Value of Large-scale, Online Student Assessments

Dr. Richard Jones

The purpose of this article is to provide a rationale for conducting large-scale, student assessments by identifying the benefits of administering the assessments in an online format and illustrating their value to all levels of the education system and the wider public.

Why Administer Large-scale Student Assessments?

Large-scale student assessments have been a fixture in education for more than a century. During that time, there have been supporters and critics of such testing. Within Canada, supporters argue that large-scale provincial assessments provide an objective measure of the health of the education system and the only way to gauge areas for improvement, particularly at the school, school board, and provincial levels. Critics argue that: formative assessment of students is the only legitimate and useful form of assessment, that large-scale summative assessments provide no real value to the education system, and they only determine which students are good at test-taking.

There are many issues that are of concern to education stakeholders. For instance, teachers' grading practices are, to a large extent, subjective and inconsistent. A student receiving an A grade in one class might get a C grade in another. Teachers may have conscious or subconscious biases for or against certain students because of (for example) their behaviour and attitudes. Furthermore, grade inflation has been occurring throughout our education systems for decades. An article by Cosh (2023)[1] , with reference to Ontario students, states, "But the last vestiges of grading standards, already untethered to any form of mandatory province-wide diploma exam, went entirely out the window during COVID, as teachers and students alike suffered a morale crisis. The result is that there are now many discouraged, angry students with grades well into the 90s who can't worm their way into prestigious university programs as they were explicitly led to expect." The bottom line is that educators apply varying standards when assessing and evaluating their students.

In Canada, provincial ministries of education tend to transition through cycles in which large-scale assessment is implemented following stakeholder outcry for more and better information about the state of the provincial education system; after a period of time the tests go out of favour due to the adoption of the next fashionable education trend (for example, relying exclusively on the whole-language approach to reading, which postulates that reading will just come naturally to students, and phonics [the relationship between sounds of spoken language and letters that represent them] drops out of fashion), as well as pressures from teachers' unions that advocate for a sole focus on formative classroom assessment. Often, after several years without large-scale assessments, jurisdictions come to the realization that they do not have credible information regarding how well their education systems are performing, and this again results in calls from education stakeholders to return to large-scale assessments – and the cycle continues.

There is no doubt that formative assessment is critically important to student success. At the classroom level, information from this type of assessment can provide the student, teacher and parent/guardian with insights about the learner's strengths and areas which require attention. However, eventually, parents and the wider public demand greater accountability in the public education system, and in order to achieve that, valid and reliable information about student achievement is considered essential to gain an understanding of how the system is performing and for effective education system improvement planning.

For education authorities and stakeholders to gain insights into how students and the education system are doing, large-scale student assessments offer the best objective measures of the learning that students have attained for accountability and system improvement purposes. At the provincial and school division levels, large-scale assessment information, together with other types of information (for example, survey data) can inform ministries of education and school divisions about what aspects of the curriculum and

---

[1] Cosh, C. (June 21, 2023). "Grade inflation kneecapped Ontario students looking to go on to higher education." National Post. Access at nationalpost.com/opinion/grade-inflation-kneecapped-ontario-students-looking-to-go-on-to-higher-education/.

implementation strategies are working and where improvement emphases should be focused. In this way, jurisdictions can evaluate the effectiveness of the education program. In the absence of such information, no one knows.

If all-student provincial assessment is conducted, results are available for every student, and the data, together with all the additional information educators' have related to the student, can form the basis for assisting them to help students identify their strengths and areas where attention is required to help them improve their learning. Including students with disabilities and English-language learners in large-scale assessments allows the measurement of how well the system is working for them and where gaps in instructional approaches and opportunities exist.

Data from large-scale provincial assessment are one important source of information for school boards to use in developing improvement plans. Provincial assessment data provide consistent reference points that can be used to complement other sources of information about student learning at the classroom, school and school board levels. Combined, this information forms the basis for conversations about how resources can be used to benefit all learners. Furthermore, typically, school board leaders are required to publish annual education reports for their school communities. Results from large-scale provincial assessments can be important sources of data for these reports.

Results from large-scale student assessments can support the jurisdiction in identifying students' academic strengths and where attention and resources are required. Ministries of education can use findings from the assessments to develop learning resources, design educator professional development programs, provide targeted funding for specific initiatives and collaborate with school boards/school communities on local strategies to ensure students learn effectively.

Furthermore, if data from large-scale student assessments are made publicly available, researchers can conduct studies to examine trends that will further support educational improvement.

The appropriate use of large-scale assessment data can, over time, result in the evolution of a culture of evidence-based decision making at all levels of the education system.

Why Online Assessments?

It must be recognized that there are some positive aspects of paper-based assessments, for instance, students are generally familiar with this traditional mode of assessment, and technical issues, often associated with online assessments, are avoided. Considering these advantages, why should one consider online assessments?

There are myriad reasons why ministries of education have either switched to digital assessments or are considering transitioning all the functions of their assessment life cycle (i.e., authoring, registering, administering, scoring, analysing, and reporting) from paper to online formats as part of their modernizing initiatives.
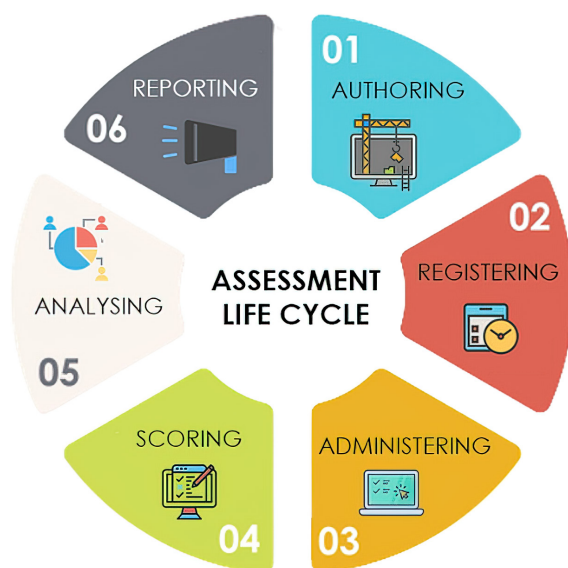

The following are several benefits of implementing online assessments.

Engagement

In recent years schools/classrooms have adopted technology-based tools for learning, and students are becoming increasingly tech-savvy preferring to do their work on computers as opposed to paper. Many believe that static, paper-based assessments are dull/boring; however, they tend to be more engaged with online assessments that contain a variety of item types and use media such as maps, diagrams, photographs, animation, video, and audio. Engagement is advanced even more when technology-enhanced items (TEIs), including game-like simulations, are used. In addition, students often experience test anxiety;

however, if they have opportunities to take online practice tests to become more comfortable with test content and presentation, anxiety can be reduced with the online format.
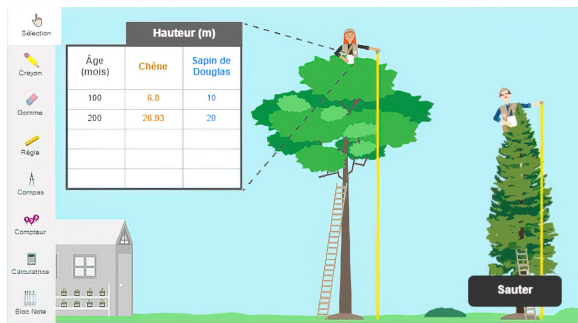
This is an example of a real-world, technology-enhanced item (TEI) used in a large-scale assessment to engage students during the assessment and gather deeper insight into key 21st century skills.

Authenticity

Authenticity in large-scale assessment is reflected by characteristics such as presenting students with real-world situations and having them demonstrate their knowledge, skills and abilities by applying their knowledge/understandings to complete complex tasks/solve problems. Unlike traditional paper-based assessments that typically emphasize the recall of facts and knowledge, online assessments can present students with realistic situations (via simulations for example) to gather information about their higher-order thinking and problem-solving skills.

Equity and Inclusion

All jurisdictions strive to include as many students as possible in their assessment programs. Paper-and-pencil assessments present barriers to some students requiring accommodations and special considerations. Equity and inclusion are important lenses through which most jurisdictions approach large-scale assessment. When assessments are offered online it is easier (compared with paper-based assessment) to adapt them to the needs of students who have, for example, visual, auditory and motor disabilities. Accessibility features include: the ability to change font size and background colour, magnification (zooming), the provision of text-to-speech and speech-to-text applications, screen readers and assistive technologies. Assessment administration via computer also facilitates adding extra time for students requiring this type of accommodation.

Item and Test Development

Compared with paper-based processes, working online makes it easier for an assessment organization to manage the tasks involved in developing items (questions) and creating assessments. For example, online workflows can be used to assign item development tasks, conduct reviews of draft items and provide approvals prior
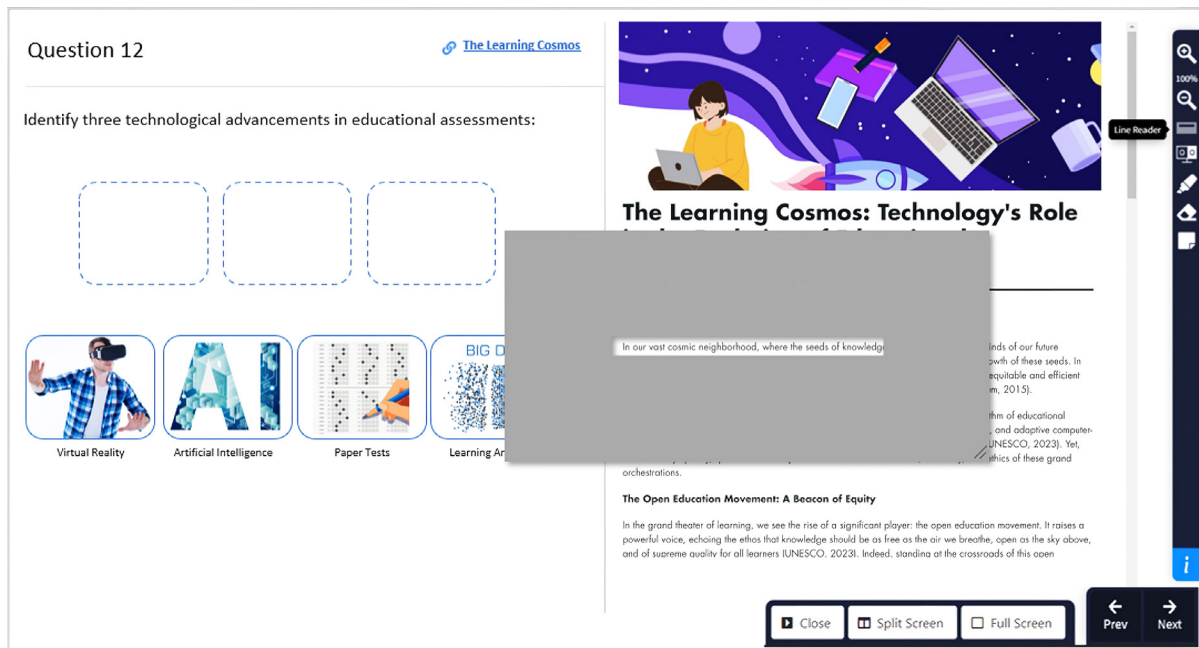
This is an example of an accommodation tool called a line reader that is used to guide students, who require the extra support, by displaying a single line at a time to read.

to adding them to an item bank (item development and reviews can be conducted electronically and remotely). Typically, all versions of an item are retained for easy reference. Once an appropriate bank of items has been developed (in terms of numbers, curriculum coverage and difficulty levels) online tools can be used to automate the creation of assessment forms. Online workflows track all steps in the development process and allow for accurate process documentation, so the organization has a clear picture of what has occurred.

Test Administration

Usually, paper-based assessments are administered at set times and locations because of the limitations resulting from their format. By definition, paper-based assessments involve printing, shipping and circulating assessment materials on paper, which can be a tedious process and be prone to error, such as misprints, missing or damaged material. Furthermore, following administration of paper-based assessments the handling of student responses on paper is cumbersome and subject to error. Online administration allows more flexibility for the scheduling and administration of assessments (multiple administrations, testing on demand and remote testing are possible). Another potential advantage of online assessments is that it allows for adaptive testing (adapting the content of the test to the knowledge/skill level of the examinee). Compared with paper-based assessment adaptive testing is both more efficient (best use of the examinees' time, as they respond to items closer to their true ability level) and effective (provides a more accurate estimate of examinees' true ability). Digital assessments provide test-takers with practice experience on the platform; test administrators have access to various supports (e.g., online, e-mail, telephone advice); and provided instructions are clear and straightforward, digital assessments significantly reduce the administrative burden for both the assessment organization, as well as teachers and administrators at the local school and school board levels.

Scoring

Paper-based scoring can be labour-intensive, inefficient and subject to error with regard to the management of scorers and related workflows, circulation and tracking of students' constructed responses, record keeping, data analytics, the use of scoring data for the supervision and moderation of scorers and security of assessment materials. With online assessments, machine-scorable items are auto-scored accurately, immediately upon submission of the assessment. For responses that require human marking, online scoring platforms allow for efficient, streamlined and standardized processes that include identifying roles (e.g., scoring supervisors, scorers, expert scorers) and permissions for their respective authorities, automated routing and workflows for all marking roles and quality assurance procedures. The use of online scoring systems saves the assessment organization a great deal of time and effort, while also enhancing the efficacy, accuracy and security of marking and results. If conducted remotely, online scoring also represents substantial cost savings compared with centralized, paper-based scoring.

Feedback/Reporting

Accurate, reliable, valid and timely results are essential to all assessment programs. Stakeholders must have confidence in the data, because typical purposes include accountability (to indicate how, for example, the education system is performing) and improvement (to identify students' needs, enhance instruction, establish and implement targeted interventions). Any question/doubt about the quality and integrity of the data/results may undermine efforts to use the data for both accountability and improvement. When assessment results are not released in a timely fashion their value is reduced for all stakeholders. In the case of large-scale, paper-based assessment, the time lapse between the administration of an assessment and examinees receiving results may be several weeks or even months. In the case of online assessments, machine-scorable items can be processed immediately and results made available very quickly (should the assessment organization wish to do so). With the more efficient online human scoring processes, results can be released more quickly than with the paper-and-pencil approach.

Furthermore, digital analysis and reporting tools can generate comprehensive, detailed reports more efficiently than is possible with paper-based assessment. Consequently, clear, thorough and timely results for students can be made available to provide useful feedback on how they are doing and where improvements are needed (for formative assessment) and for education systems, so that stakeholders can understand the strengths and challenges of the systems (accountability).

Data/Analytics

Another benefit of online assessments over the paper approach is that additional data are available to provide a fuller appreciation of the student's effort, approach and progress in responding to assessment items/tasks. For example, online assessments can identify for each student, and by item, performance data such as the amount of time on task, the sequencing of steps in responding to a challenge, items flagged to return to later and changes to responses. The teacher, analyzing these types of data/information, can determine if there are issues/problems experienced by individual students or are common to groups of them. Such process-data analytics can be used by the instructor to help students improve their learning and/or for the teacher to plan for future instruction.

Enhanced Security

Online assessment provides for greater security compared with paper-based assessment, because all examinee information, assessments and results are stored electronically. Only those people who are authorized to access given data/information are able to do so. Access management is much more challenging with paper-based assessments.

Scalability and Flexibility

Scalability and flexibility have increased with the advent of online assessment. In the case of paper-based assessments, administration occurs in physical sites such as test centres and schools. With online assessment, individual students can access assessments securely via the Internet (recently this has become a common feature as a result of COVID-19 and the closing of schools and/or the option of online learning). In remote assessment situations, assessment organizations must assess the risks related to the integrity and security of results and/or assessment content. Running secure, remote online assessments may necessitate remote proctoring alternatives.

Environment/Sustainability

Electronic assessment is more environmentally friendly and has a much smaller carbon footprint than the paper-and-pencil approach, because it involves much less paper, printing and transportation.
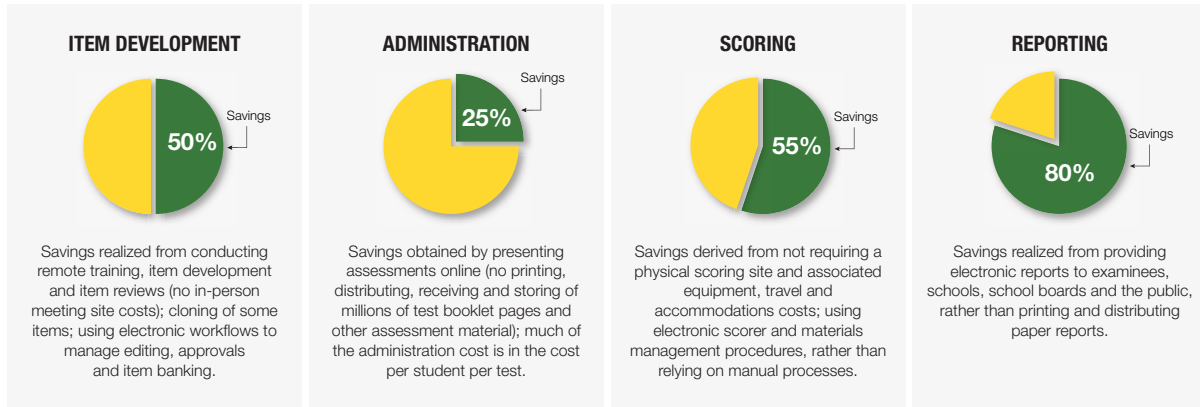
Cost Effectiveness

Throughout the entire assessment process, online assessments have the potential for cost savings as compared with paper-based assessments. For example, assessments on paper include manual processes related to: item writing, item banking, field testing, test form creation, printing, distribution (transportation), securing invigilators, administering the assessment, return of test materials (transportation), scanning/data entry of machine-scorable responses, manual processing and tracking of scoring material, storage (and eventual disposal) of assessment materials and printing and distribution of reports (not to mention staff overhead to manage all processes). Because online assessments vary so much in context (for example, number of administrations, numbers of students, number and type of items, online versus offline, distributed versus centralized scoring, use of proctors) it is difficult to provide a generalizable percentage of potential cost savings of online versus paper-based assessments. It is possible, however, to estimate cost savings associated with specific assessment scenarios.

A recent article by Jones (2021)[2] provides an example of potential cost savings given a specific assessment scenario. In this scenario, after set-up costs, the total estimated percentage savings of online versus paper-based assessment, for the test components examined, was approximately 50%. (Additional costs may be incurred depending upon client requirements and the service provider's level of involvement in activities such as item development, scoring and psychometrics.)

---

[2] Jones, R.M. (June 6, 2021). What are the Cost Savings with Computer-Based Assessment?
Access at rmjassessment.com/2021/06/06/what-are-the-cost-savings-with-computer-based-assessment/

The following charts highlight the estimated cost savings by using online assessments as compared to paper-based assessments.
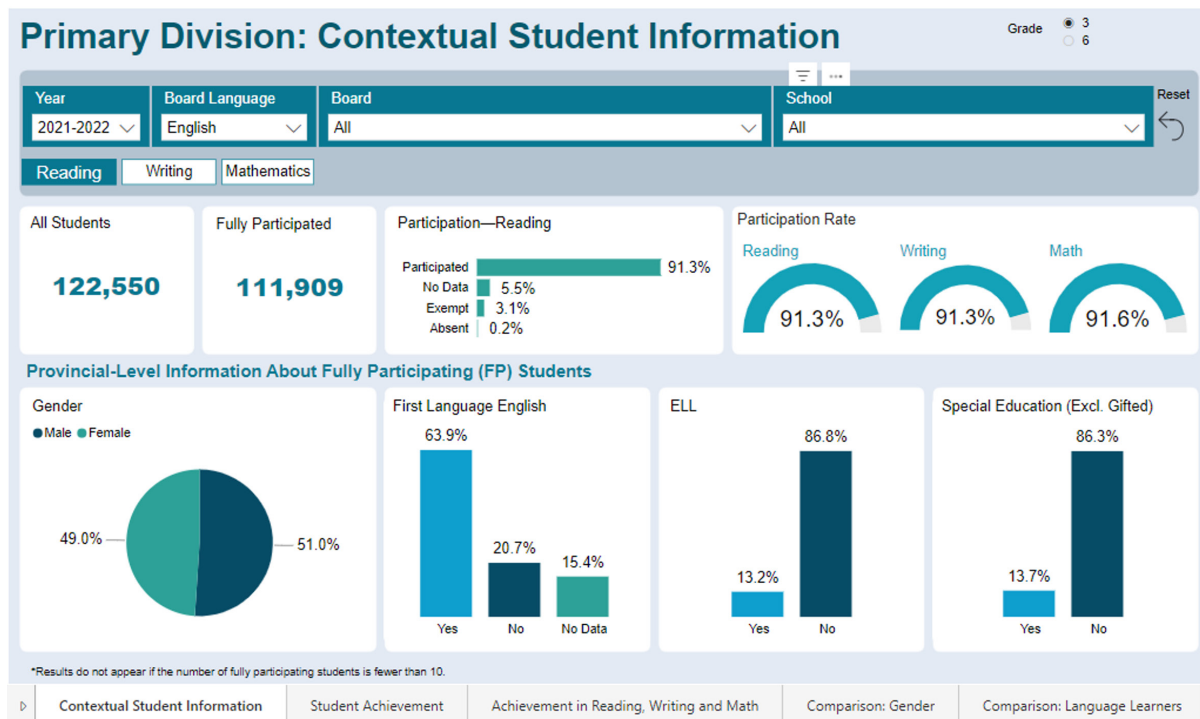
ITEM DEVELOPMENT	ADMINISTRATION	SCORING	REPORTING
50% Savings	25% Savings	55% Savings	80% Savings
Savings realized from conducting remote training, item development and item reviews (no in-person meeting site costs); cloning of some items; using electronic workflows to manage editing, approvals and item banking.	Savings obtained by presenting assessments online (no printing, distributing, receiving and storing of millions of test booklet pages and other assessment material); much of the administration cost is in the cost per student per test.	Savings derived from not requiring a physical scoring site and associated equipment, travel and accommodations costs; using electronic scorer and materials management procedures, rather than relying on manual processes.	Savings realized from providing electronic reports to examinees, schools, school boards and the public, rather than printing and distributing paper reports.

Case Study: Administering Large-scale, Online Assessments

The Education Quality and Accountability Office (EQAO) is an arm's-length agency of the Ontario provincial government that has been conducting large-scale student assessments in the province for more then two decades. The assessment program was offered in a paper-based format for many years; however, the agency has transitioned to online assessments beginning in 2022.[3]

The provincial assessments are meant to measure students' achievement at key stages of their learning as linked to the expectations of the Ontario Curriculum. The assessments include: Primary Division (Grade 3) and Junior Division (Grade 6) Assessments of Reading, Writing and Mathematics; Grade 9 Assessment of Mathematics; and, Ontario Secondary School Literacy Test (typically administered in Grade 10), which measures whether students are meeting the minimum standard for literacy (Reading and Writing) up to the end of Grade 9.

The following interactive dashboard, available on the EQAO website, provides access to data on students' achievement and learning perceptions.

Primary Division: Contextual Student Information

Grade ● 3 ○ 6

Year	Board Language	Board	School	Reset
2021-2022	English	All	All	

Reading | Writing | Mathematics

All Students	Fully Participated	Participation—Reading	Participation Rate
122,550	111,909	Participated 91.3%	Reading 91.3%	Writing 91.3%	Math 91.6%
		No Data 5.5%		
		Exempt 3.1%		
		Absent 0.2%		

Provincial-Level Information About Fully Participating (FP) Students

Gender
● Male ● Female
49.0% / 51.0%

First Language English
63.9% Yes
20.7% No
15.4% No Data

ELL
13.2% Yes
86.8% No

Special Education (Excl. Gifted)
13.7% Yes
86.3% No

*Results do not appear if the number of fully participating students is fewer than 10.

▷ Contextual Student Information | Student Achievement | Achievement in Reading, Writing and Math | Comparison: Gender | Comparison: Language Learners

---

[3] Since 2019, EQAO has engaged their assessment technology partner, Vretta, to collaborate with the agency to support the design, implementation, and operation of modernized, large-scale e-assessments.

Assessment data are used at the school, school board, and provincial levels for accountability and improvement-planning purposes. In addition to school, school board and provincial reports, the EQAO has developed (with input from provincial educators) interactive dashboards that provide ready access to data on student achievement and learning perceptions (questionnaire results). The platform is another way that the agency provides information in support of students' learning.

The following statement, provided by Dr. Pino Buffone, Director of Education, Renfrew County District School Board (RCDSB), describes how provincial assessment data has contributed to school and system improvement in their board.

"Now more than ever, provincial assessment data provide a valid and reliable, consistent and stable source of data for improvement planning processes at the school and system levels. This statement is particularly pertinent considering the volatile, uncertain, complex and ambiguous nature of our times, including the disruption to learning that occurred during the pandemic period, and both the short- and long-term consequences of that significant disruption on students, staff and school communities.

As one of the primary purposes of assessment is to guide further instruction, for educational leaders of schools and district school boards there are three key benefits to the data garnered annually through provincial assessments:

- revision and/or refinement of frameworks for learning developed and implemented in teaching and learning environments;

- coordination and delivery of priority areas for professional learning sessions with all educators; and,

- reflection upon and evolution of metrics used to measure students' progress over time.

The provincial assessment data, along with other qualitative and quantitative data sources related to achievement, equity and well-being, form pieces of a puzzle that, in effect, allow for the most reasoned and reasonable decision making possible each and every school year, especially as it pertains to the allocation of finite resources (human, material and/or fiscal).

At the RCDSB, the provincial assessment data provided by the EQAO allow school-based leaders and central staff to carefully examine strengths to celebrate, areas for growth, as well as next steps for consideration with respect to students' knowledge and understanding of concepts and skills outlined in curriculum policy documents provided by the Ontario Ministry of Education (EDU). The educators and/or representatives of these three stakeholders – EDU, EQAO and RCDSB – work collaboratively in striving to address the needs of all students, county-wide, through established improvement planning processes, annually.

More specifically, the reports provided by the EQAO highlight students' progress to date with respect to the expectations of the Curriculum. The 'strand and skill reports' for mathematics, for instance, allow educators to identify students' strengths and needs in different areas of the mathematics curriculum. Staff at the RCDSB utilize these key learnings to revise and/or refine the District's framework for 'Inspired Learning' – a framework to guide instruction that emphasizes the dynamic interplay between core skills of literacy and numeracy, global competencies of character, citizenship, collaboration, communication, creativity and critical thinking with the 'big ideas' of the Curriculum. In fact, the provincial achievement data for reading and writing have most recently supported the start of the development of an additional, more in-depth framework for 'Structured Literacy' that emphasizes the explicit instruction of oral language, phonemic awareness, phonics and spelling, fluency, vocabulary and morphology, syntax, and text comprehension and written expression.

The provincial assessment data also allow central staff of the District to coordinate and deliver professional learning sessions for all educators, including instructional strategies specific to areas of growth for students, such as the growing criticality of algebra in mathematics, for example, as seen in the most recent assessment results. These specific sessions for educators of the intermediate grades accompany the broader professional development in pedagogical areas such as Universal Design for Learning (UDL), Differentiated Instruction (DI) and Culturally-Relevant and Responsive Pedagogy (CRRP) at the primary, junior, intermediate and senior levels.

Further, the outreach sessions provided by the EQAO staff for educators of the RCDSB have allowed for the metrics associated with achievement to evolve over time, including the enhanced use of qualitative (attitudinal) data, as well as quantitative data, along with the integration of demographic data and measures of wellness. The OurSCHOOL Survey Suite, provided through The Learning Bar, as an exemplar for well-being, allows for a deeper perspective on the multitude of factors that influence students' progress over time.

The benefits of the data gathered through provincial assessments shared above, actioned through improvement planning processes at the school and system levels, ensure that provincial assessments continue to serve as a veritable value-add for the education sector, K-12."

Conclusion

Well-designed large-scale student assessments with clear purposes provide important sources of data/information for all levels of the education system (as well as the wider public) for improvement planning and accountability. Online assessments provide many advantages over the paper-based approach, and can provide more and better data, in a timely fashion, to all education stakeholders.

About the Author

Dr. Richard (Rick) Jones has extensive experience in the fields of large-scale educational assessment and program evaluation. He has worked in the assessment and evaluation field for more than 30 years. Prior to founding RMJ Assessment, his assessment consulting company, Rick held senior leadership positions with the Education Quality and Accountability Office (EQAO) in Ontario, as well as the Saskatchewan and British Columbia Ministries of Education. In these roles, he was responsible for initiatives related to student, program and curriculum evaluation; education quality indicators; school and school board improvement planning; school accreditation; and provincial, national and international assessment.

Rick began his career as an educator at the elementary, secondary and post-secondary levels. Subsequently, he was a researcher and senior manager for a multi-national corporation delivering consulting services in the Middle East.

Feel free to reach out to Rick at richard.jones@rmjassessment.com (or on LinkedIn) to learn more about the best practices in transitioning to an online assessment environment.