Blending Assessment: Shifting Pen & Paper to the Digital World
Revised LEARN (http://learnquebec.ca/) Research Briefing
Canadian eLearning Network http://CANeLearn.net
October 2016

Introduction

Assessment is central to learning: it shapes the learning experience yet is a critical and time-consuming function for teachers. Technology is seen as a solution to improve efficiencies while reducing teacher time. No matter its purpose, the use of technology in assessment must contribute positively to the quality of the learning experience first, and organizational effectiveness second. Understanding how technology can enhance assessment practices must be part of any business case made for its use, but only after consideration of its impact on learning.

Computer-assisted assessment has many potential benefits: improved efficiencies and consistencies, improved data analysis, immediate feedback for the learner, improvement in quality of the learning, and reduction in the workload of administrators and teachers. However, there are issues in accessibility, technical consistency and most importantly scalability that must be considered before adoption.

This paper provides an overview of the issues and challenges faced when implementing a program where digital technology replaces traditional pen and paper evaluation. It is intended to serve as a framework for the consideration of how to improve learning through the use of technology in both formative and summative assessment.

The report consists of five sections beginning with a brief overview of the literature on technology and assessment, followed by a discussion of issues in the use of technology for assessment. The third section offers an in-depth review of authenticity in assessment and proctoring, including a discussion of issues and strategies to address concerns. The fourth section discusses policy and practice trends in the use of technology and assessment leading to a final section summarizing issues for consideration, recommendations and a call for further research.

Literature Overview

The use of technology in education is commonly described as more than just the introduction of information and communication technologies (ICTs) in the physical or virtual classroom. Most definitions describe the use of hardware and software in the context of education theory or pedagogy. The Association for Educational Communication and Technology's (AECT, www.aect.org), a professional association of thousands of educators who work to improve instruction through technology, have developed a number of official definitions for the field of educational or instructional technology. The first was "educational technology is a complex, integrated process involving people, procedures, ideas, devices and organization for analyzing problems and devising, implementing, evaluation and managing solutions to those problems involved in all aspects of human learning" (AECT, 1977, p. 1). This was followed more recently by "instructional technology is the theory and practice of design, development, utilization, management, and evaluation of processes and resources for learning" (Seels & Richey, 1994, p. 1). Most recently, AECT adopted Januszewski and Molenda's (2008) definition of educational technology – that "educational technology is the study and ethical practice of facilitating learning and improving performance by creating, using and managing appropriate technological processes and resources" (p. 1).

Key to all of these definitions is the facilitation of learning, supporting a shared model of learning rather than a controlled transmission model of information dispensing. The use of technology in education includes both technologies' role in formal education settings, but also the informal learning that occurs in our daily lives and those of our students. This situates the use of education technology in the domains of both learning and performance assessment. This is where the use of technology in the service of assessing learning has a critical role in education.

Before exploring the role education technologies play in assessment and inherent issues, it is important to situate the discussion within the research on assessment. Some of the best research available about what aspects of pedagogical, instructional design, curriculum, and structural issues actually have meaningful impact upon student learning can be based on the research of John Hattie. For more than a decade Hattie has synthesized over 800 meta-studies that represent more than 50,000 individual studies and more than 80 million students to examine what works best in education. What is most interesting, and relevant in this context is the fact that items related to formative, summative, and competency-based assessment ranked highly (Hattie, 2009). For example, "self-reported grades" (i.e., student-generated formative and summative assessment) was ranked higher than any other aspect. Similarly, "providing formative evaluation" and "feedback" placed third and ninth (both forms of formative assessment) respectively, which was consistent with earlier findings by Black and Wiliam (1998a). Even "mastery learning" (i.e., a form of competency-based assessment) was ranked highly. What this should underscore is the importance of assessment – particularly formative assessment – in student learning. Technology can play a critical role in formative assessment by providing timely and effective feedback to students in a manner that improves learning.

There are other forms of assessment that are common within educational contexts. For example, it is not uncommon for teachers to assign larger projects to their students and use rubrics to assess their work. Similarly, teachers may also use portfolio-based forms of assessments to allow students to showcase their work over a prolonged period of time. Many of these non-standardized models of assessment have long been known to have a more positive impact on student learning (e.g., Black & Wiliam, 1998b; Calfee & Freedman, 1996; Daro, 1996; Mills, 1996). For example, the use of rubrics has also been found to have a positive impact upon student learning (Andrade, 2000; Jonsson & Svingby, 2007; Ross, 2006). Further, rubrics have also been found to be a reliable and valid way for students to undertake self-assessment (Andrade & Du, 2005; Ross, 2006); something Hattie also found to have a meaningful impact on student learning (Fachikov & Boud, 1989; Falchikov & Goldfinch, 2000; Kuncel, Crede, & Thomas, 2005; Mabe & West, 1982; Ross, 1998). Digital environments support the communication and sharing of these assessment models and can streamline and

simplify both student and teacher time and efforts.

Unfortunately, within many developed nations the issue of assessment has been closely tied to accountability and a system of determining successful and unsuccessful schools (often with fiscal penalties associated with being labeled an unsuccessful school, see Ravitch, 2011; 2013). Within this regime of accountability, assessment is limited to statewide, standardized tests. Due to the high stakes nature of large-scale standardized tests, much of the classroom based assessment has also been adapted to allow students to become accustomed to, more familiar with, and practice on these multiple-choice assessments. It is ironic that the types of assessments that Hattie and others have found to have meaningful impact on student learning and student performance, which are often used in non-accountability jurisdictions, are largely absent in much of the assessment found in the developed nations that rely upon systems of accountability.

However, even within these accountability jurisdictions, there has been some progress in how assessment has been integrated into K-12 education. For example, over the past fifteen years the field of K-12 online and blended learning has included the use of standards-based assessment to test for prior learning to allow learners to focus on content they have not mastered (Christensen, Horn, & Johnson, 2011; Moe & Chubb, 2009; Packard, 2013; Peterson, 2010; Vander Ark, 2012). Many systems are designed to allow students to focus on mastery learning. This is consistent with the current advocacy from the International Association for K-12 Online Learning on competency-based learning (Powell, 2013; Powell, Kennedy, & Patrick, 2013).

In recent years, the use of portfolios, rubrics, and self-assessment have also evolved in innovative ways – particularly within the higher education and informal learning environments. For example, digital badges have been introduced as a way to provide a form of credential for individuals that have completed a non-credit learning opportunity. Badges have been used most often in informal learning settings; however, there has been an increased use of badges in formal settings as well (e.g., as extra credit in online courses or as a way to track professional development). Within these more formal settings, the use of badges to measure and motivate learners has been shown to have positive effects on both aspects (Abramovich, Schunn, & Higashi, 2013; Denny, 2013), especially when the badges were used as a way to increase the students' own learning goals.

Another innovation in the use of assessment has come from the growth of massive open online courses (MOOCs) over the past three years. MOOCs are online courses that allow open enrolment to any learner, regardless of geographic location, and for free. This has often resulted in very high enrolments numbers (i.e., massive). Due to the large number of students within these MOOC environments, assessment has generally been undertaken using either a computer graded, standardized exam or through self- and peer-graded assessment (Karsenti, 2013). The self-assessments and peer-graded assessments have been generally accepted approaches by the faculty responsible for leading these MOOCs (Kolowich, 2013) and, interestingly, it has also been found to be a reliable measure of student learning (Balfour, 2013). These are just two examples of more innovative forms of assessment, driven by advances in educational technology, which have the potential to expand the positive aspects of assessment on student learning.

Overall, education technology has an increasingly important role in both formative and summative assessment. The issues and challenges in the implementation and use of technology in virtual and physical classrooms and schools is discussed in the next section.

Using Technology for Assessment

> Technology holds the potential to revolutionize the delivery and assessment of learning outcomes and will result in fundamental changes in how we teach; which mental processes, skills and understandings we measure; and how we make decisions about student learning. (Taylor, 2005, p. 102)

There is much promise ascribed to technology's ability to revolutionize in most human endeavours. Education technology is seen as a lightning rod to transform traditional forms of assessment and testing that hold the promise of bringing education into the 21st Century. However, like any new promise of education change, the application of technology without a well-thought out approach and adequate testing for scalability of use can be a determining factor in its positive or negative impact.

For teachers technology offers the opportunity to design more effective, interactive and engaging formative assessment and feedback. Technology can increase efficiency, thereby reducing teacher workload, and can also be used to improve the quality of learning for students. The use of educational technologies in formative assessment can enable teachers to interact with learners in direct and meaningful ways that are not attainable to them in a traditional classroom and assessment approach. Polling systems, online discussion forums, shared online communities like wikis, blogs, etc. enable a level of exchange and conversation that can allow teachers to make a more personal connection to students and provide better and timely feedback to guide learning. Digitally based feedback in self-directed, online learning spaces can open up opportunities for teachers to make better use of face-to-face classroom time for social and group learning.

While the increased use of technology may offer benefits to teachers, school administrators turn to technology to streamline efficiencies and the effectiveness of examinations. The history of large-scale testing dates back over 1000 years when paper-based examinations were introduced in China (Stobart, 2008), while the use of computers in testing dates back only mere decades. Russell (1999) found that initial research in the case of multiple-choice testing, the use of computers versus traditional paper and pencil tests found little or no variation in test results. However, given that over the past few years students are now more likely to be using technology both in their studies and to prepare for tests, recent research suggests that paper and pencil examinations can put students at a disadvantage (Ghaderi, Mogholi & Soori, 2014). In short, there is little evidence to indicate that computer-based testing puts students at any more disadvantage than paper-based testing (Seo & DeJong, 2015).

Recently computer-based testing has taken a foothold in large-scale summative testing in K-12. Many provinces have begun to phase out paper-based province-wide achievement tests in favour of computer-based examinations. Some provinces are designing these examinations in such a ways as to assess creativity, critical thinking and problem-solving skills (MacDonald, 2014; Canadian Press, 2013). Despite the promise of streamlining the efficiency of large-scale testing, to date many jurisdictions in the United States adopting large-scale computer-based testing have run into computer access issues, slow load times and servers that are so overloaded schools had to revert to paper-and-pencil tests (FairTest, 2015). In Canada, Alberta Education attempted province-wide examinations in January 2015 and Newfoundland in October 2015 with similar results (C. Oulten & M. Barbour, personal communication, January 13, 2015). Despite the issues, both states and provinces continue to move forward with the strategy. Following Alberta's and Newfoundland's lead, Ontario's Education Quality and Accountability Office is planning its first province-wide test March 31, 2016 (EQAO, 2015).

Issues raised in the initial attempts of these large-scale summative examinations include scalability, as indicated by the reports of students being disconnected from the examination's online server, but more importantly accessibility, reliability and validity of the computer-based test design. Security and effectiveness

of administrative procedures used in deployment is also a documented concern (Gilbert, Gale, Warburton & Willis, 2009). Other issues include determining what constitutes effective on-screen assessment, what human, physical, and technical resources are required to support exam delivery, and even the validation of the psychometrics and pedagogy of the examination itself. Addressing these issues is key in providing not only a successful assessment experience, but improving the overall quality of learning for all students and assessment data collected.

Authenticity and Assessment

The media abounds with stories of cheating from fraud charges laid in a cheating scandal in Waterloo Ontario (CBC, 2014) to parents scaling walls in India to help students cheat (Freeman, 2015). In a study of over 5000 graduate students in business and non-business graduate programs (McCabe, Butterfield & Trevifio, 2006), researchers found a culture of academic dishonesty, and this culture was not restricted to online but throughout the campus and programs. With the increasing use of online learning environments by teachers and instructors in both regular classrooms and designated online learning programs, challenges in preventing cheating are often at the forefront of conversations, particularly by those less experienced with online teaching and learning. A cursory consideration of online coursework, assignments and more importantly examinations can easily lead to the impression that cheating would be simple and more prevalent in virtual classrooms than in campus classrooms. However, if this were true why the continued growth of online learning in K-12 and post-secondary institutions, and in corporate training programs? If cheating were so rampant would that not immediately undervalue the credibility of any degree or certificate achieved from an online program?

The perception that cheating in online courses is greater than in face-to-face classes may arise from an increase in the availability and use of technologies that offer potential access to information on an anytime, anywhere basis. The internet, social media, and cell phones are ubiquitous, and students at any level from elementary to graduate school are well versed in using them to acquire information and answers to problems that arise in their academic and daily lives. However, countless studies have looked at cheating in online courses (Stuber-McEwen, Wiseley & Hoggatt, 2009; Watson & Sottile, 2010) and none suggest that cheating is inordinately high, concluding that cheating occurs among students in classroom-based courses at a similar rate to students enrolled in online courses. In their 2010 study, Watson and Sottile compared cheating among 635 students in online versus classroom-based courses. The results found that while students perceived their peers were far more likely to cheat online than in campus-based courses, there was actually no noticeable difference in the behaviours of online versus campus-based students. In fact, the study found lower levels of cheating in online courses than live or "on ground" courses.

With no noticeable difference between online and traditional campus-based programs regarding cheating, the question is how to prevent **any** cheating in online testing. How can programs effectively proctor and invigilate exams at a distance from the teacher or instructor? While it is not uncommon to require students to go to a testing centre for exams, this requirement undermines the convenience of distance learning and may not be feasible for some students.

Strategies to Increase Authenticity

Stemming cheating in any exam setting is an issue, and often technology is seen to be the solution. However, as with virus protection for computers, there is an endless cycle of “catch and prevent”; ingenious methods of cheating must be met with new and creative ways to prevent cheating. Recognizing that cheating can and will occur, there are a variety of strategies that can be implemented in online courses to minimize the occurrence and maintain the integrity of the course and examination. Many of these strategies are also valid for classroom testing, but learning management systems such as Blackboard, Moodle, Brightspace by D2L, ANGEL, and Sakai used in online delivery are designed to integrate many measures to deter dishonesty. Some of these strategies include creation of a large test bank, randomizing test questions and their order, restricting the amount of time for students or restricting the availability of the test. Tests may be password protected and a student’s computer/browser may be “locked down”, or entry/exit of the test may be restricted.

When developing tests, instructors can implement additional strategies such as only using automated assessments for a small percentage of the overall course grade with assignments requiring higher level

thinking making up the majority of the grade. Instructors could treat online tests as "open book" and consider developing questions such that students who know the material thoroughly can finish the test more quickly than students who need to refer to several sources to construct a response. Finally, waiting until all students have completed a test before releasing grades or feedback is a strategy to discourage the sharing of questions and/or responses with students who have yet to take the test.

These strategies are common in low-stakes testing situations. When the stakes are higher, additional measures may be implemented. New technologies are being developed that provide additional security. For example, Secureexam Student (SES) from Software Secure (www.softwaresecure.com) is a standalone test delivery solution providing access to a secure word processor or spreadsheet independent of applications installed on the student's workstation. Students access a test from their own computer, but SES protects exam integrity by preventing access to unauthorized external resources by locking down the test taker's operating system. The Respondus LockDown Browser (www.respondus.com/products/lockdown-browser/) integrates with most popular learning management systems and locks students into the test once it has been started so they are unable to print, copy, go to another URL, or access other applications. SafeExam Browser (SEB - http://safeexambrowser.org/) changes any workstation into a secure workstation, controlling access to any utilities like system functions, other websites, applications, and resources. SEB is integrated into recent versions of ILIAS and Moodle, reducing the need for specific system configurations. In addition to restricting access to resources, SEB also prevents unauthorized leaving of the test prior to completion.

Proctoring

The software strategies described above are designed for implementation on a student's personal workstation. They provide a level of control, but do not include means for monitoring student use of resources outside of the computer workstation such as another computer, tablet, phone, print materials, or even another person. Some online courses require students to take exams at campus-based or other testing centres, where groups can be monitored during the test period by one or more proctors in the same room. Taking this one step further in the online course delivery model, online proctoring is now an option that is becoming more prevalent. Some online proctoring systems monitor live feeds in real time while others record sessions for later review. Some require students to share a view of their display, others require the installation of software that restricts access to Web browsers or chat programs while a test is in progress (Kolowich, 2013). In addition, monitoring may include identity confirmation, keystroke analysis (typing patterns based on rhythm, pressure, and style), or authentication based on examination of physical features such as facial, fingerprint, palm or voice recognition.

Remote exam proctoring is a relatively young industry that is growing in importance as online enrolments and course offerings continue to increase. Remote proctoring generally refers to monitoring an exam over the Internet through a webcam and includes any processes that assist in securing a test administration event. Current online proctoring providers include:

- Kryterion Inc. (www.kryteriononline.com)
- Proctor Now (www.softwaresecure.com)
- ProctorU (www.proctoru.com)
- B Virtual (www.bvirtualinc.com)
- Tegrity (www.tegrity.com)
- Respondus (www.respondus.com)
- Microsoft Online Proctoring – Beta (https://www.microsoft.com/learning/en-us/online-proctored-exams.aspx)
- Pearson VUE (http://home.pearsonvue.com/)
- ProctorTrack (http://www.proctortrack.com/)

This rush to find technological solutions to prevent cheating has been taken up by some of the major corporations operating in the education space. Pearson Education offers its VUE testing centres worldwide, providing invigilation and proctoring services for a fee. Microsoft Online Proctoring is a companion program to the Pearson VUE testing service. While still in beta, the service allows users to take an exam in their home or office while being monitored by an offsite proctor provided by Pearson VUE's service. Users of the service are first required to install a secure browser that locks down their computer during the testing process. In addition, Pearson VUE proctors monitor the user at all times via webcam and microphone. All sessions are recorded and, should cheating be suspected, a review of the footage is used to confirm and exam results are nullified and the exam owner is notified.

Typically, online proctoring involves authentication of the test taker, a work area scan, system lockdown, observance during test completion and/or review of video capture. The authentication process verifies the identity of the test taker as the same person who is enrolled in the course. Government-issued photo IDs are held in front of the webcam for a snapshot that is matched to an ID provided with course registration. More sophisticated techniques involve fingerprint or palm scans. The work area scan enables the proctor to look for security risks such as additional monitors, computers, handheld devices, notes, or books. The student's computer system is locked down to prevent access to applications, web browsers, and search engines. Live web proctoring requires students to have a human proctor watching them through their webcam and therefore tests and proctors must be scheduled at specific times. On demand web proctoring also utilizes webcams, but requires no scheduling because the test session is recorded and reviewed later.

In online proctoring students agree to allow cameras to "watch" them while they take exams. There may or may not be a person watching in real time. If certain behaviours associated with dishonesty are observed, consequences range from further review after the test has been submitted with follow-up by the institution. Immediate behaviour correction or suspension of the test would occur as in a face-to-face proctoring situation in a classroom or testing centre. Online proctoring is only concerned with a subset of overall program vulnerabilities, threats and risks: those threats that are detectable by the online proctor and over which he/she has some control. These behaviours include colluding with others, using cheating aids, using a proxy test taker, copying answers from other test takers, capturing content by digital photography devices, or transcribing questions verbally (on paper or recording device). One major criticism of online proctoring is its reduction of human behaviour to an algorithm. Behaviours as simple as stretching, looking away from the screen, or bending down to pick up a pencil can be construed as anomalies that flag the test for a violation.

The credibility of the proctor is often an issue for monitoring remote learners if schools or programs are using a remote testing centre and services, regardless if they are online or onsite in a testing centre. Validating the credentials and integrity of offshore proctors is particularly difficult, even if they are involved in a legitimate agency or even educational institution. Accordingly schools and organizations look to policy for direction and technology for solutions. Ultimately, the determination of what is cheating is left to the proctors.

Online proctoring is intended to avoid security weaknesses of traditional proctoring such as when proctors have a stake in the outcome of testing and tests are vulnerable to compromise. On-site proctors may not receive monetary compensation (or are poorly paid), may be relatively unmotivated, and may be poorly trained. According to Caveon (www.caveon.com), a test security service, online proctors are trained 'professionals' used to improve the reliability and integrity of the examination process (Foster & Layman, 2013). However, one online proctor may be observing multiple exams from multiple institutions with different exam policies. Live proctors may be simultaneously communicating with test takers, authenticating identification, answering questions, and resolving technical issues in addition to observing behaviour. Remote proctoring is dependent on the quality of the proctors themselves and the number of live exams being proctored at the time by one individual. In fact, in many cases these online proctors are college students themselves (Bergstein, 2012).

The use of online proctoring monitoring approaches is not without its controversy. The quality of the proctors is only one issue to be considered. Students who feel that webcam monitoring is invasive have already raised privacy concerns (Singer, 2015). Personal information is disclosed to proctors during authentication. While the online proctoring companies do have privacy policies in place, according to Singer some indicate personal information may be disclosed in such events as merger, sale or acquisition. Storage of information, including video footage, presents a further problem. For the most part, the companies providing online proctoring services store information on servers located in the US. One exception is B Virtual (www.bvirtualinc.com) whose FAQ indicate that recordings for Canadian and other international clients concerned about the implications of the US Patriot Act are stored on a certified third party server in Canada.

Other privacy concerns include the inappropriateness of online proctors seeing the responses of test takers and the invasiveness of having someone watching the test takers in their homes (Kolowich, 2013). As with any security service, there is a correlation between cost and quality, (i.e., higher cost results in better security). For example, methods that employ passive proctoring (video is reviewed after test completion) are less costly than methods that utilize proctors in real time with interaction between the test taker and the proctor. The latter, however, provides a higher level of security and increased fraud detection and deterrence (Foster & Layman, 2013).

Most providers require a per test fee from students ranging from $15-$45 with the actual cost depending on the type of course, test, service, etc. At Rutgers University, students are charged $37 per course to use Proctortrack. This is in addition to a $100 tuition surcharge imposed for each online enrolment (Singer, 2015). Additional costs born by students include technology (webcams and microphones), and, in some cases, requirements include special lighting, mirrors or other reflective surfaces to show the proctor the edges of the display monitor. Costs to the institution include staff salaries for administering the online proctoring system (i.e., proctoring the proctors), and fees for signing on as a client of the service.

Finally, another key concern is false reports of impropriety. One of the monitoring software companies, Software Secure (http://www.softwaresecure.com/), classifies incidents into three tiers. Statistics indicate they note “minor suspicions” in 50 percent of exams they proctor, “intermediate suspicions” in 20-30 percent, and “major” incidents in 2-5 percent (Kolowich, 2015). If this is common among all online proctoring services, over 70% of online proctored exams are flagged as having some sort of behaviour worthy of attention. This does not always mean the students are cheating, but it does mean the institution will be notified and will have to review the recordings and follow up.

Each online proctoring company has developed its own approach. An analysis of the extent of issues surrounding student authentication and cheating will help determine if remote proctoring is the right solution. Cost-benefit and return on investment analyses identify whether the cost of implementation is warranted and results in the desired benefits.

Policy and Practice Trends in Use of Technology and Assessment

Specific policy for online examination is limited with most provincial and territorial policy related to examinations in general, without specifying the method or technology used to test students. In the case of large-scale testing students may have a choice between computer and paper-based exams, however in many cases computer based electronic exams are starting to become the only option provided to students. While most provincial and territorial ministries provide documentation on large-scale examinations specifying where exams can be written, who can administer and invigilate exams, and consequences should students' breach exam rules, in the case of computer-based testing specifications go further to describe computer and browser requirements, software requirements and security measures required. Many jurisdictions have requirements for online schools to have similar policies and documentation for computer-based testing, particularly if the student is at a distance to the teacher and school. These measures are all intended to ensure the authenticity of the test and its results.

The British Columbia Ministry of Education's 2015\16 Handbook of Procedures for the Graduation Program (2015) stipulates for students the consequences of "cheating" or being disqualified from a Provincial Exam, but does not differentiate between paper or electronic. This seems like a reasoned approach given that with constant technological change general policy statements that can be applied and interpreted in specific situations may be a more prudent and longer lasting approach. In most cases, provincial policy for both paper and electronic exams students may be disqualified if in possession of secure exam material, for communicating with other students during the exam, giving or receiving assistance or copying answers, using an unapproved piece of equipment, or accessing any materials on the local computer or internet.

Many post-secondary institutions have developed sophisticated and comprehensive exam proctoring and invigilation processes and policy. Athabasca University, one of the early leaders in online and distance education and Canada's Open University, has developed an exam invigilation network (http://registrar.athabascau.ca/exams/invignet.php) with centres across Canada. For students studying outside of Canada protocols are in place to ensure a suitable invigilator is available and in place to support examination invigilation. Similarly, Contact North, Ontario's distance education and training network, provides exam invigilation services across Ontario for a fee (Contact North, n.d.). Calgary, Alberta's SAIT Polytechnic University goes even so far as to require online exam invigilators to enter exam passwords for each student, and in the event of any disruption or issue students are required to close the test and reschedule a written exam in a supervised paper session (SAIT, 2012). Similarly, most K-12 online schools require a local test invigilator who is a professional person in the community and computer literate, ideally affiliated with an educational institution. Policies are consistent across most K-12 schools, and in fact many of these schools offer reciprocal invigilation services for the students residing in their local school community and catchment area.

Promising Trends

Promising trends is best defined as emerging changes that hold the promise to bridge the widening gap between how learning occurs, formally or not, in the real world as compared to how it occurs inside the confines of the school walls. Two strands of trends worthy of examination include the policy discourse and the observed practice. They are examined separately because educational policy does not firmly steer practice, nor does practice sufficiently inform policy. But when alternative tools fail (such as changes to the curriculum, investments in this or that technology, teacher training, etc.) changes in policy discourse remains the surest way to renew hope in another round of changes.

Policy makers feel pressured to base policy decisions on empirical evidence. To that end, they typically ask questions of the "what-have-you-done-for-me-lately?" variety, *i.e.*, they tend to ask: *what has technology done*

for learning outcomes? The available answers are equivocal at best and sobering at worst. Such empirical results tend to deter policy makers and dishearten advocates of greater use of technology by students, more so in times of financial scarcity and even scarcer patience, on the part of parents, for pedagogical experimenting. *"What-have-you-done-for-me-lately?"* questions imply three under-examined assumptions that are explored in the following section.

First, "*What have you done…"* questions assume that technology is subservient to a timeless and universal transaction of knowledge between teacher and student—should technology fail to facilitate that transaction, its use ought to be questioned. Put more directly, students have had to master irregular verbs and Pythagoras's theorem since modern schooling was conceived, and will presumably continue to do so as long as such schooling persists. If technology does not improve test scores, why bother using it?

Second, these type of questions assume that technology use outside of school walls, however widespread and growing, does not fundamentally alter the "three R's" expected at the outset of schooling. More importantly for advocates of greater use of technology by students, a recent landmark study by the OECD (2015) reaches somewhat counterintuitive conclusions that tend, at first glance, to support this second assumption. To name but two examples: "Show me a proficient technology user, and I will show you a pupil who masters the three Rs" (p.16); and "One does not need technology to teach so-called twenty-first skills—traditional pedagogies are fine" (p.17).

Lastly, such questions assume that ICT adoption by teachers is the solution to each of the aforementioned problems. This is the so-called "Trojan horse" theory (Papert, 1997) whereby fundamental educational change is smuggled past the classroom doors, hidden inside the boxes and bubble-wrap used to deliver the tools (laptops, tablets, etc.). Once the teachers adopt the technology and become adept at it, it is assumed that not only will (traditional) learning outcomes improve but also a richer type of learning (i.e., competency-based) will become the norm. If the past is prologue, advocates of greater use of technology by students ought to abandon any form of "Trojan horse" theorizing.

To summarize, the spontaneous *"What-have-you-done-for-me-lately?"* line of inquiry has so far generated answers oscillating between "losses" and "not much considering the effort". More importantly, its underlying assumptions are unlikely to improve the range of the anticipated answers in the foreseeable future, mostly because it tends to treat learning outcomes as immutable.

One promising trend would be to modify the preposition *for*, yielding: *"What-have-you-done*-to-*me-lately?"* That is to say, "what has technology done *to* learning outcomes?". Such a minor change in the trigger of the policy conversation holds the power to upend the perspective. At the very least, technology has made the mastery of three Rs, especially the first two, even more vital (OECD, 2015). It begs such questions as: how does reading today (mostly on clickable screens akin to a 3D format) compare with reading yesterday (mostly on 2D paper)? If the overlap is not 100%, as it surely isn't, what is the role of the school in bridging that gap?

More to the point, "what has technology done *to* learning outcomes?" introduces effortlessly the sub-question about assessment. Instead of asking "what has technology done lately *for* assessment?" (i.e., has it made it more efficient, less time-consuming for teachers? More cheating-proof?), a more relevant and promising question becomes "what has technology done to assessment?".

In short, technology-enhanced, as opposed to technology-based, assessment holds the power to shake some of the least desirable cornerstones of traditional schooling.

Observed Practice

Among the cornerstones of traditional schooling, make-believe and uniformity stand out. *Make-believe,* where students, especially younger ones, are extracted from the real-world to be confined behind the shielding walls of their school where they are made to believe that two neighbouring farmers would actually use geometry to settle land disputes and algebra to divide disputed cattle. *Uniformity*, where, on the basis of their birthdate, students are administered the same curriculum, at more or less the same rhythm.

The most promising trends in technology-enhanced assessment are likely to shake such foundations of traditional schooling. Immersive environments and multiplayer games hold the capacity to inject reality, virtually, inside the confines of the school walls: real-world farmers might not really draw their geometry box to settle a dispute, but what's to prevent a couple of students immersed in a problematic and highly realistic situation to do precisely that? What better way to assess their problem-solving capacities, whether in a collaborative or adversarial setup? Finally, learning analytics hold the capacity to inject personalization despite a common curriculum. Few teachers oppose, on principled grounds, the notion of differentiation. Yet even fewer know how to practice it, given the typical class composition today, especially in urban public high schools.

To summarize, two trends are highlighted:

1. On the policy discourse side, executives ought to steer away from "what has technology done *for* learning and assessment?" inquiries, towards more of the "what has technology done *to* learning and assessment?" variety, thus shedding light on new policy perspectives, hopefully more conducive to bridging the gap between school and society.
2. On the pragmatic side, we ought to steer away from technology-based assessment tools in support of traditional learning outcomes towards more technology-enhanced assessment practices which tend to mitigate the less desirable characteristics of traditional schooling, by injecting more authenticity and more individualisation.

Summary and Issues for Consideration and Future Research

Organizations and institutions need to ensure they have clear policies and protocols for managing examinations and monitoring cheating, whether delivering programs at a distance or not. To date policies have remained in the general area of disqualifications, not differentiating between campus and virtual examination spaces and should remain so. However, technology is ubiquitous and pervades all of our assessment processes. While technology is often part of a solution to cheating, more often human intervention and supervision remains the best solution. Remote proctoring is early in its development and has its limitations. Continued vigilance and policy refinement is needed to stay current on the ever spirally treadmill in our technology saturated world and learning environments. Finally, the use of technology in large-scale testing is still in early stages. Much more research is required to determine effective use of technology in both formative and summative assessment.

References

Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education?: it depends upon the type of badge and expertise of learner. *Educational Technology Research and Development, 61*, 217-232.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 13-19.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation, 10*(3), 1-11.

Association for Educational Communications and Technology. (1977). *Educational technology: Definition and glossary of terms*. Washington, DC: Author.

Balfour, S. P. (2013). Assessing writing in MOOCs: Automated essay scoring and calibrated peer review. *Research & Practice in Assessment, 8*(1), 40-48.

Bergstein, B. (2012, November 16). In online exams, big brother will be watching. MIT Technology Review. Retrieved from http://www.technologyreview.com/news/506346/in-online-exams-big-brother-will-be-watching/

Black, P., & Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148. Retrieved from http://weaeducation.typepad.co.uk/files/blackbox-1.pdf

Black, P., & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in education, 5*(1), 7-74.

British Columbia Ministry of Education. (2015). *Handbook of procedures for the graduation program 2015\2016.* Retrieved from http://www.bced.gov.bc.ca/exams/handbook/handbook_of_procedures.pdf

Calfee, R. C., & Freedman, S. W. (1996) Classroom writing portfolios: old, new, borrowed, blue. In R. Calfee & P. Perfumo (Eds.), *Writing portfolios in the classroom* (pp. 3-26). Mahwah, NJ, Lawrence Erlbaum.

Canadian Press. (2013, May 9). Alberta to overhaul standardized testing, emphasize creativity. *The Globe and Mail.* Retrieved from http://www.theglobeandmail.com/news/national/education/alberta-to-overhaul-standardized-testing-emphasize-creativity/article11821827/

CBC. (2014, December 18). Alleged exam cheating leads to fraud charges against 2 students in Waterloo. Retrieved from http://www.cbc.ca/news/canada/kitchener-waterloo/alleged-exam-cheating-leads-to-fraud-charges-against-2-students-in-waterloo-1.2877980

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2011). *Disrupting class: How disruptive innovation will change the way the world learns* (2nd ed.). New York: McGraw-Hill.

Contact North. (n.d.) Exam invigilation services and fee schedule. Retrieved from http://teachonline.ca/sites/default/files/contactNorth/files/pdf/e-exam_invigilation_services__fee_schedule_-_admin.pdf

Daro, P. (1996) Standards and portfolio assessment. In J. B. Baron & D. P. Wolf (Eds.), *Performance-based student assessment: Challenges and possibilities* (pp. 239-260). Chicago, IL, University of Chicago Press.

Denny, P. (2013, April). The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 763-772). Paris, France: Association for Computing Machinery.

EQAO Online. (n.d.). Retrieved from http://www.eqao.com/en/assessments/eqao-online/Pages/eqao-online.aspx

FairTest. (2015, September 6). *Computerized testing problems: 2013-2015*. Retrieved from http://www.fairtest.org/computerized-testing-problems-2013-2015

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395-430.

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*(3), 287-322.

Foster, D. & Layman, H. (2013). Online proctoring systems compared. Retrieved from http://www.caveon.com/blog2/wp-content/uploads/2013/03/Online-Proctoring-Systems-Compared-Mar-13-2013.pdf

Freeman, C. (2015, March 20). Parents scale school walls to help students cheat in exam in India. The Telegraph. Retrieved from http://www.telegraph.co.uk/news/worldnews/asia/india/11484667/Parents-scale-school-walls-to-help-students-cheat-in-exam-in-India.html

Ghaderi, M., Mogholi, M., & Soori, A. (2014). Comparing between computer based tests and paper-and-pencil tests. *International Journal of Education and Literacy Studies. 2(4).* Retrieved from http://journals.aiac.org.au/index.php/IJELS/article/view/1308/1308

Gilbert, L., Gale, V., Warburton, B., & Wills, G. (2009). *Report on summative e-assessment quality (REAQ).* Retrieved from http://www.jisc.ac.uk/media/documents/projects/reaqfinalreport.pdf

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

Januszewski, A., & Molenda, M. (2008). Chapter 1: Definition. In Januszewski, A., & Molenda, M. (Eds.), Educational technology: A definition with commentary. New York: Lawrence Erlbaum Associates.

JISC. (2010). *Effective assessment in a digital age*. Retrieved from http://www.jisc.ac.uk/media/documents/programmes/elearning/digiassass_eada.pdf

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144.

Karsenti, T. (2013). What the research says. *International Journal of Technologies in Higher Education, 10*(2), 23-37.

Kolowich, S. (2013, March 21). The professors who make the MOOCs. *The Chronicle of Higher Education*. Retrieved from http://publicservicesalliance.org/wp-content/uploads/2013/03/The-Professors-Behind-the-MOOC-Hype-Technology-The-Chronicle-of-Higher-Education.pdf

Kolowich, S. (2013, April 15). Behind the webcam's watchful eye, online proctoring takes hold. The Chronicle of Higher Education. Retrieved from http://chronicle.com/article/Behind-the-Webcams-Watchful/138505/

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63-82.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280.

MacDonald, M. (2014, Sept. 10). Ontario becoming first province in Canada to adopt online model for standardized tests. *National Post.* Retrieved from http://news.nationalpost.com/toronto/ontario-becoming-first-province-in-canada-to-adopt-online-model-for-standardized-tests

McCabe, D.L., Butterfield, K.D. & Trevifio, L.K. (2006). Academic dishonesty in graduate business programs: Prevalence, causes, and proposed actions. *Academy of Management Learning & Education, 5(3), p.294-305.*

Mills, R. P. (1996) Statewide portfolio assessment: The Vermont experience. In J. B. Baron & P. Wolf (Eds.), *Performance-based student assessment: Challenges and possibilities* (pp. 192-214). Chicago, IL, University of Chicago Press.

Moe, A. M., & Chubb, A. M. (2009). *Liberating learning: Technology, politics and the future of American education*. San Franciscom CA: Jossey-Bass.

Newhouse, C. P. (2013). Computer-based exams in schools: Freedom from the limitations of paper? *Research and Practice in Technology Enhanced Learning. 8(3). P.431-447.* Retrieved from http://www.apsce.net/uploaded/filemanager/96891e43-4220-4d53-99e2-0e9dabaae02b.pdf

OECD. (2015). *Students, computers and learning: Making the connection*. PISA, OECD Publishing. Retrieved from http://dx.doi.org/10.1787/9789264239555-en

Packard, R. (2013). *Education transformation: How K-12 online learning is bringing the greatest change to education in 100 years*. Hillsboro, OR: Beyond Words.

Papert, S. (1997). Why school reform Is Impossible. *The Journal of the Learning Sciences 6(4), p.417-427*.

Peterson, P. (2010). *Saving schools: From Horace Mann to virtual learning*. Cambridge, MA: Belknap Press of Harvard University Press.

Powell, A. (2013). *iNACOL's new learning models vision*. Vienna, VA: International Association for K-12 Online Learning. Retrieved from http://www.inacol.org/resource/inacols-new-learning-models-vision/

Powell, A., Kennedy, K., & Patrick, S. (2013). *Mean what you say: Defining and integrating personalized, blended and competency education*. Vienna, VA: International Association for K-12 Online Learning. Retrieved from http://www.inacol.org/resource/mean-what-you-say-defining-and-integrating-personalized-blended-and-competency-education/

Ravitch, D. (2011). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Ravitch, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. Vintage.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation, 11*(10), 1-13.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15*(1), 1-20.

Russell, M. (1999), Testing writing on computers: A follow-up study comparing performance on computer and on paper, *Educational Policy Analysis Archives,* 7(20), and Russell, M., & Haney, W. (1997), Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil, *Educational Policy Analysis Archives,* 5(1).

SAIT Polytechnic. (2012, February 28). Invigilation and security of examinations. Retrieved from http://www.sait.ca/Documents/About%20SAIT/Policies%20and%20Procedures/Academic%20Student/pdf/AC-3-3-1_InvigilationSecurityofExaminations.pdf

Sandeen, C. (2013). Assessment's place in the new MOOC world. *Research & Practice in Assessment, 8*, 5-12.

Seels, B. B., & Richey, R. C. (1994). *Instructional technology: The definition and domains of the field*. Washington, DC: Association for Educational Communications and Technology.

Seo, D.G. & DeJong, G. (2015). Comparability of online- and paper-based tests in a statewide assessment program. *Journal of Educational Computing Research. 52(1). p.88-113.*

Singer, N. (2015, April 5). Online test-takers feel anti-cheating software's uneasy glare. The New York Times.

Retrieved from http://www.nytimes.com/2015/04/06/technology/online-test-takers-feel-anti-cheating-softwares-uneasy-glare.html?_r=1

Stobart, G. (2008). Testing times, the uses and abuses of assessment. Abingdon: Routledge.

Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration. 12(3).* Retrieved from http://www.westga.edu/~distance/ojdla/fall123/stuber123.html.

Taylor, A. R. (2005). A future in the process of arrival: Using computer technologies for the assessment of student learning. Kelowna, British Columbia: Society for the Advancement of Excellence in Education. Retrieved from http://www.maxbell.org/sites/default/files/029.pdf.

Vander Ark, T. (2012). *Getting smart: How digital learning is changing the world*. San Francisco, CA: John Wiley & Sons, Inc.

Watson, G. & Sottile, J. (2010). Cheating in the Digital Age: Do Students Cheat More in Online Courses?.*Online Journal of Distance Learning Administration, 13*(1),