Executive Summary

What Is the Quality of EQAO Assessments?

Excellent!

W. Todd Rogers,
Scholar-in-Residence, EQAO,
University of Alberta

June 2013
EQAO’s Board of Directors

Dr. Brian L. Desbiens,
Chair
June 1, 2005–March 16, 2014

Jerry Ponikvar,
Vice-Chair
July 4, 1997–May 16, 2014

Roland Boudreau
June 2, 2011–June 1, 2014

Hélène Chayer
June 25, 2008–May 11, 2013

Dave Cooke
June 25, 2008–May 11, 2013

Dr. Dieudonné Detchou

Dr. Elizabeth (Lee) Ford-Jones
June 2, 2011–June 1, 2014

Abirami Jeyaratnam
October 3, 2012–October 2, 2014

Dr. Bette M. Stephenson,
Director Emeritus

EQAO’s Executive Team

Marguerite Jackson,
Chief Executive Officer

Marianne Mazzorato,
Chief Assessment Officer

Richard Jones,
Director, Assessment and Reporting

Michael Kozlow,
Director, Data and Support Services

François Lavictoire,
Executive Coordinator, French Language

Mary Low,
Executive Coordinator, Operations

Tony Saini,
Director, Corporate and Public Affairs (Acting)
What Is the Quality of EQAO Assessments?

Excellent!

W. Todd Rogers,
Scholar-in-Residence,
EQAO, University of Alberta

June 2013

Executive Summary
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>How Good Is EQAO at Doing What It Was Charged to Do? Very Good!</td>
<td>4</td>
</tr>
<tr>
<td>How Are Items Developed?</td>
<td>6</td>
</tr>
<tr>
<td>How Is Field Testing Conducted?</td>
<td>8</td>
</tr>
<tr>
<td>How Are Operational Items Selected?</td>
<td>9</td>
</tr>
<tr>
<td>How Are EQAO’s Questionnaires Constructed?</td>
<td>10</td>
</tr>
<tr>
<td>How Are Test Booklets and Questionnaires Administered?</td>
<td>12</td>
</tr>
<tr>
<td>How Are Student Responses Scored?</td>
<td>14</td>
</tr>
<tr>
<td>How Are Scored Student Responses Analyzed?</td>
<td>18</td>
</tr>
<tr>
<td>How Are Cut Scores Established?</td>
<td>20</td>
</tr>
<tr>
<td>What Steps Does EQAO Take to Ensure the Quality of Student Results?</td>
<td>22</td>
</tr>
<tr>
<td>How Are EQAO Results Reported and Used?</td>
<td>23</td>
</tr>
<tr>
<td>Summary</td>
<td>25</td>
</tr>
<tr>
<td>References</td>
<td>26</td>
</tr>
</tbody>
</table>
The processes EQAO uses to develop and score its assessments conform to well-recognized standards of educational testing and provide valid and valuable information for Ontario’s education context. The agency’s steadfast commitment to ensuring the tests reflect the curriculum accurately, provide relevant results, and are comparable from year to year means Ontarians can be confident the results are a reliable, independent measure of student achievement over time.

Dany Laveault, University of Ottawa
In response to the Royal Commission on Learning’s call for province-wide testing in 1994, the Ontario Government created the Education Quality and Accountability Office (EQAO) as an independent Crown agency in 1996. The Commission proposed that system-wide testing be conducted in Ontario to check student learning at a few critical transition points and to assure the public that there was at least one common measure of achievement for all students at these points.

EQAO develops, administers, and scores province-wide assessments of student achievement each year in both of Canada’s official languages. The results, including the change in student performance from one year to the next, are reported to the Ontario government, school boards, schools, students, parents/guardians, and the general public. EQAO’s province-wide assessments measure the cumulative knowledge and skills students have acquired in three subject areas at four key points of their journey through school:

- The Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions measure how well Grades 3 and 6 students have met the reading, writing and mathematics learner expectations outlined in *The Ontario Curriculum, Grades 1–8: Language* (revised 2006) and *The Ontario Curriculum, Grades 1–8: Mathematics* (revised 2005).

- The academic and applied versions of the Grade 9 Assessment of Mathematics measure how well students have met the learner expectations for Grade 9 outlined in *The Ontario Curriculum, Grades 9 and 10: Mathematics* (revised 2005) for the academic and applied courses.

- The Ontario Secondary School Literacy Test (OSSLT) assesses Grade 10 students’ literacy skills based on the reading and writing learner expectations across all subjects in *The Ontario Curriculum*, up to the end of Grade 9.

Approximately 1 100 000 tests are administered to about 580 000 students in approximately 4300 English-language and French-language schools in the province each year.
EQAO develops items, scores student responses, analyzes the scores, and reports the results in ways that meet standards for educational testing; provides trustworthy and useful information for the Ontario educational system; and helps and encourages school officials to work with the results for their school board and schools to enhance student learning. EQAO procedures are in keeping with both the Principles for Fair Student Assessment Practices for Education in Canada (1993) and the Standards for Educational and Psychological Testing (American Educational Research Association, National Council on Measurement in Education, & American Psychological Association, 1999).

EQAO has an Assessment Advisory Committee comprising school board members, principals, teachers, parents, and students; and a Psychometric Expert Panel, the members of which are recognized national and international experts in large-scale testing. The agency retains scholars-in-residence who act as senior measurement specialists. Seeking information from all of the above-mentioned groups on a regular basis makes EQAO unique in Canada, and helps ensure that EQAO assessment results are meaningful and can be trusted and used with confidence.
Before proceeding further, it is necessary to point out that there is potential for error whenever we measure. For example, how many of us have cut a piece of wood either too short or too long or hemmed a dress either too long or too short? What did we do just before we made the cut? We measured, but made an error. There are actually two kinds of measurement error. One type is random error, which we would get if we had many people measure the same piece of wood. Not everyone’s measurement would be exactly the same. The second type is constant or systematic error, which we would get if the tape measure began at 5 centimetres and not at 0 centimetres. All who used the tape measure would measure the length about 5 centimetres too long (for example, 96 centimetres instead of 91 centimetres). Random error affects the precision or reliability of the observed measurements. Constant error affects the accuracy of the measurements, which leads to an invalid or incorrect interpretation of the observed measurements. It is important to note that EQAO employs best practices that minimize both types of errors. That is, the results derived from the tests administered each year are precise (reliable) and can be accurately (validly) interpreted in terms of what students in Ontario know and can do.
Executive Summary—What Is the Quality of EQAO Assessments? Excellent!

New items are developed and reviewed to ensure proper coverage of the *Ontario Curriculum* expectations and that the difficulty of the assessments is similar across years, thereby allowing for the measurement of change in performance across years.

For each assessment, an item-writing committee consisting of 10 to 20 Ontario teachers and principals is established. These educators are fully trained to develop stand-alone multiple-choice items and multiple-choice items that refer to a reading selection, chart, or a data table. They also are trained to write prompts for the open-response items as well as scoring rubrics used to score students’ responses to these items. The items must be clearly referenced to the learning expectations found in *The Ontario Curriculum* for the assessment, are age and grade appropriate, and contain accurate information that does not rely on students possessing subject-specific content knowledge beyond the grade level for the assessment or on students possessing knowledge beyond a reasonable level of personal experience. Further, the items allow the widest range of students with exceptionalities (e.g., first-language learners, students with special education needs) to respond.
EQAO assessment specialists carefully review the new items and scoring rubrics to ensure that each item is stated clearly and completely, there is one correct answer for each multiple-choice item, each scoring rubric clearly reflects the range of student responses to the corresponding prompt, and each item is clearly referenced to the correct learner expectation. Revisions are made as needed. Once this is completed, members of the item-writing committee try out revised versions of their own items with their own students in their classes. They work with their students individually and have each student think out loud so that they can determine whether the student understands each item and how the student got his or her answer. The assessment specialists revise the items as needed using the results from the item tryouts.

The items are then reviewed by two independent committees. The Assessment Development Committee, comprising 8 to 12 teachers and principals who have the same qualifications as the item writers, review and discuss all new items for each assessment to ensure each item is technically correct. The Sensitivity Committee, with between 20 and 24 members who have knowledge and expertise in the areas of diversity, culture, second language acquisition, and special education, independently review and then discuss all new items to ensure that each is fair, non-offensive, and accessible to all students. Lastly, the assessment specialists revise the items using the suggestions provided by the two committees.

Clearly, EQAO employs current best practice for item development. Evidence-based design and universality are stressed, which allows assessment of the widest possible range of students. EQAO’s detailed procedures for item construction, item tryouts, and revising and reviewing items produce the most effective items possible, which is clearly commendable.
Except for new writing prompts that require long answers, the final versions of the new items are field tested each year as part of the regular, or operational, assessments.\(^1\)

Different sets of field-test items are inserted among the operational items in the current year’s assessment booklets. Usually, there are about 20 sets of field-test items for each assessment, resulting in about 20 different student booklets for each assessment. Each booklet contains the same operational items but a different set of field-test items. The operational items, which were field tested the year before, match the test specifications and have desirable measurement properties. Student scores are determined only from the operational items and not from the field-test items.

Since the field-test items are similar to the operational items, students do not know which items are operational items and which items are field-test items. This solves the problem of students not trying hard enough to answer the field-test items. Also, the booklets for an assessment are distributed in a way that ensures the sample for each booklet represents the entire provincial student population. Thus the samples for the field-test items contained in the booklets also represent the entire provincial student population. Therefore the analysis of the student responses to the field-test items provide trustworthy results.

---

\(^1\) New writing prompts that require long answers are not field tested in the administration of the operational assessments, due to the time needed for students to write their responses. Instead, they are pilot tested in a separate process. The purpose of pilot testing these items is to ensure that the prompts generate student responses at each score point in the scoring rubrics for these items. While the number of pilot-test schools is small, students in the pilot-test sample clearly represent the range of student responses in the province.
How Are Operational Items Selected?

Operational items for the current assessment are selected from the items that were field tested the previous year. The items are selected by EQAO assessment specialists and analysts in a back-and-forth manner.

They use the field-test data and select items that fit what is called the target test information function for the assessment. Using a target test information function ensures a high degree of accuracy and precision when placing students in the achievement categories (e.g., Level 1, 2, 3, or 4 for the primary, junior, and Grade 9 assessments or pass/do not pass for the OSSLT). The same target test information function for the assessment is used across years so as to be able to measure the change in student performance between one year and the next. The assessment specialists and analysts work together to create an operational form that best fits the target test information function, represents the learner expectations for the assessment and has a good range of item difficulty. They also ensure that the level of difficulty of the new operational form is comparable to that of the previous year’s operational form, and the scores around each cut score (the score used to define the low end of an achievement category) are sufficiently precise to allow accurate and precise placement of students in the achievement categories for the assessment. The Psychometric Expert Panel reviews the summary statistics for the proposed operational form and the previous year’s operational form to ensure that the forms satisfy the conditions for equating, a procedure for linking assessments across years.

Once the proposed operational form has been approved, directions are added. A unique student bar code is added to each student booklet so that students’ responses can be accurately tracked through administration, scoring, analysis, and reporting. Operational items that have been used in previous assessments are available at www.eqao.com for parents to see the kinds of items administered to their children.
EQAO develops Student Questionnaires for all of its assessments, Teacher Questionnaires for all but the OSSLT, and Principal Questionnaires for the primary and junior assessments. The information obtained from these questionnaires is used to provide a context in which to interpret the assessment results and is used in research designed to identify factors that influence student learning.

EQAO is one of the few testing agencies that administers questionnaires and that has an active, ongoing research program designed to investigate ways to improve student learning and the assessments. The first step is to conduct a literature review to identify factors that have been shown to influence student learning. Items are then created and assembled by a committee of EQAO assessment specialists and analysts, taking into account the amount of time available for students, teachers, or principals to respond to their questionnaires. Two to three outside researchers in the area of school improvement review the questionnaires to determine the breadth of coverage; item clarity; and the way students, teachers, and principals will record their responses. The questionnaires are then pilot tested on samples of students, teachers, and principals. The questionnaires are revised as needed and then reviewed and approved by EQAO’s Research Priorities Committee.
EQAO is one of the few testing agencies that administers questionnaires and that has an active, ongoing research program designed to investigate ways to improve student learning and the assessments.
How Are Test Booklets and Questionnaires Administered?

The primary and junior assessments of reading, writing, and mathematics are administered over a two-week period at the end of May and/or the beginning of June. Each assessment is divided into two parts, which are administered on separate days. One hour is allotted for each part, and additional time may be provided immediately following the allotted hour for students who need more time. The Student Questionnaire is administered after students have completed the second part of the assessment or on a separate day.

The Grade 9 Assessment of Mathematics is administered over a two-week period toward the end of the first semester (January) and a two-week period toward the end the second semester (June). Students in a full-year course write it with the students in a second-semester course. The academic and applied versions are divided into two parts, which are administered on separate days. While the estimated time to complete each part is 40 minutes, one hour is allowed for each part to accommodate students who may require extra time. The Student Questionnaire is administered after completion of the second part of the assessment or on a separate day.

The OSSLT is administered during a single morning toward the end of March or the beginning of April. The date and time are the same for all students. The OSSLT is in two parts. Students are given 75 minutes to complete each part, with a 15-minute break between the two parts. The Student Questionnaire is administered after students have completed the second part.

Teachers and principals complete their questionnaires either while the students are writing (if they are not supervising the students) or at some other time before the materials for the school are returned to EQAO.
A very detailed process is followed by EQAO, principals, and teachers to ensure that the assessments are administered in the same way across all the schools in the province. This ensures that all students have the same opportunity to show what they know and can do. EQAO provides a Teacher Bulletin in early fall that outlines what will be assessed, what resource materials are available on EQAO’s Web site (www.eqao.com), and the steps principals and teachers must follow to ensure a fair and complete administration. EQAO also provides a more detailed Administration Guide and the Guide for Accommodations, Special Provisions, and Exemptions to each school well before the assessment. The Administration Guide outlines in detail what principals and teachers must do before, during, and after the assessment is administered. The Guide for Accommodations, Special Provisions, and Exemptions provides information and directions to help principals and teachers make decisions about accommodations for students with special education needs, special provisions for students whose proficiency in the language of instruction is inadequate for them to respond in the anticipated manner, exemptions (primary, junior, and OSSLT only) and deferrals (OSSLT only).
Executive Summary—What Is the Quality of EQAO Assessments? Excellent!

Student responses to *multiple-choice items* for the primary assessment are manually entered into a computer file and are completely verified. Students’ responses to the multiple-choice items on the other assessments are read into the computer using an optical reader. EQAO conducts a quality assurance check to ensure that student responses are read with 99.5 percent accuracy.

Each year, the total number of scorers for the *open-response items* is about 3700: 1600 for the primary and junior assessments, 300 for the Grade 9 Assessment of Mathematics, and 1800 for the OSSLT. The scoring takes place in a large building with movable walls to create multiple scoring rooms. The open-response items for the OSSLT are scored over a three-week period in April; the open-response items for the other assessments are scored during the first two weeks of July.

The key tools for scoring open-response items are item-specific scoring rubrics and anchor papers. Item-specific scoring rubrics are developed from generic rubrics, which are constant across the years to facilitate linking the assessments across years. The scoring rubrics have three, four, or six ordered score codes; the highest code represents the best performance and the lowest code represents the poorest performance. A committee of eight to 25 Ontario principals and teachers identifies potential anchor papers, as well as validity papers, training papers, and qualifying-test papers for each assessment. *Anchor papers* are used to clarify or “anchor” each of the scoring categories in a scoring rubric. *Validity papers* are used during scoring to ensure that the scorers are accurately scoring the student responses. *Training papers* are used to train the scorers.
The *qualifying-test papers* are used to assess scoring leaders, scoring supervisors, and scorers to see if they are ready to begin scoring.

One long-writing item, two short reading or two short writing items, or four open-response mathematics items are scored in a scoring room under the leadership of a scoring leader, assisted by one or more scoring supervisors. Scoring leaders have subject expertise, are effective teachers of adults, and have previously been scorers. They train the scoring supervisors and scorers in their rooms so that the supervisors and scorers align their thinking and judgment to the procedures and scoring materials for the open-response item or items being scored in their room. They also oversee the scoring, ensure that the scoring materials are applied consistently, resolve issues that arise during scoring, and review and analyze daily scoring reports to ensure that the scoring in their room is accurate and precise.
Scoring supervisors are selected from a pool of experienced and proficient EQAO scorers. They assist the scoring leaders during scoring by supervising a “pod” of scorers in a scoring room. Scoring supervisors may also be asked to retrain individual scorers when necessary. Scoring leaders and scoring supervisors must take the qualifying test and attain at least an 80 percent exact and a 100 percent exact-plus-adjacent agreement with the assigned scores for the qualifying papers included in the test. “Exact agreement” means that the scores assigned to a student’s response by two scorers are exactly the same. “Exact plus adjacent” means that there is a difference of one point between the scores assigned to the student’s response by two scorers.

Scorers are current teachers and principals, retired teachers and principals, prospective teachers and, in the case of the OSSLT, individuals with at least one university degree. They are trained by their scoring leader and must attain at least a 70 percent exact match on the items included in the qualifying test. Scorers who fail the qualifying test the first time undergo further training and write a second qualifying test. Scorers who fail to pass the second qualifying test are dismissed.

Student responses to the open-response OSSLT items are scored independently by two different scorers. If the two scores agree exactly (e.g., Level 4 and Level 4), then that score is assigned to the student. If the two scores are adjacent (e.g., 3 and 4), then the higher score for reading and short writing tasks or the average of the two scores for the two long-writing tasks is assigned to the student. If the two scores are non-adjacent (e.g., 2 and 4), then the response is rescored by an expert scorer to determine the correct score for the student. Student responses to the open-response items included in the primary, junior, and Grade 9 assessments are scored once. The reason for this difference in scoring procedures for the different assessments is that passing the OSSLT is a graduation requirement, while student scores on the other assessments do not affect promotion from one grade to the next or high school graduation. To eliminate any possible
changes in scoring across two years and to ensure the consistency of provincial standards, scorers rescore the responses to the previous year’s open-response field-test items that are used in the current year’s operational form.

Scorers are monitored on a daily basis to ensure that they are scoring accurately and consistently. Scoring supervisors and leaders identify responses scored the previous day that have scoring issues. Scorers have to rescore these calibration papers to ensure that they are all scoring in the same way. Validity papers are scored daily to ensure that the scorers are scoring accurately. Scoring consistency, or interrater reliability, for the OSSLT is assessed by examining the agreement between the scores assigned by the two independent scorers. Scoring consistency for the primary, junior, and Grade 9 assessments is assessed by examining the agreement between the scores assigned by the scorers to the other open-response items in the booklets containing the validity papers. If an individual or group of scorers experienced difficulty scoring the calibration paper, the validity papers, or is scoring inconsistently, then the scoring leader or supervisor addresses the issue with the individual scorer or group of scorers.

EQAO’s scoring targets for accuracy were exceeded for all the open-response items scored in 2012 and were generally met in previous years (visit www.eqao.com to find the Technical Report for each assessment year). These results reflect the rigour of the scoring process. Parents, students, and teachers can be confident that all students receive accurate and precise scores.

During scoring, a scorer may find a student’s response to an open-response item that contains evidence that the student may be at risk (e.g., a response states or implies threats of violence to oneself or others, or possible abuse or neglect). In such cases, scorers, in consultation with the scoring site manager, are legally required to inform the Children’s Aid Society and provide them with a copy of the student’s response. Obscene, racist, or sexist content in a student’s response is reviewed by EQAO staff to determine whether or not the school or school board should be contacted. If there is any evidence in a completed assessment that may indicate some form of irregularity (e.g., many changed answers, teacher interference), EQAO staff review the booklet to determine whether or not the school should be notified. If cheating is confirmed, the student gets no score on the assessment.
The analysis is completed in three stages.

1 First, an item analysis of the current operational form is conducted to determine the difficulty and discrimination of each item. Item difficulty indicates how easy or difficult an item is, and item discrimination describes how well the item separates students who perform well from students who perform less well. There should be a range of difficulty from easy to hard, the discrimination should be at least 0.20, and the difficulty and discrimination of an item should differ very little between the operational form and when it was field tested. The difficulty and discrimination for all but one or two items in each assessment meet these criteria each year.

2 Second, the characteristics of the operational form are examined. These characteristics include average or mean score; range, or variability of the scores; and reliability of the test. Reliability is an indication of how consistently the items measure what the test is designed to measure. Higher reliability coefficients indicate lower random measurement error and higher precision. The reliabilities of EQAO’s assessments are all acceptably high (i.e., greater than or equal to 0.80).

3 Third, the scores on the current and the previous assessments are equated. The purpose of equating is to ensure that EQAO can make valid comparisons of results over years. While EQAO constructs the assessments to be as similar as possible in content and difficulty, the difficulty of the assessments typically differs somewhat. Equating adjusts for differences in difficulty, but not actual differences in student performance. To equate the achievement results for two groups of students (e.g., primary students in 2011 and primary students in 2012) it is necessary to include some common items in two consecutive assessments, so that both groups respond to the same items.
With the exception of the long-response writing items, all items in the current operational assessment were field tested in the previous year. Therefore, all students in the current year and provincially representative samples of students in the previous year responded to these items. The fact that students in both years responded to all items except the long-response writing items makes EQAO’s equating very strong. Further, the equating process\(^2\) used is acceptable and has been verified by EQAO researchers (Xie, 2007; Pang, Madera, Radwan, & Zhang, 2010). Consequently, the interpretation of change between two years is trustworthy. Thus, any changes in student assessment results from one year to the next are attributed to changes in student learning and not to changes in difficulty in the operational assessment.

\(^2\) The full process is called forward-fixed common-item parameter (FCIP) equating.
Executive Summary—What Is the Quality of EQAO Assessments? Excellent!

EQAO ensures that the performance standards and cut scores for each assessment are carefully developed. A performance standard is a clear description of what students at that level know and can do at each level of performance. Cut scores are used to delineate levels of performance identified in the performance standards. Students who score below the cut score should not be able to correctly answer most of the items associated with performance above the cut score.

Performance standards are established and cut scores are set when an assessment program first begins and they are re-established and reset each time there is a curriculum change or a change in the way the assessment is conducted. EQAO also reviews the performance standards periodically to ensure that the standards are accurate descriptions of what students know and can do and that the cut scores correspond to what students know and can do.

Ontario teachers and principals serve on panels that establish performance standards and set cut scores for EQAO’s assessments. One panel establishes the standards and a second panel sets the cut scores, with some overlap between the two panels. They have strong knowledge about the students to be assessed, familiarity with The Ontario Curriculum, and experience in writing instructional or assessment materials for students. Each panel is led by an outside expert and the assessment specialist for the assessment.
To establish the performance standards, five to 10 teachers and principals review a sample of 20 tests selected from the population of student booklets. After a training session, each panellist independently sorts the booklets into two piles. In the case of the primary, junior, and Grade 9 assessments, the two piles are sorted into two more piles to form four piles altogether. They review their final piles to see if they wish to make any changes. Two or three papers that all the panellists placed in the same pile are selected. Each panellist then independently describes what qualities were evident in these papers. This process is repeated for all the piles. These descriptions are then summarized by the outside expert. Examples of performance standards can be found by going to www.eqao.com and clicking on “Educator Resources.”

A panel of approximately 20 trained teachers and principals, led by an outside expert and the assessment specialist for the assessment, set the cut scores that separate the performance standards. First, to be sure panel members know what is on the test, they independently take the test and score it themselves, after which they discuss their experiences in taking and marking the test. They then carefully review the performance standards, paying particular attention to the differences between adjacent standards. The procedure to be used to set the cut score is then presented.

The panel members then set their initial cut scores. There is typically a lot of variation among the members’ cut scores, so the members then discuss why they set the scores they did. After their discussion, members then independently set the cut scores a second time. Typically there is still variation among the panellists’ cut scores. Therefore these scores are used to produce impact data, which show the percentages of students in each achievement category if the various cut scores were used. The panel discusses these results and then sets cut scores a third time. While these cut scores are shared with the panel, they are not discussed, since the amount of agreement among the panel members’ cut scores at this point is usually acceptable.

Each year, EQAO assesses how accurately and consistently students are placed in each of the achievement categories. The values for both accuracy and consistency are quite high, ranging from 0.89 to 0.99 across the cut scores for the nine assessments. The differences between these values and the maximum value, which is 1.00, are attributable to random error of measurement.
EQAO has several quality assurance procedures that ensure its assessments are administered consistently and fairly across the province and that the assessment results are reliable and can be validly interpreted. A random sample of schools is visited each year to determine whether the administration guidelines, the use of accommodations for students with special needs, and special provisions for first-language learners are being followed correctly. As well, the principal and teachers at every school in the province are required to report to EQAO any breaches or unanticipated circumstances that interfere with or otherwise compromise the collection of assessment information.

Once the completed assessments from each school have been received, EQAO selects a random sample of schools and examines the completed assessments from these schools for possible irregularities in administration (e.g., different writing, a large number of changed answers). Student responses to multiple-choice items are statistically analyzed to see if there are patterns that suggest the possibility of collusion among two or more students. Results for open-response items are examined for unusual changes in a school’s success rate over time.

EQAO hires a qualified external contractor to check that the EQAO analysts’ equating is correct. EQAO also examines the proportion of students in a school performing at or above the provincial standard (at or above Level 3) to see if there are any dramatic shifts over time.
Not only does EQAO report the assessment results effectively, but they also provide interpretive guides, helpful hints, videos, yearly conferences, and effective workshops throughout the school year. These make EQAO unique in Canada and represent best reporting practice.

EQAO’s two main reports at the provincial level for each year are EQAO’s Provincial Elementary School Report: Results of the Assessments of Reading, Writing and Mathematics, Primary Division (Grades 1–3) and Junior Division (Grades 4–6) and EQAO’s Provincial Secondary School Report: Results of the Grade 9 Assessment of Mathematics and the Ontario Secondary School Literacy Test. Shorter “Highlights” reports are also publicly available at www.eqao.com.

At the school-board and school levels, percentages of students meeting the provincial standard (Levels 3 and 4) over time are publicly available at www.eqao.com³ (click on “School, Board, and Provincial Results”). More detailed school-board and school results are posted on the secure section of EQAO’s Web site and only school-board and school personnel can access them. School personnel can access an interactive Web application that allows them to compare their school to similar schools and to do analyses tailored to their particular needs.

³ To prevent the identification of individual students, EQAO does not publicly release school results when fewer than six students responded to the assessment or when fewer than six students responded to the questionnaire.
At the student level, students and their parents receive Individual Student Reports (ISRs). The ISR for the primary, junior, and Grade 9 assessments shows the student’s achievement level for each component (reading, writing, and mathematics) as a small black square placed at one of five positions within that level. For example, if a student’s score was close to the cut score for that level, the square would appear at the left-most position. The ISR also includes school, school board, and provincial results to provide a context to help interpret the student’s results. The ISR for the OSSLT shows the student’s scale score and, for those students who have not passed, provides a short description of what they need to do in order to pass.

EQAO provides a number of printed resources for board members, principals, teachers, and parents to help them interpret and use the assessment and questionnaire results and, to guide follow-up activities (visit www.eqao.com). EQAO also has an effective and active workshop program designed to help school personnel use assessment results to examine and, when needed, alter instruction so as to improve student learning. In addition, EQAO provides all schools in Ontario with school success stories. The stories are a way of suggesting how school personnel can use the assessment results to improve student learning.

Clearly, the assessment results and information provided to the schools are used to benefit students. Ontario principals and teachers value EQAO’s assessment and questionnaire results provided to them each year. For example, more than 3400 elementary school principals, 8500 Grade 3 teachers, and 7300 Grade 6 teachers were surveyed in 2011. Their survey responses indicated:

- 96% of principals use EQAO achievement results and questionnaire data to guide school improvement initiatives for reading, writing, and mathematics;
- 96% of principals also use EQAO data to identify program strengths and areas for improvement in these subjects;
- 82% of Grade 3 and 80 percent of Grade 6 teachers use EQAO data to identify areas of program strength and areas for improvement in these subjects; and
- 80% of Grade 3 and 78% of Grade 6 teachers reported using EQAO data to identify how well students are meeting curriculum expectations.

(For questionnaire results, visit www.eqao.com.)
EQAO plays an important role in Ontario’s education system by providing

- an independent gauge of student learning in the core areas of reading, writing, and mathematics at four key points in a student’s journey through school,
- an indication of changes in student learning in the core subjects from one year to the next,
- important information that government, school boards, principals, and teachers can use to improve the quality of education of students, and
- a measure of the quality of our publicly funded schools across years.

EQAO’s assessment program is large and complex and EQAO takes its responsibility very seriously. The Assessments of Reading, Writing, and Mathematics, Primary and Junior Divisions; the applied and academic versions of the Grade 9 Assessment of Mathematics; and the Ontario Secondary School Literacy Test (OSSLT) are clearly based on The Ontario Curriculum and measure knowledge and skills students are expected to learn. The procedures used to develop and review assessment and questionnaire items, administer the assessments to students, score student responses, analyze and equate the results of the previous and current years, and report the results and ensure they are used are best of class procedures. EQAO’s quality assurance procedures are equally compelling, as is its willingness to examine its own procedures to ensure that the ones in use are current and best of class.

Turning to the future, in light of the rapidly growing use of technology in education and assessment, EQAO has begun carefully and systematically to implement computer-administered assessments that, based on the progress to date, will be best of class for computer-based assessments.
References


