer program, which would normally be on a remote server, such as a VLE server, so that the student cannot tamper with the grading software. Autograding, together with peer grading, is a key enabler of MOOCs, because it allows students to receive grades and feedback without tutor input. For example, it is a key element of the new MOOC-based BSc in computer science by Goldsmiths and UoL (though it should be noted that as it is a full degree, all modules also contain tutor-marked assessments).

The simplest and most common form of autograding is the MCQ test. Since MCQs have a well-defined correct answer, and they can easily be implemented digitally as check boxes, automatically grading them is straightforward. They are a very common form of assessment that underpins MOOCs and other online courses. In the BSc in computer science, for example, they are used extensively both as formative practice and summative assessment. Most VLE platforms generalise these types of quizzes to other types of questions that also have a clear correct answer and are easily assessed by a computer; for example, numerical answer questions or simple text matching. Quizzes are therefore an easily implemented and efficient form of autograding. However, they are limited in the depth of the learning outcomes they are able to assess. In most cases they are used for simple factual recall, though more sophisticated question design can improve them, for example, by careful choice of

alternative answers for MCQ. An example of the use of a quiz for more sophisticated learning is in the module 'How Computers Work' in the BSc in computer science. In the module, quizzes are used as a prompt for independent research as students are asked to use the internet to research a particular topic that has not been covered in the course (for example, a particular computer virus called Mirai) and then must complete a quiz on the topic.

Other forms of autograding require more sophisticated grading algorithms. Apart from quizzes, the most common use of autograding is for computer programming. Students upload the programs they have written and an algorithm checks this program. This is normally done by checking the output the program produces with a number of different inputs (though there are other approaches, such as analysis of the source code for correctness against the programming language's syntax). This checking process normally mirrors the testing process used in professional programming practice. It uses a type of checking software method known as unit testing, which was originally developed for professional software testing. The assessment is therefore very close to industry-standard ways of working.

Another approach to autograding used in the BSc in computer science is simulation. This is an interactive activity that models an element to be learned; for example, the workings of a computer processor or a particular algorithm. It can be a very engaging, hands-on way of learning and can also provide feedback. Simulation can be designed as an open-ended formative activity that allows students to explore the functioning of a computational system. It is also possible to design simulations that function more like a computer game, in which there are certain correct 'winning' strategies and students get more direct feedback through either a score or a 'win' condition (which can be checked automatically by the simulation software). In this latter case the simulations can be used as more summative assessment, with grades being calculated from students' scores on the simulation.

Beyond computer programming, autograding is relatively rare. It requires defining algorithms to assess a piece of work, which might be straightforward in the case of a clearly defined programming exercise as above, but in other cases can be very challenging or beyond the current technological 'state of the art'. Even in areas such as mathematics or computer programming there are many aspects that cannot easily be autograded (the 'working' and problem solving in mathematics or coding style and open-ended software design in programming). In other subject areas, such as humanities, very little can be automated. This is still a very

active research area and there are even claims of machine learning software that can grade humanities essays (though these should be judged with caution), so progress may be made in other disciplines (Arikat, 2012).

The most obvious benefit of autograding is that human tutors no longer have to mark assessments. This is a potentially significant re-education in labour for individual teachers or a considerable cost saving for institutions, though these are balanced by the much larger upfront costs of developing the autograded exercise. Developing grading software or simulations can take considerable and costly effort.

However, as with many forms of automation, autograding is not simply about reducing the cost of existing approaches to assessment, but radically changing assessment, particularly through scale. The cost and effort of marking work is a considerable bottleneck in traditional education and removing it can allow assessment to increase in scale in a number of ways. The most obvious is increasing the number of students. MOOCs can support thousands of learners at low or zero cost because all assessments are automated. Scaling up can also qualitatively change the nature of assessment. As well as efficiently assessing more students, each individual student has access to many more assessments for practice and formal evaluation. A typical course in a BSc in computer science will have several quizzes and/or programming exercises per week, allowing students to get frequent and instant feedback on their work. Since marking the work is not costly, students can also attempt an assessment multiple times and get feedback on them all, thus gradually improving their work. This enables radically new approaches to learning that would not have been possible or would have been very costly with tutor marking. For example, variants of Bloom's *mastery learning* (Bloom, 1984) discuss an approach where students can make multiple attempts at tasks, watch lectures multiple times and go through materials in a much more self-paced manner than would be possible in a traditional campus environment.

Gamification

Sleuth (Katan and Anstead, 2020) is a series of gamified (Deterding et al., 2011) code puzzles themed around a film-noir detective story. The project was developed for the introductory programming module that runs on campus at Goldsmiths College and through UoL and Coursera. Goldsmiths follows a 'learning by doing' approach to teaching programming. In the module, students build fluency in rudimentary techniques and patterns through the repeated practice of programming exercises. This raises

challenges around content generation, scalability and student motivation to which the design of Sleuth responds.

Students access Sleuth via a personalised web app. From here they can check their current grade and feedback, download puzzles and upload them to get them graded. The web app is themed as a detective agency 'Sleuth & Co', with students playing the character of a fledgling detective. They are guided by 'the Chief' who gives them feedback on individual puzzle attempts as well as their general progress in the game.

Primarily, Sleuth aims to facilitate a greater amount of practice of rudimentary programming tasks in a way that is scalable for module tutors. This is achieved through the employment of two techniques: procedural content generation and autograding. To facilitate procedural content, each student is allocated a user ID within the game. The ID is used to generate unique puzzles for students to complete, containing distinct numeric values, naming conventions and coding styles. When a student downloads a puzzle to solve, the unique version they receive is sufficiently different from other students' versions in the class so that collusion and other forms of plagiarism are eliminated from the assessment. The autograder recognises the student's user ID and marks the correct variation of the puzzle. Uploading another student's work would be rejected by the system as the variations in the code would not match the ones given to the colluding student.

The downloaded puzzle comprises a sketch template (a set of files defining a program in the programming environment employed) containing a unique puzzle task written as text comments and starter code for the student to complete. Included in these comments is a reference that identifies the puzzle variant to the autograder. Students attempt the puzzle and upload it for grading. They receive immediate feedback from 'the Chief', which includes any compile or runtime errors and tells them what parts of the task they have achieved and what parts they still need to work on. Students get five attempts to complete a puzzle, after which 'the Chief' suspends them from that particular task for one hour. On returning, students must start afresh by downloading a new variant. Through this design, students are provided with as much practice as they need to master topics without placing extra burden on the teacher.

The design also aims to provide differentiated outcomes for students of varying experience levels, which is carried out through the arrangement of puzzles and the scoring system. The puzzles are arranged into 16 cases, each based on a particular topic from the syllabus and consisting of four stages. Students can attempt the cases in any order and need not complete them before starting another. However, the stages of each case progress

in order of difficulty and are unlocked in sequence as the student solves them. While more experienced students might complete the higher stages of most cases, less experienced students can still practise and achieve in all areas of the course by completing the lower stages of each case. Having built up their confidence, these same students might return to harder stages that they had previously abandoned. These types of behaviour are further supported through the scoring system. Students' grading comprises a 'rookie' score and 'pro' score. The 'rookie' score is made up of the average of the first nine cases that are made available to students from the start of the course. At the midpoint of the course, the students 'go pro'. Their 'rookie' score is frozen and the remaining seven cases are released. The 'pro' score is made up of the average of all the cases. This means that students are rewarded for completing their work in a timely manner but are also rewarded for continuing work on unfinished rookie cases after the midpoint deadline.

The aim to create an engaging environment in Sleuth that would motivate students to practise their code rudiments is in part achieved through the instant feedback and summative scoring design. However, the unit also used game-like theming to amuse students and arouse their curiosity. Each case tells a different story set around the criminal residents of Console City. As students solve more cases, they uncover further connections and evidence of a criminal conspiracy. Additionally, graphical content for the puzzles was produced by a professional illustrator.

At the time of completing this chapter (October 2022), Sleuth has so far run four times on campus and eight times online for approximately 5,000 degree-level learners. The results have been encouraging. In the initial on-campus run, students made a total of 42,534 code submissions – an average of 138 per student over a ten-week period. Despite perceiving the task's level as between fair and difficult, the class's achievement was high with a median grade of 90.67 per cent (quartile 1 mark: 75.79; quartile 3 mark: 96.49). These figures demonstrate that Sleuth has facilitated a very different environment for students to learn in. Such submission quantities are beyond the capacity of any team of human graders that could be resourced and the motivational, gamified aspects of the design have engendered altogether different levels of student engagement.

Another advantage of autograded assessment is that rich data can be collected about how students are responding to the assignments. Sleuth records the details of every student submission, including the submitted code, grade and feedback. In reviewing the data, several areas of improvement for subsequent iterations of Sleuth have been identified. Despite the conceived behaviour in the level design, less experienced

students tend to persist in their attempts at the more difficult case stages instead of moving on to other cases. With this in mind, there is an aim to increase the interventions of 'the Chief', who will increase suspension times and recommend alternative cases for students to try. By reviewing the average number of attempts per puzzle it is possible to identify those cases that students find particularly challenging. For these cases, the unit plans to experiment with different forms of feedback to improve student performance. Currently, a significant proportion of students achieve full marks for the assignment. While passable at this level, a lack of differentiation at the top end of grades may be considered undesirable in assessment in HE. Therefore, an area of improvement might be to raise the difficulty level of the final stages of cases. It would be interesting here to see what proportion of students still continue to achieve the highest grades, given the automated feedback and unlimited attempts.

Conclusion

Assessment in online and distance education has largely attempted to reproduce the forms of assessment used in traditional on campus settings in digital form. While this reproduction of existing forms is typical of initial experiments with new technology, assessment seems to be a particularly conservative area. This is due in part to concerns about the rigour and fairness of assessments, with new approaches being viewed with suspicion and traditional approaches such as paper exams viewed (perhaps overgenerously) as a rigorous 'gold standard'. This concern is held by many academics and universities and, in many countries, paper exams are enshrined in educational regulations.

While the examples gathered in this chapter pre-date the COVID-19 pandemic, we have seen similar tensions play out in the context of the 'pivot online', which saw campus-based teaching around the world replaced with 'emergency remote teaching' (Hodges et al., 2020). In many cases, traditional, in-person timed exams have been substituted with online, sometimes 'proctored' timed exams; however, the switch to online learning and assessment has also initiated a wider uptake of alternative forms of assessment. The case study in Chapter 12 explores many of the practical issues in how UoL moved to online assessment in 2020. It remains to be seen how many of the new practices that have been developed as a result of the pandemic will be retained and to what extent. However, there are encouraging signs that practices in both distance and campus-based assessment are evolving to take advantage of distinctly

digital opportunities now that these are becoming more widely available and better understood.

Recent years have seen considerable innovations in digitally supported assessments. These have been driven in large part by the MOOC model of online education, which is characterised by the ability to scale to very large numbers of students with very limited input from tutors, resulting in a focus on automated and peer assessment. These MOOC-style approaches are now starting to be used in credit-bearing and degree programmes, allowing for greater scale and efficiency. However, their use in degrees raises new challenges due to concerns of rigour and quality of assessment. While automated and peer assessments can be used to evaluate many important learning outcomes in many subject areas, they also have considerable limitations. Automated assessments can work well for technical subjects such as mathematics and computer science, but there are many deeper aspects of work that cannot be assessed in this way. Many disciplines, such as humanities, may have little scope for using automated assessments. Peer assessment is more flexible, but the quality of assessment is limited by students' own prior conceptions.

Students are also likely to question the validity of peer grading in high-stakes assessments. For these reasons, it is important to balance automated assessments with human-graded assessments. For example, UoL's BSc in computer science makes extensive use of both automated and peer assessment but has a policy of requiring that every module also includes human-graded assessments and that peer-graded assessments do not count towards final course grades.

However, many of the concerns listed here also relate to the summative function of assessment and there is a strong element to formative assessment that drives learning. This is where a typical MOOC style of assessment can be particularly valuable, since the scale it provides is not simply in terms of the number of students that can be assessed but the frequency of assessment for each student. Having several small assessments every week and giving feedback on them would be unfeasible for human tutors, but it becomes possible with MOOC-style assessment. Students are therefore able to test their learning and get fine-grained feedback on their learning more frequently. When used well, there is not simply a quantitative increase in feedback but a qualitative change in the nature of learning, as exemplified by the aforementioned Sleuth example. By offering instant feedback for many small exercises, Sleuth has been able to shift the style of learning towards one of intensive practice and sustained engagement, which are both vital when learning programming.

Peer assessment also has considerable formative benefits. It allows students to engage in tasks that are too complex for automated assessment without being restricted by the bottlenecks that arise in tutor grading. However, there are also important benefits not related to performing the task but to the fact that students are assessing the same task. Evaluating one's own work and that of others is a vital skill in academic and professional settings. On the one hand, an education system based purely on tutor grading is unlikely to develop these skills in students as they can feel reliant on others for feedback. The requirement to assess peers, on the other hand, ensures that students develop evaluation skills and engage deeply with the marking criteria to develop critical thinking and improve their own work.

The new approaches to digitally supported assessment that are currently emerging therefore have the potential to result in major changes in how students learn in distance settings, supporting more intensive engagement and deeper self-evaluation among other things. These changes will not result from simply viewing digital technology as a way of making traditional assessment more efficient or from a naive techno-optimism that sees any use of technology as automatically resulting in improved learning. They will result from the conscious use of technology to enable new pedagogies centred on improving students' learning.

References

Arikat, Y. M. (2012) 'Subtractive neuro-fuzzy modeling techniques applied to short essay auto-grading problem'. In *6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT) 2012*, Sousse, Tunisia. IEEE, 889–95. Accessed 18 July 2022. https://doi.org/10.1109/SETIT.2012.6482032.

Bali, M. (2014) 'MOOC pedagogy: Gleaning good practice from existing MOOCs'. *MERLOT Journal of Online Learning and Teaching*, 10 (1), 44–56.

Biggs, J. (2003) 'Aligning teaching and assessing to course objectives'. *Teaching and Learning in Higher Education: New Trends and Innovations*, 2 (April), 13–17.

Bloom, B. S. (1984) 'The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring'. *Educational Researcher*, 13 (6), 4–16. Accessed 18 July 2022. https://doi.org/10.3102/0013189X013006004.

Daradoumis, T., Bassi, R., Xhafa, F. and Caballé, S. (2013) 'A review on massive e-learning (MOOC) design, delivery and assessment'. In *Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*. 28–30 October 2013, Compiègne, France. IEEE, 208–13. Accessed 18 July 2022. https://doi.org/10.1109/3PGCIC.2013.37.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K. and Dixon, D. (2011) 'Gamification: Using game-design elements in non-gaming contexts'. *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. New York: ACM, 2425–8. Accessed 18 July 2022. https://doi.org/10.1145/1979742.1979575.

Glance, D. G., Forsey, M. and Riley, M. (2013) 'The pedagogical foundations of massive open online courses'. *First Monday*, 18 (5), 1–10.

Havemann, L. and Sherman, S. (2017) *Assessment, Feedback and Technology: Contexts and case studies in Bloomsbury*. London: Bloomsbury Learning Environment. Accessed 18 July 2022. https://doi.org/10.6084/m9.figshare.5315224.

Hodges, C., Moore, S., Lockee, B., Trust, T. and Bond, A. (2020) 'The difference between emergency remote teaching and online learning'. *Educause Review*, 27 March. Accessed 18 July 2022. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning.

Katan, S. and Anstead, E. (2020) 'Work in progress: Sleuth, a programming environment for testing gamification'. In *IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 1503–7. Accessed 18 July 2022. https://doi.org/10.1109/EDUCON45650.2020.9125098.

Rofe, J. Simon (2011) 'The IR model and e-moderating'. In G. Salmon (ed.), *E-moderating: The key to teaching and learning online* (3rd edn). London: Routledge Falmer.

Salmon, G. (2002) *E-tivities: The key to active online learning*. New York and London: Routledge Falmer.

Weitz, N. and Seddon, K. (2017) 'Assessment, technology and innovation in distance learning in the Bloomsbury Learning Environment institutions'. In L. Havemann and S. Sherman (eds), *Assessment, Feedback and Technology: Contexts and case studies in Bloomsbury*. London: Bloomsbury Learning Environment, 23–37. Accessed 18 July 2022. https://doi.org/10.6084/m9.figshare.5315224.

Zimmerman, B. (1990) 'Self-regulated learning and academic achievement: An overview'. *Educational Psychologist*, 25, 3–17. Accessed 18 July 2022. https://doi.org/10.1207/s15326985ep2501_2.