Accepted for presentation at EdMedia 2015 and publication in the *Proceedings of World Conference on Educational Media &Technology 2015.*

The Transition to Online Marking in Large Classes

Andrew Eberhard
University of Auckland Business School
New Zealand
a.eberhard@auckland.ac.nz

Don Sheridan
The University of Auckland Business School
New Zealand
d.sheridan@auckland.ac.nz

Abstract: The prospect of business as usual given shrinking budgets, increased enrolments and an emphasis on an integrated approach to instructional design, delivery and finally assessments creates opportunities to consider systemic solutions such as digitizing students' work and then trialing new workflows and processes. This paper describes three innovative methods for digitizing students' work, the technologies involved in marking, and then provides a summary of lessons learned.

The Context

This innovation originated in the University of Auckland Business School. We offer a 3 year Bachelor of Commerce degree. The majority of students will have 2 majors, and many have conjoint degrees spread over 4 years, such as Bachelor of Science in Computer Science. We are among 1% of all business schools that have three accreditations and our departments are individually rated among the top 100 worldwide. Our budget is about 60% of the 'Big 8' 'sandstone' universities in Australia, so we must be frugal and well organized.

As a consequence our undergraduate courses have large classes that are manageable if divided into streams. Our 2 largest lecture theatres seat approximately 600 students. Thus first year information science, accounting and business classes, offered over 2 semesters per year, have combined enrollments of ~2,500 each. All these classes have centralized coordination so that teaching teams use the same teaching / learning resources as well as assessments, assignments and final examinations. Lecture recordings are provided for all first year classes. Second and third year classes become smaller as students select their majors and specializations within majors. Class sizes in accounting, a popular major, remain much higher than others throughout the curriculum.

The Course

The introductory information systems course is offered each semester during the academic year and has ~2200 students divided into groups of up to 550 each (based upon the lecture theatres available). It is a compulsory course for first year students where the majority (90%) will not be majoring in the subject. In addition, while the entry requirements for Commerce are the highest in the country, it is possible that students have focused on gaining the best marks on national exams while not being required or motivated to gain any sophisticated skills with personal computers. Thus, they fit the demographic of a 'digital native' but there is a significant variance in technological skills.

The course was rated among the least popular among students until 2009 and through passionate lecturers and tutors, and a complete re-design is amongst the most popular in 2012. One aspect of its success lies in the design / relevance of assessments, and the rapid and thorough feedback that accompanies them.

A Research Informed Innovation

Research articles have suggested there are efficiencies in marking large volumes of assessments and final examinations if they could be first converted to a digital format. (Duffin 2013) (Massey 2014) (Tossell 2014)

Research also points to learning improvements that may occur if we provide thoroughly informative forms of feedback as quickly as possible following assignment submissions. (Hattie 2005, 2009) (Petty 2014) The institution aspires to provide such feedback within 2 weeks of the submission deadline but there no evidence to indicate the degree to which this is achieved.

In the context of providing rapid and thorough feedback research evidence points to the value of associating rubrics with the assignment when it is provided to students, discussing the assignment and the marking rubric with the students, and finally using these rubrics as part of the marking process. (Stevens & Levi 2005) (Auvinen 2009) (Sheridan & Gardner 2012)

We were also keen to implement a system that allowed for simultaneous monitoring of markers to determine if issues were emerging in the interpretation of the solutions guide or rubric that was being applied. Such a system would enable real-time detection of a problem and retrospective corrections without the necessity of returning to the physical or digital image of the question itself. (Solano-Flores, Schneider & Timms 1999)

Moving to Online Marking Phase One

Online marking required the hard copy assessments to be scanned and digitized in the most efficient manner of the total elapsed time for marking was to be significantly shorter than the existing manual workflow. This was accomplished by:

1. Designing a cover sheet that could be read by an optical mark reader (OMR).
2. Designing the assessment so that the task or question and answer / diagram were always on the same page of the assessment document.
3. Clearly marking the boundaries for the student's response. Students were cautioned not to 'write outside the lines' or the scanner could miss their work.
4. Asking students not to take the document apart (remove the staple) because the pages must remain in sequence.
5. Providing a space at the top of each page for the students to write their ID. This was another precaution should the document become damaged.

The assessments were processed using a scanner built into our leased photocopiers. We soon learned the scanner must have a single pass, two-sided read capability if the process is to work without damaging the sheets. The photocopier's firmware created a PDF file and was networked to the tutor's desktop computer. The ReMark ™ (http://www.gravic.com/remark/) software was used to read the OMR sector on the cover sheet (student's ID) and subsequently create a file (with filename = ID) containing the cover page and the next "n" pages. The students' scanned assignments or tests were subsequently copied to a server to be accessed by the marking teams.

Various marking solutions were trialed including PDF annotation using iPads and other tablets, as well as Google forms (used as a rubric marking sheet). The benefits of using iPads and/or tablet pcs to annotate the PDFs were mainly around ease of use. The markers found it very easy to open the documents and write comments and marks directly onto the script. We used Adobe Pro (https://www.adobe.com/nz/products/acrobatpro.html) for this on the PCs and PDF Expert (https://readdle.com/products/pdfexpert5) on the iPads. However, there was a significant challenge around capturing the marking data as it proved extremely cumbersome to have to open each file and capture the marking data manually. In addition, there were not enough iPads/tablets to go around due to financial constraints.

Another marking solution we trialed was to mark using Google Forms (this solution didn't depend on scanning of the student's work). The main benefit of this solution was that we immediately captured the marking data as the markers progressed. This meant that we could do on the fly moderation by putting the data into statistical packages and checking for significant differences. Consistency between markers is a huge concern when you have a large marking team and this really helped ensure the course coordinator of marker consistency. We could also largely negate student concerns about consistency by pointing to our data analysis. Another benefit of using Google forms is that we could embed a marking rubric with feedback into the form. This also helped with consistency. This solution was also free which was a big plus for us. On the downside there were a few data entry errors as our markers had to manually enter the student ID of the scripts they were marking. There were workarounds for these but the effort involved meant that we just lived with the errors and manually fixed them later. The main downside though was the negative perception by students that we weren't giving them much feedback because there was no written feedback on the actual script (not even a tick). This was despite the fact that they received a report with feedback based on the rubric. They really seemed wedded to the "ink" on the page.

Modifications to Phase One

While the workflow was efficient, the experienced marking teams were not convinced the process was significantly better than their previous, blue-collar method where each document was physically manipulated. On the positive side it was suggested that inefficiencies in manually calculating the sum of the marks for each question was time consuming and there was always a risk of making a mistake due to fatigue. This could be avoided by 'flipping' the innovation. Thus the process was altered by placing an OMR (bubble scale) on the bottom of each page. The markers then read the student's answer and filled in the bubble matching the mark awarded. Once all of the assignments had been marked, the staples were removed and the documents scanned. This time the ReMark™ software encoded the student's ID and accumulated all of the marks for each question. The student's work was assigned a filename = ID and a record of the student's marks by question was created in a separate file.

Feedback: Timely & Thorough

Hattie (2009) and others place great emphasis on the timely and thoroughness feedback. To this end the feedback provided to students was beyond their experience in any other classes because:

1. Three days after the submission date or test, the results were emailed to the students
2. The email package included:
    a. A general comment about the difficulty of the assignment or test based upon the overall class statistics, then the student's ranking in the class, i.e. 275/1063
    b. The student was advised of their potential to do well overall or cautioned to work harder in order to pass the course. Those that were identified as "at risk" were asked to see the course coordinator to devise a study plan for the rest of the course.
    c. Graphs were provided to illustrate the distribution for the total assessment marks and marks for the sub-modules (multiple-choice, short answer, etc.). In this context the student's marks relative to the class average for each module and their ranking in that module was provided.
    d. Feedback on the short answers was mapped to the student's mark on the rubric:
        i. A range of marks for a given rubric triggered a related range of advice for the student.
        ii. If the overall mark on the short answer fell below a specified threshold then actual examples of students' good answers were provided (with their permission).

The email was generated by using a Microsoft Word mail merge using conditional rules. This generated a highly personalized set of feedback using the marking data.

Moving to Online Marking Phase Two

While the 'home-grown' solutions generated positive outcomes, we became aware of commercial solutions that possibly provided a better integrated workflow and the potential of generating many useful reports from the data generated throughout the marking process. (Examsoft http://learn.examsoft.com/) (Crowdmark http://crowdmark.com/higher-ed/) (Callibrand http://www.calibrand.com/) (Transforming Assessment http://www.transformingassessment.net/) Based upon our experience to date, Crowdmark appeared to have a comprehensive solution for marking large classes and a cost-free trial was easily organized. What follows is not an endorsement of Crowdmark, per se, but an example of the workflow and analytics that we believe should be considered when reviewing the marketplace. The following steps were followed in the Crowdmark evaluation:

1. Writing the assessment or exam questions followed the conventional workflow up to the point when the output was pasted into the Crowdmark template.
    a. If 'n' versions of a M/C portion of the test were required then these were created as separate copies for the Crowdmark template
2. These versions are then pasted into the vendor's template and the tutor decides how many copies of each version are required
3. The Crowdmark system generates 'n' unique copies of the assessment, each unique copy has an associated QR code for the document and each page in the document. This collection of documents forms a large *.PDF that can be forwarded to the university's printer or sent to a photocopier. In our experience if it's sent to the university's printer it's returned as 'n' documents stapled and ready to go.
4. The students fill in their name, ID and email address on the cover sheet and complete the assessment.
5. The assessments are scanned and uploaded to the Crowdmark cloud.

6. The tutor downloads a list of the students enrolled in the course.
7. The tutor manually assigns the student's email address to a scanned document. (This is reasonably fast as the search engine has a 'look ahead' function so an email address is quickly identified.)
8. The marking team is now able to simultaneously access all assessments. (Markers can be allocated to mark a specific question or the entire test as decided by the course coordinator.)
9. Marks are automatically accumulated on each assessment.
10. Comments may be added by the markers on individual student's papers. These comments may become somewhat standardized (based on the rubrics used) so the markers have a range of comments to apply as they see fit.
11. Markers can see other markers results and comments in real-time as the process continues. Markers can be anywhere, anytime (as in Step One) but being able to see other marker's activities is useful.
12. The course coordinator can see real time graphs of overall student performance which is useful for moderating marking.

Lessons Learned

The marking team for the first year information systems class, comprising the same number of individuals, experienced all of the phases described (above). Based upon: 1) manually shuffling paper, 2) marking with scanned documents and then 3) a compromise of using hard copy and having the marks computed with a scanned solution or 4) Crowdmark, - the marking team's preference was overwhelmingly in favor of Crowdmark. Marking was fast. Comments could be typed/pasted directly onto the paper and there was no shuffling of paper required. This was despite a large degree of reluctance to use Crowdmark from the team. Markers had been concerned that they would struggle to use the system and weren't comfortable with the thought of having to sit in front of a computer for an extended period marking. Option 3 is the second most favored.

Students were split between options 3 and 4. The main difference from their perspective was that with option 3 they received a PDF file of their actual test answers via email whereas with option 4 their test paper was only returned as a webpage link and they could not easily download this to convert to a neat physical artifact.

Outcomes

- In classes with very large enrolments we can mark examinations in parallel using marking teams (without the need for moving physical documents).
- In the case of essay questions, we have also used rubrics with our markers to monitor both the validity of the rubric and the reliability/speed of the markers. (For example we can provide the same set of examination papers to all markers at the outset to confirm marking standard.)
- In the case of mid-semester tests we have scanned 1000 booklets, marked and sent feedback to the students within two days. (Examination on a Friday, marking bee over the weekend, and feedback sent Sunday night.)
- Formative assessment is much easier to manage. Our tutors have created feedback messages based on the scanned results which are then sent immediately to the associated student.
- We have the option of sending the assessment, and feedback to students in a digital format and/or making their original hard copy document available. In any even we retain a digital repository of all assessments.
- Another significant benefit is we reduce the prospect of post-test cheating because we have a digital copy of the test so if a student tampers with their answers after the test and ask for a remark we can check the original first. (Anecdotally we have often suspected that this happened but it is hard to prove.)
- We will relieve the Evaluation and Scanning Centre of an immense load in so far as it is possible.
- We continue to use Registry procedures for supervised evaluations, but we are now able to scan final examinations for subsequent marking and provision to all students who request a copy.

Scope of Widespread Adoption

We are all familiar with the phrase "herding cats" when applied to convincing academics about a new workflow much less a comprehensive innovation. Thus, the prospect that the entire faculty might adopt a digital solution for marking assessments is certainly remote. Nonetheless the volume of hardcopy now managed manually is surprisingly large and remains invisible / unquantified, i.e. just a normal overhead.

This project may stimulate a consideration of what can be saved in real terms (staff hours) as against what a cloud solution like Crowdmark might charge. If we are providing timely and thorough feedback to students these innovations may result in lower failure rates. (Hattie, 2009) The additional benefits of increasing service levels to students and providing quality assurance throughout the evaluation processes are a plus; noted by the student in class surveys and possibly helpful to the administration in the next accreditation cycle. Finally, the paradigm shift will be BYOD and e-exams? (Hillier, 2014) (Schultz & Apostolopoulous, 2014) (Botelho, 2014) (Hillier & Fleck, 2013)

Potential Scanning Requirements for B&E										
Degree	Students	Class/Sem	# Assign	# Pages	# Tests	# Pages	Final	#Pages	#TotDocs	#TotPages
U/G	4878	8	1	10	1	10	1	15	117,072.00	4,097,520.00
P/G T	663	12	2	8	1	10	1	10	31,824.00	891,072.00
P/G R	42	4	1	20	0	0	1	20	336.00	13,440.00
									149,232.00	5,002,032.00

References: Cited and Consulted

Ahn, J. Y., Han, K. S., Choi, S. H., & Mun, G. S. (2014). Designing a personalized exam system to enhance students' understanding. *ICIC Express Letters*, *8*, 349–355.

Auvinen, T., Karavirta, V., & Ahoniemi, T. (2009, July). Rubyric: an online assessment tool for effortless authoring of personalized feedback. In *ACM SIGCSE Bulletin* (Vol. 41, No. 3, pp. 377-377). ACM.

Botelho, S. (2014). *Large-scale, fully online BYOD final exams: Not your parents' multiple choice | University Business Magazine*. *University Business UBC*. Retrieved December 12, 2014, from http://www.universitybusiness.com/article/large-scale-fully-online-byod-final-exams-not-your-parents-multiple-choice

Duffin, C. (2013). Two-thirds of exams marked online. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/education/educationnews/10103816/Two-thirds-of-exams-marked-online-report-finds.html

Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research, 43*(6), 387-425.

Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement.* New York, NY: Routledge.

Hillier, M. (2014). Final Answer: The Question of Summative eAssessment. In M. Hillier (Ed.), *e-Assessment Scotland 2014*. Dundee: Teaching and Educational Development Institute Publications. Retrieved from http://espace.library.uq.edu.au/view/UQ:342026#.VITdBIclmZE.mendeley

Hillier, M. (2014). UQ: BYOD for e-Exam Trial outcomes. Retrieved from http://espace.library.uq.edu.au/view/UQ:342018#.VITbKjF6A_4.mendeley

Hillier, M., & Fluck, A. (2013). Arguing again for e-exams in high stakes examinations, 385–396. Retrieved from http://espace.library.uq.edu.au/view/UQ:320481#.VITgZ3oH_NE.mendeley

Hillier, M., & Tran, L. (2014). Student perceptions of e-exams: hopes and fears. In *Higher Education Research and Development Society of Australasia (HERDSA)*. HERDSA.

Massey University. (2014). *An Introduction to e-Assessment - Online Assignment Marking and Management*. Retrieved from http://www.massey.ac.nz/massey/staffroom/teaching-and-learning/centres_tl/ctl/professional-development/stream-professional-development/online-assignment-management---being-more-efficient-and-effective/online-assignment-management---being-more-efficient-and-effective_home.cfm

Petty, G. (2011). *Teachers toolbox by Mediamerge*. Retrieved December 09, 2014, from http://www.teacherstoolbox.co.uk/index.html

Schultz, A., & Apostolopoulos, N. (2014). 10 Years of e-Exams at Freie Universitat Berlin: an Overview. In *Transforming Assessments*. Retrieved from http://transformingassessment.com/eAS_2014/events_19_september_2014.php

Sheridan, D. & Gardner, L. (2012). From Cellulose to Software: The Evolution of a Marking System. In T. Amiel & B. Wilson (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2012 (pp. 454-461). Chesapeake, VA: AACE. Retrieved December 10, 2014 from http://www.editlib.org/p/40784.

Solano-Flores, G., Raymond, B., Schneider, S. A., & Timms, M. (1999). {Management of scoring sessions in alternative assessment: the computer-assisted scoring approach}. *Computers & Education*, *33*(1). Retrieved from http://www.sciencedirect.com/science/article/B6VCJ-3XMGN91-4/2/ad87d0adb7f78886fc206c633eb67e1b

Stevens, D. D., & Levi, A. (2005). *Introduction to rubrics: an assessment tool to save grading time, convey effective feedback, and promote student learning*. *Academy of Management Learning Education* (Vol. 8, p. 131). Stylus Pub. Retrieved from http://books.google.com/books?id=LIxWgDn8_N0C

Tossell, I. (2014). Software takes the pain out of marking exams. *The Globe & Mail*. Retrieved from http://www.theglobeandmail.com/report-on-business/small-business/starting-out/software-takes-the-pain-out-of-marking-exams/article16920652/

University of Auckland. (2014). *An online Assessment Tool*. *Teaching_and_Learning _Hub*. Retrieved from https://www.fmhshub.auckland.ac.nz/menu0_17.html

Walker, D. (2012). The strategic importance of policy development to successful e-assessment implementation. *International Journal of E-Assessment*, *2*. Retrieved from http://journals.sfu.ca/ijea/index.php/journal/article/view/36

White, M. (2012). E-assessment: Accessibility in practice: A case study of the Life in the UK Test. *International Journal of E-Assessment*, *2*. Retrieved from http://journals.sfu.ca/ijea/index.php/journal/article/view/38