**TECHNOLOGY-ENABLED ASSESSMENT IN LARGE CLASSES: LEARNING FROM SUCCESS AND FAILURE**

**ABSTRACT**

Large class numbers make it challenging for educators to provide feedback and mid-term assessment to students, particularly in the situations of relatively limited teaching support. Developments in technology have raised hopes that technology-enabled approaches can solve feedback and assessment challenges arising in large classes. In this session, we seek to share our positive and negative experience of using these approaches to enhance the undergraduates’ learning experience. We will discuss specific tools that help: a) to automate the grading process; b) to use input from peers to assist a lecturer in grading; and c) to transform group-level grades into individual-level grades.

**Keywords: large group teaching, assessment, technology**

**INTRODUCTION**

The number of students in third level education has been growing for the last decade and is predicted to grow further. Many business schools globally have been facing the same trend. These trends translate into large classes for most core business modules. For instance, for few last academic years, the number of students registered for undergraduate business courses in our school varied around 700-800 for some first year modules and 400-500 for some second year modules.

As educators, we seek to deliver high quality of teaching and positive learning experience to our students. For this, having a portfolio of different assessment and feedback approaches is vital. However, large class numbers make it very challenging for instructors to provide feedback and mid-term assessment to students, particularly in the situations of relatively limited teaching resources and support. Contemporary developments in technology have raised hopes that technology-enabled approaches can solve feedback and assessment challenges arising in large classes.

In the proposed discussion session we seek to share our positive and negative experiences of using a variety of approaches to enhance the undergraduates’ learning experience through technology-enabled assessment and feedback tools. For the purposes of this session large classes are defined as classes with more than 100 students.

**THEORETICAL BACKGROUND AND REVIEW OF EXISTING PRACTICE**

Research on technology-enabled assessment is a sub-stream of the growing literature on technology-enabled teaching and learning. The focus of this review is on assessment and related feedback through technology. This literature is still at the early stages of development (Pesare, Roselli, Rossano, & Di Bitonto, 2015); however, we have been able to identify several key themes within it.

The first theme addresses the value of technology-enabled assessment from the conceptual point of view (e.g. Scardamalia, Bransford, Kozma, & Quellmalz, 2012; Shephard, 2009). This literature predominantly sees technology-enabled assessments in a positive light, and argues that technology broadens participation in higher education (Shephard, 2009) and creates opportunities for new types of assessment approaches (Webb, Gibson, & Forkosh-Baruch, 2013).

The second theme looks at the implementation of technology-enabled assessments and the difference between them and the traditional (paper-based) tests (e.g. Bayazit & Askar, 2012; Boeve, Meijer, Albers, Beetsma, & Bosler, 2015). A separate sub-stream within this literature addresses the assessment issues in the context of Massive Open Online Courses (MOOCs), with one example of these issues being the multiple-account cheating (Northcutt, Ho, & Chuang, 2016). While large classes in a traditional university are different from MOOCs, some of the insights on how students might exploit technology in online assessments are relevant for this setting. Overall, the studies within the second theme found that students’ performance does not differ significantly between technology-enabled and paper-based tests. However, students consistently found that their focus and concentration on the task was lower in technology-enabled assessments relative to traditional paper-based tests.

The third theme assesses the challenges on introducing technology-enabled assessments from the infrastructural and institutional point of view (e.g. Voce, 2015). This literature argues that the opportunities offered by technology may not be fully exploited due to the conflict between the needs of students and the needs of educators (Malau-Aduli, Assenheimer, Choi-Lundberg, & Zimitat, 2014), and due to poor understanding of the rules and responsibilities in the use of technology-enabled assessment (Voce, 2015). One key insight from this stream of research is that technology does not create value to students and educators unless it is grounded in sound pedagogical approaches and supported by appropriate resources at the organisational level.

In order to complement these insights with the understanding of the range of existing technologies, we engaged with colleagues in various universities around the globe. We attended a number of workshops on technological solutions for education, including the ones organised at Teaching and Learning Conference, and by Management Education and Development division at the Academy of Management Annual Meeting. These interactions exposed us to both various pedagogical approaches and software solutions available on the market.

On the basis of our overview and discussions, we structured technology-enabled assessment into three main types, according to how they address the challenge of large class assessment.

The first type seeks to assist lecturer by automating the grading process. It tests each student’s knowledge on individual basis and allocates an individual mark. The simplest form of this type of technology is the machine that grades standard paper-based MCQ answer sheets. More advanced technology includes automatically graded quizzes by using Moodle/Blackboard platform, and immediate response software such as Socrative. This type of technologies enable lecturer to give immediate performance feedback to (almost) unlimited number of students. However, due to natural limitation of full automation, these technologies to the moment can deal only with multiple choice questions that are commonly criticised for their inability to test some of the important learning outcomes, such as the students’ “ability to organize and present ideas” (Piontek, 2008: 3).

The second type uses input from peers to assist a lecturer in grading and providing feedback to students. Technology is used to make peer grading anonymous, to structure peer feedback and to manage the process of assigning reviewers and delivering their feedback. The idea is that involving peers in assessment can diminish the workload for instructor and thus enable to use in large classes assignments that could not be otherwise graded due to teaching resource limitations. Peer feedback and grades can be used both as they are and as a tentative grade, which instructors might use as a starting point to their own grading. Examples of technology that enables this include PeerMark module of Turnitin software and Moodle workshop module of Moodle.

The third type seeks to transform group-level grades (grades for group projects) into individual-level grades, using peer evaluation to adjust the grades for group work for individual contribution. Such approach aims to alleviate grading workload in large classes by limiting teacher grading to groups only and yet to account for individual performance. Numerous software products have this functionality, such as CATME (Hrivnak, 2013), SparkPlus, Moodle workshop, PEAR, and WebPA.

**GENERAL DISCUSSION OVERVIEW**

How do these different technologies work in the context of continuous assessment in large classes? We tested four specific tools from three categories outlined above and would like to share our experiences in this discussion session. We tested these tools in classes of different size and level and had some variations of how exactly the tools were applied. In this session, we will share what guided our choices, what challenges we faced along the way and what finally worked well and what did not. After presenting our experience, we will invite all participants of the discussion session to share their experiences and suggestions on how these tools can be used to enhance learning in large classes. More specifically, we will present the following technology-enabled assessment solutions:

Automated grading of individual assignments:

Online Quizzes via Moodle (run outside the classroom)

We used Moodle functionality to conduct automatically graded quizzes in three undergraduate modules that covered the cohort of 700 students in first year and the cohort of 450 students in second year.

In-class Quizzes via Socrative

Socrative immediate response system was used for running in-class quizzes in the cohort of 900 first year students and in the cohort of 550 second year students.

There were some variations in the way these methods were applied: we used either textbook-associated question bank to source MCQs or created our own questions; we used “closed book” approach and alternatively allowed students to consult with each other. Finally, we piloted a number of approaches that aim to improve the ability of MCQs to test for deep learning such as the use of vignettes to situate conceptual knowledge in real-life context (Malau-Aduli et al., 2014) or use of paired MCQs (Ventouras, Triantis, Tsiakas, & Stergiopoulos, 2010).

Technologies for peer feedback and grading:

PeerMark

We tested the PeerMark tool in two modules. One was a first year undergraduate module with 700 students, and the second one was a second year undergraduate module with 450 students.

Technologies to transform group-level grades into individual ones:

CATME

In previous academic year, CATME assessment of groupwork was used to assess over 1100 students in seven business modules across all years of our undergraduate program. The largest of these modules had 550 students (85 groups of 5-6 students). In the forthcoming Spring semester, one of the co-authors is going to use CATME for managing groupwork with the cohort of 820 students.

We have presented some aspects of these experiments in local teaching workshops, however, these experiences have not been previously systematically analysed and contrasted. Therefore, our proposal makes a unique contribution to iMOBTS.

**SESSION DESCRIPTION**

The running order of the discussion session is summarized below. The planned length of the session is 90 minutes.

* **Technology-enabled assessment in large classes: Three main approaches**

**Presenters:** Olga Ryazanova and Tatiana Andreeva (10 min)

* **Engaging students through MCQ quizzes run in class using Socrative**

**Presenter:** Olga Ryazanova (10 min)

* **Running online MCQ quizzes outside the classroom using Moodle**

**Presenters:** Tatiana Andreeva, Ruifang Wang (10 min)

*Discussion break: should students be allowed to use lecture notes and other teaching materials while completing online MCQ quizzes? (15 min)*

* **Peer grading using PeerMark**

**Presenter:** Tatiana Andreeva (10 min)

*Discussion break: how do we motivate students to engage in high-quality peer evaluation? (15 min)*

* **From group grade to individual grade using CATME**

**Presenters:** Olga Ryazanova, Ruifang Wang (10 min)

*Discussion break: how do we help weaker students engage with technology? (10 min)*

Presentation slides and the record of the key points of discussions will be made available to all participants after the end of the session, to support the implementation of discussed pedagogical approaches.

**REFERENCES**

Bayazit, A. and Aşkar, P., 2012. Performance and duration differences between online and paper–pencil tests. *Asia Pacific Education Review*, *13*(2), pp.219-226.

Boevé, A.J., Meijer, R.R., Albers, C.J., Beetsma, Y. and Bosker, R.J., 2015. Introducing computer-based testing in high-stakes exams in higher education: Results of a field experiment. *PloS one*, *10*(12), p.e0143616.

Hrivnak, G.A., 2013. CATME smarter teamwork. *Academy of Management Learning & Education*, *12*(4), pp.679-681.

Malau-Aduli, B.S., Assenheimer, D., Choi-Lundberg, D. and Zimitat, C., 2014. Using computer-based technology to improve feedback to staff and students on MCQ assessments. *Innovations in Education and Teaching International*, *51*(5), pp.510-522.

Northcutt, C.G., Ho, A.D. and Chuang, I.L., 2016. Detecting and preventing “multiple-account” cheating in massive open online courses. *Computers & Education*, *100*, pp.71-80.

Shephard, K., 2009. e is for exploration: Assessing hard‐to‐measure learning outcomes. *British Journal of Educational Technology*, *40*(2), pp.386-398.

Pesare, E., Roselli, T., Rossano, V. and Di Bitonto, P., 2015. Digitally enhanced assessment in virtual learning environments. *Journal of Visual Languages & Computing*, *31*, pp.252-259.

Piontek, M.E., 2008. Best practices for designing and grading exams. *University of Michigan CRTL* *Occasional Paper*, (24). Available from http://www.crlt.umich.edu/publinks/CRLT\_no24.pdf , last accessed on 09/03/2017.

Scardamalia, M., Bransford, J., Kozma, B. and Quellmalz, E., 2012. New assessments and environments for knowledge building. In *Assessment and teaching of 21st century skills* (pp. 231-300). Springer Netherlands.

Ventouras, E., Triantis, D., Tsiakas, P. and Stergiopoulos, C., 2011. Comparison of oral examination and electronic examination using paired multiple-choice questions. *Computers & Education*, *56*(3), pp.616-624.

Voce, J., 2015. Reviewing institutional policies for electronic management of assessment. *Higher Education*, *69*(6), pp.915-929.

Webb, M., Gibson, D. and Forkosh‐Baruch, A., 2013. Challenges for information technology supporting educational assessment. *Journal of Computer Assisted Learning*, *29*(5), pp.451-462.