

EQAO's Technical Report

for the 2009–2010 Assessments

*Assessments of Reading, Writing and Mathematics,
Primary Division (Grades 1–3) and Junior Division (Grades 4–6),
Grade 9 Assessment of Mathematics and
Ontario Secondary School Literacy Test*



About the Education Quality and Accountability Office

The Education Quality and Accountability Office (EQAO) is an independent provincial agency funded by the Government of Ontario. EQAO's mandate is to conduct province-wide tests at key points in every student's primary, junior and secondary education and report the results to educators, parents and the public.

EQAO acts as a catalyst for increasing the success of Ontario students by measuring their achievement in reading, writing and mathematics in relation to *Ontario Curriculum* expectations. The resulting data provide a gauge of quality and accountability in the Ontario education system.

The objective and reliable assessment results are evidence that adds to current knowledge about student learning and serves as an important tool for improvement at all levels: for individual students, schools, boards and the province.

EQAO's Technical Report for the 2009–2010 Assessments: Assessments of Reading, Writing and Mathematics, Primary Division (Grades 1–3) and Junior Division (Grades 4–6), Grade 9 Assessment of Mathematics and Ontario Secondary School Literacy Test

**Education Quality and
Accountability Office**



2 Carlton Street, Suite 1200, Toronto ON M5B 2M9

Telephone: 1-888-327-7377 Web site: www.eqao.com

© 2011 Queen's Printer for Ontario

TABLE OF CONTENTS

CHAPTER 1: PURPOSE AND OVERVIEW OF THIS REPORT 1

THE EQAO ASSESSMENT PROGRAM: PRIMARY (GRADES 1–3), JUNIOR (GRADES 4–6), GRADE 9 AND THE ONTARIO SECONDARY SCHOOL LITERACY TEST.....	1
---	---

CHAPTER 2: ASSESSMENT DESIGN AND DEVELOPMENT 3

ASSESSMENT FRAMEWORKS	3
ASSESSMENT BLUEPRINTS.....	3
TEST CONSTRUCTION: SELECTING ITEMS FOR THE OPERATIONAL FORM	X
ITEM DEVELOPMENT	3
Item Developers	4
Training for Item Developers	4
EQAO Education Officer Review	5
Item Tryouts	5
THE ASSESSMENT DEVELOPMENT AND SENSITIVITY REVIEW COMMITTEES	5
The EQAO Assessment Development Committees	6
The EQAO Sensitivity Committee	6
FIELD TESTING	7
QUESTIONNAIRES	7

CHAPTER 3: TEST ADMINISTRATION AND PARTICIPATION 9

ASSESSMENT ADMINISTRATION	9
The Administration Guides	9
Support for Students with Special Education Needs and English Language Learners: the Guides for Accommodations and Special Provisions.....	9
QUALITY ASSURANCE	11

CHAPTER 4: SCORING 12

THE RANGE-FINDING PROCESS	12
Pre-Range Finding.....	12
Range Finding	13
Overview of the Range-Finding Process	14
PREPARING TRAINING MATERIALS FOR SCORING	14
FIELD-TEST SCORING	14
Training Field-Test Scoring Leaders and Scorers	15
Scoring Open-Response Field-Test Items	15
Developing Additional Scoring Training Materials Before Scoring Operational Items.....	15
SCORING OPEN-RESPONSE OPERATIONAL ITEMS	16
Scoring Rooms for Scoring Open-Response Operational Items.....	16
Training for Scoring Open-Response Operational Items.....	17
Training of Scoring Leaders and Scoring Supervisors for Scoring Open-Response Operational Items	17
Training of Scorers for Scoring Open-Response Operational Items	18
PROCEDURES AT THE SCORING SITE.....	19
Students at Risk	19
Inappropriate Content, Cheating and Other Issues	19
Ongoing Daily Training	20
Daily Scoring-Centre Reports for Monitoring the Quality of Open-Response Item Scoring	20
Required Actions: Consequences of the Review and Analysis of Daily Scoring-Centre Data Reports	22
Auditing.....	23
SCORER RELIABILITY AND VALIDITY	23
Scoring Validity	23
Scorer Reliability.....	25

CHAPTER 5: EQUATING 26

IRT MODELS	26
EQUATING DESIGN	27
EQUATING SAMPLES.....	28
CALIBRATION	28
THE ASSESSMENTS OF READING, WRITING AND MATHEMATICS: PRIMARY AND JUNIOR DIVISIONS.....	28
Description of the IRT Model	28
Equating Sample: Exclusion Rules.....	28
Equating Steps.....	29
Eliminating Items and Collapsing of Score Categories	30
Equating Results.....	30
THE GRADE 9 ASSESSMENT OF MATHEMATICS	33
Description of the IRT Model	33
Equating Sample.....	33
Equating Steps.....	34
Eliminating Items and Collapsing of Score Categories	34
Equating Results.....	35
THE ONTARIO SECONDARY SCHOOL LITERACY TEST (OSSLT).....	36
Description of the IRT Model	36
Equating Sample.....	36
Equating Steps.....	37
Scale Score	37
Eliminating Items and Collapsing of Score Categories	37
Equating Results.....	37
REFERENCES	38

CHAPTER 6: REPORTING RESULTS 39

REPORTING THE RESULTS OF THE ASSESSMENTS OF READING, WRITING AND MATHEMATICS, PRIMARY AND JUNIOR DIVISIONS	39
REPORTING THE RESULTS OF THE GRADE 9 ASSESSMENT OF MATHEMATICS	40
REPORTING THE RESULTS OF THE OSSLT	41
INTERPRETATION GUIDES.....	42

CHAPTER 7: STATISTICAL AND PSYCHOMETRIC SUMMARIES 43

THE ASSESSMENTS OF READING, WRITING AND MATHEMATICS, PRIMARY AND JUNIOR DIVISIONS.....	43
Classical Test Theory (CTT) Analysis	43
Item Response Theory (IRT) Analysis.....	44
Descriptive Item Statistics for Classical Test Theory (CTT) and Item Response Theory (IRT)	58
THE GRADE 9 ASSESSMENT OF MATHEMATICS	60
Classical Test Theory (CTT) Analysis	60
Item Response Theory (IRT) Analysis.....	60
Descriptive Item Statistics for Classical Test Theory (CTT) and Item Response Theory (IRT)	66
THE ONTARIO SECONDARY SCHOOL LITERACY TEST (OSSLT).....	68
Classical Test Theory (CTT) Analysis	68
Item Response Theory (IRT) Analysis.....	68
Descriptive Item Statistics for Classical Test Theory (CTT) and Item Response Theory (IRT)	71
DIFFERENTIAL ITEM FUNCTIONING (DIF)	71
The Primary- and Junior-Division Assessments.....	73
The Grade 9 Mathematics	75
The OSSLT.....	77
DECISION ACCURACY AND CONSISTENCY	78
Accuracy.....	78
Consistency	79
Estimation from One Test Form.....	80
The Primary and Junior Assessments	81

The Grade 9 Assessment of Mathematics	81
The OSSLT.....	82
REFERENCES	82

APPENDIX 4.1: SCORING VALIDITY AND INTERRATER RELIABILITY 83

APPENDIX 7.1: SCORE DISTRIBUTIONS AND ITEM STATISTICS FOR THE 2010 EQAO ASSESSMENTS 98

CHAPTER 1: PURPOSE AND OVERVIEW OF THIS REPORT

The EQAO Assessment Program: Primary (Grades 1–3), Junior (Grades 4–6), Grade 9 and the Ontario Secondary School Literacy Test

In order to fulfill its mandate, EQAO conducts four province-wide assessments: the Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions; the Grade 9 Assessment of Mathematics (academic and applied); and the Ontario Secondary School Literacy Test (OSSLT). All four assessments are conducted annually and involve all students in the specified grades in all publicly funded schools in Ontario, as well as a number of students in private schools that use *The Ontario Curriculum*. For example, students enrolled in inspected private schools are among those who write the OSSLT, as it is a graduation requirement for all students who wish to receive the Ontario Secondary School Diploma (OSSD).

EQAO assessments are developed in keeping with the *Principles for Fair Student Assessment Practices for Education in Canada* (1993), a document widely endorsed by Canada's psychometric and education communities. The assessments measure how well students are achieving selected expectations outlined in *The Ontario Curriculum*. The assessments contain performance-based activities requiring written responses to open-response questions as well as multiple-choice questions, through which students demonstrate what they know and can do in relation to the curriculum expectations measured. One version of each assessment is developed for English-language students and another version is developed for French-language students. Both versions have the same number of items and kinds of tasks, but reflect variations in the curricula for the two languages.

The assessments provide individual student, school, school board and province-wide results on student achievement of selected *Ontario Curriculum* expectations. Every year, EQAO posts school and board results on its Web site (www.eqao.com) for public access. EQAO publishes annual provincial reports in English and in French for education stakeholders and the general public, which are available on its Web site. The assessment results provide valuable information that supports improvement planning by schools, school boards and the Ontario Ministry of Education.

The annual Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions, measure how well elementary school students have met the reading, writing and mathematics curriculum expectations assessed by EQAO and outlined in *The Ontario Curriculum, Grades 1–8: Language* (revised 2006) and *The Ontario Curriculum, Grades 1–8: Mathematics* (revised 2005). The reading component assesses students on their skill at understanding explicit and implicit information and ideas in a variety of text types required by the curriculum. The reading component also requires students to make connections between what they read and their own personal knowledge and experience. The writing component assesses students on their skill at organizing main ideas and supporting details using correct spelling, grammar and punctuation in a variety of written communication forms required by the curriculum. The mathematics component assesses students on their knowledge and skill across the five mathematical strands in the curriculum: number sense and numeration, measurement, geometry and spatial sense, patterning and algebra, and data management and probability.

The Grade 9 Assessment of Mathematics measures how well students have met the expectations outlined in *The Ontario Curriculum, Grades 9 and 10: Mathematics* (revised 2005). The Grade 9 assessment is conducted twice annually to accommodate secondary schools with semestered programs. EQAO develops separate versions of the Grade 9 Assessment of Mathematics for students in academic and applied courses. Students in Grade 9 academic mathematics are assessed on their knowledge and skill across the four mathematical strands in the curriculum: number sense and algebra, linear relations, analytic geometry, and measurement and geometry. Students in Grade 9 applied mathematics are assessed on their knowledge and skill across the three mathematical strands in the curriculum: number sense and algebra, linear relations, and measurement and geometry. A parallel form of the Grade 9 Assessment of Mathematics is developed for each of the two courses (i.e., academic and applied), with one form administered toward the end of the first semester and the second administered toward the end of the second semester.

The OSSLT is administered annually and assesses Grade 10 students' literacy skills based on reading and writing curriculum expectations across all subjects in *The Ontario Curriculum* up to the end of Grade 9. The reading component assesses students on their skill at understanding explicit and implicit information and ideas in a variety of text types required by the curriculum. It also assesses students on their ability to make connections between what they read and their own personal knowledge and experience. The writing component assesses students on their skill at organizing main ideas and supporting details using correct spelling, grammar and punctuation for communication in written forms required by the curriculum. Successful completion of the OSSLT is one of the 32 requirements for the OSSD.

EQAO education officers involve educators across the province in most aspects of EQAO assessments, including

- design and development of items and item-specific scoring rubrics;
- review of items for curriculum content and sensitivity;
- administration of the assessments in schools;
- scoring student responses to open-response items and
- reporting assessment results.

Educators are selected to participate in EQAO activities based on the following criteria:

- geographic location (to represent the northern, southern, eastern and western parts of the province);
- the representation of rural and urban regions;
- current elementary and secondary experience (teachers, administrators, subject experts and consultants);
- expertise in assessment, evaluation and large-scale assessment and
- previous experience working with EQAO.

Educators interested in taking part in the EQAO assessments can apply by visiting the EQAO Web site, www.eqao.com, and clicking "Careers and Educator Opportunities."

CHAPTER 2: ASSESSMENT DESIGN AND DEVELOPMENT

Assessment Frameworks

EQAO posts the current framework for each large-scale assessment on its Web site to provide educators, students, parents and the general public with a detailed description of the assessment, including an explanation of how it relates to *Ontario Curriculum* expectations. The English-language and French-language frameworks for the four EQAO assessments can be found on www.eqao.com.

Assessment Blueprints

EQAO assessment blueprints are used to develop multiple-choice and open-response items for each assessment so that each year the assessment has the same characteristics. This consistency in assessment design ensures that the number and types of items, the relationship to *Ontario Curriculum* expectations (or “curriculum coverage”) and the difficulty of the assessments are comparable each year. It should be noted that not all expectations can be measured in a large-scale assessment. Measurable curriculum expectations are clustered by topic, and items are then mapped to these clusters of expectations. Not all of the measurable expectations in a cluster are measured in any one assessment; however, over time all measurable expectations in a cluster are assessed.

The blueprints can be found in EQAO’s assessment frameworks. A more detailed version of the blueprints is provided to item developers.

Test Construction: Selecting Items for the Operational Form

Operational items are selected from the items that have been field tested in previous assessments. The collected operational items in an assessment constitute the operational form (or “operational assessment” or “operational test”). The operational form is the portion of items that are scored for inclusion in the reporting of student results, contrary to field-test items. Several important factors are taken into consideration when items are selected for an operational form:

- **Data:** The data for individual items, groups of items and test characteristic curves (based on selected items) need to indicate that the assessment items are fair and comparable in difficulty to those on previous assessments.
- **Educator Perspective:** The items selected for an assessment are reviewed to ensure that they reflect the blueprint for the assessment and are balanced for aspects such as subject content, gender representations and provincial demographics (e.g., urban or rural, north or south).
- **Curriculum Coverage:** It is important to note that while items are mapped to clusters of curriculum expectations and not all expectations within a cluster are measured in any one assessment, over time, all measurable expectations in a cluster are included on an assessment.

Sample assessments are available on www.eqao.com.

Item Development

New items are developed and field tested each year before becoming operational items in future assessments. Educators from across the province assist EQAO with all aspects of the development of the assessments, including

- finding or developing reading selections appropriate for the applicable grade levels;
- developing multiple-choice and open-response reading and writing or mathematics items and item-specific scoring rubrics for open-response items;
- trying out items as they are being developed and
- reviewing reading selections, items and item-specific scoring rubrics for curriculum content and possible bias against subgroups of students (e.g., students with special education needs, English language learners, students of a particular gender or ethnic or racial background).

Item Developers

EQAO recruits and trains experienced educators in English and French language (reading and writing) and mathematics to participate in its item-writing committees. The item-writing committee for each assessment comprises 10–20 educators who serve for terms of one to five years. Committee members meet twice a year to write and revise items, discuss results of item tryouts and review items that will be considered for use in subsequent operational assessments.

Item developers construct multiple-choice items in reading and writing or mathematics; open-response items in reading or mathematics; and open-response writing prompts for short- and long-writing tasks. All are referenced to *Ontario Curriculum* expectations and matched to the blueprints for the individual assessments. Item developers are provided with a copy of the *Development Specifications Guide for EQAO Assessments* to assist them in the development of multiple-choice and open-response items and writing prompts.

Item writers for EQAO assessments are selected based on their

- expert knowledge and recent classroom experience in language (English or French) or mathematics education, or both;
- familiarity with and knowledge of the elementary or secondary school curricula in Ontario (especially in language or mathematics);
- familiarity with the cross-curricular literacy requirements for elementary and secondary education in Ontario (especially for the OSSLT);
- expertise and experience in the application of elementary and secondary literacy and mathematics rubrics based on the achievement charts in *The Ontario Curriculum* (to identify varying levels of student performance);
- excellent written communication skills;
- comfort using word-processing software (and, for writers of mathematics items, mathematics software);
- experience in writing instructional or assessment materials for students;
- a proven track record of working collaboratively with others and accepting instruction and feedback and
- access to grade and subject classrooms to conduct item tryouts.

Training for Item Developers

The field-test materials for 2009–2010 were developed by EQAO in partnership with educators from across Ontario. EQAO led a two-day workshop for item developers and spent half a day introducing item developers to the criteria for item writing. EQAO provided an overview of the assessments, including a description of the frameworks, and provided details on the elements of

effective item writing. The remaining time involved a guided item-writing session structured by EQAO education officers.

Each item developer was assigned to write items based on the blueprint for the specific assessment. After the session, the draft materials were reviewed by education officers and returned to the developers, who conducted cognitive labs of the items (i.e., trials with their own students). The results of the cognitive labs were discussed, and further changes to the items were made in a second two-day workshop. EQAO further refined and edited the items after the workshop and prepared the items for committee reviews in preparation for selection of items for the matrix field test.

EQAO Education Officer Review

When the first draft of the items and item-specific scoring rubrics is developed by the item developers, the items and rubrics are reviewed by EQAO education officers. The education officers ensure that each item is referenced correctly in terms of curriculum expectations and difficulty levels. For the multiple-choice items, the education officers consider the clarity and completeness of the stem, the integrity of the correct answer and the plausibility of the three incorrect options. For the open-response items, the education officers consider the correspondence between the items and their scoring rubrics to determine if the items will elicit the range of responses expected.

Item Tryouts

After the initial review of first-draft items by the education officers, item writers use cognitive labs in their own classes and schools to try, with students, the items they have developed. These item tryouts allow item writers to see if their items are working as intended. The student responses are used to inform the editing and refining of stems of multiple-choice items, multiple-choice options, open-response items and item-specific scoring rubrics for open-response items. The results of these item tryouts are provided to EQAO education officers to help them review, revise and edit the items.

The Assessment Development and Sensitivity Review Committees

EQAO recruits and trains Ontario educators with expertise in English and French language, mathematics and equity issues to participate in its Assessment Development and Sensitivity committees. All field-test and operational assessment materials that appear on EQAO assessments are reviewed by these committees.

The goal of these committees is to ensure that items on the Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions, the Grade 9 Assessment of Mathematics and the OSSLT assess literacy and mathematics standards based on *Ontario Curriculum* expectations and that these items are appropriate, fair and accessible to the broadest range of students in Ontario.

Assessment Development and Sensitivity committee members are selected based on their

- knowledge and experience respecting the elementary or secondary school curricula in Ontario;
- expert knowledge and recent classroom experience in at least one of the following:

- o language education (English or French) across the curriculum;
- o mathematics education and
- o equity issues in education (e.g., cultural issues, issues relating to students with special education needs, English language learners) to ensure that no particular group of students is advantaged or disadvantaged by any part of an assessment (e.g., reading selections, writing prompts, mathematics contexts) and
- experience with large-scale assessments (such as previous work with EQAO) or related expert knowledge.

The EQAO Assessment Development Committees

The Assessment Development Committee for each subject in each assessment comprises 10–12 Ontario educators who serve for terms of one to five years. Members meet once a year to provide expert advice from a specialized content and assessment perspective on the quality and fairness of materials being proposed for EQAO assessments and to ensure that all field-test and operational items appropriately assess standards of literacy and mathematics based on *Ontario Curriculum* expectations.

The members of the Assessment Development Committee possess expertise in and current experience with the curriculum and students in at least one of the subjects in the grade being assessed:

- language or mathematics in the primary division (Grades 1– 3) for the primary assessment, administered in Grade 3;
- language or mathematics in the junior division (Grades 4– 6) for the junior assessment, administered in Grade 6;
- mathematics in the intermediate division (Grades 7–10) for the Grade 9 assessment, administered in Grade 9 or
- literacy across the curriculum to the end of Grade 9 for the OSSLT, administered in Grade 10.

The members of the Assessment Development Committee work collaboratively under the guidance of EQAO education officers to ensure that the materials (e.g., reading selections, reading, writing and mathematics items and writing prompts) for a particular assessment are appropriate to the age and grade of the students, the curriculum expectations being measured and the purpose of the assessment. They make suggestions for the inclusion, exclusion or revision of items.

The EQAO Sensitivity Committee

The Sensitivity Committee, which considers all four EQAO assessments, comprises 20–24 Ontario educators who serve for terms of one to five years. About 6–8 members meet in focused subgroups once a year to make recommendations that will assist EQAO in ensuring the fairness of all field-test and operational items being proposed for its assessments. They provide expert advice from a specialized equity perspective to ensure that assessment materials are fair for a wide range of students. The members of the Sensitivity Committee possess expertise in and current experience with equity issues in education (issues related to the diversity of Ontario students, students with special education needs and English language learners).

The members of the Sensitivity Committee work collaboratively under the guidance of EQAO education officers to review assessment materials (e.g., reading selections, items) in various stages of development to ensure that no particular group of students is unfairly advantaged or disadvantaged on any item. They make suggestions for the inclusion, exclusion or revision of items.

Field Testing

Field testing of assessment materials ensures that assessment items selected for future operational assessments are psychometrically sound and fair for all students. Field testing also provides data to equate each year's assessment with the previous year's assessment, so that assessment results can be validly compared over time. Only items found to be acceptable based on field-test results are used operationally in EQAO assessments.

EQAO uses a matrix-sample design in which newly developed items are embedded as field-test items in each assessment. Scores on the field-test items are not used in determining student, school, school board or provincial results. The field-test items are arranged in the student booklets according to psychometric principles to ensure that valid and reliable data are obtained for each field-test item. The field-test items are divided into subsets that are inserted into each assessment, among the operational items, to guarantee that they are attempted by a representative sample of students. Since the field-test items are like the operational items, the students do not know whether they are responding to a field-test item or an operational item. This similarity is meant to counter the low motivation that students may feel when they know that items are field-test items and therefore do not count toward their score. The number of field-test items in each subset ensures that no more than 20% of the items in an assessment are field-test items.

All items, except for the long-writing tasks on the primary- and junior-division assessments and the OSSLT, are field tested this way. A single long-writing task takes more time for students to complete than the combination of all multiple-choice and short-writing tasks. Long-writing tasks are therefore not embedded as field-test items with operational items. Long-writing prompts go through a rigorous process of committee reviews, and, for the OSSLT, field trials are conducted as part of the item development process to ensure their appropriateness. Long-writing tasks are not used for equating.

Questionnaires

EQAO constructs student, teacher and principal questionnaires to collect information on instructional practices and attitudes related to literacy and mathematics so that EQAO results can be used to make recommendations to improve student learning. Some of the contextual data derived from these questionnaires are included in EQAO's annual provincial reports and are used to inform the recommendations in these reports.

The Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions, include student, teacher and principal questionnaires. Students are asked about their age; the language spoken at home; their attitudes toward reading, writing and mathematics; the help available at home with reading, writing and mathematics; their use of computers at school for reading, writing and mathematics; and their use of calculators at school for mathematics learning.

Teachers are asked about the size and makeup of their class (e.g., single, combined or multi-grade); the usefulness and accessibility of the resources in their classroom and the frequency with which they are used; their collaboration with colleagues to plan and discuss curriculum and assessment; their use of assessment-related materials from the EQAO Web site; and their gender, experience, qualifications and participation in professional development. Principals are asked about the school's characteristics (e.g., grades taught, enrolment, student absenteeism and number of staff); their gender, education and participation in professional development; their knowledge about instruction and assessment and use of data for school improvement planning; their communication with parents and staff about EQAO testing; and their use of the *EQAO Guide to School and Board Improvement Planning* (2005), copies of which are available from www.eqao.com.

The Grade 9 Assessment of Mathematics also includes student and teacher questionnaires. The student questionnaires include questions on students' attitudes toward mathematics; their access to a computer, calculator and graphing calculator at home; their time spent on mathematics homework; their age and language spoken at home; their attendance record for mathematics class; and the number of elementary schools attended. The teacher questionnaires include questions about the availability and use of computers and graphing calculators for Grade 9 mathematics students in the school; the instructional and assessment practices used in mathematics classrooms in the school; their collaboration with colleagues to plan and discuss curriculum and assessment; their use of assessment-related materials from the EQAO Web site; and their gender, experience, mathematics qualifications and participation in professional development.

The OSSLT includes a student questionnaire that asks students about their access to a computer at home; the amount of time spent reading different types of materials outside school; their access to reading materials and the language spoken at home; and the time spent on different forms of writing outside school.

CHAPTER 3: TEST ADMINISTRATION AND PARTICIPATION

Assessment Administration

To ensure consistent and fair practice across the province in the administration of the assessments, EQAO publishes an administration guide and a guide for accommodations and special provisions for English language learners annually for each assessment. The guides can be found on www.eqao.com.

The Administration Guides

The administration guide for each EQAO assessment describes in detail the administration procedures that principals and teachers must follow to ensure that the administration of the assessment is consistent and fair for all students in the province. Each school is sent copies of the English- or French-language administration guide for training teachers to administer the assessment. The guide outlines in detail what is expected of educators involved in the administration, including

- the procedures to follow (e.g., preparation of materials for distribution to students, do's and don'ts during the assessment);
- what to say to students (e.g., instructions for presenting the assessment) and
- the professional responsibilities of all school staff involved in the assessment.

During the assessment, students answer multiple-choice questions and write their responses to open-response items. Students must work independently in a quiet environment and be supervised at all times.

Support for Students with Special Education Needs and English Language Learners: the Guides for Accommodations and Special Provisions

The guide for each assessment provides information and directions to assist principals and teachers in making decisions about

- accommodations for students with special education needs;
- special provisions for English language learners and
- the exemption (primary, junior and OSSLT only) or deferral (OSSLT only) of students.

Students with special education needs are allowed accommodations, and English language learners are provided with special provisions, to ensure that they can participate in the assessment and demonstrate the full extent of their skills. In cases where the list of accommodations and special provisions does not address a student's needs, exemption from participation in an assessment is allowed (for the primary and junior assessments only). Each year, EQAO reviews and updates these accommodations and provisions to ensure that they reflect Ministry of Education guidelines and new developments in the support available for students.

The guides for accommodations and special provisions also clarify the expectations for the documentation of accommodations, special provisions, exemptions and deferrals for students receiving them. The guides are based on three Ontario Ministry of Education policy documents: *Individual Education Plans: Standards for Development, Program Planning, and*

Implementation; English Language Learners / ESL and ELD Programs and Services: Policies and Procedures for Ontario Elementary and Secondary Schools, Kindergarten to Grade 12 (2007); and *Policy/Program Memorandum 127* (for Grade 9 and OSSLT only), available at www.edu.gov.on.ca.

Definition of “Accommodations”

Accommodations are defined in the accommodations guides (modified from *Policy/Program Memorandum 127*) as follows.

“Accommodations” are supports and services that enable students with special education needs to demonstrate their competencies in the skills being measured by the assessment. Accommodations change only the way in which the assessment is administered or the way in which a student responds to the components of the assessment. It is expected that accommodations will not alter the content of the assessment or affect its validity or reliability.

On the other hand, “modifications,” which are not allowed, are changes to content and to performance criteria. Modifications are not permitted, because they affect the validity and reliability of the assessment results.

Clarification of instructions for all students is permitted prior to the assessment. Clarification of questions during the assessment (e.g., rewording or explaining) is not allowed.

Special Version Assessments for Accommodated Students

EQAO provides the following special versions of the assessments to accommodate the special education needs of students:

- contracted Braille version plus a set of regular-print booklets for the scribe’s use
- uncontracted Braille version (for primary, junior and OSSLT only) plus a set of regular-print booklets for the scribe’s use
- large-print version—white paper
- large-print version—blue, green or yellow paper
- regular-print version—blue, green or yellow paper
- audio CD version plus a set of regular-print booklets
- audio CD version plus a set of large-print booklets
- one single-sided hard copy to be scanned for use with assistive devices and technology such as text-to-speech software, plus the required sets of regular-print booklets

Special Provisions for English Language Learners

“Special provisions” are adjustments for English language learners to the setting or timing of an assessment. These provisions do not affect the validity or reliability of the assessment results for these students.

Exemptions (Primary, Junior and OSSLT Only)

If a Grade 3 or 6 student is unable to participate in all or part of an assessment, even given accommodations or special provisions, the student may be exempted at the discretion of his or

her school principal. A Grade 3 or 6 student *must* be exempted, however, if for reading, a teacher or another adult must read to him or her and, for mathematics, if mathematics terms have to be defined for him or her.

All students working toward a Grade 9 academic- or applied-level mathematics credit must participate in the Grade 9 assessment.

If a student's Individual Education Plan (IEP) states that he or she is not working toward an OSSD, the student may be exempted from the OSSLT.

Deferrals (OSSLT Only)

All Ontario secondary school students are expected to write the OSSLT in their Grade 10 year. However, this requirement can be deferred for one year (every year until graduation) when a student is working toward the OSSD, if one of the following applies:

- the student has been identified as exceptional by an Identification, Placement and Review Committee (IPRC) and is not able to participate in the assessment, even with the permitted accommodations;
- the student has not yet acquired the reading and writing skills appropriate for Grade 9;
- the student is an English language learner and has not yet acquired a level of proficiency sufficient to participate in the test, or
- the student is new to the board and requires accommodations that cannot yet be provided.

All deferred students who wish to graduate with the OSSD must eventually complete the OSSLT requirement.

Quality Assurance

EQAO has established quality assurance procedures to help ensure that its assessments are administered consistently and fairly across the province and that the data produced is valid and reliable. EQAO follows a number of procedures to ensure that parents, educators and the public have confidence in the validity and reliability of the results reported:

- Quality assurance monitors: EQAO contracts quality assurance monitors to visit and observe the administration of the assessments (in a random sample of schools) to determine the extent to which EQAO guidelines are being followed.
- Database analyses: EQAO conducts two types of statistical analysis of student response data. The first analysis identifies student response patterns that suggest the possibility of collusion between two or more students. The second examines unusual changes in the proportion of primary, junior and Grade 9 students in a school performing at or above the provincial standard (Level 3) over time and overall patterns of school results for open-response items. In the case of the OSSLT, the data are analyzed for unusual changes in a school's rate of success.
- Examination of test materials: Following each assessment, EQAO looks for evidence of possible irregularities in its administration. This is done through an examination of test materials from a random sample of schools prior to scoring.

CHAPTER 4: SCORING

EQAO follows rigorous scoring procedures to ensure that its assessment results are valid and reliable. All responses to open-response field-test and operational reading and mathematics items, as well as writing prompts, are scored by trained scorers. The responses to multiple-choice items, except on the primary assessment, are captured by a scanner. For multiple-choice items on the primary assessment, students fill in the circle corresponding to their response, and their choices are keyed in manually.

Item-specific scoring rubrics and anchors are the key tools used for scoring open-response reading, writing and mathematics items. Anchors illustrate the descriptors for each code in the rubrics. In order to maintain consistency across items and years, item-specific rubrics for open-response items are based on generic rubrics. Each student response to an open-response item is scored according to its best match with one of the code descriptors in the rubric for the item and its anchors. EQAO scoring rubrics describe work at different codes or score points; each code represents a different quality of student performance. The anchors are chosen and validated by educators from across the province during the range-finding process. Scorers are trained to refer constantly to the anchors to ensure consistent scoring. The rubric codes are related to, but do not correspond to, the levels of achievement outlined in the achievement charts in the Ministry of Education curriculum documents.

The generic rubrics used to create item-specific rubrics for each assessment are included in each framework document on www.eqao.com.

The main stages of the scoring process are outlined below.

The Range-Finding Process

Range finding is used to define the range of acceptable performances for each code or score point in each scoring rubric that is used to train scorers. The process is completed in two stages: pre-range finding and range finding.

Range finding for open-response reading and mathematics items and short-writing prompts uses student field-test responses and occurs prior to field-test scoring. Field-test scoring follows operational scoring for the primary, junior and Grade 9 assessments. Field-test scoring for the OSSLT occurs during the summer, after operational scoring has finished.

The long-writing prompts on the OSSLT are field tested with a limited number of students. As a result, range finding for long-writing tasks uses student responses to operational items and occurs just prior to operational scoring.

Pre-Range Finding

During pre-range finding, practising educators work with EQAO staff to select responses that represent the full range of codes or score points for each item or prompt. These responses are used by the range-finding committee. An overview of the process is provided below, though a

few minor variations of this process occur across assessments and between field-test and operational range finding:

1. EQAO education officers are responsible for pre-range finding.
2. Once student booklets arrive at EQAO from schools, a purposeful, demographically representative sample of about 500 student responses for each open-response field-test reading or mathematics item, short-writing task and operational long-writing task is set aside for pre-range finding.
3. Education officers read through 25 booklets, or more if necessary, to see if there is a range of responses and if the item or prompt worked with students. The pre-range finding process for items or tasks does not proceed unless there is a range of responses.
4. Typically, booklets are sorted into four piles based on the range of responses: approximately 20 low, 20 medium, 20 high and 25 of mixed range. The booklets chosen for the piles represent the full range of student responses, including off-topic, incorrect, typical and unusual responses. The mixed pile is determined after the other three piles.
5. Items and tasks that have been left unanswered (“blanks”) or that are difficult to read due to poor handwriting or light ink are not selected for pre-range finding.
6. A cover sheet for each range, showing item, task and booklet numbers, is printed and labelled “high,” “medium,” “low” or “mixed.”

Range Finding

During the range-finding process, subject experts from the Ontario education system meet to make recommendations about high-quality scoring tools and training materials for scorers, in order to ensure the accurate, reliable and consistent scoring of open-response items on EQAO assessments. These experts select representative samples of student responses to define and illustrate the range of student performance within the scoring rubric codes and to provide consensus on the coding of student responses used to train scorers of open-response items.

Range-finding committees consisting of 8–25 Ontario educators meet up to three times a year to make recommendations about student responses that will be used as anchors during scoring. They also discuss other possible responses to be used as training materials for scorers (e.g., as training papers, qualifying test responses and possible responses for calibration activities).

The qualifications for range-finding committee members include

- expertise and experience in the application of rubrics based on the achievement charts in *The Ontario Curriculum* (to identify varying levels of student performance in language and mathematics);
- the ability to explain clearly and concisely the reasons why a student response is at one of the codes in a rubric and
- expertise in and current experience with the curriculum and the grades being assessed.

Members of the range-finding committees

- use their scoring expertise to assign the appropriate generic rubric or item-specific rubric codes to a set of student responses for each group of assessment items;
- share the codes they have assigned with the other members of the committees;

- work collaboratively with the other members of the committees, under the guidance of an EQAO education officer, to reach consensus on appropriate codes for each student response used to train scorers;
- make recommendations for refinements to the item-specific rubrics and
- suggest wording for the annotations explaining the codes assigned.

Overview of the Range-Finding Process

1. Range-finding committee members (including subject experts and current classroom teachers) are recruited and selected for each assessment.
2. Range-finding committee meetings are facilitated by EQAO education officers. After thorough training, the committees are often divided into groups of three or four members.
3. Each group discusses a set of items, prompts and associated item-specific scoring rubrics and recommends appropriate responses to be used as anchors, training papers and qualifying test items to train scorers for each task. The discussions focus on
 - the content and requirements of each item or task;
 - consensus scores by the group for student responses and
 - scoring rules, as required, to ensure consistent scoring of each item or task.

Preparing Training Materials for Scoring

EQAO education officers prepare materials to train scorers for scoring both field-test and operational open-response items. They consider all recommendations and scoring decisions reached during the range-finding process and make final decisions about which student responses will be used for anchors, scorer training, qualifying tests and monitoring the validity (accuracy) and reliability (consistency) of scoring.

Training materials include

- generic or item-specific rubrics;
- anchors that are a good (or “solid”) representation of the codes in the scoring rubrics;
- training papers that represent both solid score-point responses and unusual responses (e.g., shorter than average, atypical approaches, a mix of very low and very high attributes);
- annotations for each anchor and training paper used;
- solid score-point responses for one or more qualifying tests;
- responses to be used for ongoing training during the daily calibration activity (operational scoring only) and
- solid responses used for monitoring validity (operational scoring only).

Field-Test Scoring

Field-test scoring occurs following operational scoring. Since field-test items are to be used in future assessments, they are scored according to the same high standards applied to the scoring of operational items. To ensure the consistency of year-to-year scoring and to reduce the time required for training, the most reliable and productive scoring leaders and scorers of operational items are selected to score field-test items similar to the operational items they have already scored. Education officers arrange for sufficient copies of materials to train the scorers of field-test items. All training materials are kept secure.

Training Field-Test Scoring Leaders and Scorers

Field-test scorers and leaders are trained on the scoring requirements of field-test items, tasks and item-specific rubrics in order to produce valid and reliable item- and task-specific data for operational test construction.

Scoring leaders for each scoring room (designated according to open-response reading and mathematics items and short-writing tasks) are trained by EQAO education officers. These scoring leaders then train scorers. Training includes

- an introduction to the purpose of field-test scoring;
- an explanation of the need to report suspected abuse to the Children's Aid Society;
- a grounding in field-test scoring procedures (using the first item or task and its scoring rubric, anchors and training papers);
- a qualifying test on the first item or task (when field-test scoring does not immediately follow operational scoring) and
- an introduction to subsequent items and tasks and their scoring rubrics, anchors and training papers prior to scoring them.

Standards for passing the qualifying test are the same as those for scoring operational items.

Scoring Open-Response Field-Test Items

A sample of approximately 1200 demographically representative English- and 500 French-language student responses for each field-test item or prompt is scored.

In-depth training for the first item or prompt is provided to scorers by their scoring leader. For the OSSLT, when field-test scoring does not immediately follow operational scoring, scorers write a qualifying test on the first item or prompt before scoring begins. There are no qualifying tests for subsequent items and prompts. Scorers are trained on each item and complete the scoring of one item before proceeding to the next.

Item-analysis statistical reports are prepared following field-test scoring. These reports, together with scorer comments related to field-test item performance, are used to inform test construction.

Developing Additional Scorer-Training Materials Before Scoring Operational Items

When the full range of training materials has not been used for field-test scoring of open-response reading and mathematics items or writing tasks, EQAO develops additional scoring materials using the original range-finding data or field-test scoring data. In the latter case, education officers collect student responses in bundles of high, medium, low and mixed range, so that range finders can select additional scorer-training materials (e.g., anchors, training papers or qualifying tests) for operational scoring.

Education officers are responsible for arranging all of the materials required to train the scorers who are to score operational items.

Scoring Open-Response Operational Items

EQAO has rigorous policies and procedures for the scoring of operational assessment items and tasks to ensure the reliability of assessment results.

The primary, junior and Grade 9 assessments are scored by qualified Ontario educators. The primary and junior assessments are scored by educators representing all the primary and junior grades. The Grade 9 Assessment of Mathematics is scored by educators with expertise in mathematics and experience working with Grade 9 students. Scoring provides teachers with valuable professional development in the area of understanding curriculum expectations and assessing student achievement.

The OSSLT is scored before the end of the school year. EQAO recruits as many teacher-scorers (i.e., members of the Ontario College of Teachers) as possible and fills the complement of required scorers with retired educators and qualified non-educators (or “other-degree scorers”). As part of the initial screening process administered by the contractor that recruits the other-degree scorers, applicants write a test to ensure that they have sufficient proficiency in English or French to score the test effectively.

Scoring Rooms for Scoring Open-Response Operational Items

A limited set of the operational assessment items is scored in a scoring room under the leadership of a scoring leader. He or she trains all the scoring supervisors and scorers in the room. Scoring leaders, with the assistance of the scoring supervisors, manage the training, scoring and retraining of the scorers. All scorers are trained to use the EQAO scoring guide (rubrics and anchors) for each item they score. Following training, scorers must pass a qualifying test. The validity (accuracy) and reliability (consistency) of scoring is tracked daily at the scoring site, and retraining occurs when required. All scoring procedures are conducted under the supervision of EQAO’s program managers and education officers.

Scorers sit alone and score individually. Scorers can discuss anomalous responses with their scoring leader or supervisor.

Operational open-response reading, writing and mathematics items for the primary and junior assessments and operational mathematics items for the Grade 9 assessment are single scored. To measure overall scorer reliability (consistency), the scores from sets of validity booklets are used in scoring those items that are not part of the validity process based on validity papers. More information about scoring validity is provided later in this chapter.

In scoring, “exact agreement” means that the code or score point assigned to an open-response item by a pair of scorers is exactly the same. “Adjacent” means that there is a difference of one score point between the codes assigned to an open-response item by a pair of scorers. “Non-adjacent” means that there is a difference of more than one score point between the codes assigned to an open-response item by a pair of scorers.

Each open-response reading item and writing task on the OSSLT is scored by two trained scorers independently, using the same rubric. A “blind scoring” model is used: that is, scorers do not

know what score has been assigned by the other scorer. The routing system automatically ensures that responses are read by two different scorers. If the two scores are in exact agreement, that score is assigned to the student. If the two scores are adjacent, the higher score (for reading and short-writing tasks) or the average of the two scores (for news reports and paragraphs expressing an opinion) is assigned to the student. If the two scores are non-adjacent, the response is scored again by an expert scorer, to determine the correct score for the student. This rigour ensures that parents, students and teachers can be confident that all students have received valid scores.

Training for Scoring Open-Response Operational Items

The purpose of training is to develop a clear and common understanding of the scoring materials so that each scoring leader, scoring supervisor and scorer applies the scoring materials in the same way, resulting in valid (accurate) and reliable (consistent) student scores.

Training of Scoring Leaders and Scoring Supervisors for Scoring Open-Response Operational Items

Scoring leaders must have subject expertise and be, first and foremost, effective teachers of adults. They must encourage scorers to abandon preconceived notions about scoring procedures and align their thinking and judgment to the procedures and scoring materials for the items being scored. The responsibilities of scoring leaders include

- training all scoring supervisors and scorers in the applicable room;
- overseeing the scoring of items;
- ensuring that scoring materials are applied consistently and
- resolving issues that arise during scoring.

Scoring leaders are also responsible for reviewing and analyzing daily data reports to ensure that a high quality of scoring occurs in their scoring room.

Scoring supervisors are selected from a pool of experienced and proficient EQAO scorers. Scoring supervisors assist scoring leaders and ensure that their assigned scorers are qualified and are scoring accurately. Scoring supervisors may also be asked to retrain individual scorers when necessary.

The training for scoring leaders and scoring supervisors is conducted before scoring begins. EQAO education officers train scoring leaders and oversee the training of scoring supervisors. Supervisor training is substantially similar to the training and qualifying for scorers. The only difference is that supervisors receive additional training regarding scoring materials, room-management problems and issues that may arise during scoring. For Grade 9 scoring, an EQAO education officer trains the scoring leaders and supervisors assigned to one room at the same time.

Following training and prior to scoring, scoring leaders and scoring supervisors must pass a qualifying test that involves scoring 14–20 student responses for the items to be scored in their room. The items included in the qualifying test are selected during the range-finding process. Scoring leaders and supervisors must attain at least an 80% exact and a 100% exact-plus-

adjacent match with the expertly assigned scores. Scoring leaders or supervisors who fail the qualifying test may not continue in the role of leader or supervisor.

Training of Scorers for Scoring Open-Response Operational Items

The purpose of training for open-response operational items is to ensure that all scorers become experts in scoring specific items or subsets of items. All operational items require a complete set of scoring materials: item-specific rubrics, anchors (real student responses illustrating work at each code in the rubric) and their annotations, training papers, a qualifying test, validity papers (primary, junior, OSSLT) or validity booklets (Grade 9) and items for the daily calibration activity.

To obtain high levels of validity (accuracy) and reliability (consistency) during scoring, EQAO adheres to stringent criteria for selecting, training and qualifying scorers. Various other quality control procedures, as outlined below, are used during the scoring process to identify scorers who need to be retrained or dismissed from scoring.

All the scorers in one room are trained to score using the same scoring materials. These scoring materials are approved by EQAO and cannot be altered. During training, scorers are told they may have to adjust their thinking about scoring student performance in a classroom setting in order to accept EQAO's standards and practices for its assessments.

Training for scorers on the limited number of open-response items scored in a room takes approximately half a day and includes

- general instructions about the security, confidentiality and suitability of the scoring materials;
- instructions on entering scores into the Personal Digital Assistant (PDA) used to collect scoring data. For instance,
 - prior to entering scores, scorers scan the unique student booklet barcodes into the PDA (which has a built-in barcode scanner) in order to link student names to their corresponding scores and
 - scorers enter their scores for student responses into the PDA, then synchronize the PDA in a cradle connected to a laptop, which uploads the data to a server;
- a thorough review and discussion of the scoring materials for each item to be scored (the item, generic or item-specific rubrics, anchors and their annotations):
 - emphasis is placed on the scorer's understanding of how the responses differ in incremental quality and how each response reflects the description of its code on the rubric and
 - the anchors consist of responses that are typical of each achievement level (rather than unusual or uncommon) and solid (rather than controversial or "borderline") and
- the scoring of a series of validity papers, consisting of selected expertly scored student responses:
 - validity papers typically contain responses that are solid as well as responses that are unusual and less solid (e.g., shorter than average, employing atypical approaches, containing both very low and very high attributes);
 - scorers will first score the responses then compare their scores with the annotated, correct scores provided;
 - scorers will then discuss the annotation and attributes of each correct response with their scoring leader and supervisor and

- o scorers will internalize the rubric during this process and adjust their individual scoring to conform to it.

Scorers are also trained to

- read responses in their entirety prior to making any scoring decisions;
- view responses as a whole rather than focusing on particular details such as spelling;
- remain objective and fair and view the whole response through the filter of the rubric and
- score all responses in the same way, to avoid adjusting their scoring to take into account a characteristic they assume about a student (e.g., special education needs, being an English language learner).

Following training and prior to scoring, scorers must pass a qualifying test consisting of 14–20 student responses to all the items to be scored in a room. These items are selected during the range-finding process as examples of solid score points for rubrics. Scorers must attain at least a 70% exact match with the expertly assigned score. This ensures that scorers have understood and can apply the information they received during training. Scorers who fail the qualifying test the first time may undergo further training and write the test a second time. Scorers who fail to pass the qualifying test a second time are dismissed.

Procedures at the Scoring Site

Students at Risk

On occasion, a student's response to an open-response question will contain evidence that he or she may be at risk (e.g., content that states or implies threats of violence to oneself or others, or possible abuse or neglect). Copies of student responses that raise concerns are sent to the student's local Children's Aid Society. It is the legal responsibility and duty of scorers, in consultation with their site managers, to inform the Children's Aid Society of such cases.

Inappropriate Content, Cheating and Other Issues

Student responses to open-response questions occasionally contain inappropriate content or evidence of possible teacher interference or other issues. Booklets containing any such issues are sent to the exceptions room for consideration.

Offensive Content

Obscene, racist or sexist content in student response booklets is reviewed by EQAO staff to determine whether the school should be contacted. If the offensive content warrants it, EQAO will notify the school.

Cheating

When there is any evidence in a booklet that may indicate some form of irregularity (e.g., many changed answers, teacher interference), the booklet is reviewed by EQAO staff to determine whether the school should be notified. In cases where cheating is confirmed, no scores are provided for the student.

Damaged or Misprinted Booklets

In very few cases, booklets given to students are torn, stapled incorrectly or have missing pages or a defaced barcode that cannot be scanned. In such cases, students are not penalized. These damaged booklets are further reviewed by EQAO staff to determine whether the results in these booklets should be pro-rated based on the results in booklets unaffected by such problems.

Ongoing Daily Training

Scoring leaders provide clarification on scoring of specific items and key elements of item-specific rubrics in their scoring rooms. EQAO conducts morning and afternoon training to refresh scorers' understanding of the scoring materials and to ensure that they apply the scoring materials accurately and consistently from one day to the next, and before and after lunch breaks.

Daily Morning Review of Anchors

Scoring leaders begin each day with a review of all or a portion of the rubrics and anchors. The purpose of the review is to refocus scorers and highlight any section of the rubrics that require attention. This review is more comprehensive after a weekend break (or following any extended break).

Daily Afternoon Calibration Activity

Scorers begin each afternoon by scoring one or more of the selected calibration items (expertly scored student responses that were challenging to score). Calibration items facilitate the review of and response to scoring issues raised by scorers or the daily scoring data reports. Scorers score and record the scores for the calibration items. Scoring leaders review the calibration item scores and provide scorers with an explanation of the issues raised and clear information and guidance on the correct way to score these items. Individual scorers or groups of scorers that encounter difficulty with daily afternoon calibration items can address their issues with their scoring leader or scoring supervisor.

Daily Scoring-Centre Reports for Monitoring the Quality of Open-Response Item Scoring

Scoring leaders and supervisors receive daily data reports showing daily and cumulative validity, reliability and productivity data for individual scorers and for groups of scorers in their room. These data reports are described below.

Daily and Cumulative Validity

During scoring, EQAO tracks the validity (accuracy) of scorers through the use of expertly scored validity papers. Scorers score up to 10 validity papers a day. They ensure that scorers are giving correct and accurate scores by comparing them with those assigned during the range-finding process. Scoring leaders and supervisors use the report data to determine whether scorers are drifting from the scoring standards (provided by scorer training and training materials) and whether any retraining is required. During scoring, all scorers must maintain a minimum accuracy rate on the validity papers. The accuracy rates are as follows: 75% exact and 95% adjacent for three-point rubrics, 70% exact and 95% exact-plus-adjacent agreement for four-point rubrics, 65% exact and 95% exact-plus-adjacent agreement for five-point rubrics and 60% exact and 95% exact-plus-adjacent agreement for six-point rubrics.

“Exact agreement” means that the code or score point assigned to an open-response item by a pair of scorers is exactly the same. “Adjacent” means that there is a difference of one score point between the codes assigned to an open-response item by a pair of scorers. “Non-adjacent” means that there is a difference of more than one score point between the codes assigned to an open-response item by a pair of scorers.

The data reports summarize daily and cumulative levels of agreement (exact, adjacent and high or low non-adjacent agreement) on validity papers with pre-set scores. “Adjacent-low” means that the score assigned to a certain response by a scorer is one point below the score assigned by the expert panel. “Adjacent-high” means that the score is one point above the score given by the expert panel, and “non-adjacent” means that the difference between the scores assigned by the scorer and the expert panel is greater than one score point.

The reports also include a cumulative-trend review and are summarized by item or item set, rubric, room, group and scorer. Scorers are listed from low to high validity. Scorers not meeting the exact-agreement requirement are highlighted in the report.

Accuracy is measured primarily by the use of validity metrics. The daily data reports for scorers who pass the qualifying test after retraining are carefully monitored to ensure that they continue to meet standards. If, after a minimum of 10 validity items, a scorer falls below the required exact-plus-adjacent-agreement standards, the scorer receives retraining (including a careful review of the anchors). If retraining does not correct the situation, the scorer may be dismissed. The scores of dismissed scorers are audited and, if necessary, changed.

Daily and Cumulative Mean-Score and Score-Point Distribution

Daily and cumulative mean-score and score-point distribution data reports are used to monitor room and individual scorer drift. They confirm validity and guide ongoing training (based on calibration items) at both the individual and room levels.

These reports identify and summarize (by item or item set, room, group, rubric and scorer) the daily and cumulative mean score and the distribution of assigned score points.

Daily and Cumulative Reliability

Open-response primary- and junior-division and Grade 9 assessment items in reading, writing and mathematics are single scored. To measure overall scorer reliability (consistency) during scoring, scores from sets of validity booklets are used for items that are not part of the validity process. All open-response OSSLT items are routed for a second scoring, which is used to monitor interrater reliability. Scoring leaders and supervisors use this secondary source of data, in conjunction with validity data, to determine whether individual scorers or groups of scorers require retraining.

The reports identify and summarize daily and cumulative levels of interrater agreement, including exact, adjacent and high and low non-adjacent agreement. The reports are summarized by item or item set, room, group, rubric and scorer, and scorers are listed from low to high reliability. Scorers not meeting the exact-agreement requirement are highlighted in the report.

Daily and Cumulative Productivity

During scoring, EQAO tracks scoring-centre productivity and monitors progress through daily productivity reports to ensure that all scoring will be completed during the scoring session. The reports show the number and percentage of responses for which the scoring is complete. These reports, which are provided to scoring leaders and supervisors, report daily and cumulative productivity. The reports also track the productivity of each scorer to ensure that daily targets and minimums are met. Productivity targets and minimums are set for each room, taking into consideration the subset of items being scored. Productivity targets for scoring supervisors are 50–60% of scorer targets.

The reports are summarized by room, group and individual scorer and include the daily and cumulative number of student responses scored and a cumulative-trend review. The reports list scorers from low to high productivity. Scorers not meeting the minimum productivity rate for the room in which they are scoring are highlighted in the report. Scoring leaders and supervisors review the data highlighted in this report to determine whether retraining is required for any scorer.

Scoring completion reports also compare, on a daily and cumulative basis, the number of scorings completed, with completion targets for the scoring room.

Aggregated Daily and Cumulative Individual Scorer Data

These reports combine validity data with secondary data for each scorer. The reports include not only daily and cumulative validity data but also daily and cumulative reliability, mean score and productivity data. The reports list scorers from low to high validity. Scorers not meeting the exact-agreement requirement of 75% on three-point rubrics, 70% on four-point rubrics or 60% on six-point rubrics are highlighted in this report. As such, this report assists the scoring leaders in identifying the scorers that require retraining.

Required Actions: Consequences of the Review and Analysis of Daily Scoring-Centre Data Reports

Scoring leaders are responsible for the daily review and analysis of all scoring-centre data reports to ensure the quality of the scoring in their scoring room. EQAO personnel (the chief assessment officer, director of assessment and reporting and education officers) also review the daily reports and work with scoring leaders to identify

- individual scorers who need retraining
- groups of scorers who need retraining
- calibration items that will ensure quality scoring
- issues arising that require additional training for an entire room
- productivity issues

Scoring leaders share the data and discuss data-related issues with the appropriate scoring supervisors so that interventions can be planned. The following occurs when a scorer is not meeting the validity metrics:

- The scorer is retrained and re-qualified if the exact-plus-adjacent standard is not met.

- The scorer is retrained and participates in recalibration if the exact-agreement requirement is not met.

Scorers, as well as their leaders and supervisors, are required to demonstrate their ability to score student responses accurately and consistently throughout training, qualification and scoring. Scoring supervisors and scorers must meet EQAO standards for validity, reliability and productivity in order to continue. If a scoring supervisor or scorer does not meet one or more of these standards, he or she will receive retraining. If his or her scoring does not improve, the scoring supervisor or scorer may be dismissed. All retraining is documented by scoring leaders and supervisors.

Auditing

EQAO audits individual student score sheets (i.e., student records showing the scores assigned to selected open-response items) for inconsistencies that may indicate incomplete scoring. Any booklet scored entirely blank is rerouted for a second scoring.

Scorer Validity and Reliability

The procedures used for estimating the validity and reliability of EQAO assessments are summarized below. The estimates of validity and interrater reliability are presented in Appendix 4.1. Two sets of results are reported for each writing prompt: one for topic development and one for conventions.

Scoring Validity

As described earlier in this chapter, scoring validity is assessed by having scorers assign scores to validity papers and validity booklets, which are student responses that have been scored by an expert panel. For the primary and junior assessments and for the OSSLT, a set of five validity papers is prepared, copied and distributed to all scorers each morning and afternoon. In addition, the original student booklets that these validity papers were copied from are used as blind validity booklets and circulated to provide additional validity material for the scorers. For Grade 9, only blind validity booklets are used, and they are circulated as frequently as possible so that most scorers can score at least one validity booklet per day. The sets of validity papers are not used for Grade 9 because high levels of scorer consistency have been achieved over the years through the use of the blind validity booklets only.

Validity is assessed by examining the agreement between the scores assigned by the scorers and those assigned by the expert panel. The following six indices are computed: percentage of exact agreement, percentage of exact-plus-adjacent agreement, percentage of adjacent agreement, percentage of adjacent-low agreement, percentage of adjacent-high agreement and percentage of non-adjacent agreement.

“Adjacent-low” means that the score assigned to a certain response by a scorer is one point below the score assigned by the expert panel. “Adjacent-high” means that the score is one point above the score given by the expert panel, and “non-adjacent” means that the difference between the scores assigned by the scorer and the expert panel is greater than one score point.

The Assessments of Reading, Writing and Mathematics: Primary and Junior Divisions

There are 10, three and eight open-response items for the reading, writing and mathematics components of the assessments, respectively. Four-point scoring rubrics are used for reading and mathematics. For the writing components, there are two short-writing prompts and one long-writing prompt that are scored for topic development and use of conventions. A four-point scoring rubric is used for topic development and a three-point scoring rubric for conventions. The scoring validity estimates for reading, writing and mathematics for the primary and junior divisions are presented in Tables 4.1.1–4.1.12 of Appendix 4.1. The statistics are provided for each item and for the aggregate of the items for each assessment. For writing, the aggregate statistics for short-writing prompts, long-writing prompts and all prompts are provided separately.

In 2009–2010, the EQAO target of 95% exact-plus-adjacent agreement was met for all items in the reading components of the assessments, for all but one item in the mathematics components and for 23 of the 24 rubrics for the writing components of the assessments. The aggregate validity estimates for exact-plus-adjacent agreement ranged from 98.2–99.7%.

The Grade 9 Assessment of Mathematics: Academic and Applied

The Grade 9 Assessment of Mathematics has separate English-language and French-language versions for students enrolled in academic and applied courses. The assessment is administered in January for students in mathematics courses in the first semester and in June for students in second-semester and full-year courses. The scoring validity estimates for the Grade 9 Assessment of Mathematics are presented in Tables 4.1.13–4.1.16 of Appendix 4.1 for both administrations. The tables present statistics for each open-response item and the aggregate for open-response items for each administration. They also include aggregate statistics across the winter and spring administrations, because both were scored during the same scoring session in July 2010. Seven questions were scored for each administration using four-point rubrics for a total of 56 questions across both administrations.

The EQAO target of 95% exact-plus-adjacent agreement was met for all items and the aggregate validity estimates ranged from 98.8–99.8%.

The Ontario Secondary School Literacy Test (OSSLT)

The scoring validity estimates for the OSSLT are reported in Tables 4.1.17–4.1.20 of Appendix 4.1. For each test, four reading items were scored with three-point rubrics and two long-writing prompts were scored with a six-point rubric for topic development and a four-point rubric for conventions. Two short-writing prompts were scored with a three-point rubric for topic development and a two-point rubric for conventions, which were combined into a five-point rubric for the purposes of validity statistics. Aggregate statistics are provided separately for reading items, short-writing prompts, long-writing prompts and all writing prompts.

In 2009–2010, the EQAO target of 95% exact-plus-adjacent agreement was met for the four reading items on both the English-language test and on the French-language test. This target was met for all short-writing prompts, for all conventions rubrics for the long-writing prompts and for three out of four topic-development rubrics for the long-writing prompts. The percentage of exact-plus-adjacent agreement was 92.1 for one English-language topic-development rubric. The

aggregate validity estimates ranged from 96.9–99.8%.

Scorer Reliability

Test reliability is affected by different sources of measurement error. In the case of open-response items, inconsistency in scoring is the source of error. As part of its attempt to minimize this type of error for the primary, junior and Grade 9 assessments, EQAO tracks reliability using scores assigned to items in the pre-selected blind validity booklets that are not part of the validity process.

The process for determining the reliability of open-response scoring for the OSSLT does not require that selected student booklets be brought back to a scoring room (or “reinserted”). All student responses to open-response items are scored automatically by at least two scorers. Scoring reliability is determined from the scores assigned by the two independent scorers for each student response.

The percentage of agreement between the scores awarded by a pair of scorers is known as the interrater reliability. Four indices are used to identify the interrater reliability: percentage of exact agreement, percentage of exact-plus-adjacent agreement, percentage of adjacent agreement and percentage of non-adjacent agreement.

“Exact agreement” means that the code or score point assigned to an open-response item by a pair of scorers is exactly the same. “Adjacent agreement” means that there is a difference of one score point between the codes assigned to an open-response item by a pair of scorers. “Non-adjacent agreement” means that there is a difference of more than one score point between the codes assigned to an open-response item by a pair of scorers. Scoring reliability estimates for the primary, junior and Grade 9 assessments and the OSSLT are presented in Tables 4.1.21–4.1.40 of Appendix 4.1.

The Assessments of Reading, Writing and Mathematics: Primary and Junior Divisions

In 2009–2010, the EQAO target of 95% exact-plus-adjacent agreement was met for 37 out of 40 reading items, 23 out of 24 rubrics for writing and 28 out of 31 mathematics items. The aggregate reliability estimates ranged from 96.6–99.4%.

The Grade 9 Assessment of Mathematics: Academic and Applied

In 2009–2010, the EQAO target of 95% exact-plus-adjacent agreement was met for 47 out of 56 items, and the aggregate reliability estimates ranged from 95.4–98.3%.

The Ontario Secondary School Literacy Test (OSSLT)

The EQAO target of 95% exact-plus-adjacent agreement was met for all reading items, for three out of four short-writing prompts, for all conventions rubrics for the long-writing prompts and for none of the topic-development rubrics for the long-writing prompts. The aggregate reliability estimates ranged from 93.4–97.8%.

CHAPTER 5: EQUATING

For security purposes, EQAO constructs different assessments every year while ensuring that content and statistical specifications are similar to those of the assessments from previous years. Despite such efforts to ensure similarity, assessments from year to year may differ somewhat in their difficulty. To account for this, EQAO uses a process called equating, which adjusts for differences in difficulty between assessments from year to year (Kolen & Brennan, 2004). Equating ensures that students in one year are not given an unfair advantage over students in another and that reported changes in achievement levels are due to differences in student performance and not differences in assessment difficulty.

From time to time, the Ministry of Education makes modifications to *The Ontario Curriculum*, and EQAO assessments are modified accordingly in content and length. The new assessments differ in content and statistical specifications from those constructed in previous years, prior to the curriculum revisions. In such cases, EQAO uses a process called scaling to link the previous years' assessments with the current year's modified ones.

The processes used in equating and scaling are similar, but their purposes are different. Equating is used to adjust for differences in difficulty among assessments that are similar in content and statistical specifications. Scaling is used to link two assessments that are different in content and statistical specifications (Kolen & Brennan, 2004). Since there were no significant changes to the test specifications from 2008–2009 to 2009–2010, only equating procedures were used in 2009–2010.

The following sections describe the Item Response Theory (IRT) models, equating design, equating samples and calibration procedures used during the 2009–2010 school year for the various EQAO assessments.

IRT Models

Item-response models define the relationship between an unobserved construct or ability (θ) and the probability (P) of a student correctly answering a dichotomously scored item. For polytomously scored items, the models define the relationship between an unobserved construct or ability and the probability of a student receiving a particular score on the item. The Three-Parameter Logistic (3PL) model and the Generalized Partial Credit (GPC) model are the general models used by EQAO to estimate the parameters of multiple-choice and open-response items and the ability parameters. The 3PL model (see Yen & Fitzpatrick, 2006, for example) is given by Equation 1:

$$P_i(\theta) = c_i + (1 - c_i) \frac{\exp^{Da_i(\theta - b_i)}}{1 + \exp^{Da_i(\theta - b_i)}} , \quad (1)$$

where $P_i(\theta)$ is the probability of a student (with ability θ) answering item i correctly;
 a_i is the slope parameter for item i ;
 b_i is the difficulty parameter for item i ;

c_i is the pseudo-guessing parameter for item i and
 D is a scaling constant equal to 1.7.

The GPC model (Muraki, 1997) is given by Equation 2:

$$P_{ih}(\theta) = \frac{\exp\left[\sum_{v=0}^h Da_i(\theta - b_i + d_h)\right]}{\sum_{c=0}^{M_i} \exp\left[\sum_{v=0}^c Da_i(\theta - b_i + d_h)\right]}, \quad h = 0, 1, \dots, M_i, \quad (2)$$

where $P_{ih}(\theta)$ is the probability of a student with ability θ choosing the h th score category for item i ;
 a_i is the slope parameter for item i ;
 b_i is the difficulty parameter for item i ;
 d_h is the category parameter for category h of item i ;
 D is a scaling constant equal to 1.7 and
 M_i is the maximum score on item i .

Equating Design

The forward fixed common-item-parameter non-equivalent group design is used to equate EQAO assessments over different years. Common items are sets of items that are identical in two assessments and are used to create a common scale for all the items in the assessments. These common items are selected from the field-test items administered in one year and used as operational items in the next. There are several steps in the forward fixed common-item design:

1. Operational item parameters in the current year's assessments are calibrated.
2. Operational items and field-test items from the previous year are brought forward to the current year's assessments and recalibrated:
 - This is done by fixing the parameters of the items common to the two years at the values obtained in Step1.
 - This process places the item parameters from the two years on the same scale.
3. Recalibrated parameters for the operational items from the previous year are then used to rescore the corresponding equating sample:
 - For the OSSLT, the theta value of the cut point (corresponding to the percentage of successful students in the previous year) is then identified and applied to the current year's test-score distribution to obtain the percentage of successful and unsuccessful students for the current year.
 - For the primary, junior and Grade 9 assessments, the theta values of the cut points (corresponding to the percentage of students at each performance level) are identified and then applied to the current year's test-score distribution to obtain the percentage of students at each performance level for the current year.

Equating Sample

For each assessment, EQAO uses a set of exclusion rules to select a calibration and equating sample. The exclusion rules ensure that the equating sample is representative of the population of students who wrote the assessment under typical administration conditions. While the exclusion rules are similar for all assessments, there are some differences. Therefore, the exclusion rules are provided below in the description of the equating conducted for each assessment.

Calibration Sample

Calibration is the process of estimating the item parameters that determine the relationship between ability and the probability of answering a multiple-choice item correctly or receiving a particular score on a polytomously scored open-response item. For each assessment, the calibration of the items for the English-language and the French-language populations is conducted separately. The calibrations are conducted using the program PARSCALE 4.1 (Muraki & Bock, 2003). Before equating is conducted, the difficulty and slope parameter estimates of common items are plotted to identify “outliers”—items showing large differences in their difficulty or slope estimates between the two years. Once potential outliers are identified, sensitivity analysis is conducted to compare results (i.e., the theta cut values) obtained by including and excluding such outlier items in the equating process. The outliers are eliminated only if the theta cut values are very different.

The Assessments of Reading, Writing and Mathematics: Primary and Junior Divisions

Description of the IRT Model

A modified 3PL model (Equation 1) is used for multiple-choice items on the primary- and junior-division assessments, with the pseudo-guessing parameter fixed at $0.20 [1/(k + 1)]$, where k is the number of options] to reflect the possibility of students with a very low ability to correctly answer an item. The GPC model (Equation 2) is used for open-response items.

Equating Sample: Exclusion Rules

The following categories of students were excluded from the equating samples for 2008–2009 and 2009–2010:

1. students who were not attending a publicly funded school;
2. students who were home-schooled;
3. French Immersion students;
4. students who completed no work on an item (by booklet and item type);
5. students receiving accommodations;
6. students who were exempted and
7. students who did not attempt at least one question in each significant part of the test.

Using these exclusion rules, three student samples were obtained, as presented in Table 5.1:

1. students from the 2008–2009 population who responded to both the 2009–2010 operational-test items and the field-test items that had been brought forward to form the 2009–2010 operational assessments and who were not excluded by the rules stated above (calibration sample);

2. students from the 2008–2009 population who wrote the operational assessment and who were not excluded by the rules stated above (equating sample) and
3. students from the 2009–2010 population who wrote the operational assessment and who were not excluded by the rules stated below (calibration and equating samples).

Table 5.1 Number of Students in the Calibration and Equating Samples for the 2008–2009 and 2009–2010 Primary- and Junior-Division Assessments (English and French)

Assessment	Number of Students		
	2009–2010 Calibration and Equating Samples	2008–2009 Calibration Sample	2008–2009 Equating Sample
Primary Reading (English)	105 254	29 817	104 824
Junior Reading (English)	114 849	32 398	117 646
Primary Reading (French)	6 072	2 356	5 954
Junior Reading (French)	5 677	2 702	5 772
Primary Writing (English)	105 747	48 133	105 315
Junior Writing (English)	115 099	53 279	117 901
Primary Writing (French)	6 118	3 441	6 018
Junior Writing (French)	5 697	3 470	5 795
Primary Mathematics (English)	102 176	87 373	102 187
Junior Mathematics (English)	116 598	96 544	120 155
Primary Mathematics (French)	6 146	5 137	6 028
Junior Mathematics (French)	5 733	5 855	5 864

Equating Steps

In equating the 2008–2009 and 2009–2010 tests, the forward fixed common-item-parameter non-equivalent group design was implemented as follows:

1. The 2009–2010 operational items were calibrated independently to obtain item parameter estimates and student-ability scores for the 2009–2010 calibration and equating sample.
2. The 2008–2009 operational items were calibrated (using the calibration sample) together with the field-test items that were brought forward to the 2009–2010 operational assessments. In this calibration, the item parameter estimates of the field-test items were fixed at the values obtained from the 2009–2010 calibration runs (Step 1).
3. The 2008–2009 equating sample was scored using the operational item parameter estimates obtained in Step 2.
4. The percentage of students at each achievement level was determined for the 2008–2009 equating sample from the levels assigned in 2008–2009. The theta value of the cut points that replicated this distribution was identified for each boundary (0/1, 1/2, 2/3 and 3/4).
5. These theta values were then used as the cut points for 2009–2010.
6. The operational item parameter estimates of 2009–2010 obtained in Step 1 were used to score the full student population.
7. The cut-score points identified in Step 4 were applied to the 2009–2010 student theta values, students were assigned to levels, and the percentage of students at each performance level was determined.

Eliminating Items and Collapsing of Score Categories

For the primary- and junior-division assessments, six multiple-choice items and five open-response items across the 12 assessment components (English and French reading, writing and mathematics) were excluded from equating. These items were either modified between field- and operational-test administrations or were not field tested in the 2008–2009 assessment year. The number of items not used in the equating process and the number of items dropped from each assessment component are presented in Table 5.2.

Table 5.2 Number of Items Excluded from the Equating Process and Dropped from the Primary- and Junior-Division Assessments (2009–2010)

Assessment	No. of Items Excluded from Equating*		No. of Items Dropped from the Assessment
	Multiple-Choice	Open-Response	
Primary Reading (English)	1	0	0
Junior Reading (English)	0	0	0
Primary Reading (French)	2	0	0
Junior Reading (French)	2	0	0
Primary Writing (English)	0	1	0
Junior Writing (English)	0	1	0
Primary Writing (French)	0	1	0
Junior Writing (French)	0	1	0
Primary Mathematics (English)	1	1	1
Junior Mathematics (English)	0	0	0
Primary Mathematics (French)	0	0	0
Junior Mathematics (French)	0	0	0

*Long-writing prompts for the current year are not field tested in the previous year's operational test, so they are never used in the equating link. As such, they have been included in the number of items not used in equating.

Equating Results

The results of the equating process for the reading and mathematics components of the assessments are provided in Tables 5.3–5.6, and the results for the writing assessments are in tables 5.7 and 5.8. The theta cut scores and the number of students at each achievement level in 2008–2009 and 2009–2010 are reported for both English-language and French-language students. For example, the theta cut scores for the reading component of the English-language primary-division assessment were 1.16 for Levels 3 and 4, –0.40 for Levels 2 and 3, –1.48 for Levels 1 and 2 and –2.48 for “not enough evidence for Level 1” (NE1) and Level 1.

Since the 2008–2009 and 2009–2010 student thetas are on the same scale, the theta cut scores in the following tables apply to the assessments for both years.

Table 5.3 Equating Results for Reading: Primary Division (English and French)

	Primary Reading (English)			Primary Reading (French)		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		104 824	105 254		5954	6072
Level 4		7.9%	9.1%		19.7%	21.9%
Level 3	1.16	58.9%	59.9%	0.71	49.7%	50.4%
Level 2	-0.40	26.0%	24.5%	-0.50	25.0%	23.6%
Level 1	-1.48	6.3%	5.6%	-1.71	5.3%	3.8%
NE1	-2.48	1.0%	0.9%	-2.96	0.3%	0.3%
% of Students at or Above the Provincial Standard		66.8%	69.0%		69.4%	72.3%

Table 5.4 Equating Results for Reading: Junior Division (English and French)

	Junior Reading (English)			Junior Reading (French)		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		117 646	114 849		5387	5677
Level 4		12.2%	12.6%		16.2%	19.5%
Level 3	1.03	64.3%	66.3%	0.79	65.6%	65.1%
Level 2	-0.72	20.1%	18.1%	-0.94	17.3%	14.8%
Level 1	-1.86	3.3%	2.9%	-2.36	0.8%	0.7%
NE1	-3.07	0.2%	0.1%	-3.32	0.1%	0.0%
% of Students at or Above the Provincial Standard		76.5%	78.9%		81.9%	84.6%

Table 5.5 Equating Results for Mathematics: Primary Division (English and French)

	Primary Mathematics (English)			Primary Mathematics (French)		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		102 187	102 176		6028	6146
Level 4		13.8%	13.5%		13.5%	17.9%
Level 3	1.05	61.8%	62.9%	0.87	55.6%	57.4%
Level 2	-0.68	22.7%	21.7%	-0.63	27.6%	22.4%
Level 1	-2.08	1.7%	1.7%	-1.98	3.1%	2.2%
NE1	-2.78	0.1%	0.2%	-2.80	0.3%	0.2%
% of Students at or Above the Provincial Standard		75.6%	76.5%		69.0%	75.2%

Table 5.6 Equating Results for Mathematics: Junior Division (English and French)

	Junior Mathematics (English)			Junior Mathematics (French)		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		120 155	116 598		5864	5733
Level 4		14.7%	15.6%		37.4%	40.5%
Level 3	0.99	53.8%	51.2%	0.24	45.6%	45.1%
Level 2	-0.43	26.6%	28.5%	-1.02	16.8%	14.0%
Level 1	-1.58	4.8%	4.6%	-2.43	0.2%	0.2%
NE1	-3.09	0.1%	0.1%	-2.72	0.1%	0.2%
% of Students at or Above the Provincial Standard		68.5%	66.8%		83.0%	85.6%

Since the percentages of students in the 2008–2009 equating sample are very low at NE1 for the writing component of the French-language primary-division assessment and at NE1 and Level 1 for the writing component of the French-language junior-division assessment (<0.02%), the relevant cut points were determined by an alternative method for the 2009–2010 assessments:

- Primary French (0/1 cut point):
 - To determine the 0/1 cut point for the 2009–2010 assessment, the ratio of the distance between the 0/1 and 1/2 cuts to the distance between the 1/2 and 2/3 cuts on the 2008–2009 scale was applied to the 2009–2010 scale.
- Junior French (0/1 and 1/2 cut points):
 - First, to determine the 1/2 cut point for the 2009–2010 assessment, the ratio of the distance between the 1/2 and 2/3 cuts to the distance between the 2/3 and 3/4 cuts on the 2008–2009 scale was applied to the 2009–2010 scale.
 - Second, to further determine the 0/1 cut point for the 2009–2010 assessment, the ratio of the distance between the 0/1 and 1/2 cuts to the distance between the 1/2 and 2/3 cuts was applied to the 2009–2010 scale.

Table 5.7 Equating Results for Writing: Primary Division (English and French)

	Primary Writing (English)			Primary Writing (French)		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		105 316	105 747		6018	6118
Level 4		4.9%	5.1%		18.5%	23.5%
Level 3	1.45	68.0%	70.1%	0.65	61.3%	61.0%
Level 2	-0.62	27.0%	24.8%	-0.92	20.0%	15.3%
Level 1	-3.11	0.1%	0.1%	-2.62	0.2%	0.3%
NE1	-3.41	0.1%	0.0%	-3.52	0.0%	0.0%
% of Students at or Above the Provincial Standard		72.9%	75.2%		79.7%	84.5%

Table 5.8 Equating Results for Writing: Junior Division (English and French)

	Junior Writing (English)			Junior Writing (French)		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		117 902	115 099		5795	5697
Level 4		10.0%	10.3%		16.4%	17.0%
Level 3	1.14	64.2%	67.4%	0.91	67.2%	68.2%
Level 2	-0.67	25.5%	22.1%	-0.94	16.4%	14.8%
Level 1	-2.72	0.3%	0.2%	-3.31	0.0%	0.1%
NE1	-3.49	0.1%	0.0%	-3.69	0.0%	0.0%
% of Students at or Above the Provincial Standard		74.2%	77.7%		83.6%	85.2%

The Grade 9 Assessment of Mathematics

Description of the IRT Model

The 3PL model (Equation 1) and GPC model (Equation 2) were used to estimate item and ability parameters. For the Grade 9 academic and applied mathematics assessments, the 3PL model was modified by fixing the pseudo-guessing parameter at $0.20 [1/(k + 1)]$, where k is the number of options] for multiple-choice items, to reflect the possibility of students with a very low ability to correctly answer an item.

The academic and applied versions of the mathematics assessment are administered twice in one school year—in winter and in spring. The winter and spring assessments for each version have a set of common items and a set of unique items. The common items are used for equating across the winter and spring administrations.

Equating Sample

Equating samples for 2008–2009 and 2009–2010 were identified using a common set of selection rules. The equating samples for the academic and the applied courses were selected separately. However, the selection and exclusion rules for both courses and across the two years were the same. Students were excluded if they

- did not attend a publicly funded school;
- were home-schooled;
- completed no work on an item (by booklet and item type);
- received accommodations, except for English language learners who received the accommodation for setting, and
- did not attempt at least one question in each significant part of the test.

Table 5.9 presents the number of students in the Grade 9 equating and calibration samples for the 2008–2009 and the 2009–2010 assessments.

Table 5.9 Number of Grade 9 Students in the Equating Samples

Version	2008–2009				2009–2010	
	English		French		English	French
	Calibration Sample	Equating Sample	Calibration Sample	Equating Sample	Calibration and Equating Sample	Calibration and Equating Sample
Academic	77 424	98 438	3 403	3 856	98 943	4 025
Applied	34 143	41 114	1 189	1 289	40 330	1 268

Equating Steps

The calibration and equating of the 2008–2009 and 2009–2010 assessments for the English-language and the French-language student populations in both the academic and applied courses were conducted using the following steps:

1. A concurrent calibration was conducted simultaneously for the 2009–2010 winter and spring samples.
2. Calibration and equating were conducted for the 2008–2009 winter and spring samples (including the field-test items that were brought forward to the 2009–2010 operational test). In this calibration, the item parameter estimates of the field-test items that were brought forward to the 2009–2010 operational tests were fixed at the values obtained from the 2009–2010 calibration runs (Step 1). The parameter estimates of the 2008–2009 operational items that were repeated on the 2009–2010 operational tests were also fixed at the values obtained from the 2009–2010 calibration runs (Step 1). The 2008–2009 equating samples (operational items only) were scored using the scaled 2008–2009 operational item parameter estimates.
3. The percentage of students in each achievement level was determined for the 2008–2009 equating sample from the levels assigned in 2008–2009. The theta-value cut points that replicated this distribution were identified for each boundary (0/1, 1/2, 2/3 and 3/4).
4. These theta values were then used as the cut points for the 2009–2010 assessments.
5. The item parameter estimates for the 2009–2010 operational items obtained in Step 1 were used to score the full student population.
6. The cut-score points identified in Step 4 were applied to the 2009–2010 student theta values, students were assigned to levels, and the percentage of students at each performance level was determined.

Eliminating Items and Collapsing of Score Categories

For the Grade 9 mathematics assessments, 10 multiple-choice items were excluded from the equating process. Out of the 10 items, six were modified from the field-test to operational-test administrations, and four items were replaced with the 2010 matrix items. No items were dropped from any of the assessments (see Table 5.10).

Table 5.10 Number of Items Excluded from the Equating Process and Dropped: Grade 9 (2009–2010)

Assessment Version	No. of Items Excluded from Equating		No. of Items Dropped from the Assessment
	Multiple-Choice	Open-Response	
Applied, Winter (English)	0	0	0
Applied, Spring (English)	0	0	0
Academic, Winter (English)	0	0	0
Academic, Spring (English)	4	0	0
Applied, Winter (French)	3	0	0
Applied, Spring (French)	1	0	0
Academic, Winter (French)	1	0	0
Academic, Spring (French)	1	0	0

Equating Results

The equating results for the applied version of the Grade 9 mathematics assessments are summarized in Table 5.11. The results for the academic version of the Grade 9 mathematics assessments are summarized in Table 5.12. The theta cut scores and number of students in 2008–2009 and 2009–2010 at each achievement level are reported for both the English-language and French-language students. For example, the equated cut scores for the English-language applied version of the assessment were 1.35 for Levels 3 and 4, 0.13 for Levels 2 and 3, –0.89 for Levels 1 and 2 and –1.63 for “below Level 1” and Level 1.

Table 5.11 Equating Results for the Grade 9 Applied Mathematics Assessment

	English Applied			French Applied		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		41 114	40 330		1 289	1 268
Level 4		5.8%	7.4%		5.4%	6.1%
Level 3	1.35	36.5%	37.2%	1.46	37.1%	32.7%
Level 2	0.13	38.5%	37.8%	0.19	40.2%	45.3%
Level 1	-0.89	14.6%	14.2%	-0.96	13.7%	14.3%
Below Level 1	-1.63	4.6%	3.5%	-1.68	3.7%	1.6%
% of Students at or Above the Provincial Standard		42.3%	44.6%		42.4%	38.8%

Table 5.12 Equating Results for the Grade 9 Academic Mathematics Assessment

	English Academic			French Academic		
	Theta Cut Scores	Students at Each Level for Each Equating Sample		Theta Cut Scores	Students at Each Level for Each Equating Sample	
		2009	2010		2009	2010
Number of Students		98 438	98 943		3 856	4 025
Level 4		8.2%	10.4%		3.5%	3.3%
Level 3	1.24	70.6%	73.3%	1.73	65.9%	68.4%
Level 2	-0.95	15.1%	12.1%	-0.57	20.9%	18.8%
Level 1	-1.58	5.9%	4.1%	-1.23	9.6%	9.2%
Below Level 1	-2.66	0.1%	0.1%	-2.36	0.1%	0.2%
% of Students at or Above the Provincial Standard		78.9%	83.7%		69.5%	71.7%

The Ontario Secondary School Literacy Test (OSSLT)

Description of the IRT Model

In contrast to the primary-division, junior-division and Grade 9 assessments, both the a -parameter and the c -parameter (see Equation 1) were fixed for the OSSLT, yielding a modified Rasch model for multiple-choice items. The a -parameter for all multiple-choice and open-response items was set at 0.588. The pseudo-guessing parameter for multiple-choice items was set at 0.20 $[1/(k + 1)]$, where k is the number of options that reflect the possibility of students with a very low ability to correctly answer an item]. The GPC model (see Equation 2), with a constant slope parameter of 0.5888, was used to estimate the item and ability parameters for open-response items.

Equating Sample

First-time eligible students from both publicly funded and private schools were selected for the 2008–2009 and 2009–2010 equating samples. The following categories of students were excluded from the equating samples:

- students with no work or incomplete work on a major section of the test;
- students receiving the following accommodations on: assistive devices and technology, sign language, Braille, an audio recording or verbatim reading of the test, a computer, audio- or video -recorded responses and scribing;
- previously eligible students;
- students who were exempted, deferred or taking the Ontario Secondary School Literacy Course (OSSLC) and
- students who were home-schooled.

Table 5.13 presents the number of first-time eligible students in the OSSLT equating samples for the 2008–2009 and 2009–2010 tests.

Table 5.13 Number of First-Time Eligible OSSLT Students in the Equating Samples

OSSLT	2009		2010	
	English	French	English	French
First-Time Eligible	140 646	5164	137 608	4928

Equating Steps

The following steps were implemented to calibrate and equate the 2009 and 2010 OSSLT:

1. The parameter estimates of the operational items administered in 2010 were calibrated using the 2010 equating sample.
2. The operational items that formed the 2009 test and the field-test items brought forward to the 2010 test were recalibrated using the 2009 equating sample. In this calibration, the parameter estimates of the common items were fixed at the values obtained in Step 1.
3. The operational item parameter estimates of the 2009 test, obtained in Step 2, were used to score the 2009 equating sample data.
4. The percentage of successful students was determined for the 2009 equating sample from the student results reported in 2009. The theta-value cut point that replicated this percentage was identified for the distribution of scores in the 2009 equating sample.
5. This theta value was applied to student scores for the 2010 assessment to determine which students would be successful. The results are presented in Table 5.14.

Scale Score

The reporting scale scores for the 2010 OSSLT, which range from 200 to 400, were generated using a linear transformation. The slope and intercept were obtained by fixing two points: the theta value -4.0 was fixed at the lowest value of the scale score (200), and the theta cut score obtained from the equating steps was fixed at the scale score of 300.

Eliminating Items and Collapsing of Score Categories

Three multiple-choice items in the English-language OSSLT and six multiple-choice items in the French-language OSSLT were excluded from the equating process due to modifications made to the items between the 2009 field-test and the 2010 operational-test administrations.

The score 1.0 in the scoring rubric for long-writing prompts was collapsed with the score 1.5 for topic development and the use of conventions in both the English- and the French-language tests.

Equating Results

The equating results based on the equating samples for the OSSLT are summarized in Table 5.15. The theta cut score and the percentages of successful and unsuccessful students in 2009 and 2010 are reported for English-language and French-language students. For example, the equated cut score for the English-language test was -0.77 . The percentage of successful students in the equating samples was 87.2% in 2009 and 86.2% in 2010.

Table 5.14 Equating Results for the OSSLT

	English-Language			French-Language		
	Theta Cut Point 2009–2010	Equating Sample 2009	Equating Sample 2010	Theta Cut Point 2009–2010	Equating Sample 2009	Equating Sample 2010
No. of Students		140 646	137 608		5164	4928
% Successful	-0.77	87.2%	86.2%	-1.07	86.0%	85.2%
Unsuccessful		12.8%	13.8%		14.0%	14.8%

References

- Kolen, M. J. & Brennan, R. L. (2004). *Test Equating, Scaling, and Linking: Methods and Practices* (2nd ed.). New York: Springer-Verlag.
- Muraki, E. (1997). A generalized partial credit model. In W. J. Van der Linden, & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 153-164). New York: Springer-Verlag.
- Muraki, E. & Bock, R.D. (2003). PARSCALE: IRT item analysis and test scoring for rating-scale data. (Version 4.1) [Computer Software]. Chicago, IL: Scientific Software International.
- Yen, W. M. & Fitzpatrick, A. R. (2006). Item response theory. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp.111-153). Westport, CT: American Council on Education and Praeger Publishers.

CHAPTER 6: REPORTING RESULTS

EQAO assessment results are reported at the student, school, school board and provincial levels.

EQAO publishes an annual provincial report for education stakeholders and the general public. The reports for the 2009–2010 English-language assessments are available on www.eqao.com:

- *EQAO's Provincial Elementary School Report on the Results of the 2009–2010 Assessments of Reading, Writing and Mathematics, Primary Division (Grades 1–3) and Junior Division (Grades 4–6)* and
- *EQAO's Provincial Secondary School Report on the Results of the Grade 9 Assessment of Mathematics and the Ontario Secondary School Literacy Test, 2009–2010.*

Corresponding reports for the French-language assessments are also available.

EQAO posts school and board results on www.eqao.com for public access. However, EQAO does not publicly release school or board results when the number of students having written an assessment therein is small enough that individual students could be identified (i.e., fewer than 15 students for achievement results and fewer than six students for questionnaire results).

Two types of aggregate results are reported for schools, boards and the province:

1. the percentage of all students enrolled in Grades 3 and 6, students enrolled in Grade 9 academic and applied mathematics courses and students eligible to write the OSSLT and
2. the percentage of students who participated in each assessment.

More detailed school and board results are posted on the secure section of the EQAO Web site and are available only to school and school board personnel through user identification numbers and passwords. These reports include the results for small *n*-counts that are suppressed in the public reports and additional achievement results for sub-groups of the student population (i.e., English language learners, students with special education needs, French Immersion students in Grade 3). Results for male and female students are included in both the public and secure reports. In addition, schools and school boards receive data files with individual student achievement results for all their students and data files with aggregated results for each school, board and the province.

Reporting the Results of the Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions

The four achievement levels defined in *The Ontario Curriculum* are used by EQAO to report on student achievement in reading, writing and mathematics. Level 3 has been established as the provincial standard. Levels 1 and 2 indicate achievement below the provincial standard, and Level 4 indicates achievement above the provincial standard. There are three reporting categories in addition to the four performance levels, as follows:

- not enough evidence for Level 1 (NE1)
- no data
- exempt

Two sets of results are reported: those based on all students and those based on participating students. Students without data and exempted students are not included in the calculation of results for participating students.

The provincial, board and school reports provide the following:

- overall jurisdictional results for each component;
- longitudinal data showing jurisdictional results over time;
- overall jurisdictional results for each component by gender and other relevant characteristics (e.g., English language learners, special education needs);
- results by exceptionality category and results for students receiving accommodations;
- areas of strength with respect to the curriculum and areas for improvement;
- cohort-tracking results from Grade 3 to Grade 6;
- contextual data and
- student questionnaire results.

Results for the teacher and principal questionnaires are reported at the board and provincial levels. Some results for the questionnaires are included in the provincial reports. Full results for all questionnaires are posted on EQAO's public Web site.

In addition, schools receive Item Information Report: Student Rosters, which provide item results for each student who has completed each assessment and summary item statistics for the school, school board and province. The data for individual students are also provided in data files. Cohort-tracking reports for schools and boards show a comparison of Grade 3 and Grade 6 results for the same students.

Directors of education are provided with graphs that show the number and percentage of schools with achievement levels in specified categories (e.g., 75% of their students having achieved the provincial standard). The directors also receive school lists with achievement results over time, for convenient reference.

The Individual Student Report (ISR) for students in Grades 3 and 6 shows the overall achievement level for each component (reading, writing and mathematics) at one of five positions. For example, Level 1 includes the sub-categories 1.1, 1.3, 1.5, 1.7 and 1.9. The five sub-categories are created from the distribution of student theta scores. Through the equating process described in Chapter 5, four theta values are identified that define the cut points between adjacent achievement levels (i.e., NE1 and Level 1, Level 1 and Level 2, Level 2 and Level 3 and Level 3 and Level 4). The width of each sub-category in a given level is determined by the range of theta values represented in the level, divided by five. These results are designated in student data files accordingly as 1.1, 1.3, 1.5, 1.7, and so on, to 4.9. School, school board and provincial results are included on the ISR to provide a context for interpreting student results. For students in Grade 6, the assessment results they achieved in Grade 3, if available, are printed on their ISR.

Reporting the Results of the Grade 9 Assessment of Mathematics

Reporting for the Grade 9 mathematics assessment is very similar to that for the primary- and junior-division assessments. The same four achievement levels with five sub-categories are used

to report student achievement. The set of reports described above for the primary- and junior-division assessments is also produced for the Grade 9 assessment, with the exception of the cohort-tracking report. However, there are some differences in the reports for the Grade 9 assessment. For instance, the option to exempt students from the Grade 9 mathematics assessment was removed in 2007. Moreover, the reporting category “not enough evidence for Level 1” is called “below Level 1” for the Grade 9 assessment. In addition to the disaggregations identified for the primary- and junior-division assessments, results for the Grade 9 mathematics assessment are reported for Semester 1, Semester 2 and the full year. Furthermore, the ISR for Grade 9 does not include school and board results, and there is no principal questionnaire for the Grade 9 assessment. Summary cohort-tracking results from Grade 6 to Grade 9 are provided at the school, board and provincial levels, but the Grade 6 results for individual students are not recorded on the Grade 9 ISR.

Reporting the Results of the OSSLT

For the OSSLT, EQAO reports only two levels of achievement: successful and unsuccessful. A successful result on the OSSLT (or the successful completion of the OSSLC) is required to meet the literacy requirement for graduation. Students must achieve a minimum theta score to receive a successful result on the OSSLT. The process for establishing this minimum score is described in Chapter 5. EQAO provides feedback to unsuccessful students to assist them in working to achieve the minimum score.

As with the other assessments, EQAO reports results for all and for participating students. Students are considered to be “not participating” if they were deferred, opted to take the OSSLC or have no data for the current administration. Students who are not working toward the OSSD are exempt from the OSSLT and are not included in either reported population. Aggregated results are reported separately for first-time eligible students and previously eligible students. Previously eligible students are those who were unsuccessful on a previous administration, were deferred from a previous administration or arrived in an Ontario school during their Grade 11 or 12 year.

The OSSLT provincial, board and school reports provide the following:

- overall successful and unsuccessful jurisdictional results;
- overall successful and unsuccessful jurisdictional results by gender and other characteristics (e.g., English language learners, special education needs);
- results by exceptionality category and results for students receiving accommodations;
- results by type of English- or French-language course—academic, applied, locally developed, English as a second language (ESL) or English literacy development (ELD), “actualisation linguistique en français” (ALF) or “programme d’appui aux nouveaux arrivants” (PANA);
- longitudinal data showing jurisdictional results over time;
- areas of strength relative to the curriculum and areas for improvement;
- cohort-tracking results from Grade 6 to the OSSLT and
- results for the student questionnaire and contextual data.

In addition, schools receive the student rosters, which provide item results for each student who completed the test and summary item statistics for the school, board and province.

The OSSLT ISR provides the following:

- the statement of a successful or unsuccessful result;
- the student's scale score;
- the median scale score for the school and province and
- feedback for students on areas of strength and areas for improvement.

Each unsuccessful student is informed that a successful result requires a scale score of 300.

Interpretation Guides

Guides for interpreting results are included in the school and board reports, and released test items and scoring guides are posted on the EQAO Web site.

EQAO also produces the following resource documents:

- "Using Data to Promote Student Success: A Brief Guide to Assist School Administrators in Interpreting Their Data";
- the "Guide to Using the Item Information Report: Student Roster";
- "EQAO Questionnaire Data: Using Contextual Data to Inform Improvement Planning" and
- the "Technical Paper: School and School Board Profiles of Strengths and Areas for Improvement."

CHAPTER 7: STATISTICAL AND PSYCHOMETRIC SUMMARIES

A variety of statistical and psychometric analyses were conducted for the 2009–2010 assessments. The results from these analyses are summarized in this chapter, including results for Classical Test Theory (CTT), Item Response Theory (IRT), Differential Item Functioning (DIF) and decision accuracy and consistency. All IRT item parameter estimates were obtained from the calibration process used for the equating samples (described in Chapter 5). Detailed data for individual items appear in Appendix 7.1.

The Assessments of Reading, Writing and Mathematics, Primary and Junior Divisions

Classical Test Theory (CTT) Analysis

Table 7.1 presents descriptive statistics, Cronbach's alpha estimates of test score reliability and the standard error of measurement (SEM) for the English-language and French-language versions of the primary and junior assessments. The test means (converted to percentages) ranged from 59.0% (for junior English-language mathematics) to 71.2% (for junior French-language writing).

Reliability and the corresponding SEMs refer to the precision of test scores, with higher reliability coefficients and lower SEMs indicating higher levels of precision. For the primary and junior assessments, Cronbach's alpha estimates range from 0.85 to 0.87 for reading, 0.79–0.81 for writing and 0.88–0.89 for mathematics. The corresponding standard errors of measurement range from 4.7% to 5.0% of the possible maximum score for reading, 6.7–7.0% of the possible maximum score for writing and 5.9–6.4% of the possible maximum score for mathematics. The reliability coefficients for writing are a little lower than those for reading and mathematics, and the corresponding SEMs are a little higher. This is attributable, in part, to the smaller number of writing items and the subjectivity in scoring writing performance.¹ Taking these two factors into account, the obtained reliability coefficients and standard errors of measurement are acceptable and indicate that the test scores from these assessments provide a satisfactory level of precision.

¹For example, if the number of writing items was increased to match the number of reading items, Cronbach's alpha estimate for the French-language version of the primary writing assessment would be about 0.91.

Table 7.1 Test Descriptive Statistics, Reliability and Standard Error of Measurement: Primary and Junior Divisions

Assessment	No. of Items	Item Type		Max. Score	No. of Students	Min.	Max.	Mean	SD	Alpha	SEM
		MC	OR								
Primary Reading (English)	36	26	10	66	119 249	0	66	39.99	9.09	0.87	3.28
Junior Reading (English)	36	26	10	66	130 171	0	66	40.61	8.60	0.87	3.10
Primary Reading (French)	36	26	10	66	6 588	5	62	41.21	8.16	0.85	3.16
Junior Reading (French)	36	26	10	66	6 145	5	63	43.25	8.11	0.85	3.14
Primary Writing (English)	14	8	6*	29	119 966	0	29	19.27	4.57	0.81	1.99
Junior Writing (English)	14	8	6*	29	130 465	0	29	19.17	4.52	0.81	1.97
Primary Writing (French)	14	8	6*	29	6 595	1	29	17.37	4.40	0.79	2.02
Junior Writing (French)	14	8	6*	29	6 169	0	29	20.64	4.42	0.81	1.93
Primary Mathematics (English)	35	28	7	56	123 568	0	56	34.41	10.09	0.88	3.49
Junior Mathematics (English)	36	28	8	60	129 882	0	60	35.39	11.56	0.89	3.83
Primary Mathematics (French)	36	28	8	60	6 612	4	60	39.16	10.16	0.88	3.52
Junior Mathematics (French)	36	28	8	60	6 160	2	60	38.18	10.29	0.88	3.56

Note. MC = multiple choice; OR = open response; SD = standard deviation; Alpha = Cronbach's alpha; SEM = standard error of measurement.

*Short writing and long writing.

Item Response Theory (IRT) Analysis

In the IRT analysis, item parameters are estimated from the student responses used in the equating calibration sample (see Chapter 5). The estimated item parameters are then used to score all student responses. The descriptive statistics for the IRT scores reported in Table 7.2 refer to the total population. The mean student ability scores are less than zero, and the standard deviations are less than one (due to the inclusion of all students). The item parameter estimates for individual items are presented in Appendix 7.1.

Table 7.2 Test Descriptive Statistics of IRT Scores: Primary and Junior Divisions

Assessment	No. of Students	Min.	Max.	Mean	SD
Primary Reading (English)	119 249	-3.85	3.49	-0.11	1.00
Junior Reading (English)	130 171	-3.93	3.55	-0.13	1.00
Primary Reading (French)	6 588	-3.88	2.86	-0.07	0.96
Junior Reading (French)	6 145	-3.84	2.77	-0.08	0.97
Primary Writing (English)	119 966	-3.53	2.42	-0.06	0.93
Junior Writing (English)	130 465	-3.68	2.37	-0.13	0.97
Primary Writing (French)	6 595	-3.67	2.78	-0.05	0.94
Junior Writing (French)	6 169	-3.73	2.12	-0.08	0.95
Primary Mathematics (English)	123 568	-3.69	2.60	-0.09	0.99
Junior Mathematics (English)	129 882	-3.68	2.62	-0.10	0.99
Primary Mathematics (French)	6 612	-3.68	2.55	-0.06	0.97
Junior Mathematics (French)	6 160	-3.66	2.65	-0.06	0.97

Note. SD = standard deviation.

The Test Characteristic Curves (TCCs) and the distribution of student thetas are provided in Figures 7.1–7.12. The TCCs slope upward from the lower left to the upper right. These curves can be used to translate a student ability score in IRT to a student ability score in CTT, as indicated by the left vertical axis. For example, a primary student with a theta score of -1.0 in English-language reading is expected to have an observed score of about 50%. The right vertical scale indicates the percentage of students at each theta value, and the theta cut points used to assign students to performance levels are marked on the graphs. The Test Information Functions (TIFs), which indicate where the components of each assessment contain the most information, are provided in Figures 7.13–7.24. For example, the maximum information provided by the reading component of the English-language primary-division assessment is at approximately theta -0.80 . The precision of the scores is greatest at this point. The theta cut points used to assign students to achievement levels are also marked to show the amount of information at each cut point.

Figure 7.1 Test Characteristic Curve and Distribution of Thetas for Reading: Primary Division (English)

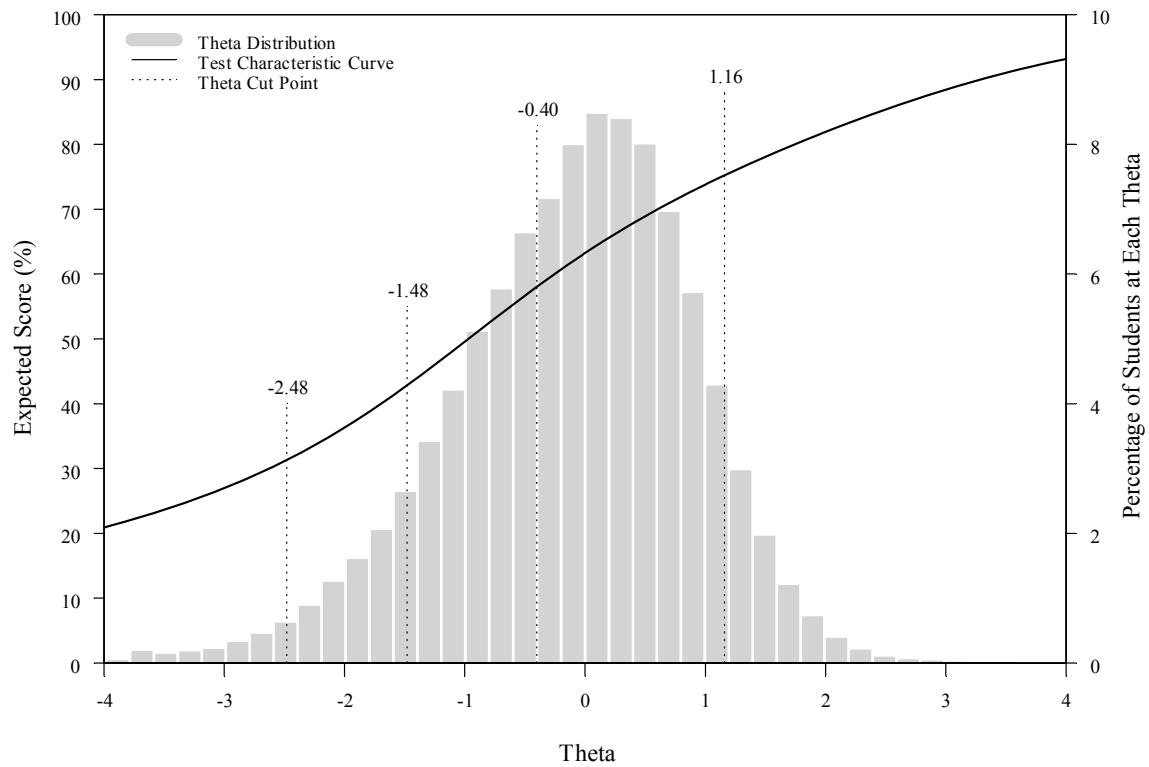


Figure 7.2 Test Characteristic Curve and Distribution of Thetas for Reading: Junior Division (English)

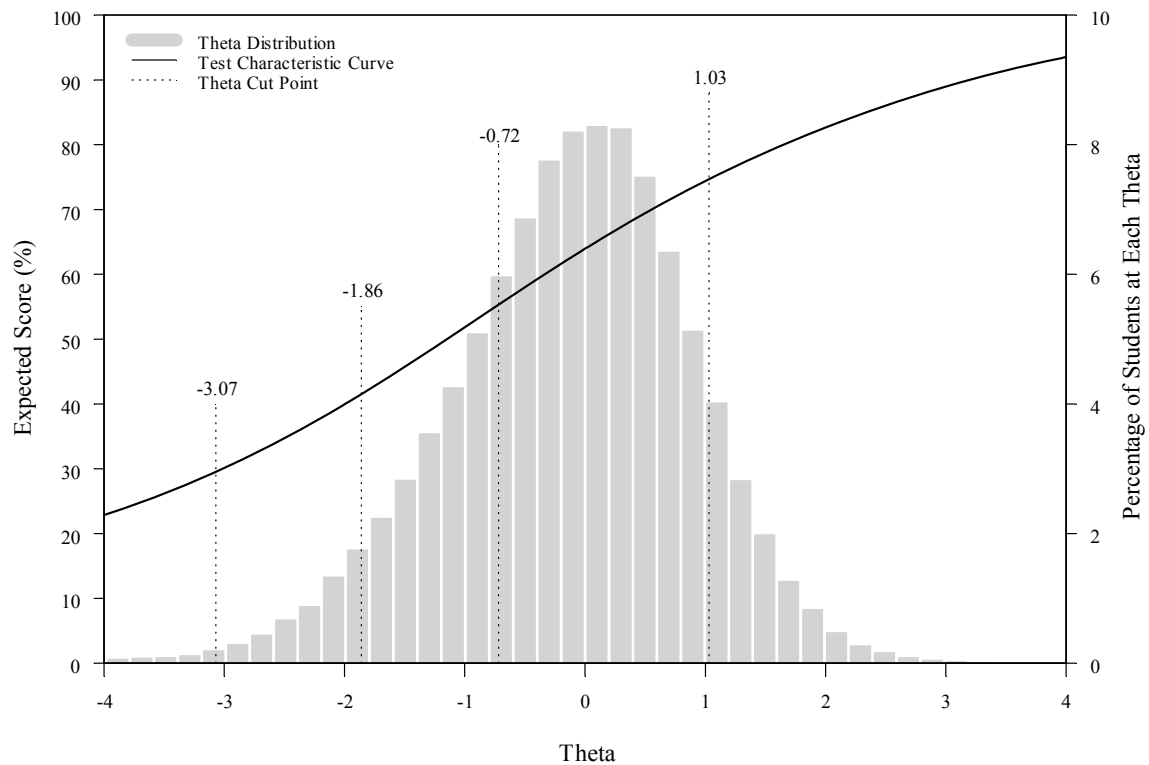


Figure 7.3 Test Characteristic Curve and Distribution of Thetas for Reading: Primary Division (French)

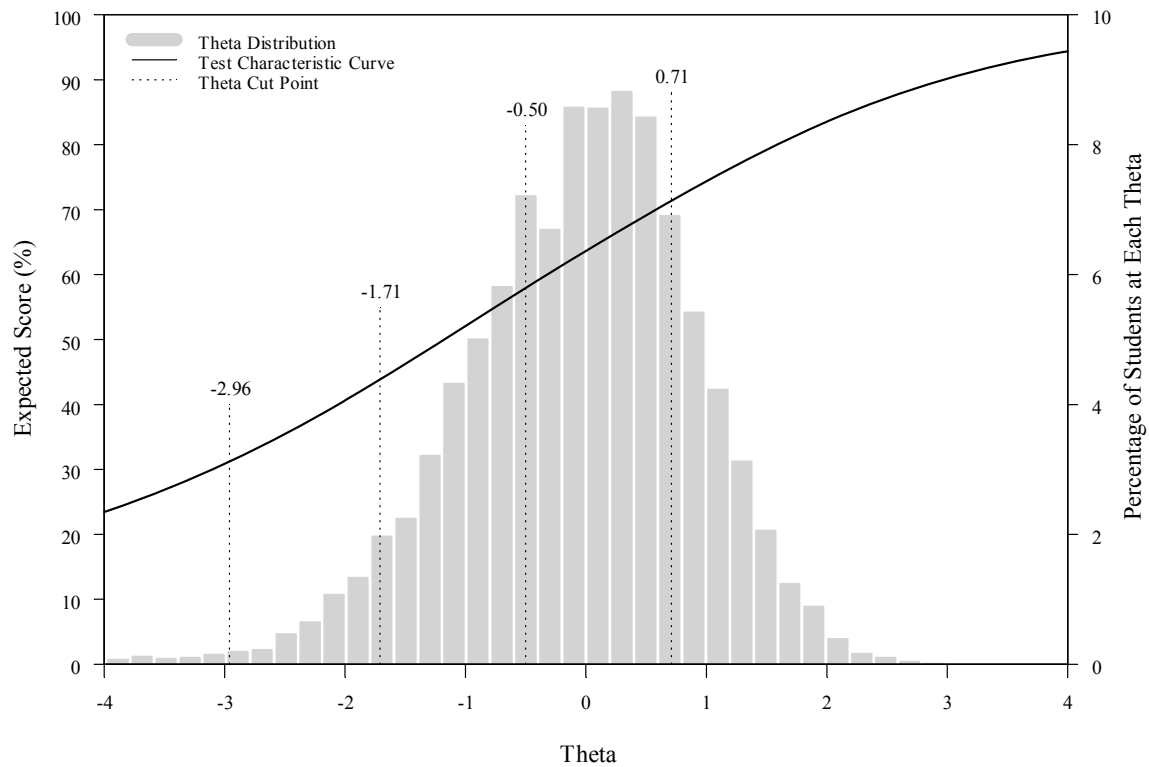


Figure 7.4 Test Characteristic Curve and Distribution of Thetas for Reading: Junior Division (French)

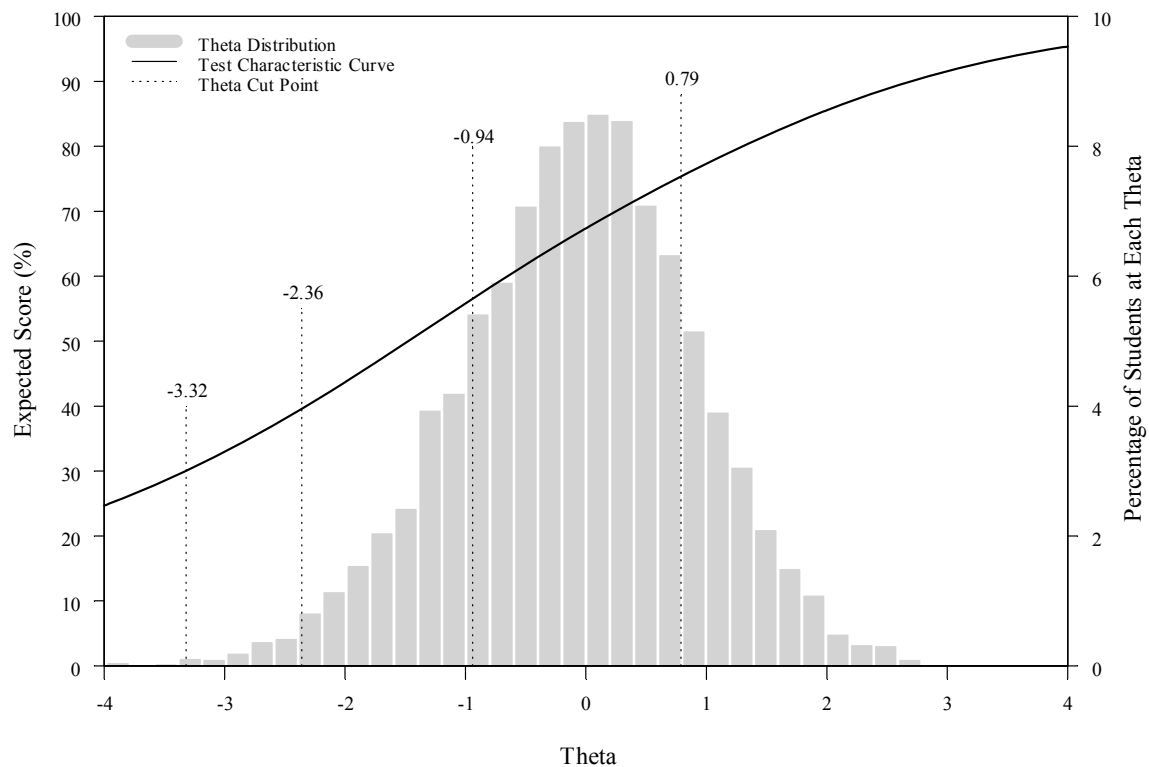


Figure 7.5 Test Characteristic Curve and Distribution of Thetas for Writing: Primary Division (English)

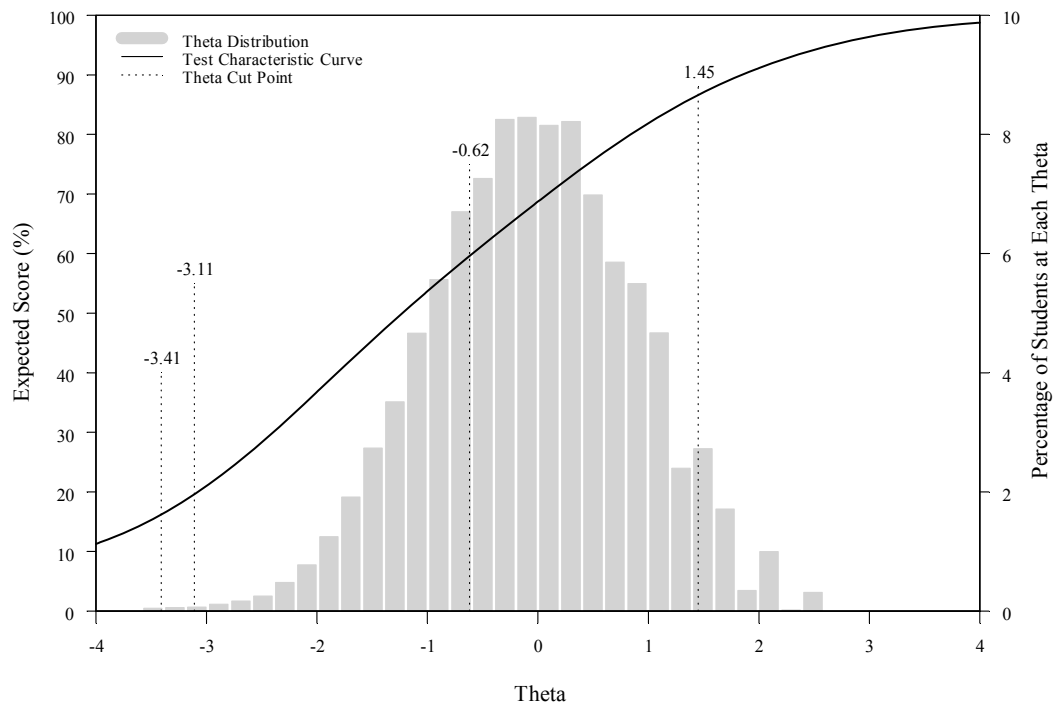


Figure 7.6 Test Characteristic Curve and Distribution of Thetas for Writing: Junior Division (English)

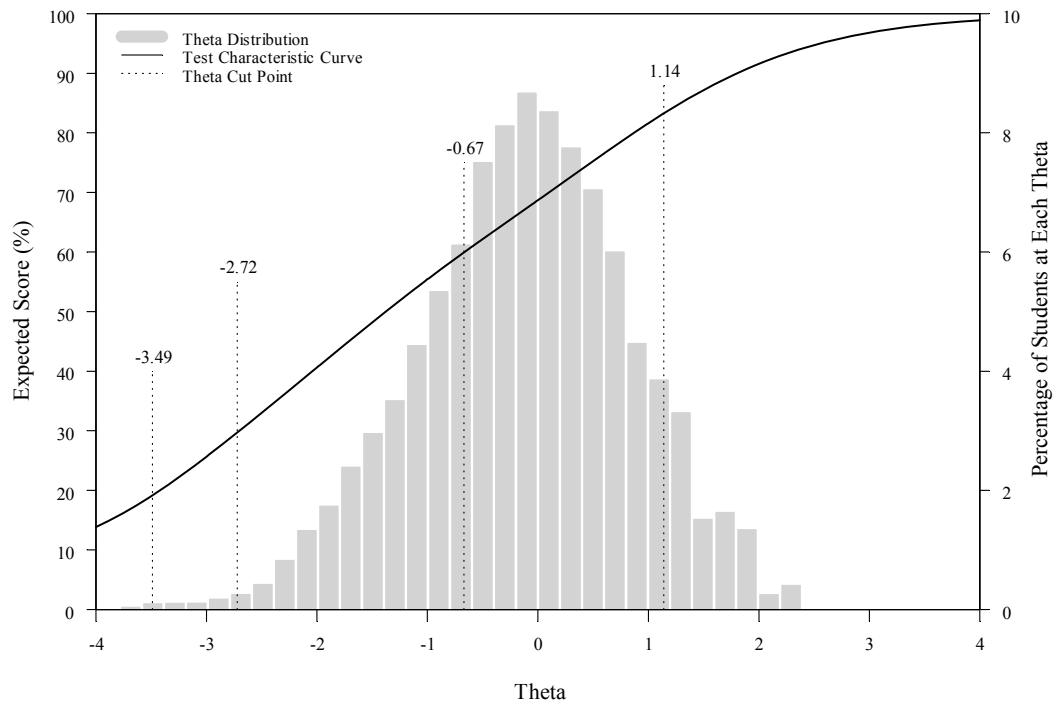


Figure 7.7 Test Characteristic Curve and Distribution of Thetas for Writing: Primary Division (French)

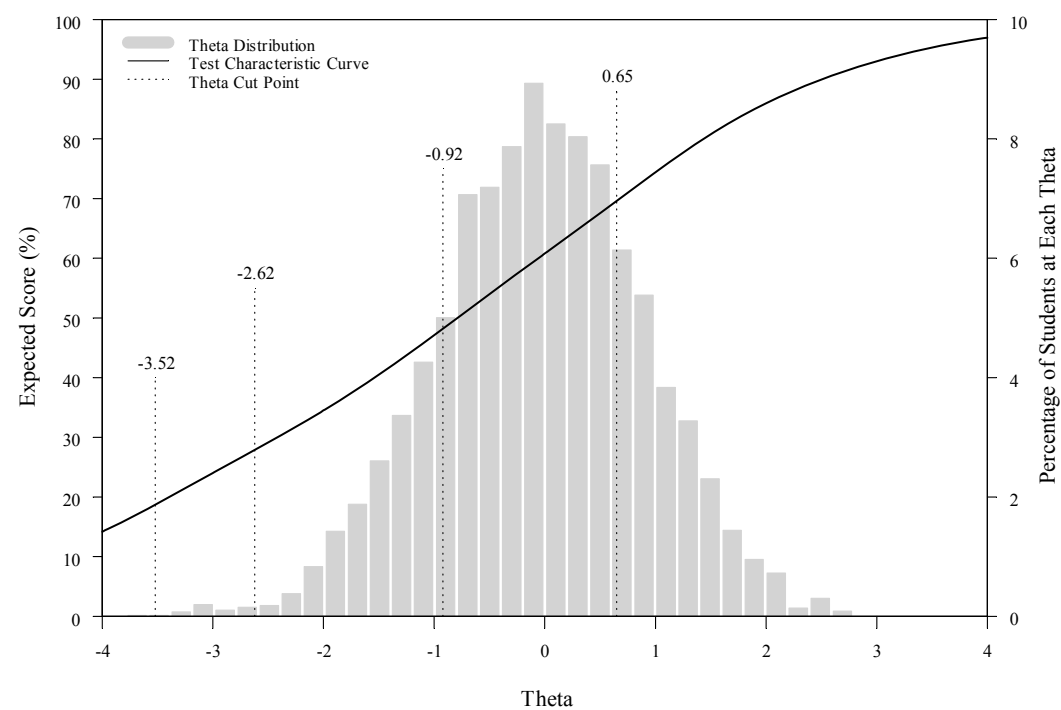


Figure 7.8 Test Characteristic Curve and Distribution of Thetas for Writing: Junior Division (French)

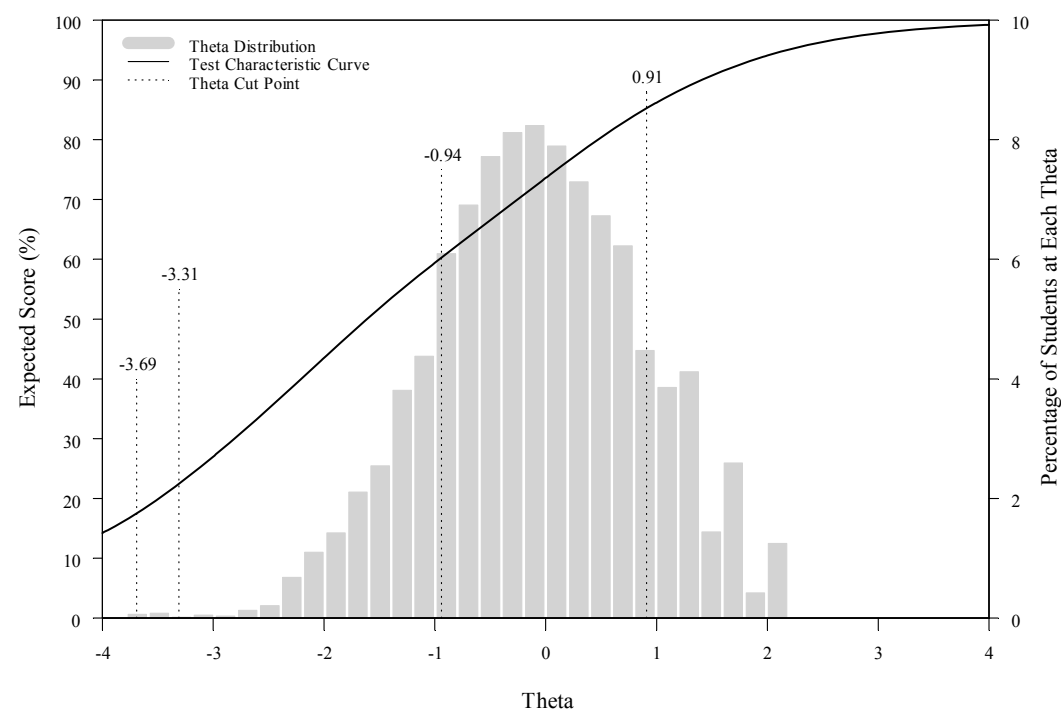


Figure 7.9 Test Characteristic Curve and Distribution of Thetas for Mathematics: Primary Division (English)

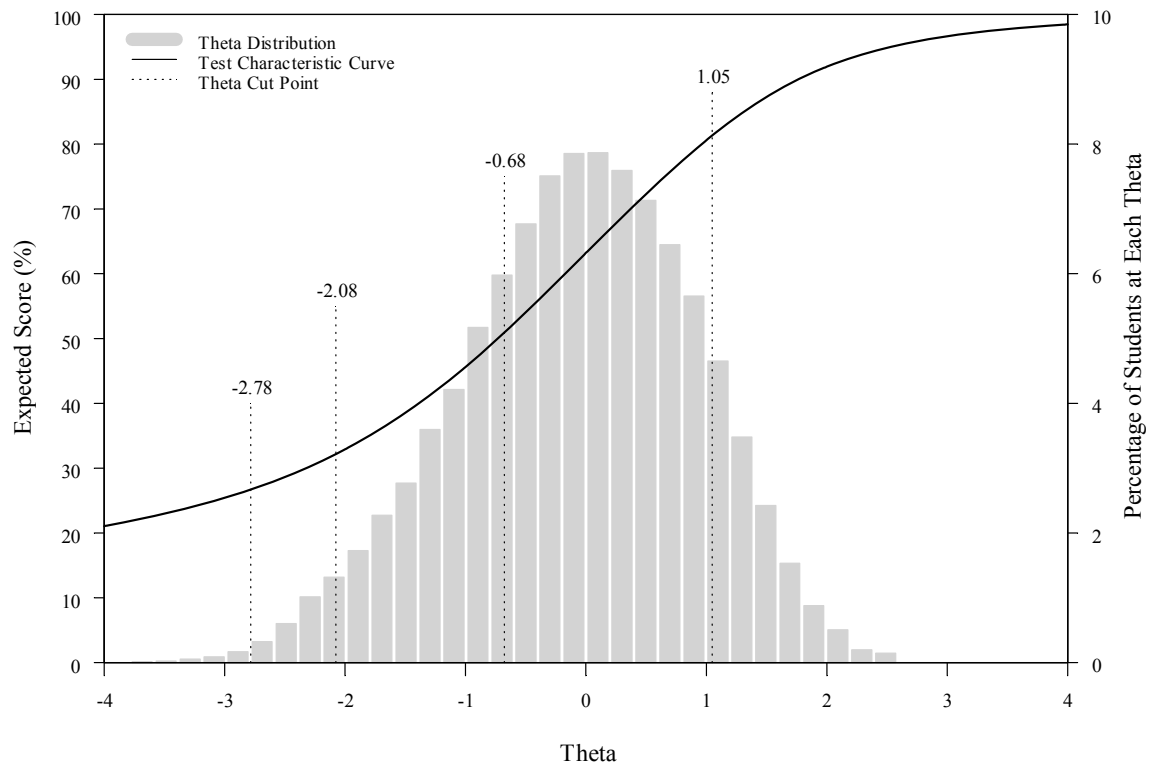


Figure 7.10 Test Characteristic Curve and Distribution of Thetas for Mathematics: Junior Division (English)

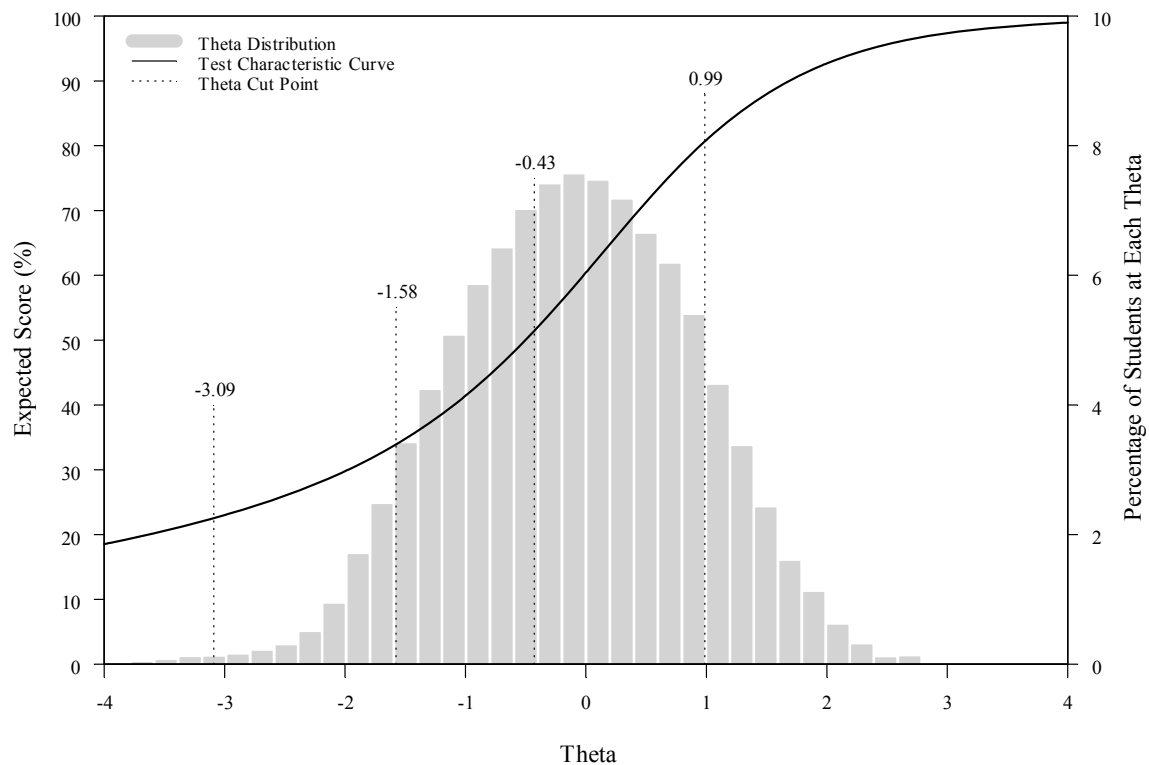


Figure 7.11 Test Characteristic Curve and Distribution of Thetas for Mathematics: Primary Division (French)

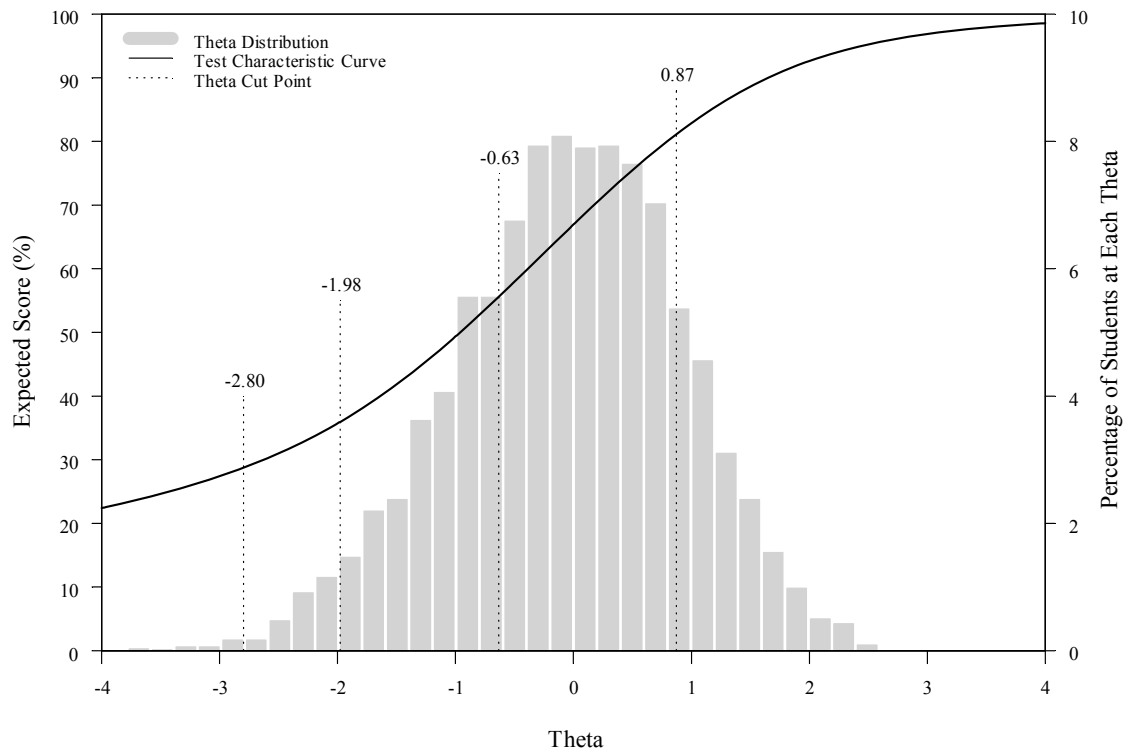


Figure 7.12 Test Characteristic Curve and Distribution of Thetas for Mathematics: Junior Division (French)

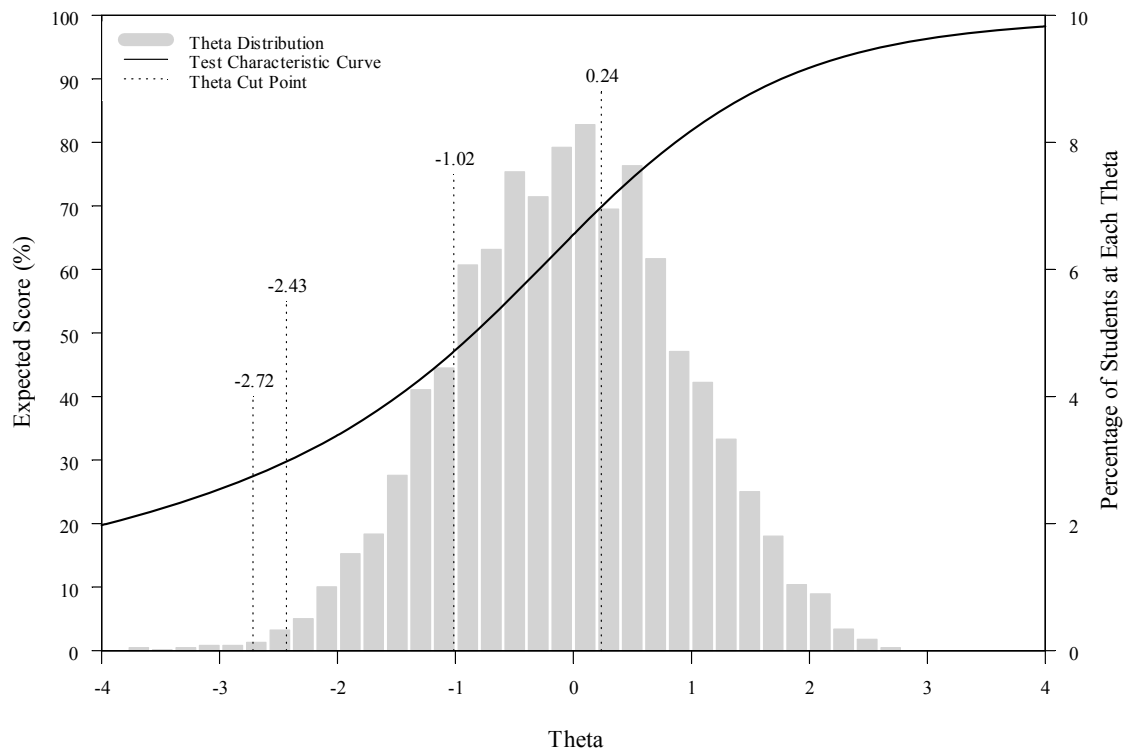


Figure 7.13 Test Information Function for Reading: Primary Division (English)

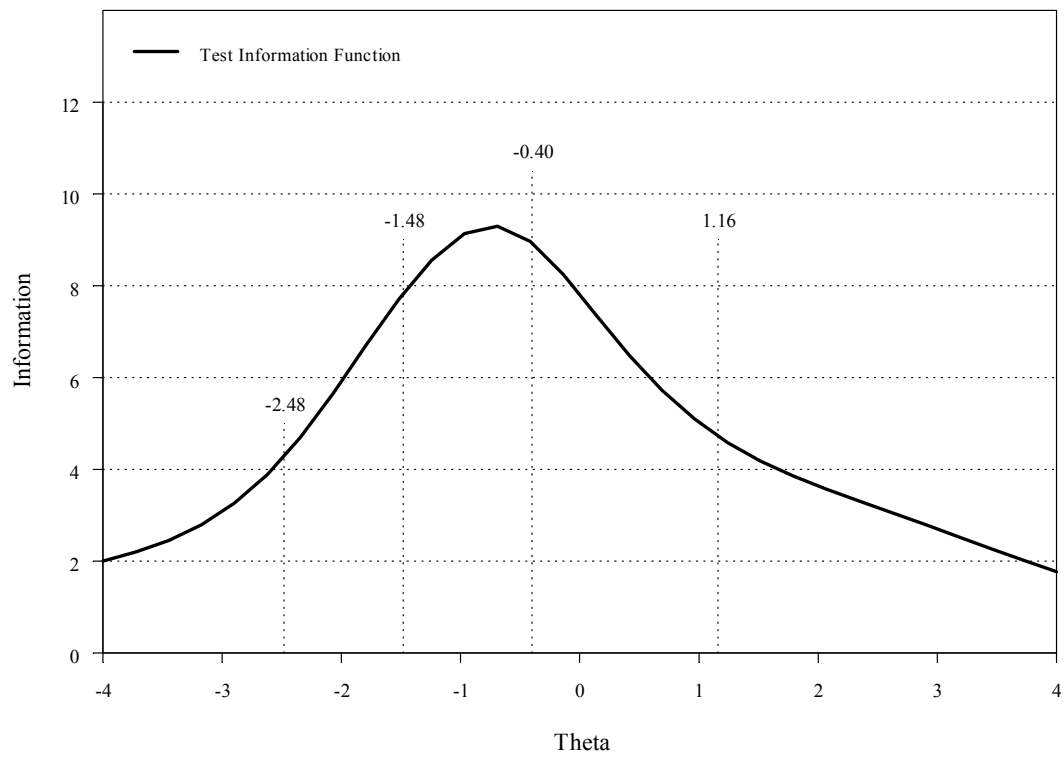


Figure 7.14 Test Information Function for Reading: Junior Division (English)

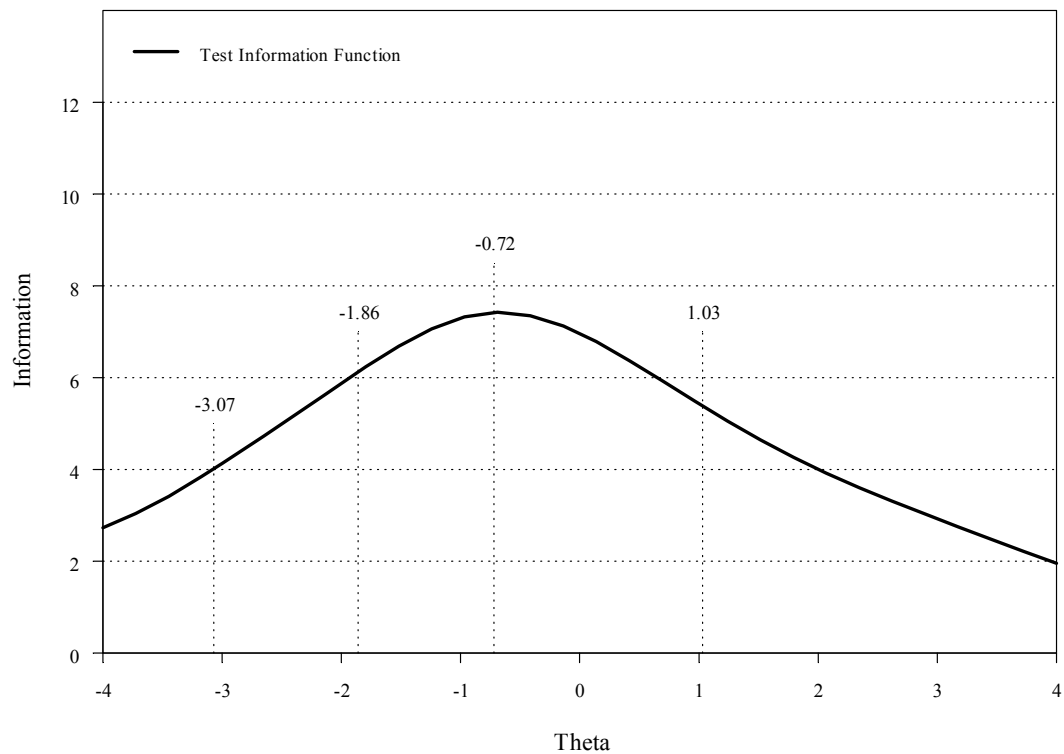


Figure 7.15 Test Information Function for Reading: Primary Division (French)

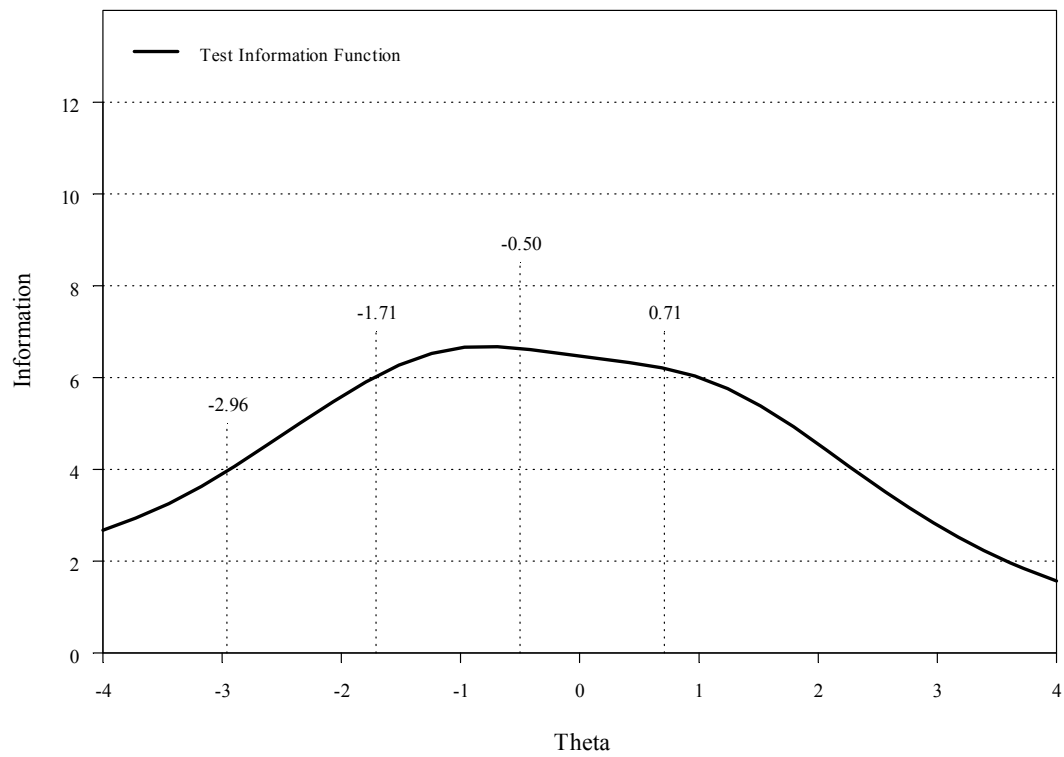


Figure 7.16 Test Information Function for Reading: Junior Division (French)

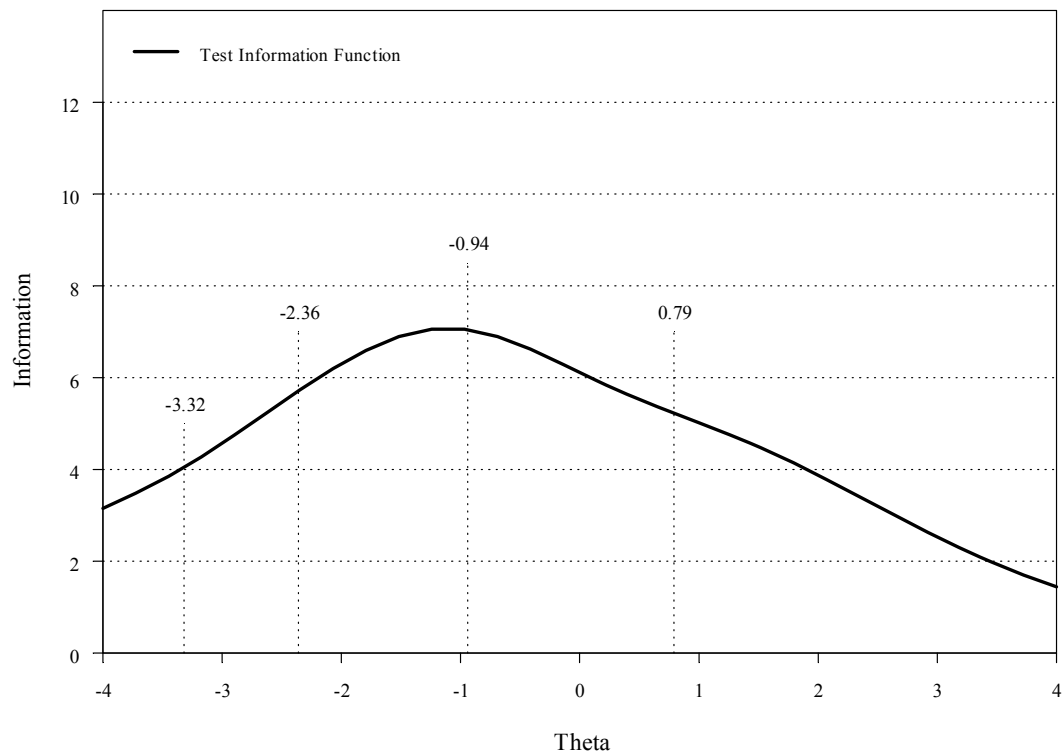


Figure 7.17 Test Information Function for Writing: Primary Division (English)

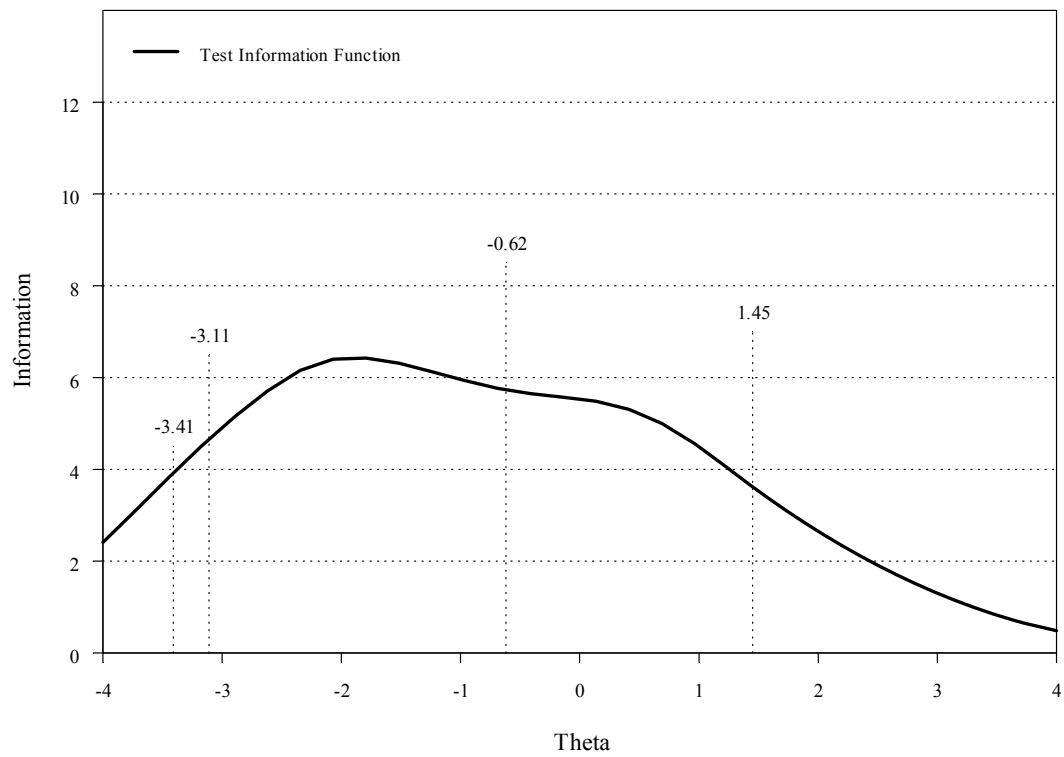


Figure 7.18 Test Information Function for Writing: Junior Division (English)

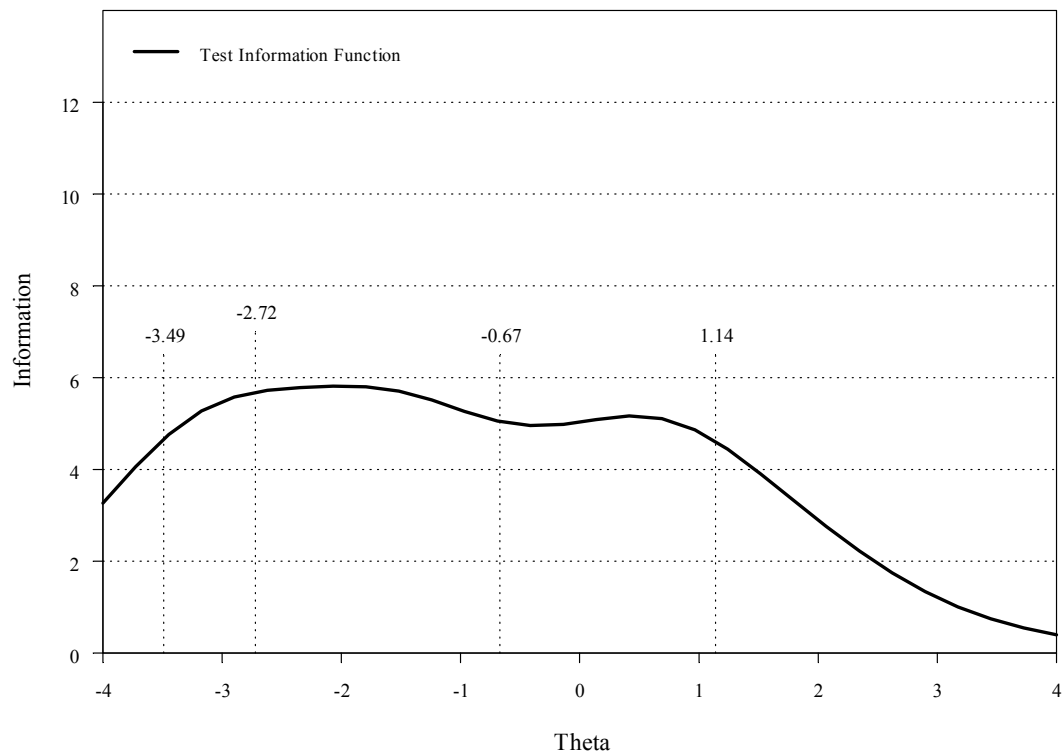


Figure 7.19 Test Information Function for Writing: Primary Division (French)

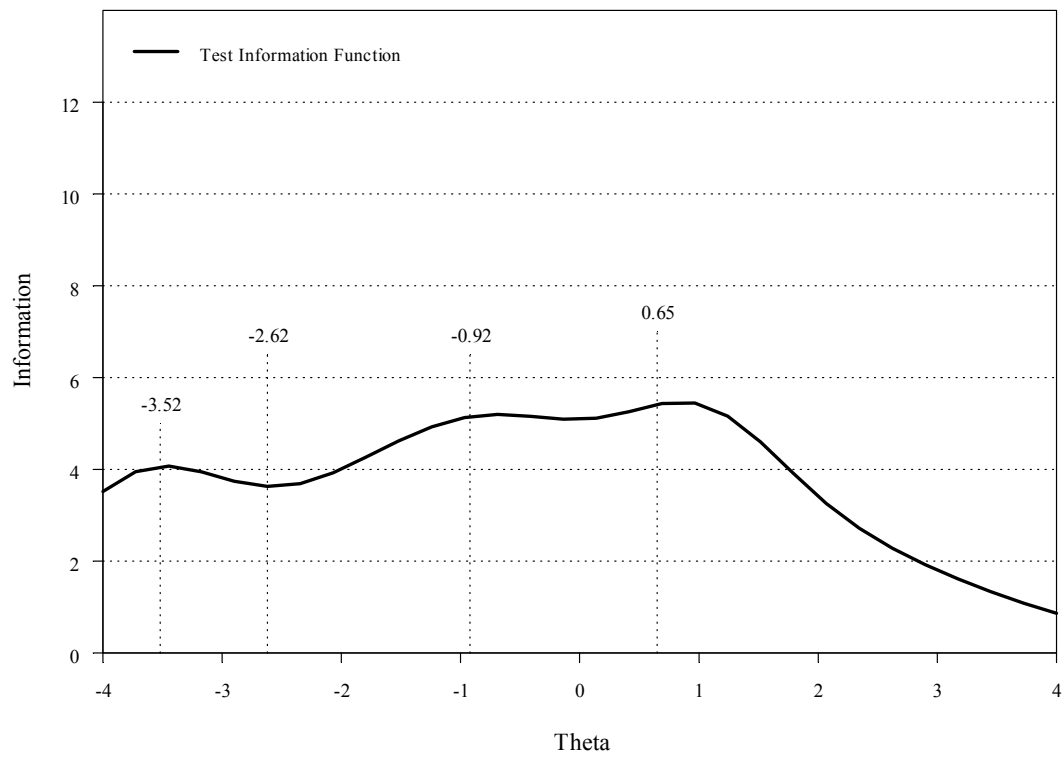


Figure 7.20 Test Information Function for Writing: Junior Division (French)

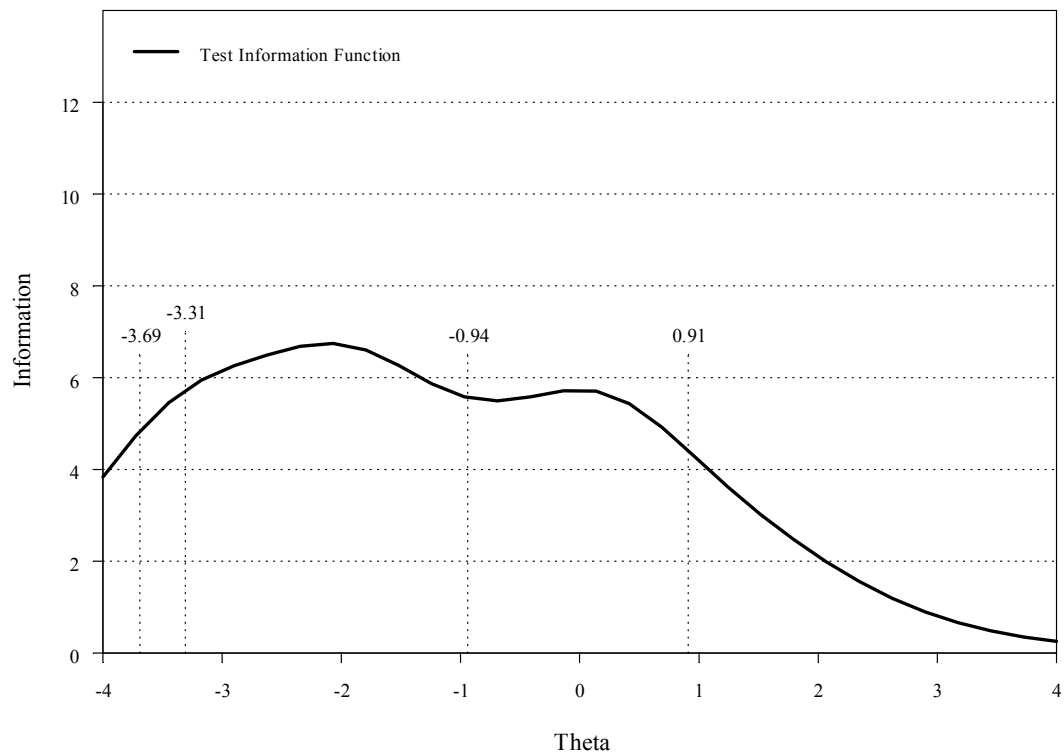


Figure 7.21 Test Information Function for Mathematics: Primary Division (English)

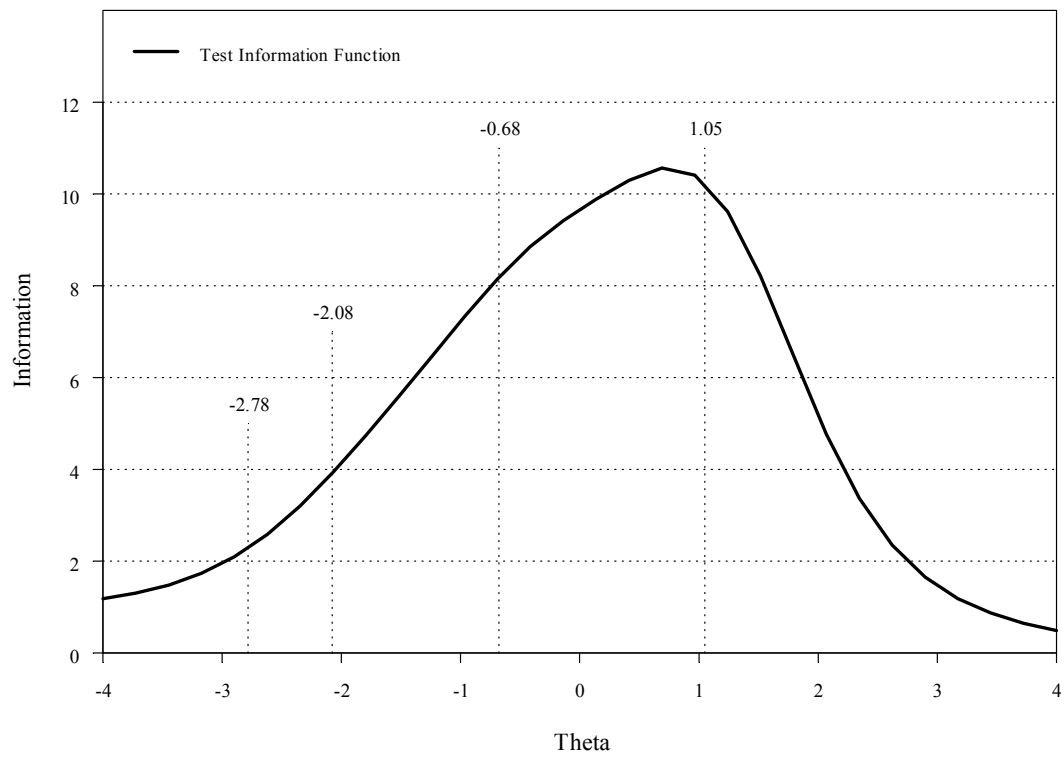


Figure 7.22 Test Information Function for Mathematics: Junior Division (English)

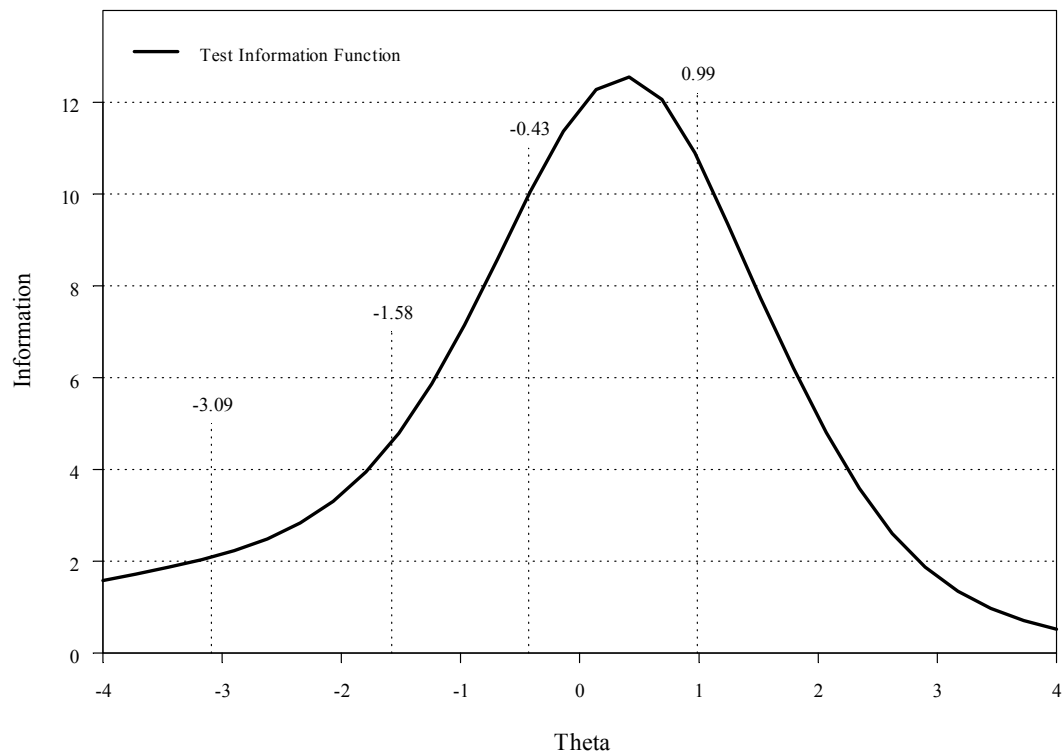


Figure 7.23 Test Information Function for Mathematics: Primary Division (French)

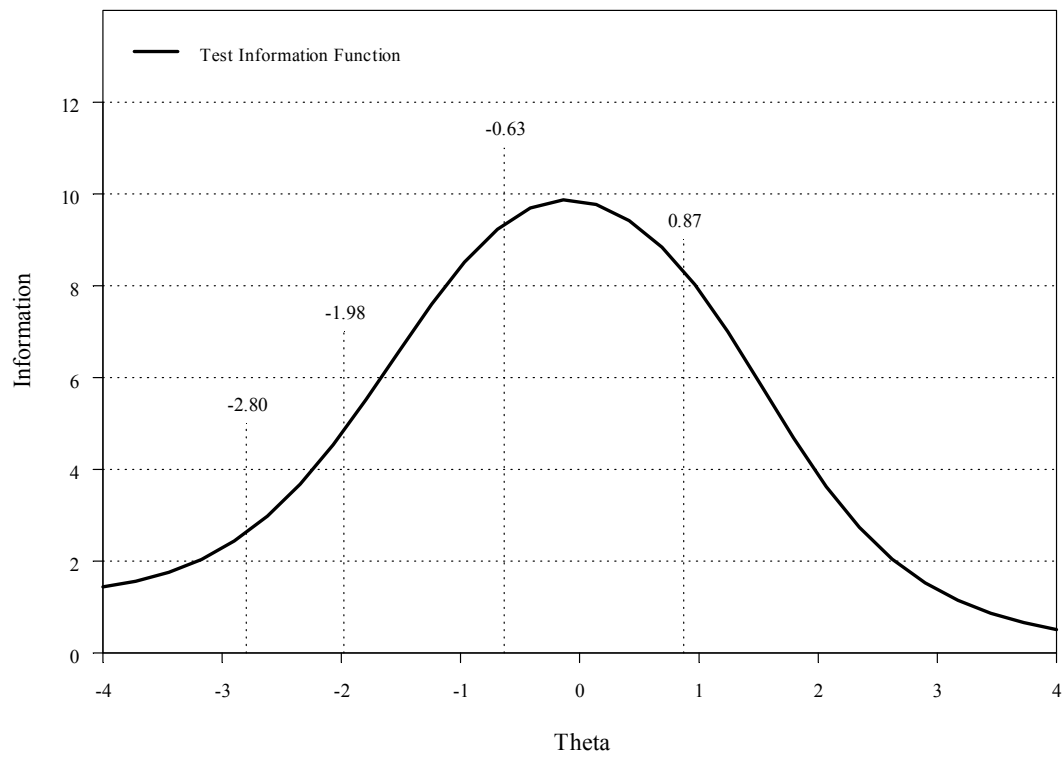
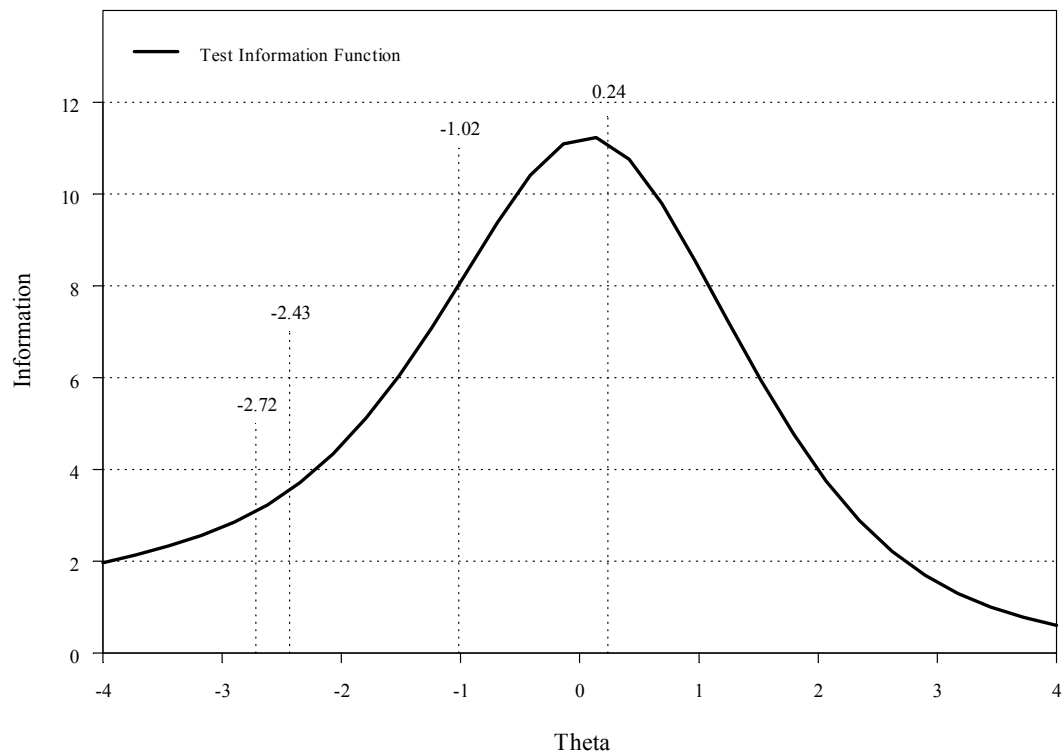


Figure 7.24 Test Information Function for Mathematics: Junior Division (French)



Descriptive Item Statistics for Classical Test Theory (CTT) and Item Response Theory (IRT)

Table 7.3 contains a summary of both the CTT and IRT descriptive item statistics for the items included in the English-language and French-language versions of the primary- and junior-division assessments. These statistics were computed using the equating sample (see Chapter 5). As expected, there is an inverse relationship between the CTT item difficulty estimates and the IRT location parameter estimates, due to the difference between the difficulty definitions in the two approaches. In contrast, there is a positive relationship between the CTT item-total correlation estimates and the IRT slope parameter estimates. Statistics for individual items are presented in Appendix 7.1.

Table 7.3 Descriptive Statistics of CTT Item Statistics and IRT Item Parameter Estimates: Primary and Junior Divisions

Assessment	No. of Items	Descriptive Statistics	CTT Item Statistics		IRT Item Parameters*	
			Item Difficulty	Item-Total Correlation	Slope	Location
Primary Reading (English)	36	Min.	0.43	0.26	0.19	-2.77
		Max.	0.92	0.52	1.37	0.68
		Mean	0.68	0.39	0.66	-0.90
		SD	0.15	0.07	0.25	0.76
Junior Reading (English)	36	Min.	0.41	0.26	0.36	-2.94
		Max.	0.97	0.54	1.08	0.49
		Mean	0.69	0.38	0.62	-1.04
		SD	0.15	0.07	0.17	0.88
Primary Reading (French)	36	Min.	0.35	0.23	0.29	-2.56
		Max.	0.96	0.53	1.03	1.20
		Mean	0.67	0.35	0.62	-0.90
		SD	0.15	0.07	0.19	0.91
Junior Reading (French)	36	Min.	0.33	0.16	0.29	-4.03
		Max.	0.97	0.48	0.93	1.76
		Mean	0.70	0.35	0.61	-1.16
		SD	0.17	0.09	0.18	1.20

Note. SD = standard deviation.

*The guessing parameter was set at a constant of 0.2 for multiple-choice items.

Table 7.3 Descriptive Statistics of CTT Item Statistics and IRT Item Parameter Estimates: Primary and Junior Divisions (continued)

Assessment	No. of Items	Descriptive Statistics	CTT Item Statistics		IRT Item Parameters*	
			Item Difficulty	Item-Total Correlation	Slope	Location
Primary Writing (English)	14	Min.	0.49	0.25	0.42	-2.09
		Max.	0.89	0.59	1.07	0.74
		Mean	0.69	0.43	0.73	-0.98
		SD	0.12	0.13	0.18	0.79
Junior Writing (English)	14	Min.	0.48	0.23	0.37	-2.87
		Max.	0.95	0.61	1.05	0.87
		Mean	0.69	0.42	0.69	-1.12
		SD	0.13	0.16	0.22	1.04
Primary Writing (French)	14	Min.	0.48	0.19	0.28	-1.88
		Max.	0.87	0.59	1.16	0.17
		Mean	0.64	0.40	0.66	-0.71
		SD	0.11	0.15	0.29	0.67
Junior Writing (French)	14	Min.	0.63	0.23	0.34	-2.61
		Max.	0.86	0.60	1.10	-0.45
		Mean	0.73	0.42	0.69	-1.32
		SD	0.07	0.16	0.27	0.54
Primary Mathematics (English)	35	Min.	0.24	0.27	0.27	-2.81
		Max.	0.93	0.69	1.46	1.60
		Mean	0.61	0.46	0.83	-0.49
		SD	0.18	0.09	0.31	1.12
Junior Mathematics (English)	36	Min.	0.15	0.21	0.22	-3.66
		Max.	0.93	0.71	1.33	1.74
		Mean	0.58	0.45	0.70	-0.41
		SD	0.14	0.12	0.25	1.07
Primary Mathematics (French)	36	Min.	0.31	0.24	0.33	-2.70
		Max.	0.90	0.66	1.10	1.36
		Mean	0.65	0.44	0.72	-0.69
		SD	0.15	0.10	0.23	1.08
Junior Mathematics (French)	36	Min.	0.43	0.19	0.23	-3.41
		Max.	0.82	0.65	1.10	0.90
		Mean	0.63	0.45	0.69	-0.55
		SD	0.11	0.11	0.22	0.94

Note. SD = standard deviation.

*The guessing parameter was set at a constant of 0.2 for multiple-choice items.

The Grade 9 Assessment of Mathematics

Classical Test Theory (CTT) Analysis

Table 7.4 presents descriptive statistics, Cronbach's alpha estimates of test score reliability and the SEM for the English-language and French-language applied and academic versions of the Grade 9 mathematics assessment. In 2009–2010 the mean percentages ranged from 52.87% (French, spring) to 54.13% (English, spring) for applied mathematics. The mean percentages ranged from 59.02% (French, spring) to 65.71% (English, spring) for academic mathematics.

Cronbach's alpha estimates ranged from 0.81 to 0.84 for applied mathematics and from 0.84 to 0.87 for academic mathematics. The corresponding SEMs were likewise similar, ranging from 7.08 to 7.57% of the possible maximum score for applied mathematics and 6.82–7.29% of the possible maximum score for academic mathematics. The obtained reliability coefficients and SEMs are acceptable, which indicates that the test scores from these assessments provide a satisfactory level of precision.

Table 7.4 Test Descriptive Statistics, Reliability and Standard Error of Measurement: Grade 9 Mathematics

Assessment	No. of Items	Item Type		Possible Max. Score	No. of Students	Min.	Max.	Mean	SD	Cronbach's Coefficient Alpha	SEM
		MC	OR								
Applied, Winter (English)	31	24	7	52	20 508	0	52	28.05	9.34	0.83	3.85
Applied, Spring (English)	31	24	7	52	24 555	1	52	28.15	9.23	0.84	3.68
Academic, Winter (English)	31	24	7	52	44 259	2	52	32.50	9.62	0.85	3.71
Academic, Spring (English)	31	24	7	52	56 177	1	52	34.17	9.68	0.87	3.55
Applied, Winter (French)	31	24	7	52	431	6	50	28.14	9.19	0.83	3.78
Applied, Spring (French)	31	24	7	52	1 041	4	52	27.49	9.03	0.81	3.90
Academic, Winter (French)	31	24	7	52	1 559	4	52	30.84	9.93	0.85	3.79
Academic, Spring (French)	31	24	7	52	2 439	7	52	28.94	9.48	0.84	3.85

Note. MC = multiple choice; OR = open response; SD = standard deviation; SEM = standard error of measurement.

Item Response Theory (IRT) Analysis

In the IRT analysis, item parameters were estimated from the student responses used in the equating calibration sample (see Chapter 5). The estimated item parameters were then used to

score all the student responses. The descriptive statistics reported in Table 7.5 refer to the total population of students. The mean student ability scores range from -0.14 to 0.03 for the applied version of the mathematics assessment and from -0.07 to 0.02 for the academic version of the assessment. The means that differ from zero and the standard deviations that differ from one are due to the inclusion of all students. The item parameter estimates for individual items are presented in Appendix 7.1.

Table 7.5 Descriptive Statistics of IRT Scores: Grade 9 Mathematics

Assessment	No. of Students	Min.	Max.	Mean	SD
Applied, Winter (English)	20 508	-3.31	2.80	-0.14	0.97
Applied, Spring (English)	24 555	-3.36	2.85	-0.03	0.99
Academic, Winter (English)	44 259	-3.38	2.51	-0.07	0.96
Academic, Spring (English)	56 177	-3.50	2.40	0.02	0.95
Applied, Winter (French)	431	-2.81	2.33	-0.06	0.98
Applied, Spring (French)	1 041	-2.93	2.90	-0.07	0.94
Academic, Winter (French)	1 559	-2.65	2.65	-0.04	0.95
Academic, Spring (French)	2 544	-2.94	2.71	0.00	0.93

Note. SD = standard deviation.

The TCCs and the distribution of student thetas are displayed in Figures 7.25 and 7.28. The TCCs follow the expected S-shaped distribution. The right vertical scale indicates the percentage of students at each theta value, and the theta cut points for assigning students to performance levels are marked on all the graphs. The TCCs for all the winter and spring administrations either overlap or are very close to each other. This indicates that the winter and spring applied and academic versions of the assessment had the same difficulty level. The TIFs, which are displayed in Figures 7.29–7.32, indicate that the academic version provided more information than the applied version. Further, both the applied and academic versions provided most of their information between the $2/3$ and $3/4$ cut points.

Figure 7.25 Test Characteristic Curves and Distributions of Thetas: Grade 9 Applied Math (English)

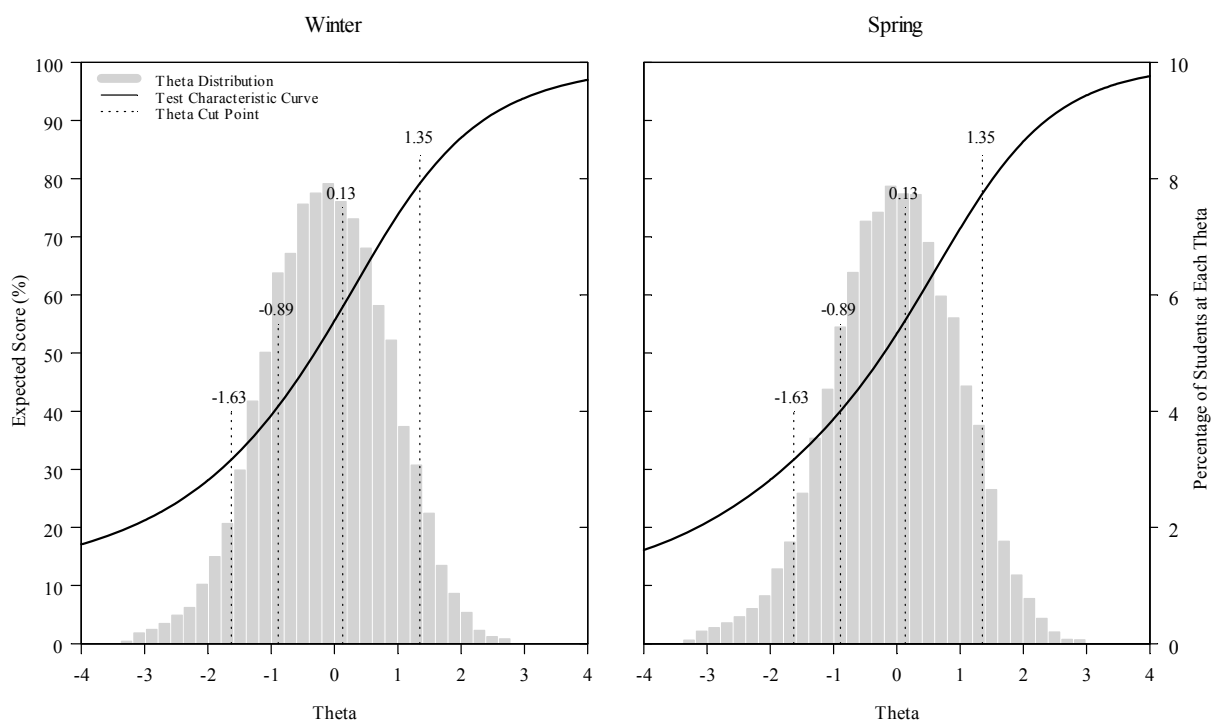


Figure 7.26 Test Characteristic Curves and Distributions of Thetas: Grade 9 Academic Math (English)

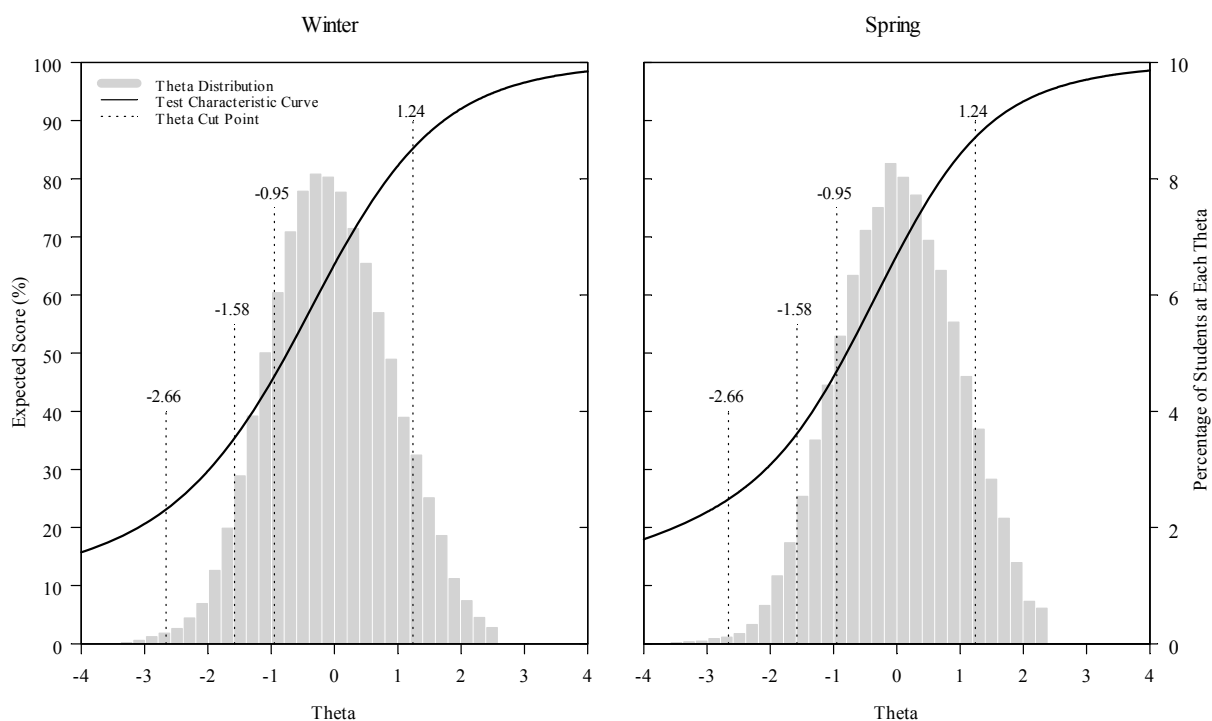


Figure 7.27 Test Characteristic Curves and Distributions of Thetas: Grade 9 Applied Math (French)

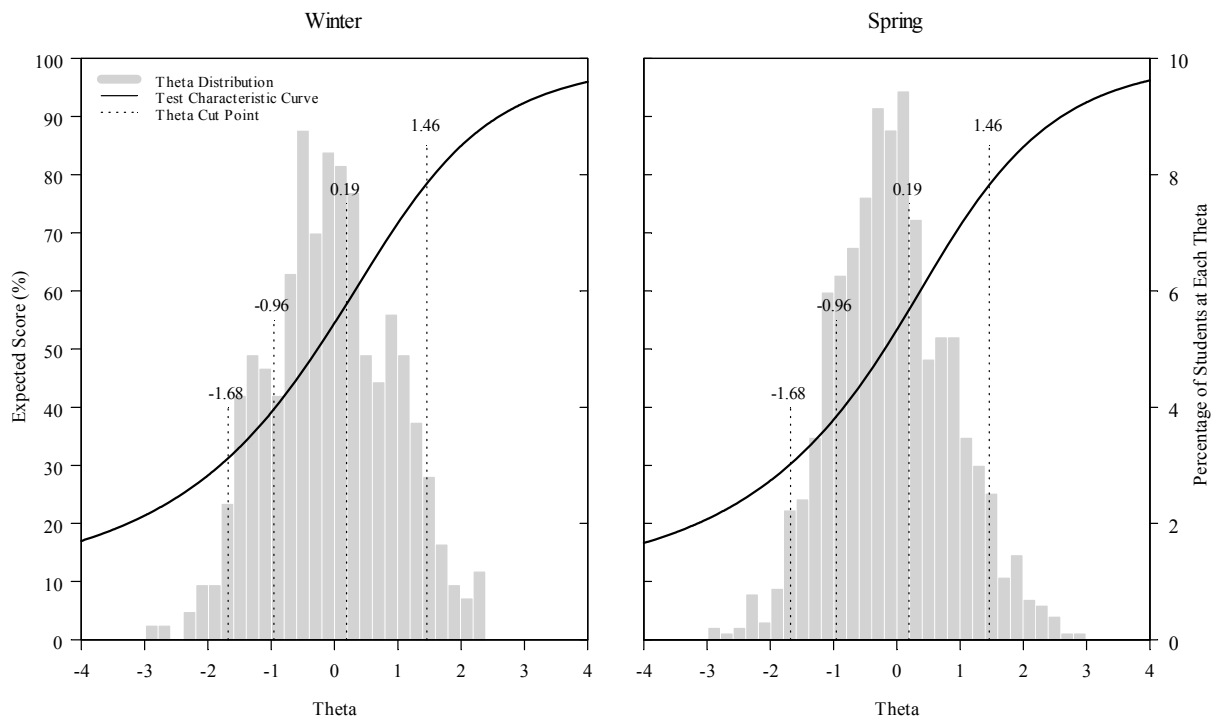


Figure 7.28 Test Characteristic Curves and Distributions of Thetas: Grade 9 Academic Math (French)

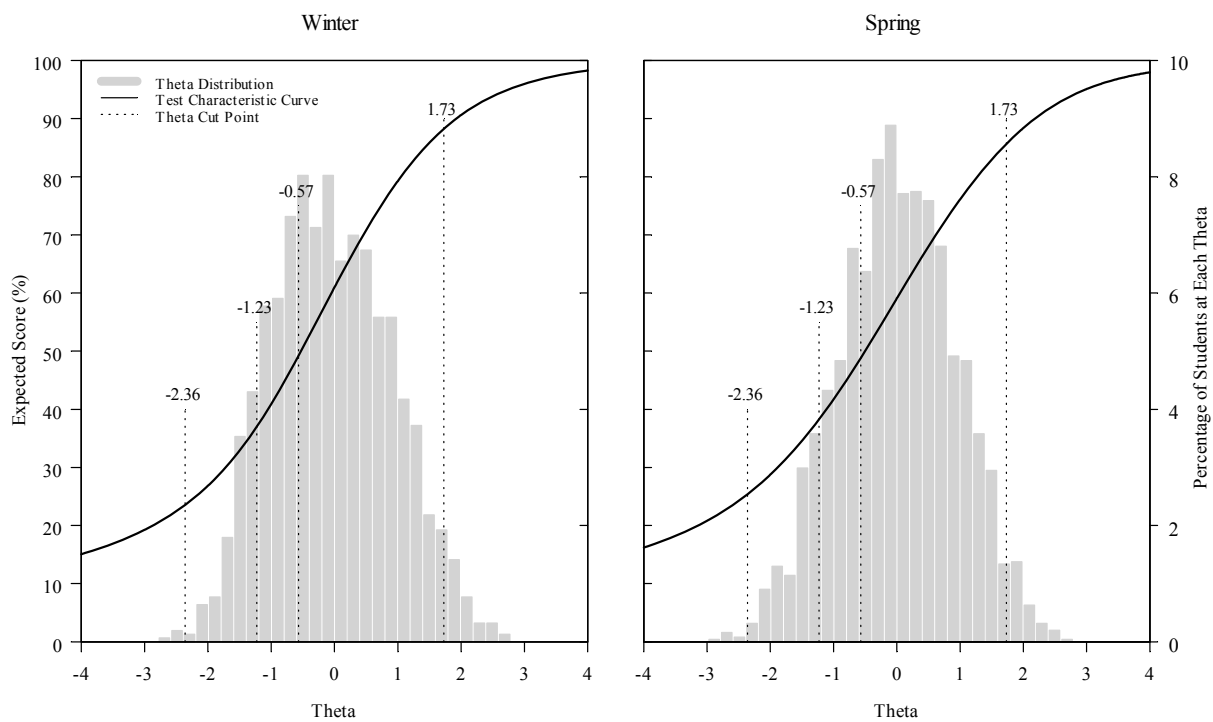


Figure 7.29 Test Information Functions: Grade 9 Applied Math (English)

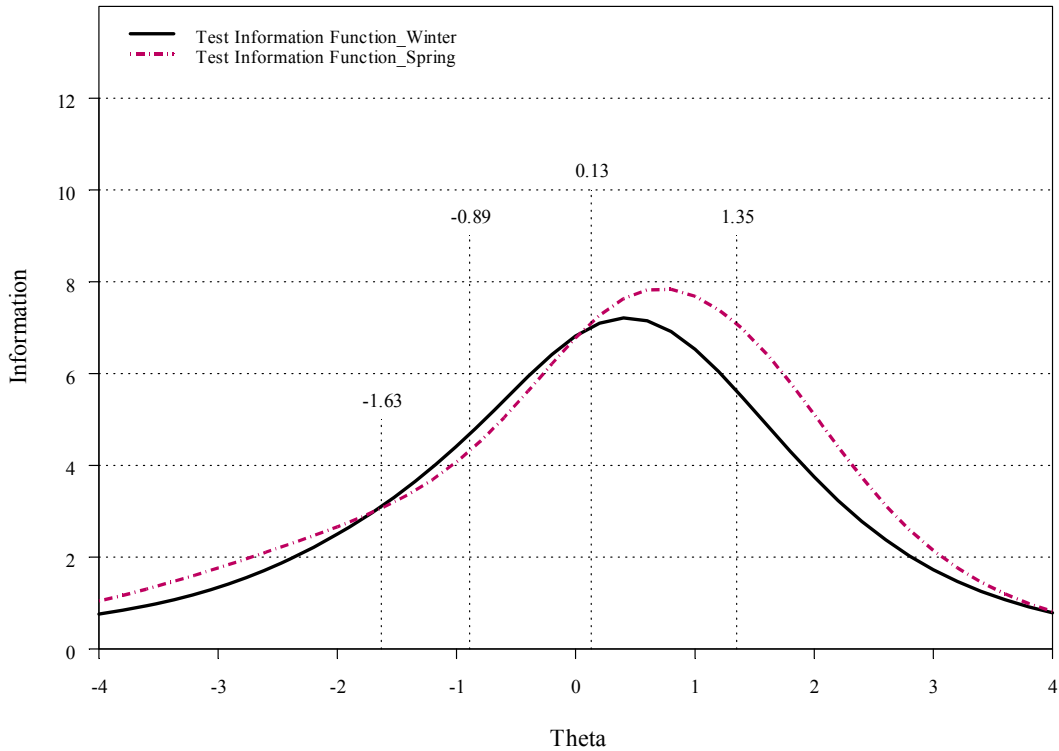


Figure 7.30 Test Information Functions: Grade 9 Academic Math (English)

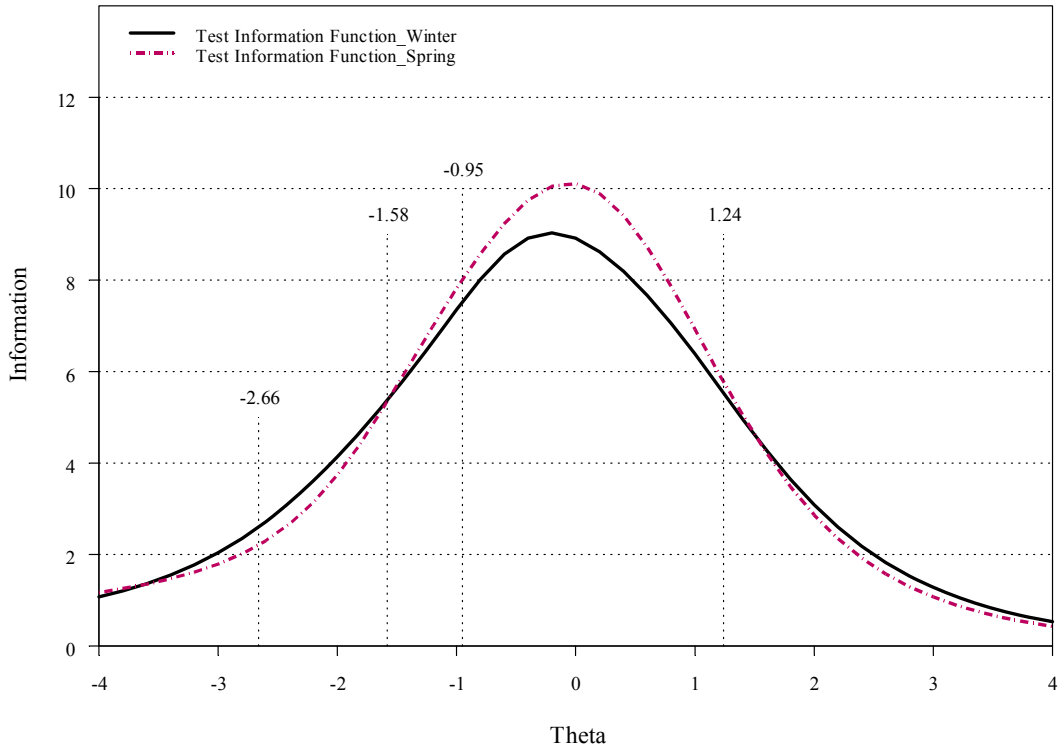


Figure 7.31 Test Information Functions: Grade 9 Applied Math (French)

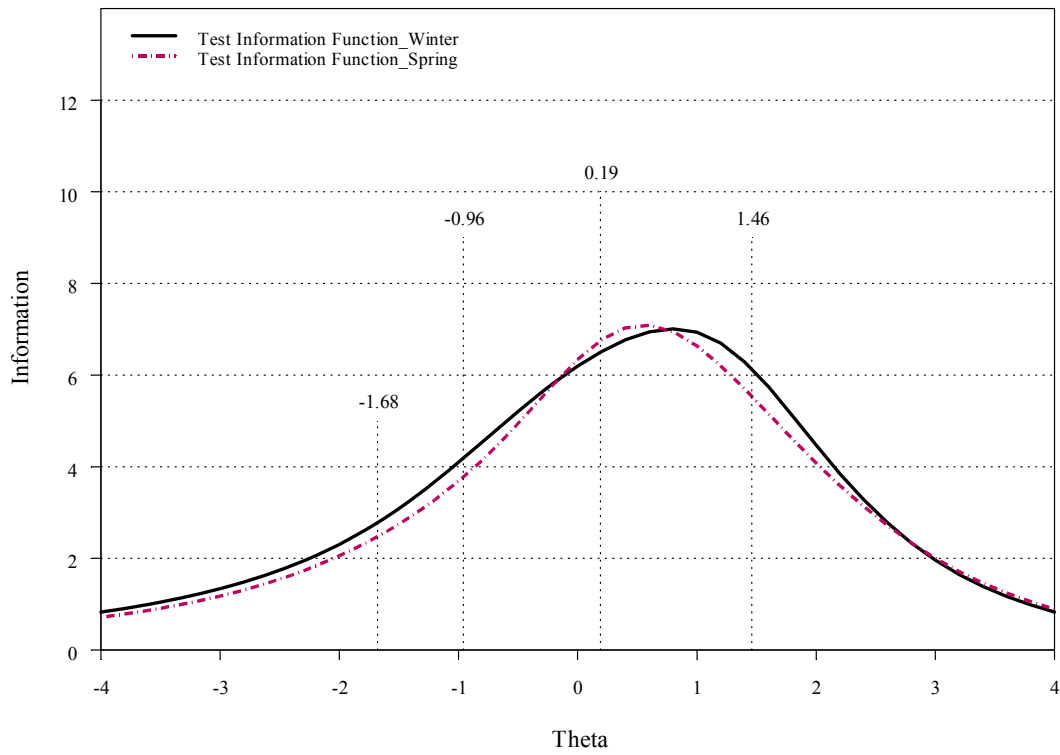
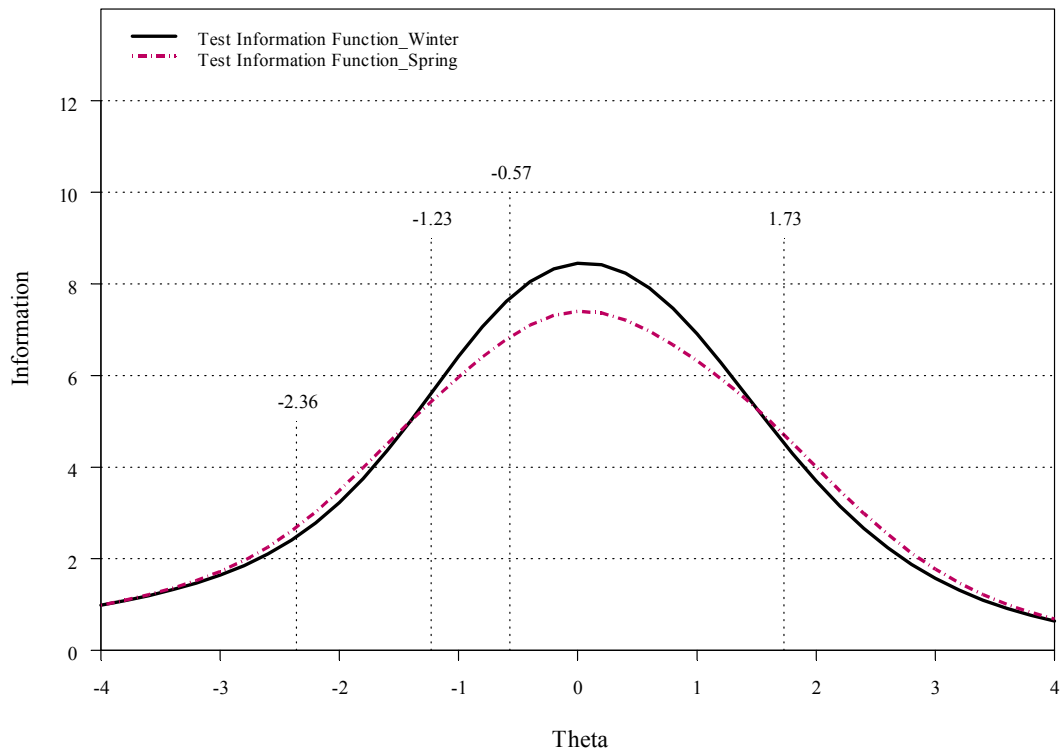


Figure 7.32 Test Information Functions: Grade 9 Academic Math (French)



Descriptive Item Statistics for Classical Test Theory (CTT) and Item Response Theory (IRT)

Table 7.6 contains a summary of both the CTT and IRT item statistics for the items on the Grade 9 mathematics assessment. Both classical item statistics and IRT item statistics were computed using the equating sample. As with the primary- and junior-division assessments, care must be taken to avoid matching the minimum CTT p-value with the minimum IRT location parameter estimate, and the item-total correlation with the slope estimate. As expected, there is an inverse relationship between the CTT item difficulty estimates and the IRT location parameter estimates, and there is a positive relationship between the CTT item-total correlation estimates and the IRT slope parameter estimates. The item difficulty and location parameter estimates are in an acceptable range. Likewise, the point-biserial correlation coefficients are, for the most part, within an acceptable range, though values less than 0.20 are not ideal and indicate possible flaws in test items. The statistics for individual items are presented in Appendix 7.1.

Table 7.6 Descriptive Statistics of CTT Item Statistics and IRT Item Parameter Estimates: Grade 9 Mathematics

Assessment	No. of Items	Descriptive Statistics	CTT Item Statistics		IRT Item Parameters*	
			Item Difficulty	Item-Total Correlation	Location	Slope
Applied, Winter (English)	31	Min.	34.50	0.17	-2.89	0.25
		Max.	87.27	0.55	1.93	0.91
		Mean	57.43	0.33	-0.15	0.55
		SD	13.49	0.10	0.99	0.18
Applied, Spring (English)	31	Min.	26.91	0.17	-2.67	0.28
		Max.	93.18	0.54	1.93	1.23
		Mean	57.65	0.34	-0.12	0.61
		SD	16.83	0.10	1.16	0.23
Academic, Winter (English)	31	Min.	39.32	0.22	-2.15	0.36
		Max.	87.25	0.63	1.08	1.17
		Mean	62.27	0.36	-0.38	0.61
		SD	12.71	0.11	0.86	0.21
Academic, Spring (English)	31	Min.	44.51	0.25	-2.10	0.36
		Max.	88.07	0.57	0.85	1.22
		Mean	65.69	0.40	-0.53	0.70
		SD	11.59	0.09	0.77	0.26
Applied, Winter (French)	31	Min.	33.88	0.17	-1.41	0.13
		Max.	76.31	0.59	2.20	1.13
		Mean	54.89	0.34	0.10	0.57
		SD	11.60	0.10	0.89	0.24
Applied, Spring (French)	31	Min.	29.17	0.13	-1.74	0.21
		Max.	80.55	0.52	2.20	1.18
		Mean	53.69	0.32	0.18	0.56
		SD	11.33	0.09	0.88	0.24
Academic, Winter (French)	31	Min.	37.79	0.23	-1.52	0.31
		Max.	78.00	0.56	1.04	0.98
		Mean	58.31	0.38	-0.11	0.61
		SD	9.59	0.10	0.69	0.17
Academic, Spring (French)	31	Min.	30.62	0.23	-1.87	0.28
		Max.	90.71	0.61	1.48	1.10
		Mean	60.97	0.35	-0.26	0.59
		SD	14.58	0.10	0.88	0.20

Note. SD = standard deviation.

*The guessing parameter was set at a constant of 0.2 for multiple-choice items.

The Ontario Secondary School Literacy Test (OSSLT)

Classical Test Theory (CTT) Analysis

Table 7.7 presents descriptive statistics, Cronbach's alpha estimates of test-score reliability and the SEM for the first-time eligible students who wrote the English-language and French-language OSSLT. The test means (as percentages) for first-time eligible students are 78.6% for English-language students and 78.5% for French-language students.

Cronbach's alpha estimates are 0.89 and 0.87 for English and French, respectively. The corresponding SEMs are 3.9% and 4.0% of the possible maximum score. The obtained reliability coefficients and SEMs are acceptable and indicate that test scores from these assessments are at a satisfactory level of precision.

Table 7.7 Test Descriptive Statistics, Reliability and Standard Error of Measurement: OSSLT (First-Time Eligible Students)

Language	No. of Items	Item Type				Possible Max. Score	No. of Students	Min.	Max.	Mean	SD	R	SEM
		MC	OR	SW	LW								
English	47	39	4	2	2	81	142 955	2.0	81.0	63.7	9.44	0.89	3.17
French	47	39	4	2	2	81	5 353	20.0	81.0	63.6	9.07	0.87	3.22

Note. MC = multiple choice; OR = open response (reading); SW = short writing; LW = long writing; N = number of students; R = Cronbach's alpha; SD = standard deviation; SEM = standard error of measurement.

Item Response Theory (IRT) Analysis

In the IRT analysis, item parameters were estimated from the student responses used in the equating calibration sample (see Chapter 5). The estimated item parameters were then used to score all student responses. The descriptive statistics reported in Table 7.8 are for all first-time-eligible students in the provincial population. Due to the inclusion of all students for both the English- and French-language tests, the mean of student ability scores is not exactly zero and the standard deviation is less than one. The item parameter estimates for individual items are presented in Appendix 7.1.

Table 7.8 Descriptive Statistics for IRT Scores: OSSLT (First-Time Eligible Students)

Language	No. of Students	Min.	Max.	Mean	SD
English	142 955	-3.83	2.92	0.17	0.93
French	5 353	-3.51	2.66	-0.20	0.93

The TCCs and the distribution of student thetas are displayed in Figures 7.33 and 7.34 for the English-language and French-language students, respectively. The TCCs follow the expected S-shaped distribution. The distribution of student thetas is plotted on the TCC graphs, with the right vertical scale indicating the percentage of students at each theta value. The TIF plots for the English-language and French-language tests are shown in Figures 7.35 and 7.36, respectively. The theta cut point for assigning students to the successful and unsuccessful levels of performance is marked on each plot.

Figure 7.33 Test Characteristic Curve and Distribution of Theta: OSSLT (English)

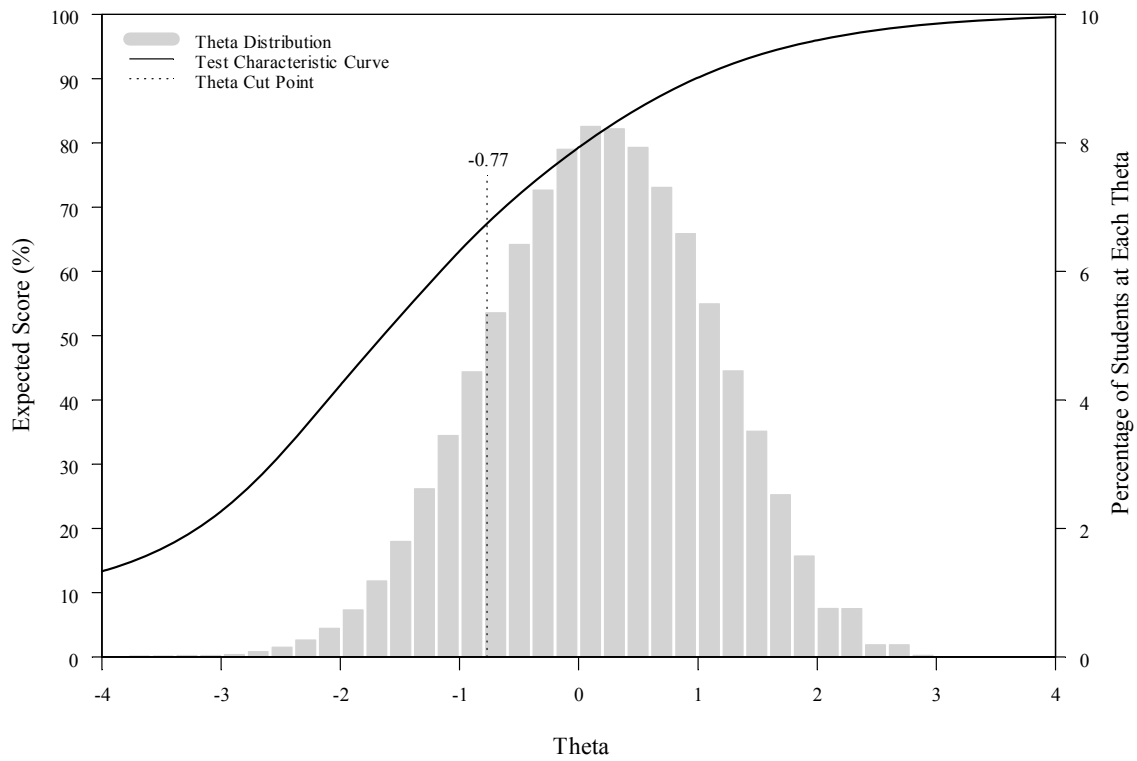


Figure 7.34 Test Characteristic Curve and Distribution of Theta: OSSLT (French)

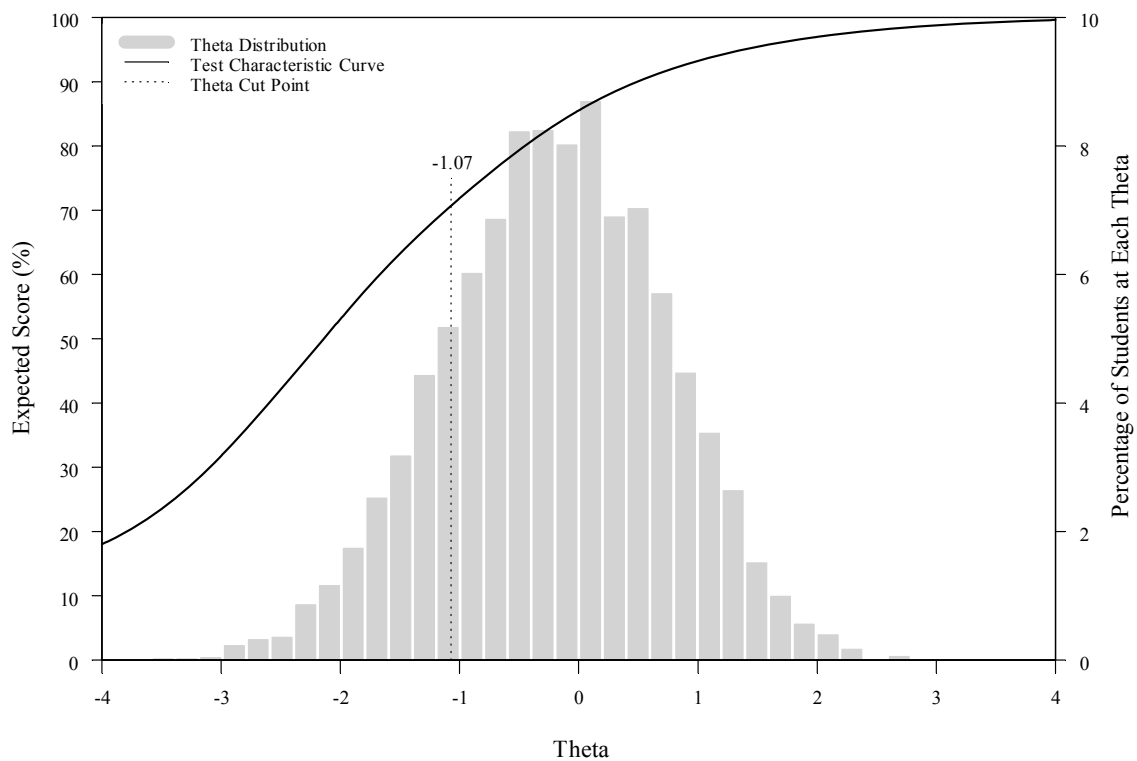


Figure 7.35 Test Information Function: OSSLT (English)

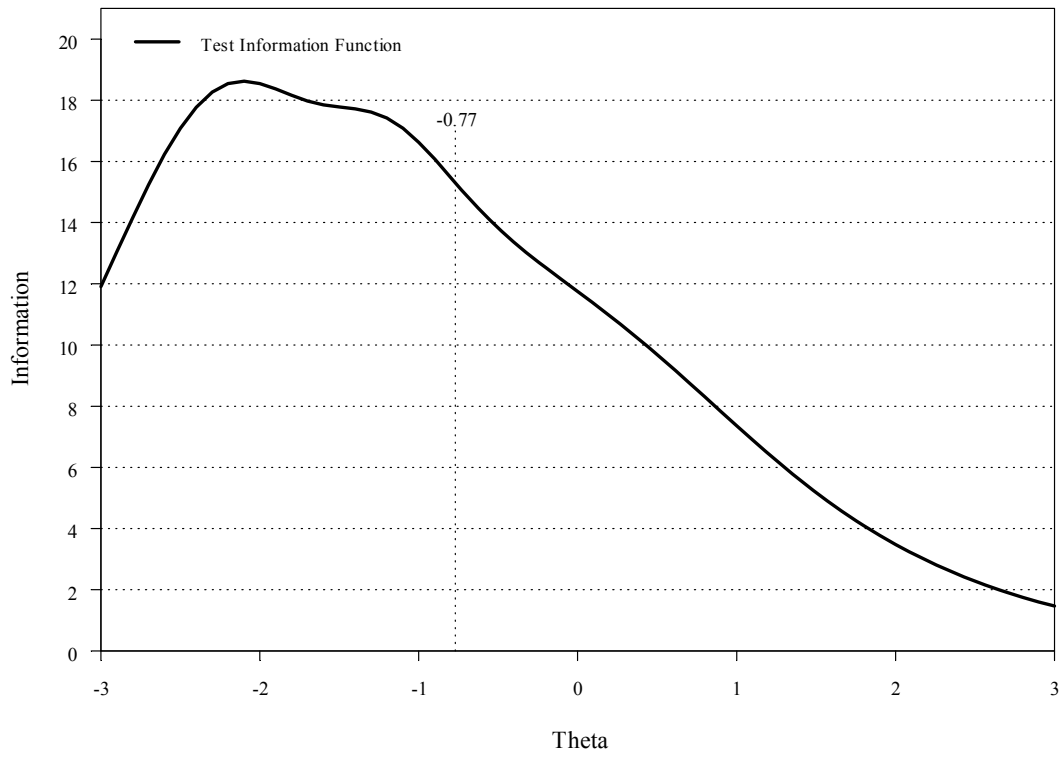
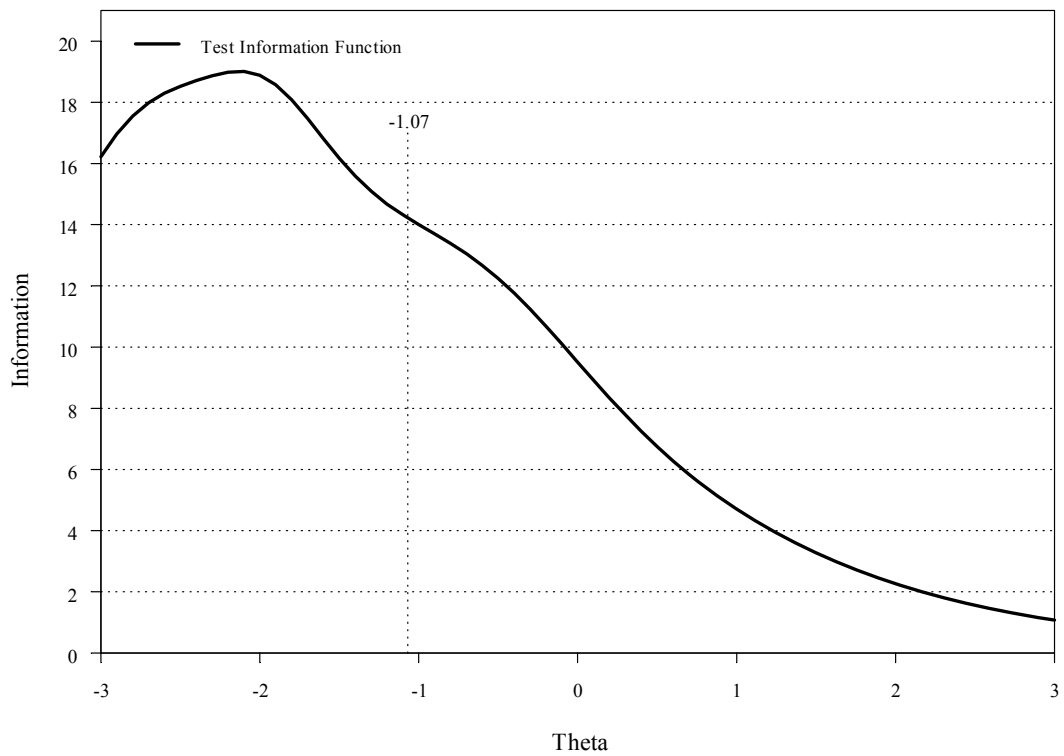


Figure 7.36 Test Information Function: OSSLT (French)



Descriptive Item Statistics for Classical Test Theory (CTT) and Item Response Theory (IRT)

Table 7.9 contains a summary of both the CTT and IRT item statistics for the multiple-choice items on the OSSLT. As with the primary- and junior-division assessments, care must be taken to avoid matching the minimum CTT p -value with the minimum IRT location estimate, and the item-total correlation with the slope estimate. As expected, there is an inverse relationship between the CTT item difficulty estimates and the IRT location parameter estimates due to the difference between the definitions of difficulty in the two approaches. Unlike the primary, junior and Grade 9 assessments, the a -parameter was set at a constant value for all items. Hence, it is not possible to determine the nature of the relationship between the CTT item-total correlations and the IRT slope parameter estimates. However, the low values for the minimum point-biserial correlations suggest that some of the items did not reach the desired level (0.20). The item difficulty values were within an acceptable range. Presented in Appendix 7.1 are the statistics for individual items, the distribution of score points and threshold parameters for the open-response items and the analysis results for differential item functioning for all items.

Table 7.9 Descriptive Statistics of CTT Item Statistics and IRT Item Parameter Estimates: OSSLT

Language	No. of Items	Descriptive Statistics	CTT Item Statistics		IRT Item Parameters*
			Item Difficulty	Item-Total Correlation	
English	47	Min.	0.38	0.18	-3.09
		Max.	0.99	0.55	1.71
		Mean	0.77	0.34	-1.10
		SD	0.13	0.09	1.00
French	47	Min.	0.21	0.13	-4.05
		Max.	0.98	0.60	3.03
		Mean	0.78	0.32	-1.68
		SD	0.16	0.11	1.30

Note. SD = standard deviation.

*The slope was set at 0.588 for all items, and the guessing parameter was set at a constant of 0.20 for all multiple-choice items.

Differential Item Functioning (DIF)

One goal of test development is to assemble a set of items that provides an estimate of a student's ability that is as fair and accurate as possible for all groups within the student population. Differential Item Functioning (DIF) statistics are used to identify items on which students with the same level of ability but from different identifiable groups have different probabilities of answering correctly (e.g., girls and boys, second language learners, or "SLLs," in English- or French-language schools). If an item is more difficult for one subgroup than for another, the item may be measuring something other than what it intends. However, it is important to recognize that DIF-flagged items may measure actual differences in relevant knowledge or skill (i.e., item impact) or statistical Type I error. Therefore, items identified through DIF statistics must be reviewed by content experts and bias and sensitivity committees to determine the possible sources and interpretations of differences in achievement.

EQAO examined the 2010 assessments for gender- and SLL-based DIF, using the Simultaneous Item Bias Test (SIBTEST) for multiple-choice items and Poly-SIBTEST for open-response

items. In all analyses, males and non-SLLs were the reference group, and females and SLLs were the focal, or studied, group. The valid subtest for each SIBTEST and Poly-SIBTEST run consisted of all the items that were not the suspect item. To estimate the expected score (ES) for a suspect item that would be conditional on the primary dimension for the reference group (R) and the focal group (F), the SIBTEST and Poly-SIBTEST runs used the total test score for the matching subtest of items as defined below:

$$ES_R(t) = \sum_{k=1}^m kP_{Rk}(t) \text{ and } ES_F(t) = \sum_{k=1}^m P_{Fk}(t), \quad (3)$$

where $P_{Gk}(t)$ is the empirical proportion of examinees in group G ($G = R$ or F) that have obtained score k on the studied item and have valid subtest score t .

SIBTEST and Poly-SIBTEST use a regression-correction formula to correct for bias due to measurement error in the matching subtest that would indicate DIF favouring the group with the higher ability, when, in fact, no DIF exists (Shealy & Stout, 1993). The regression-correction formula adjusts the mean item scores to their unbiased estimates by computing a regression equation for each group that estimates a true score on the matching subtest as a function of the matching-subtest observed score. This function is computed for each group using the estimated means, variances and reliability coefficients of the scores on the matching subtest. The resulting adjusted expected item scores, $ES_R^*(t)$ and $ES_F^*(t)$, are used for comparing group performance on the item to estimate the DIF index:

$$\hat{\beta} = \sum_{t=1}^T [ES_R^*(t) - ES_F^*(t)] \frac{N_R(t) - N_F(t)}{N}, \quad (4)$$

where T is the maximum score on the matching subtest;
 $N_R(t)$ is the number of examinees in the reference group (R) with matching subtest score t ; $N_F(t)$ is the number of examinees in the focal group (F) with matching subtest score t ;
 N is the total number of examinees and
 $\hat{\beta}$ is interpreted as the difference in the expected total score between the focal and reference groups.

Two statistics were produced by SIBTEST and Poly-SIBTEST: the $\hat{\beta}$, which is an indicator of effect size describing the size of the difference, and a p -value, which estimates the probability that an observed difference occurred by chance. For the 2010 EQAO assessments, the guidelines proposed by Roussos and Stout (1996) for SIBTEST were used to classify the multiple-choice items for which $\hat{\beta}$ was significantly different from zero (with a significance level set to 0.05): items with $\hat{\beta}$ less than 0.059 were considered to have negligible or A-level DIF, items with $\hat{\beta}$

equal to or greater than 0.059 but less than 0.088 were considered to have moderate or B-level DIF and items with $\hat{\beta}$ equal to or greater than 0.088 were considered to have large or C-Level DIF. In the case of open-response items, the same guidelines as those used with SIBTEST results were applied to poly-SIBTEST results to classify the items.

For each assessment, except for the French-language Grade 9 versions, two random samples of 2000 examinees were selected from the population of students who took the test under standard administration conditions (i.e., with no accommodations).² The samples were stratified according to gender or second language learner (SLL) status. The term “second language learner” is used to represent English language learners for the English-language assessments and students in the ALF/PANA program for the French-language assessments. The use of two samples provided an estimate of the stability of the results in a cross-validation process. Items that were classified as having B-level or C-level DIF in both samples were considered DIF items.

The item-level results are provided in Appendix 7.1. The results in each table are from the two random samples, which include the value of $\hat{\beta}$, the significance level and the severity of DIF. Negative estimates of $\hat{\beta}$ indicate that the girls outperformed the boys or that the SLLs outperformed the non-SLLs; positive $\hat{\beta}$ estimates indicate that the boys outperformed the girls or that the non-SLLs outperformed the SLLs.

The Primary- and Junior-Division Assessments

For the reading, writing and mathematics components of the 2010 primary- and junior-division assessments for both languages, the summaries of the number of items that showed statistically significant gender-based DIF with at least a B-level or C-level effect size in both the two samples are reported in Tables 7.10 and 7.11. The numbers in the boys and girls columns indicate the number of DIF items favouring boys and girls.

Table 7.10 Number of B-Level Gender-Based DIF Items: Primary and Junior Assessments

Component	Primary English		Junior English		Primary French		Junior French	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Reading <i>k</i> = 36	0 (MC) 0 (OR)	1 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 1 (OR)	0 (MC) 2 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)
Writing <i>k</i> = 14	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)
Math <i>k</i> = 35*	0 (MC) 1 (OR)	0 (MC) 0 (OR)	1 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)

Note. MC = multiple choice; OR = open response.

*The total number of items for primary English mathematics was 36, but one item was dropped.

²During test administration, some students receive accommodations for various reasons and do not write the test under standard conditions. These students were not included in the DIF analyses.

Table 7.11 Number of C-Level Gender-Based DIF Items: Primary and Junior Assessments

Component	Primary English		Junior English		Primary French		Junior French	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Reading <i>k</i> = 36	0 (MC) 1 (OR)	0 (MC) 0 (OR)	1 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)
Writing <i>k</i> = 14	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)
Math <i>k</i> = 35*	0 (MC) 1 (OR)	0 (MC) 2 (OR)	0 (MC) 2 (OR)	0 (MC) 1 (OR)	0 (MC) 3 (OR)	0 (MC) 1 (OR)	0 (MC) 1 (OR)	0 (MC) 1 (OR)

Note. MC = multiple choice; OR = open response.

*The total number of items for primary English mathematics was 36, but one item was dropped.

Of the 343 items comprising the primary- and junior-division assessments, 22 items were found to have gender-based DIF. The majority of the DIF items had C-level DIF (15) and seven items had B-level DIF. Of the three multiple-choice items showing B- or C-level DIF, two favoured the boys, while one favoured the girls. Of the 19 open-response DIF items, 11 favoured the boys, while eight favoured the girls. Overall, more items were found to favour boys than girls. The mathematics component of the English-language junior-division and French-language primary-division assessments had the largest number of gender-based DIF items (four each).

The summaries for SLL-based DIF are reported in Tables 7.12 and 7.13. Fifteen out of the 343 items across all 12 assessments showed SLL-based DIF, with six items having B-level DIF and nine items having C-level DIF. Of the five multiple-choice items showing B- or C-level DIF, three favoured non-SLL students and two favoured SLL students. For the 10 open-response DIF items, four favoured non-SLL students and six favoured SLL students. The reading component of the English-language junior-division assessment was found to have the largest number of SLL-based DIF items (five).

Table 7.12 Number of B-Level SLL-Based DIF Items: Primary and Junior Assessments

Component	Primary English		Junior English		Primary French		Junior French	
	Non-SLLs	SLLs	Non-SLLs	SLLs	Non-SLLs	SLLs	Non-SLLs	SLLs
Reading <i>k</i> = 36	0 (MC) 0 (OR)	0 (MC) 0 (OR)	1 (MC) 0 (OR)	1 (MC) 0 (OR)	1 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)
Writing <i>k</i> = 14	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)
Math <i>k</i> = 35*	0 (MC) 1 (OR)	0 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)

Note. SLL = Second-language learner; MC = multiple choice; OR = open response.

*The total number of items for primary English mathematics was 36, but one item was dropped.

Table 7.13 Number of C-Level SLL-Based DIF Items: Primary and Junior Assessments

Component	Primary English		Junior English		Primary French		Junior French	
	Non-SLLs	SLLs	Non-SLLs	SLLs	Non-SLLs	SLLs	Non-SLLs	SLLs
Reading <i>k</i> = 36	1 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 1 (OR)	0 (MC) 2 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 1 (OR)
Writing <i>k</i> = 14	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	1 (MC) 0 (OR)
Math <i>k</i> = 35*	0 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 1 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)	0 (MC) 0 (OR)

Note. SLL = Second language learner; MC = multiple choice; OR = open response.

*The total number of items for primary English mathematics was 36, but one item was dropped.

All items identified as having B-level or C-level DIF on the primary- and junior-division assessments were reviewed by the assessment team. Since the reviewers did not identify any apparent bias in the content of these items, they were not removed from the scoring process.

The Grade 9 Mathematics Assessment

The gender- and SLL-based DIF results for the academic and applied versions of the Grade 9 assessment are provided in Tables 7.14–7.17. For gender-based DIF, it was not possible to have two random samples for the French-language academic and applied versions of the Grade 9 assessment due to the small number of participating students. The number of participating French-language SLL students was also too small to conduct SLL-based DIF analysis.

Table 7.14 Number of B-Level Gender-Based DIF Items: Grade 9 Applied and Academic Mathematics

Assessment	English		French	
	Boys	Girls	Boys	Girls
Applied, Winter <i>k</i> = 31	1 (MC)	2 (MC)	0 (MC)	0 (MC)
	0 (OR)	0 (OR)	0 (OR)	0 (OR)
Applied, Spring <i>k</i> = 31	1 (MC)	1 (MC)	3 (MC)	0 (MC)
	1 (OR)	0 (OR)	0 (OR)	0 (OR)
Academic, Winter <i>k</i> = 31	1 (MC)	0 (MC)	3 (MC)	1 (MC)
	0 (OR)	0 (OR)	0 (OR)	0 (OR)
Academic, Spring <i>k</i> = 31	2 (MC)	2 (MC)	2 (MC)	0 (MC)
	0 (OR)	0 (OR)	0 (OR)	0 (OR)

Note. MC = multiple choice; OR = open response.

Table 7.15 Number of C-Level Gender-Based DIF Items: Grade 9 Applied and Academic Mathematics

Assessment	English		French	
	Boys	Girls	Boys	Girls
Applied, Winter <i>k</i> = 31	1 (MC)	0 (MC)	2 (MC)	2 (MC)
	2 (OR)	2 (OR)	0 (OR)	2 (OR)
Applied, Spring <i>k</i> = 31	0 (MC)	0 (MC)	1 (MC)	1 (MC)
	1 (OR)	1 (OR)	0 (OR)	0 (OR)
Academic, Winter <i>k</i> = 31	0 (MC)	0 (MC)	1 (MC)	0 (MC)
	1 (OR)	1 (OR)	2 (OR)	1 (OR)
Academic, Spring <i>k</i> = 31	0 (MC)	0 (MC)	0 (MC)	0 (MC)
	2 (OR)	1 (OR)	1 (OR)	3 (OR)

Note. MC = multiple choice; OR = open response.

Of the 248 items on the Grade 9 assessments, 19 multiple-choice items and one open-response item showed B-level gender-based DIF. Of the 19 B-level gender-based multiple-choice DIF items, 13 favoured boys and six favoured girls. The B-level gender-based open-response item favoured boys. Twenty-eight items showed C-level gender-based DIF, of which eight were multiple-choice items and 20 were open-response items. Of the eight C-level multiple-choice items, five favoured boys and three favoured girls. Of the 20 C-level open-response items, nine favoured boys and 11 favoured girls.

Eight to 13 multiple-choice items were used in both versions of the winter and spring assessments. Ten items showed gender-based DIF. Of the 10 items, four items showed gender-based DIF in both administrations; two items showed DIF in the winter administration, and four items showed DIF in the spring administration.

Table 7.16 Number of B-Level SLL-Based DIF Items: Grade 9 Applied and Academic Mathematics

Assessment	Applied		Academic	
	Non-SLLs	SLLs	Non-SLLs	SLLs
Winter <i>k</i> = 31	3 (MC)	1 (MC)	0 (MC)	1 (MC)
	0 (OR)	0 (OR)	0 (OR)	0 (OR)
Spring <i>k</i> = 31	0 (MC)	0 (MC)	1 (MC)	2 (MC)
	0 (OR)	1 (OR)	1 (OR)	0 (OR)

Note. MC = multiple choice; OR = open response.

Table 7.17 Number of C-Level SLL-Based DIF Items: Grade 9 Applied and Academic Mathematics

Assessment	Applied		Academic	
	Non-SLLs	SLLs	Non-SLLs	SLLs
Winter k = 31	1 (MC)	1 (MC)	3 (MC)	2 (MC)
	1 (OR)	2 (OR)	3 (OR)	2 (OR)
Spring k = 31	2 (MC)	1 (MC)	3 (MC)	3 (MC)
	2 (OR)	2 (OR)	2 (OR)	3 (OR)

Note. MC = multiple choice; OR = open response.

There are 10 B-level SLL-based DIF items. Of the 10 items, eight are multiple-choice items, with four items favouring boys, and two are open-response items, with one item favouring boys and one item favouring girls. Thirty-three items (16 multiple choice and 17 open response) showed C-level SLL-based DIF. Of the 16 multiple-choice items, eight favoured non-SLL students and eight favoured SLL students. Of the 17 open-response items, eight favoured non-SLL students and nine favoured SLL students. Five common items showed SLL-based DIF in both administrations, and one common item showed DIF in the spring.

All Grade 9 assessment items identified as B-level or C-level DIF items were reviewed by the assessment team. Since the reviewers did not identify any apparent bias in the content of these items, they were not removed from the scoring process.

The OSSLT

Gender-based DIF results for the OSSLT are presented in Table 7.18 for B-level items; no items were identified as showing C-level DIF.

Table 7.18 Number of B-Level Gender-Based DIF Items: OSSLT

English (k = 51)		French (k = 51)	
Males	Females	Males	Females
2 (MC)	2 (MC)	1 (MC)	1 (MC)
2 (OR)		1 (OR)	1 (OR)

Note. MC = multiple choice; OR = open response.

There were six B-level DIF items on the English-language version of the OSSLT. Four of the DIF items, of which two were open-response reading items, favoured boys, and two of the DIF items favoured girls.

There were four B-level DIF items on the French-language version of the OSSLT. Two of these DIF items favoured boys and two favoured the girls. One open-response reading item exhibited B-level DIF favouring girls, and one long-writing task exhibited B-level DIF favouring boys for topic development.

Table 7.19 Number of SLL-Based DIF Items: OSSLT (English)

English (k = 51)			
B-Level DIF		C-Level DIF	
Non-SLLs	SLLs	Non-SLLs	SLLs
1 (MC)	4 (MC) 1 (OR)	5 (MC) 2 (OR)	4 (MC) 2 (OR)

Note. MC = multiple choice; OR = open response.

DIF analysis was not conducted for SLL students taking the French-language version of the OSSLT due to the small number of students in this group. For the English-language version of the OSSLT (see Table 7.19), four multiple-choice items and one short-writing task exhibited B-level DIF favouring SLLs for topic development. One multiple-choice item exhibited B-level DIF favouring non-SLLs.

Four multiple-choice items, one open-response reading item and one long-writing task exhibited C-level DIF favouring SLLs for the use of conventions. Five multiple-choice items and one long-writing task exhibited C-level DIF favouring non-SLLs for topic development and the use of conventions.

All OSSLT items that were identified as exhibiting B-level or C-level DIF were reviewed by the assessment team. Since the reviewers did not identify any apparent bias in the content of these items, they were not removed from the scoring process.

Decision Accuracy and Consistency

The four achievement levels defined in *The Ontario Curriculum* are used by EQAO to report student achievement in reading, writing and mathematics for the primary- and junior-division assessments and in the academic and applied versions of the Grade 9 mathematics assessment. Level 3 has been established as the provincial standard. Levels 1 and 2 indicate achievement below the provincial standard, and Level 4 indicates achievement above the provincial standard. In addition to these four performance levels, students who lack enough evidence to achieve Level 1 are placed at NE1. (Students without data and exempted students are not included in the calculation of results for participating students.) Through the equating process described in Chapter 5, four theta values are identified that define the cut points between adjacent levels (NE1 and Level 1, Level 1 and Level 2, Level 2 and Level 3 and Level 3 and Level 4). In the case of the OSSLT, EQAO reports only two levels of achievement: successful and unsuccessful. Thus, the OSSLT has one cut score.

Two issues that arise when students are placed into categories based on assessment scores are accuracy and consistency.

Accuracy

The term “accuracy” refers to the extent to which classifications based on observed student scores agree with classifications based on true scores. While observed scores include measurement error, true scores do not. Thus, classification decisions based on true scores are true or correct classifications. In contrast, classification decisions based on observed scores or derived

from observed scores are not errorless. Since the errors may be positive, zero or negative, an observed score may be too low, just right or too high. This is illustrated in Table 7.20 for classifications in two adjacent categories (0 and 1).

Table 7.20 Demonstration of Classification Accuracy

		Classification Based on True Scores		Row Margins
		0	1	
Classification Based on Observed Scores	0	p_{00}	p_{01}	$p_{1.}$
	1	p_{10}	p_{11}	$p_{2.}$
Column Margins		$p_{.1}$	$p_{.2}$	1.00

The misclassifications, p_{01} and p_{10} , are attributable to the presence of measurement error. The sum of p_{00} and p_{11} equals the rate of classification accuracy, which should be high or close to 1.00.

Consistency

The term “consistency” refers to the extent to which classifications based on observed student scores on one form of the assessment agree with the classifications of the observed scores of the same students on a parallel form. In contrast to accuracy, neither set of observed scores on the two interchangeable tests is errorless. Some student scores on one test will be higher than their scores on the second test. For other students, their scores will be equal, and for other students still, their scores on the first test will be lower than their scores on the second. The differences, when they occur, may be so large that they lead to different or inconsistent classifications. The classification based on the observed score could be lower, the same as or higher than the classification based on the second score. This is illustrated in Table 7.21 for classifications in two adjacent categories (0 and 1).

Table 7.21 Demonstration of Classification Consistency

		Classification Based on Observed Scores 2		Row Margins
		0	1	
Classification Based on Observed Scores 1	0	p_{00}	p_{01}	$p_{1.}$
	1	p_{10}	p_{11}	$p_{2.}$
Column Margins		$p_{.1}$	$p_{.2}$	1.00

The different classifications, p_{01} and p_{10} , are attributable to the presence of measurement error. The sum of p_{00} and p_{11} equals the rate of classification consistency, which should be high or close to 1.00.

Estimation from One Test Form

There are several procedures for estimating decision accuracy and decision consistency. The procedure developed by Livingston & Lewis (1995) is used by EQAO because it yields estimates of both accuracy and consistency and allows for both multiple-choice and open-response items. Further, this procedure is commonly used in large-scale assessment programs.

The Livingston-Lewis procedure uses the classical true score, τ , to determine classification accuracy. The true score, corresponding to an observed score X , is expressed as a proportion on a scale of 0 to 1:

$$\tau_p = \frac{\xi_f(X) - X_{\min}}{X_{\max} - X_{\min}}, \quad (5)$$

where τ_p is the proportional true score;

$\xi_f(X)$ is the expected value of a student's observed scores across $f \rightarrow \infty$

interchangeable forms and

$X_{\min}$ and $X_{\max}$ are, respectively, the minimum and maximum observed scores.

Decision consistency is estimated using the joint distribution of reported performance-level classifications on the current test form and performance-level classifications on the alternate or parallel test form. In each case, the proportion of performance-level classifications with exact agreement is the sum of the entries shown in the diagonal of the contingency table representing the joint distribution.

The Livingston-Lewis procedure requires the creation of an effective test length to model the complex data. The effective test length is determined by the “number of discrete, dichotomously scored, locally independent, equally difficult test items necessary to produce total scores having the same precision as the scores being used to classify the test takers” (Livingston & Lewis, 1995, p. 180). The formula for determining the effective test length is

$$\tilde{n} = \frac{(\hat{\mu}_X - X_{\min})(X_{\max} - \hat{\mu}_X) - r_{XX} \hat{\sigma}_X^2}{\hat{\sigma}_X^2 (1 - r_{XX})}, \quad (6)$$

where $\tilde{n}$ is the effective test length rounded to the nearest integer;

$\hat{\mu}_X$ is the mean of the observed scores;

$\hat{\sigma}_X^2$ is the unbiased estimator of the variance of the observed scores and

r_{XX} is the reliability of the observed scores.

The third step of the method requires that the observed scores in the original scale score for test X be transformed into a new scale score X' :

$$X' = \tilde{n} \frac{X - X_{\min}}{X_{\max} - X_{\min}} . \quad (7)$$

The distribution of true scores is estimated by fitting a four-parameter beta distribution, the parameters of which are estimated from the observed distribution of X' . In addition, the distribution of conditional errors is estimated by fitting a binomial model with regard to X' and $\tilde{n}$. Both classification accuracy and classification consistency can then be determined by using these two distributions. The results for each are then adjusted so that the predicted marginal category proportions match those for the observed test. The computer program BB-CLASS (Brennan, 2004) was used to determine these estimates.

The Primary and Junior Assessments

The classification indices for the primary- and junior-division assessments are presented in Table 7.22. The table includes the overall classification indices (i.e., across the five achievement levels) and the indices for the cut point at the provincial standard (i.e., classifying students into those who met the provincial standard and those who did not, using the Level 2-Level 3 cut). As expected, the indices for overall classification are lower than those for the provincial standard.

Table 7.22: Classification Accuracy and Consistency Indices: Primary- and Junior-Division Assessments

Assessment	Overall Accuracy	Overall Consistency	Accuracy at the Provincial Standard Cut	Consistency at the Provincial Standard Cut
Primary Reading (English)	0.78	0.70	0.90	0.86
Junior Reading (English)	0.81	0.73	0.91	0.87
Primary Reading (French)	0.76	0.66	0.90	0.85
Junior Reading (French)	0.82	0.74	0.91	0.87
Primary Writing (English)	0.85	0.79	0.88	0.83
Junior Writing (English)	0.81	0.74	0.88	0.83
Primary Writing (French)	0.79	0.70	0.90	0.86
Junior Writing (French)	0.81	0.74	0.89	0.85
Primary Mathematics (English)	0.82	0.75	0.91	0.87
Junior Mathematics (English)	0.80	0.71	0.90	0.86
Primary Mathematics (French)	0.81	0.73	0.91	0.87
Junior Mathematics (French)	0.82	0.75	0.92	0.89

The Grade 9 Assessment of Mathematics

The classification indices for the Grade 9 assessment are presented in Table 7.23. As is the case for the primary and junior assessments, the overall classification indices are lower than those for the provincial standard.

Table 7.23 Classification Accuracy and Consistency Indices: Grade 9 Mathematics

Assessment	Overall Accuracy	Overall Consistency	Accuracy at the Provincial Standard Cut	Consistency at the Provincial Standard Cut
Applied, Winter (English)	0.69	0.58	0.87	0.82
Applied, Spring (English)	0.70	0.59	0.88	0.83
Academic, Winter (English)	0.82	0.76	0.91	0.86
Academic, Spring (English)	0.84	0.79	0.92	0.88
Applied, Winter (French)	0.69	0.57	0.88	0.84
Applied, Spring (French)	0.69	0.58	0.88	0.83
Academic, Winter (French)	0.79	0.72	0.89	0.84
Academic, Spring (French)	0.80	0.74	0.89	0.84

The OSSLT

The classification indices for the English-language and French-language versions of the test are presented in Table 7.24. They indicate high accuracy and consistency for both versions.

Table 7.24 Classification Accuracy and Consistency Indices: OSSLT

Assessment	Accuracy (Successful or Unsuccessful)	Consistency (Successful or Unsuccessful)
English	0.93	0.89
French	0.92	0.88

References

- Brennan, R. L. (2004). BB-CLASS: A computer program that uses the beta-binomial model for classification consistency and accuracy [Computer software]. Iowa City, IA: The University of Iowa.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.
- Roussos, L. A., & Stout, W. F. (1996). Simulation studies of the effects of small sample size and studied item parameters on SIBTEST and Mantel-Haenszel type I error performance. *Journal of Educational Measurement*, 33, 215–230.
- Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika*, 58, 159–194.

APPENDIX 4.1: SCORING VALIDITY AND INTERRATER RELIABILITY

This appendix presents validity and interrater reliability estimates for the scoring of all open-response questions.

Validity: the Primary and Junior Assessments

Table 4.1.1 Validity Estimates for Reading: Primary Division (English)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
EQR31615.141	1	5	1 616	96.0	71.3	24.7	14.9	9.8	4.0
EQR31614.141	1	6	1 437	97.5	74.1	23.4	15.2	8.2	2.4
EQR31660.146	1	11	1 330	99.8	88.6	11.2	8.6	2.6	0.2
EQR31661.146	1	12	1 382	99.8	90.4	9.4	5.1	4.3	0.1
EQR31697.150	1	18	1 294	99.3	89.0	10.3	9.7	0.6	0.7
EQR31696.150	1	19	1 301	97.9	90.7	7.2	5.1	2.1	2.2
EQR31809.159	2	11	1 637	98.0	70.3	27.7	12.0	15.7	2.0
EQR31810.159	2	12	1 691	98.8	78.4	20.4	19.3	1.1	1.3
EQR31733.154	2	22	1 488	99.9	97.1	2.8	1.5	1.3	0.0
EQR31735.154	2	23	1 378	99.9	96.2	3.7	2.5	1.2	0.1
Aggregate			14 554	98.6	84.0	14.7	9.7	4.9	1.4

Table 4.1.2 Validity Estimates for Reading: Junior Division (English)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
EQR31874.162	1	5	1 722	99.4	76.6	22.8	19.1	3.7	0.6
EQR31875.162	1	6	1 617	99.2	65.4	33.8	12.8	21.0	0.8
EQR31919.167	1	11	1 671	99.3	82.4	16.9	11.8	5.1	0.6
EQR31920.167	1	12	1 660	98.8	72.2	26.6	17.0	9.6	1.1
EQR31957.171	1	18	1 748	99.3	76.9	22.4	16.0	6.4	0.7
EQR31955.171	1	19	1 706	99.1	78.0	21.1	18.1	3.0	0.8
EQR32031.177	2	11	1 407	99.7	79.7	20.0	8.7	11.3	0.3
EQR32029.177	2	12	1 471	99.6	78.1	21.5	7.5	14.0	0.4
EQR31974.173	2	22	1 692	99.8	81.9	17.9	10.6	7.3	0.2
EQR31973.173	2	23	1 577	98.3	70.8	27.5	15.3	12.2	1.6
Aggregate			16 271	99.2	76.2	23.0	13.9	9.2	0.7

Table 4.1.3 Validity Estimates for Reading: Primary Division (French)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
OQLF31051.103	1	5	62	100.0	53.2	46.8	0.0	46.8	0.0
OQLF31052.103	1	6	41	100.0	85.4	14.6	2.4	12.2	0.0
OQLF31069.105	1	11	141	100.0	92.9	7.1	2.8	4.3	0.0
OQLF31071.105	1	12	165	99.9	71.5	28.4	14.5	13.9	0.0
OQLF31078.106	1	18	146	100.0	84.2	15.8	6.2	9.6	0.0
OQLF31080.106	1	19	131	98.5	69.5	29.0	27.5	1.5	1.5
OQLF20445.060	2	11	116	100.0	93.1	6.9	2.6	4.3	0.0
OQLF20448.060	2	12	145	99.3	74.5	24.8	13.8	11.0	0.7
OQLF20549.069	2	22	100	100.0	94.0	6.0	6.0	0.0	0.0
OQLF20550.069	2	23	89	100.0	75.3	24.7	21.3	3.4	0.0
Aggregate			1136	99.7	79.9	19.8	10.7	9.1	0.3

Table 4.1.4 Validity Estimates for Reading: Junior Division (French)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
OQLF31251.116	1	5	202	100.0	78.7	21.3	19.3	2.0	0.0
OQLF31252.116 _(A)	1	6	180	99.4	85.6	13.8	9.4	4.4	0.6
OQLF31260.117	1	11	130	99.2	87.7	11.5	9.2	2.3	0.8
OQLF31262.117	1	12	130	100.0	77.7	22.3	16.9	5.4	0.0
OQLF31309.120	1	18	70	100.0	87.1	12.9	4.3	8.6	0.0
OQLF31288.120	1	19	169	97.0	75.1	21.9	10.1	11.8	3.0
OQLF31215.113	2	11	127	99.9	81.1	18.8	9.4	9.4	0.0
OQLF31216.113	2	12	167	99.5	71.3	28.2	12.6	15.6	0.6
OQLF31305.122	2	22	182	100.0	83.0	17.0	12.6	4.4	0.0
OQLF31307.122	2	23	155	98.7	72.3	26.4	21.9	4.5	1.3
Aggregate			1512	99.3	79.4	19.9	13.2	6.7	0.7

Table 4.1.5 Validity Estimates for Writing: Primary Division (English)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
11042_T	2	24	1841	99.5	78.4	21.1	12.7	8.4	0.5
11042_V	2	24	1719	100.0	93.5	6.5	2.9	3.5	0.0
Aggregate Long Writing			3560	99.7	85.7	14.0	8.0	6.0	0.3
EQWE20246_T	1	13	1260	99.5	70.7	28.8	8.6	20.2	0.5
EQWE20246_V	1	13	1379	99.8	72.4	27.3	4.9	22.4	0.2
EQWE30427_T	2	17	1537	99.7	75.9	23.8	12.6	11.2	0.3
EQWE30427_V	2	17	1561	99.9	75.0	24.9	3.7	21.2	0.1
Aggregate Short Writing			5737	99.7	73.7	26.1	7.5	18.6	0.3
Aggregate All Items			9297	99.7	78.3	21.4	7.7	13.8	0.3

Table 4.1.6 Validity Estimates for Writing: Junior Division (English)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
10715_T	2	24	2393	98.4	72.4	26.0	11.3	14.7	1.6
10715_V	2	24	2297	99.9	84.3	15.5	8.0	7.5	0.1
Aggregate Long Writing			4690	99.1	78.3	20.9	9.7	11.2	0.9
EQWE30498_T	1	13	1933	99.2	78.2	21.1	10.9	10.2	0.8
EQWE30498_V	1	13	1445	99.8	75.9	23.9	5.3	18.6	0.2
EQWE30504_T	2	17	1854	99.2	76.0	23.2	19.3	3.9	0.8
EQWE30504_V	2	17	1888	99.8	82.1	17.7	5.8	12.0	0.2
Aggregate Short Writing			7120	99.5	78.2	21.3	10.6	10.7	0.5
Aggregate All Items			11810	99.4	78.2	21.1	10.2	10.9	0.7

Table 4.1.7 Validity Estimates for Writing: Primary Division (French)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
11248_T	2	24	168	92.9	68.5	24.4	17.9	6.5	7.1
11248_V	2	24	191	100.0	81.7	18.3	11.5	6.8	0.0
Aggregate Long Writing			359	96.7	75.5	21.2	14.5	6.7	3.3
OQEF30521_T	1	13	155	98.7	65.8	32.9	25.2	7.7	1.3
OQEF30521_V	1	13	143	100.0	77.6	22.4	10.5	11.9	0.0
OQEF30523_T	2	17	128	97.7	80.5	17.2	12.5	4.7	2.3
OQEF30523_V	2	17	146	100.0	87.0	13.0	3.4	9.6	0.0
Aggregate Short Writing			572	99.1	77.4	21.7	13.1	8.6	0.9
Aggregate All Items			931	98.2	76.7	21.5	13.6	7.8	1.8

Table 4.1.8 Validity Estimates for Writing: Junior Division (French)

Item Code	Booklet	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
12542_T	2	24	219	99.5	76.7	22.8	1.8	21.0	0.5
12542_V	2	24	176	100.0	79.5	20.5	8.0	12.5	0.0
Aggregate Long Writing			395	99.7	78.0	21.8	4.6	17.2	0.3
12462_T	1	13	202	99.0	63.4	35.6	14.4	21.3	1.0
12462_V	1	13	94	100.0	78.7	21.3	10.6	10.6	0.0
OQEF30601_T	2	17	71	98.6	49.3	49.3	22.5	26.8	1.4
OQEF30601_V	2	17	64	100.0	84.4	15.6	7.8	7.8	0.0
Aggregate Short Writing			431	99.3	67.5	31.8	13.9	17.9	0.7
Aggregate All Items			826	99.5	72.5	27.0	9.4	17.6	0.5

Table 4.1.9 Validity Estimates for Mathematics: Primary Division (English)

Item Code	Sequence	No. of Scores	% Exact Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
EQME20812	7	768	99.6	92.8	6.8	1.7	5.1	0.4
EQME31587	8	636	99.7	93.1	6.6	3.4	3.1	0.3
EQME31516	9	970	100.0	94.7	5.3	2.8	2.5	0.0
EQME10075	25	723	99.6	89.1	10.5	0.3	10.2	0.4
EQME31536	26	579	98.2	91.1	7.1	1.8	5.3	1.8
EQME31560	27	921	99.9	95.8	4.1	3.0	1.1	0.1
EQME31588	28	739	99.4	79.2	20.2	9.3	11.0	0.6
Aggregate		5336	99.6	91.1	8.4	3.2	5.2	0.4

Table 4.1.10 Validity Estimates for Mathematics: Junior Division (English)

Item Code	Sequence	No. of Scores	% Exact Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
OQMF20505	7	946	100.0	93.1	6.9	1.2	5.7	0.0
OQMF20508	8	900	99.9	97.2	2.7	2.1	0.6	0.1
OQMF31597	9	888	99.9	86.4	13.5	7.3	6.2	0.1
OQMF10063	10	829	100.0	93.2	6.8	3.7	3.0	0.0
OQMF10057	26	768	99.8	92.3	7.5	0.6	6.9	0.2
OQMF20517	27	713	93.5	81.7	11.7	9.0	2.7	6.5
OQMF31582	28	774	99.9	84.6	15.2	8.8	6.4	0.1
OQMF20535	29	786	99.5	93.1	6.4	3.2	3.2	0.5
Aggregate		6604	99.2	90.5	8.7	4.4	4.3	0.8

Table 4.1.11 Validity Estimates for Mathematics: Primary Division (French)

Item Code	Sequence	No. of Scores	% Exact Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
EQME31744	7	87	100.0	86.2	13.8	12.6	1.1	0.0
EQME21011	8	70	100.0	82.9	17.1	8.6	8.6	0.0
EQME21009	9	79	100.0	87.3	12.7	6.3	6.3	0.0
EQME31702	10	59	100.0	93.2	6.8	1.7	5.1	0.0
EQME21015	26	71	98.6	75.0	23.6	13.9	9.7	1.4
EQME21056	27	64	100.0	78.1	21.9	9.4	12.5	0.0
EQME31662	28	70	97.3	86.7	10.7	4.0	6.7	2.7
EQME10275	29	61	100.0	77.0	23.0	8.2	14.8	0.0
Aggregate		561	99.5	83.4	16.1	8.3	7.8	0.5

Table 4.1.12 Validity Estimates for Mathematics: Junior Division (French)

Item Code	Sequence	No. of Scores	% Exact Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
OQMF10112	7	72	100.0	77.8	22.2	20.8	1.4	0.0
OQMF20720	8	56	100.0	92.9	7.1	7.1	0.0	0.0
OQMF20782	9	87	100.0	88.5	11.5	1.1	10.3	0.0
OQMF10206	10	48	96.9	55.4	41.5	16.9	24.6	3.1
OQMF20756	26	70	100.0	92.9	7.1	5.7	1.4	0.0
OQMF20768	27	74	100.0	75.7	24.3	10.8	13.5	0.0
OQMF31641	28	68	100.0	79.4	20.6	8.8	11.8	0.0
OQMF20812	29	78	100.0	87.2	12.8	11.5	1.3	0.0
Aggregate		553	99.7	82.2	17.5	10.0	7.6	0.3

Validity: the Grade 9 Assessment of Mathematics (Academic and Applied)

Table 4.1.13 Validity Estimates for Grade 9 Applied Mathematics (English)

Administration	Item Code	Sequence	Item No.	No. of Scores	% Exact Plus Adjnt.	% Exact	% Adjnt.	% Adjnt.-Low	% Adjnt.-High	% Non-Adjnt.
Winter	EQME31763	9	8	158	98.7	84.8	13.9	13.9	0.0	1.3
	EQME31784	NR	9	206	99.5	92.7	6.8	3.4	3.4	0.5
	EQME31847	NR	10	197	100.0	95.9	4.1	3.6	0.5	0.0
	EQME31869	NR	11	208	99.5	90.9	8.7	5.3	3.4	0.5
	EQME31802	NR	26	212	98.1	89.6	8.5	4.2	4.2	1.9
	EQME21454	30	28	207	100.0	95.7	4.3	3.4	1.0	0.0
	EQME31826	23	29	194	99.5	92.8	6.7	4.1	2.6	0.5
Aggregate				1382	99.3	92.0	7.4	5.1	2.2	0.7
Spring	EQME31863	NR	8	178	99.4	97.2	2.2	0.0	2.2	0.6
	EQME31769	31	9	179	99.4	82.7	16.8	15.1	1.7	0.6
	EQME31786	NR	10	196	97.4	85.7	11.7	9.7	2.0	2.6
	EQME31842	8	11	208	95.2	81.3	13.9	11.5	2.4	4.8
	EQME21441	22	26	272	100.0	90.4	9.6	6.3	3.3	0.0
	EQME31828	NR	28	248	99.2	91.5	7.7	7.3	0.4	0.8
	EQME31865	21	29	284	100.0	76.8	23.2	3.9	19.4	0.0
Aggregate				1565	98.8	86.2	12.6	7.4	5.2	1.2
Aggregate Across Administrations				2947	99.1	89.1	10.0	6.3	3.7	0.9

Table 4.1.14 Validity Estimates for Grade 9 Academic Mathematics (English)

Administration	Item Code	Sequence	Item No.	No. of Scores	% Exact Plus Adjnt.	% Exact	% Adjnt.	% Adjnt.-Low	% Adjnt.-High	% Non-Adjnt.
Winter	EQME31964	NR	8	386	99.5	92.0	7.5	4.4	3.1	0.5
	EQME31894	NR	9	334	99.1	85.6	13.5	7.2	6.3	0.9
	EQME21275	22	10	312	99.4	89.4	9.9	3.8	6.1	0.6
	EQME31988	NR	11	349	100.0	98.0	2.0	1.4	0.6	0.0
	EQME31960	31	26	561	96.8	80.6	16.2	13.0	3.2	3.2
	EQME31906	14	28	525	99.8	81.5	18.3	8.0	10.3	0.2
	EQME31957	NR	29	577	99.0	79.5	19.4	13.9	5.5	1.0
Aggregate				3044	98.9	85.4	13.5	8.3	5.2	1.1
Spring	EQME21239	23	8	377	99.5	81.7	17.8	14.9	2.9	0.5
	EQME31927	NR	9	372	97.8	71.5	26.3	14.2	12.1	2.2
	EQME31980	6	10	403	99.5	91.1	8.4	4.5	4.0	0.5
	EQME31917	NR	11	365	98.9	89.9	9.0	3.3	5.8	1.1
	EQME31946	NR	26	545	99.3	87.5	11.7	9.0	2.8	0.7
	EQME31939	13	28	506	99.8	91.5	8.3	2.6	5.7	0.2
	EQME31963	30	29	546	100.0	89.2	10.8	5.7	5.1	0.0
Aggregate				3114	99.3	86.6	12.7	7.5	5.3	0.7
Aggregate Across Administrations				6158	99.1	86.0	13.1	7.9	5.2	0.9

Table 4.1.15 Validity Estimates for Grade 9 Applied Mathematics (French)

Administration	Item Code	Sequence	Item No.	No. of Scores	% Exact Plus Adjnt.	% Exact	% Adjnt.	% Adjnt.-Low	% Adjnt.-High	% Non-Adjnt.
Winter	OQMF21308	8	8	26	100.0	100.0	0.0	0.0	0.0	0.0
	OQMF31838	0	9	29	96.6	93.1	3.4	3.4	0.0	3.4
	OQMF10313	19	10	21	100.0	85.7	14.3	9.5	4.8	0.0
	OQMF21363	0	11	36	100.0	94.4	5.6	5.6	0.0	0.0
	OQMF31765	0	26	31	100.0	83.9	16.1	12.9	3.2	0.0
	OQMF31793	0	28	30	100.0	86.7	13.3	13.3	0.0	0.0
	OQMF21420	30	29	32	100.0	100.0	0.0	0.0	0.0	0.0
Aggregate				205	99.5	92.2	7.3	6.3	1.0	0.5
Spring	OQMF21364	18	8	91	98.9	86.8	12.1	6.6	5.5	1.1
	OQMF21334	9	9	84	100.0	94.0	6.0	2.4	3.6	0.0
	OQMF21379	0	10	76	100.0	85.5	14.5	0.0	14.5	0.0
	OQMF21392	29	11	52	100.0	98.1	1.9	1.9	0.0	0.0
	OQMF31841	0	26	90	100.0	76.7	23.3	0.0	23.3	0.0
	OQMF21311	0	28	53	100.0	84.9	15.1	3.8	11.3	0.0
	OQMF21348	0	29	77	96.1	92.2	3.9	2.6	1.3	3.9
Aggregate				523	99.2	87.8	11.5	2.5	9.0	0.8
Aggregate Across Administrations				728	99.4	90.0	9.4	4.4	5.0	0.6

Table 4.1.16 Validity Estimates for Grade 9 Academic Mathematics (French)

Administration	Item Code	Sequence	Item No.	No. of Scores	% Exact Plus Adjnt.	% Exact	% Adjnt.	% Adjnt. Low	% Adjnt. High	% Non-Adjnt.
Winter	OQMF31881	22	8	102	100.0	98.0	2.0	2.0	0.0	0.0
	OQMF20904	0	9	135	100.0	85.2	14.8	3.0	11.9	0.0
	OQMF20926	30	10	140	100.0	99.3	0.7	0.0	0.7	0.0
	OQMF31940	0	11	122	100.0	77.9	22.1	0.0	22.1	0.0
	OQMF20912	8	26	212	99.5	94.3	5.2	5.2	0.0	0.5
	OQMF31921	15	28	215	99.5	97.2	2.3	1.9	0.5	0.5
	OQMF31902	0	29	210	100.0	91.0	9.0	8.6	0.5	0.0
Aggregate				1136	99.8	92.3	7.5	3.4	4.0	0.2
Spring	OQMF31881	22	8	187	100.0	98.4	1.6	0.5	1.1	0.0
	OQMF10375	31	9	178	100.0	93.8	6.2	1.1	5.1	0.0
	OQMF20916	16	10	199	100.0	98.0	2.0	2.0	0.0	0.0
	OQMF31898	0	11	191	97.4	90.1	7.3	5.2	2.1	2.6
	OQMF20905	7	26	230	100.0	98.3	1.7	1.3	0.4	0.0
	OQMF31874	0	28	217	100.0	93.1	6.9	6.9	0.0	0.0
	OQMF31899	0	29	230	96.1	95.7	0.4	0.4	0.0	3.9
Aggregate				1432	99.0	95.4	3.6	2.5	1.1	1.0
Aggregate Across Administrations				2568	99.4	93.9	5.6	3.0	2.6	0.6

Validity: the Ontario Secondary School Literacy Test

Table 4.1.17 Validity Estimates for Reading: OSSLT (English)

Item Code	Section	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
EQRE32149.185	I	6	7 042	98.5	81.7	16.8	9.0	7.8	1.5
EQRE32099.180	V	7	7 869	99.0	82.3	16.6	5.1	11.5	1.0
EQRE32177.188	IX	6	7 797	99.4	72.3	27.2	15.3	11.8	0.6
EQRE32178.188	IX	7	7 538	99.2	79.3	19.9	4.8	15.2	0.8
Aggregate			30 246	99.0	78.8	20.2	8.6	11.6	1.0

Table 4.1.18 Validity Estimates for Writing: OSSLT (English)

Item Code	Section	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
1W3_066T	IV	1	18 258	92.1	53.4	38.7	23.8	14.9	7.9
1W3_066V	IV	1	18 258	99.3	70.4	28.9	12.9	16.0	0.7
1W2_059T	VI	1	23 122	98.5	69.3	29.3	16.1	13.2	1.5
1W2_059V	VI	1	23 122	99.5	77.7	21.8	11.1	10.7	0.5
Aggregate Long Writing			82 760	97.5	68.4	29.2	15.7	13.5	2.5
EQWE30565T	III	1	8 631	99.3	77.6	21.6	13.3	8.3	0.7
EQWE30574T	X	1	9 416	99.5	85.0	14.6	6.4	8.2	0.5
Aggregate Short Writing			18 047	99.4	81.5	18.0	9.7	8.3	0.6
Aggregate All Items			100 807	97.9	70.7	27.2	14.6	12.5	2.1

Table 4.1.19 Validity Estimates for Reading: OSSLT (French)

Item Code	Section	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
OQLF31468.128	I	6	178	100.0	94.9	5.1	3.4	1.7	0.0
OQLF31506.125	V	7	180	99.4	98.9	0.6	0.0	0.6	0.6
OQLF31402.130	IX	6	191	100.0	83.8	16.2	8.4	7.9	0.0
OQLF31401.130	IX	7	113	100.0	93.8	6.2	4.4	1.8	0.0
Aggregate			662	99.8	92.6	7.3	4.1	3.2	0.2

Table 4.1.20 Validity Estimates for Writing: OSSLT (French)

Item Code	Section	Sequence	No. of Scores	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Adjacent-Low	% Adjacent-High	% Non-Adjacent
2É3_029T	IV	1	781	95.3	60.2	35.1	9.6	25.5	4.7
2É3_029V	IV	1	781	96.0	65.6	30.5	11.9	18.6	4.0
2É2_028T	VI	1	1130	96.3	65.6	30.7	11.8	18.9	3.7
2É2_028V	VI	1	1130	99.3	73.4	25.9	12.8	13.1	0.7
Aggregate Long Writing			3822	96.9	66.8	30.1	11.7	18.5	3.1
OQEF30690T	III	1	298	99.3	80.5	18.8	10.7	8.1	0.7
OQEF30684T	X	1	336	95.5	72.3	23.2	9.8	13.4	4.5
Aggregate Short Writing			634	97.3	76.2	21.1	10.3	10.9	2.7
Aggregate All Items			4456	97.0	68.1	28.9	11.5	17.4	3.0

Interrater Reliability: the Primary and Junior Assessments

Table 4.1.21 Interrater Reliability Estimates for Reading: Primary Division (English)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
EQRE31615.141	1	5	1 024	93.0	62.0	31.0	7.0
EQRE31614.141	1	6	1 058	93.6	65.3	28.3	6.4
EQRE31660.146	1	11	974	97.6	76.8	20.8	2.4
EQRE31661.146	1	12	949	98.6	78.4	20.2	1.4
EQRE31697.150	1	18	1 113	97.8	71.3	26.4	2.2
EQRE31696.150	1	19	1 111	97.7	73.7	23.9	2.3
EQRE31809.159	2	11	1 265	98.1	68.9	29.2	1.9
EQRE31810.159	2	12	1 240	97.0	68.6	28.4	3.0
EQRE31733.154	2	22	865	98.7	84.2	14.6	1.3
EQRE31735.154	2	23	922	98.2	82.1	16.1	1.8
Aggregate			10 521	97.0	72.6	24.4	3.0

Table 4.1.22 Interrater Reliability Estimates for Reading: Junior Division (English)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
EQRE31874.162	1	5	1 436	97.1	67.5	29.6	2.9
EQRE31875.162	1	6	1 440	97.6	70.3	27.3	2.4
EQRE31919.167	1	11	1 458	98.6	65.9	32.7	1.4
EQRE31920.167	1	12	1 463	97.0	64.7	32.3	3.0
EQRE31957.171	1	18	1 474	97.8	68.4	29.4	2.2
EQRE31955.171	1	19	1 467	97.9	68.1	29.8	2.1
EQRE32031.177	2	11	1 665	97.2	85.4	11.8	2.8
EQRE32029.177	2	12	1 663	97.4	79.7	17.7	2.6
EQRE31974.173	2	22	1 689	97.8	73.7	24.1	2.2
EQRE31973.173	2	23	1 748	96.9	71.7	25.2	3.1
Aggregate			15 503	97.5	71.9	25.6	2.5

Table 4.1.23 Interrater Reliability Estimates for Reading: Primary Division (French)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
OQLF31051.103	1	5	35	100.0	57.1	42.9	0.0
OQLF31052.103	1	6	45	100.0	77.8	22.2	0.0
OQLF31069.105	1	11	113	99.1	59.3	39.8	0.9
OQLF31071.105	1	12	101	99.0	80.2	18.8	1.0
OQLF31078.106	1	18	71	100.0	71.8	28.2	0.0
OQLF31080.106	1	19	80	97.5	56.2	41.2	2.5
OQLF20445.060	2	11	67	100.0	82.1	17.9	0.0
OQLF20448.060	2	12	56	98.2	60.7	37.5	1.8
OQLF20549.069	2	22	56	98.2	83.9	14.3	1.8
OQLF20550.069	2	23	62	98.4	67.7	30.6	1.6
Aggregate			686	99.0	69.5	29.4	1.0

Table 4.1.24 Interrater Reliability Estimates for Reading: Junior Division (French)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
OQLF31251.116	1	5	226	97.8	61.1	36.7	2.2
OQLF31252.116(A)	1	6	235	96.6	55.7	40.9	3.4
OQLF31260.117	1	11	67	98.5	76.1	22.4	1.5
OQLF31262.117	1	12	67	94.0	58.2	35.8	6.0
OQLF31309.120	1	18	112	98.2	62.5	35.7	1.8
OQLF31288.120	1	19	65	98.5	70.8	27.7	1.5
OQLF31215.113	2	11	154	100.0	68.8	31.2	0.0
OQLF31216.113	2	12	136	95.6	62.5	33.1	4.4
OQLF31305.122	2	22	83	98.8	66.3	32.5	1.2
OQLF31307.122	2	23	94	95.7	61.7	34.0	4.3
Aggregate			1239	97.4	62.9	34.5	2.6

Table 4.1.25 Interrater Reliability Estimates for Writing: Primary Division (English)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
11042 T	2	24	1100	98.6	72.1	26.5	1.4
11042 V	2	24	1215	99.0	77.4	21.6	1.0
Aggregate Long Writing			2315	98.8	74.9	24.0	1.2
EQWE20246 T	1	13	1538	96.0	59.9	36.1	4.0
EQWE20246 V	1	13	1479	98.2	70.9	27.3	1.8
EQWE30427 T	2	17	900	97.9	75.4	22.4	2.1
EQWE30427 V	2	17	888	99.3	82.2	17.1	0.7
Aggregate Short Writing			4805	97.7	70.3	27.3	2.3
Aggregate All Items			7120	98.0	71.8	26.2	2.0

Table 4.1.26 Interrater Reliability Estimates for Writing: Junior Division (English)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
10715 T	2	24	2194	95.4	68.6	26.7	4.6
10715 V	2	24	2218	97.8	69.5	28.3	2.2
Aggregate Long Writing			4412	96.6	69.1	27.5	3.4
EQWE30498 T	1	13	1508	97.3	72.1	25.3	2.7
EQWE30498 V	1	13	1673	98.8	76.2	22.7	1.2
EQWE30504 T	2	17	1637	97.3	63.7	33.7	2.7
EQWE30504 V	2	17	1623	98.0	68.8	29.1	2.0
Aggregate Short Writing			6441	97.9	70.2	27.7	2.1
Aggregate All Items			10853	97.3	69.7	27.6	2.7

Table 4.1.27 Interrater Reliability Estimates for Writing: Primary Division (French)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
11248_T	2	24	108	97.2	63.9	33.3	2.8
11248_V	2	24	100	99.0	71.0	28.0	1.0
Aggregate Long Writing			208	98.1	67.3	30.8	1.9
OQEF30521_T	1	13	110	98.2	66.4	31.8	1.8
OQEF30521_V	1	13	114	100.0	71.1	28.9	0.0
OQEF30523_T	2	17	202	99.0	68.3	30.7	1.0
OQEF30523_V	2	17	193	100.0	77.2	22.8	0.0
Aggregate Short Writing			619	99.4	71.2	28.1	0.6
Aggregate All Items			827	99.0	70.3	28.8	1.0

Table 4.1.28 Interrater Reliability Estimates for Writing: Junior Division (French)

Item Code	Booklet	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
12542_T	2	24	99	100.0	59.6	40.4	0.0
12542_V	2	24	124	98.4	73.4	25.0	1.6
Aggregate Long Writing			223	99.1	67.3	31.8	0.9
12462_T	1	13	45	95.6	55.6	40.0	4.4
12462_V	1	13	97	100.0	59.8	40.2	0.0
OQEF30601_T	2	17	37	86.5	32.4	54.1	13.5
OQEF30601_V	2	17	41	100.0	58.5	41.5	0.0
Aggregate Short Writing			220	96.8	54.1	42.7	3.2
Aggregate All Items			443	98.0	60.7	37.2	2.0

Table 4.1.29 Interrater Reliability Estimates for Mathematics: Primary Division (English)

Item Code	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
EQME20812	7	948	98.5	84.3	14.2	1.5
EQME31587	8	1030	98.3	75.9	22.4	1.7
EQME31516	9	866	97.7	80.1	17.6	2.3
EQME10075	25	1392	94.8	71.5	23.3	5.2
EQME31536	26	1445	98.1	78.6	19.4	1.9
EQME31560	27	1291	98.8	76.8	22.0	1.2
EQME31588	28	1341	94.3	78.2	16.1	5.7
Aggregate		8313	97.1	77.6	19.5	2.9

Table 4.1.30 Interrater Reliability Estimates for Mathematics: Junior Division (English)

Item Code	Sequence	No. of Pairs	% Exact Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
OQMF20505	7	1421	98.7	92.0	6.7	1.3
OQMF20508	8	1441	98.6	79.5	19.2	1.4
OQMF31597	9	1423	98.2	70.8	27.4	1.8
OQMF10063	10	1479	98.2	81.5	16.8	1.8
OQMF10057	26	1683	99.5	96.4	3.1	0.5
OQMF20517	27	1730	97.6	83.7	13.9	2.4
OQMF31582	28	1685	94.8	70.6	24.2	5.2
OQMF20535	29	1690	98.0	80.9	17.2	2.0
Aggregate		12552	97.9	82.0	15.9	2.1

Table 4.1.31 Interrater Reliability Estimates for Mathematics: Primary Division (French)

Item Code	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
EQME31744	7	138	96.4	78.3	18.1	3.6
EQME21011	8	145	100.0	88.3	11.7	0.0
EQME21009	9	141	97.9	86.5	11.3	2.1
EQME31702	10	151	95.4	77.5	17.9	4.6
EQME21015	26	180	97.2	87.8	9.4	2.8
EQME21056	27	186	95.7	73.7	22.0	4.3
EQME31662	28	181	97.2	86.7	10.5	2.8
EQME10275	29	185	96.2	72.4	23.8	3.8
Aggregate		1307	96.9	81.2	15.8	3.1

Table 4.1.32 Interrater Reliability Estimates for Mathematics: Junior Division (French)

Item Code	Sequence	No. of Pairs	% Exact Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
OQMF10112	7	178	95.5	75.3	20.2	4.5
OQMF20720	8	187	96.3	88.2	8.0	3.7
OQMF20782	9	170	95.9	78.8	17.1	4.1
OQMF10206	10	182	91.2	65.9	25.3	8.8
OQMF20756	26	200	99.0	74.5	24.5	1.0
OQMF20768	27	201	97.0	87.6	9.5	3.0
OQMF31641	28	202	97.5	87.6	9.9	2.5
OQMF20812	29	196	99.5	80.1	19.4	0.5
Aggregate		1516	96.6	79.9	16.6	3.4

Interrater Reliability: the Grade 9 Assessment of Mathematics (Academic and Applied)

Table 4.1.33 Interrater Reliability Estimates for Grade 9 Applied Mathematics (English)

Administration	Item Code	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
Winter	EQME31763	9	568	95.6	88.7	6.9	4.4
	EQME31784	NR	558	97.3	83.3	14.0	2.7
	EQME31847	NR	560	96.4	76.1	20.4	3.6
	EQME31869	NR	559	97.3	90.2	7.2	2.7
	EQME31802	NR	179	95.0	90.5	4.5	5.0
	EQME21454	30	184	96.7	89.7	7.1	3.3
	EQME31826	23	183	96.2	79.2	16.9	3.8
Aggregate			2791	96.5	85.0	11.6	3.5
Spring	EQME31863	NR	279	98.6	89.6	9.0	1.4
	EQME31769	31	279	97.5	84.9	12.5	2.5
	EQME31786	NR	269	96.3	74.7	21.6	3.7
	EQME31842	8	268	95.5	73.9	21.6	4.5
	EQME21441	22	219	95.0	79.5	15.5	5.0
	EQME31828	NR	224	99.6	89.7	9.8	0.4
	EQME31865	21	217	99.1	82.5	16.6	0.9
Aggregate			1755	97.3	82.1	15.3	2.7
Aggregate Across Administrations			4546	96.9	83.5	13.4	3.1

Table 4.1.34 Interrater Reliability Estimates for Grade 9 Academic Mathematics (English)

Administration	Item Code	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
Winter	EQME31964	NR	4541	98.7	90.7	8.0	1.3
	EQME31894	NR	4562	93.8	67.2	26.6	6.2
	EQME21275	22	4581	97.4	83.0	14.3	2.6
	EQME31988	NR	4558	98.5	86.0	12.5	1.5
	EQME31960	31	1133	98.2	91.2	7.1	1.8
	EQME31906	14	1152	96.6	77.3	19.4	3.4
	EQME31957	NR	1134	98.4	80.9	17.5	1.6
Aggregate			21661	97.2	81.9	15.3	2.8
Spring	EQME21239	23	608	97.5	77.5	20.1	2.5
	EQME31927	NR	609	97.4	75.7	21.7	2.6
	EQME31980	6	600	99.0	82.8	16.2	1.0
	EQME31917	NR	616	98.9	89.6	9.3	1.1
	EQME31946	NR	2248	95.1	65.3	29.8	4.9
	EQME31939	13	2262	97.8	80.9	17.0	2.2
	EQME31963	30	2243	98.5	90.0	8.6	1.5
Aggregate			9186	97.4	79.4	18.0	2.6
Aggregate Across Administrations			30847	97.3	80.7	16.6	2.7

Table 4.1.35 Interrater Reliability Estimates for Grade 9 Applied Mathematics (French)

Administration	Item Code	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
Winter	OQMF21308	8	23	91.3	69.6	21.7	8.7
	OQMF31838	NR	23	91.3	82.6	8.7	8.7
	OQMF10313	19	25	96.0	84.0	12.0	4.0
	OQMF21363	NR	22	100.0	90.9	9.1	0.0
	OQMF31765	NR	21	100.0	81.0	19.0	0.0
	OQMF31793	NR	20	95.0	85.0	10.0	5.0
	OQMF21420	30	19	94.7	68.4	26.3	5.3
Aggregate	Aggregate		153	95.4	80.4	15.0	4.6
Spring	OQMF21364	18	45	95.6	71.1	24.4	4.4
	OQMF21334	9	45	95.6	73.3	22.2	4.4
	OQMF21379	0	46	91.3	71.7	19.6	8.7
	OQMF21392	29	49	98.0	93.9	4.1	2.0
	OQMF31841	NR	33	100.0	90.9	9.1	0.0
	OQMF21311	NR	40	95.0	67.5	27.5	5.0
	OQMF21348	NR	37	94.6	81.1	13.5	5.4
Aggregate			295	95.6	78.3	17.3	4.4
Aggregate Across Administrations			448	95.5	79.3	16.2	4.5

Table 4.1.36 Interrater Reliability Estimates for Grade 9 Academic Mathematics (French)

Administration	Item Code	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
Winter	OQMF31881	22	66	100.0	92.4	7.6	0.0
	OQMF20904	NR	65	100.0	93.8	6.2	0.0
	OQMF20926	30	64	100.0	84.4	15.6	0.0
	OQMF31940	NR	69	95.7	84.1	11.6	4.3
	OQMF20912	8	169	99.4	92.9	6.5	0.6
	OQMF31921	15	170	94.7	85.3	9.4	5.3
	OQMF31902	NR	171	100.0	90.6	9.4	0.0
Aggregate			774	98.3	89.3	9.0	1.7
Spring	OQMF31881	22	104	99.0	76.9	22.1	1.0
	OQMF10375	31	103	93.2	85.4	7.8	6.8
	OQMF20916	16	102	96.1	83.3	12.7	3.9
	OQMF31898	NR	105	90.5	77.1	13.3	9.5
	OQMF20905	7	70	97.1	90.0	7.1	2.9
	OQMF31874	NR	76	100.0	92.1	7.9	0.0
	OQMF31899	NR	69	100.0	88.4	11.6	0.0
Aggregate			629	96.2	83.9	12.2	3.8
Aggregate Across Administrations			1403	97.3	86.6	10.6	2.7

Interrater Reliability: the Ontario Secondary School Literacy Test (OSSLT)

Table 4.1.37 Interrater Reliability Estimates for Reading: OSSLT (English)

Item Code	Section	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
EQRE32149.185	I	6	184 712	97.2	64.4	32.8	2.8
EQRE32099.180	V	7	184 712	97.6	74.4	23.2	2.4
EQRE32177.188	IX	6	184 705	98.0	61.8	36.2	2.0
EQRE32178.188	IX	7	184 706	98.5	63.3	35.2	1.5
Aggregate			738 835	97.8	66.0	31.9	2.2

Table 4.1.38 Interrater Reliability Estimates for Writing: OSSLT (English)

Item Code	Section	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
1W3_066T	IV	1	184 711	87.8	42.8	45.0	12.2
1W3_066V	IV	1	184 711	98.0	57.1	40.8	2.0
1W2_059T	VI	1	184 704	93.3	48.5	44.8	6.7
1W2_059V	VI	1	184 704	98.6	61.2	37.4	1.4
Aggregate Long Writing			738 830	94.4	52.4	42.0	5.6
EQWE30565T	III	1	184 713	97.2	63.3	33.9	2.8
EQWE30574T	X	1	184 706	98.2	63.9	34.3	1.8
Aggregate Short Writing			369 419	97.7	63.6	34.1	2.3
Aggregate All Items			1 108 249	95.5	56.2	39.4	4.5

Table 4.1.39 Interrater Reliability Estimates for Reading: OSSLT (French)

Item Code	Section	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
OQLF31468.128	I	6	6 318	98.6	74.4	24.3	1.4
OQLF31506.125	V	7	6 318	95.5	85.0	10.5	4.5
OQLF31402.130	IX	6	6 317	96.4	62.2	34.2	3.6
OQLF31401.130	IX	7	6 317	97.0	62.6	34.4	3.0
Aggregate			25 270	96.9	71.1	25.8	3.1

Table 4.1.40 Interrater Reliability Estimates for Writing: OSSLT (French)

Item Code	Section	Sequence	No. of Pairs	% Exact-Plus-Adjacent	% Exact	% Adjacent	% Non-Adjacent
2É3_029T	IV	1	6 318	89.6	45.2	44.3	10.4
2É3_029V	IV	1	6 318	98.0	61.4	36.5	2.0
2É2_028T	VI	1	6 318	88.8	44.4	44.4	11.2
2É2_028V	VI	1	6 318	97.2	57.7	39.6	2.8
Aggregate Long Writing			25 272	93.4	52.2	41.2	6.6
OQEF30690T	III	1	6 318	95.2	58.7	36.5	4.8
OQEF30684T	X	1	6 318	92.4	53.3	39.1	7.6
Aggregate Short Writing			12 636	93.8	56.0	37.8	6.2
Aggregate All Items			37 908	93.5	53.5	40.1	6.5

APPENDIX 7.1: SCORE DISTRIBUTIONS AND ITEM STATISTICS FOR THE 2010 EQAO ASSESSMENTS

This appendix presents the classical item statistics and IRT item parameter estimates for the operational items and the DIF statistics for individual items with respect to gender and students who are second language learners. For the French-language versions of the Grade 9 Assessment of Mathematics and the OSSLT, DIF analysis for SLLs was not conducted due to the small number of students in the French-language SLL population.

Classical item statistics and IRT item parameter estimates are combined into tables for each assessment: Tables 7.1.1–7.1.24 for the primary- and junior assessments, Tables 7.1.49–7.1.64 for the Grade 9 Assessment of Mathematics and Tables 7.1.7 –7.1.82 for the OSSLT. The distribution of score points and the step parameter estimates are also provided for open-response items.

DIF statistics for individual items are shown in Tables 7.1.25–7.1.48 for the primary- and junior-division assessments, Tables 7.1.65–7.1.76 for the Grade 9 Assessment of Mathematics and Tables 7.1.83–7.1.85 for the OSSLT.

The Primary and Junior Assessments

Classical Item Statistics and IRT Item Parameters

Table 7.1.1 Item Statistics: Primary Reading (English)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQRE31619.141	1	1	R2.0	I	4	0.62	0.35	-0.22	0.67
EQRE31617.141	1	2	R1.0	C	4	0.87	0.37	-1.77	0.81
EQRE31621.141	1	3	R3.0	I	3	0.77	0.33	-1.28	0.58
EQRE31622.141	1	4	R3.0	C	2	0.76	0.38	-1.00	0.74
EQRE31615.141	1	5	R1.0	C	4*	0.44 (1.76)	0.42	-0.07	0.39
EQRE31614.141	1	6	R1.0	C	4*	0.51 (2.04)	0.46	-0.41	0.44
EQRE31666.146	1	7	R3.0	E	3	0.71	0.31	-1.03	0.46
EQRE31667.146	1	8	R3.0	I	4	0.72	0.41	-0.73	0.77
EQRE31662.146	1	9	R1.0	I	4	0.71	0.32	-0.83	0.56
EQRE31664.146	1	10	R2.0	C	3	0.68	0.39	-0.56	0.72
EQRE31660.146	1	11	R1.0	C	4*	0.47 (1.87)	0.45	-0.04	0.43
EQRE31661.146	1	12	R2.0	I	4*	0.44 (1.75)	0.42	-0.17	0.38
EQRE31702.150	1	14	R3.0	C	2	0.75	0.40	-0.99	0.74
EQRE31698.150	1	15	R1.0	I	1	0.81	0.35	-1.45	0.63
EQRE31700.150	1	16	R2.0	C	3	0.48	0.26	0.68	0.52
EQRE31703.150	1	17	R3.0	E	2	0.71	0.44	-0.64	0.90
EQRE31697.150	1	18	R1.0	E	4*	0.61 (2.45)	0.52	-1.02	0.49
EQRE31696.150	1	19	R1.0	C	4*	0.54 (2.16)	0.30	-2.21	0.19
EQRE31819.159	2	1	R2.0	C	1	0.68	0.32	-0.64	0.55
EQRE31816.159	2	2	R1.0	C	2	0.85	0.35	-1.77	0.69
EQRE31824.159	2	3	R3.0	I	4	0.89	0.41	-1.81	0.93
EQRE31817.159	2	4	R2.0	I	2	0.73	0.41	-0.84	0.80
EQRE31822.159	2	5	R3.0	C	2	0.87	0.48	-1.53	1.24
EQRE31814.159	2	6	R1.0	E	3	0.76	0.43	-0.94	0.89
EQRE31811.159	2	7	R1.0	I	4	0.68	0.40	-0.47	0.91
EQRE31812.159	2	8	R1.0	I	1	0.57	0.31	0.07	0.60
EQRE31820.159	2	9	R3.0	I	3	0.75	0.51	-0.69	1.37
EQRE31818.159	2	10	R2.0	C	4	0.77	0.44	-0.88	0.97
EQRE31809.159	2	11	R1.0	C	4*	0.53 (2.12)	0.50	-0.49	0.45
EQRE31810.159	2	12	R2.0	I	4*	0.53 (2.12)	0.51	-0.35	0.47
EQRE31736.154	2	18	R2.0	I	1	0.89	0.31	-2.31	0.64
EQRE31737.154	2	19	R1.0	I	4	0.89	0.40	-1.90	0.87
EQRE31738.154	2	20	R1.0	I	2	0.92	0.31	-2.77	0.59
EQRE31740.154	2	21	R3.0	I	2	0.61	0.38	-0.19	0.68
EQRE31733.154	2	22	R1.0	C	4*	0.65 (2.61)	0.43	-1.32	0.36
EQRE31735.154	2	23	R2.0	C	4*	0.43 (1.71)	0.43	-0.02	0.43

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. C = connections; E = explicit; I = implicit; R = reading; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.2 Distribution of Score Points and Step Parameters for Open-Response Items: Primary Reading (English)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
EQRE31615.141	1	5	% of Students	0.64	1.82	40.41	38.62	15.78	2.73
			Parameters			-5.28	-0.14	1.69	3.46
EQRE31614.141	1	6	% of Students	0.68	1.27	21.98	50.23	22.31	3.53
			Parameters			-4.62	-1.57	1.31	3.25
EQRE31660.146	1	11	% of Students	1.32	1.64	29.67	47.73	16.97	2.66
			Parameters			-4.34	-0.93	1.74	3.36
EQRE31661.146	1	12	% of Students	1.53	1.00	43.11	34.66	15.99	3.72
			Parameters			-5.45	0.19	1.53	3.07
EQRE31697.150	1	18	% of Students	0.68	0.47	13.77	31.82	45.52	7.73
			Parameters			-4.61	-1.73	-0.44	2.69
EQRE31696.150	1	19	% of Students	0.95	0.36	35.53	25.21	22.02	15.92
			Parameters			-11.55	0.87	0.51	1.33
EQRE31809.159	2	11	% of Students	0.80	1.64	24.44	37.33	30.51	5.29
			Parameters			-4.41	-1.02	0.46	3.02
EQRE31810.159	2	12	% of Students	0.96	1.49	21.36	41.03	31.99	3.16
			Parameters			-4.14	-1.32	0.48	3.60
EQRE31733.154	2	22	% of Students	0.67	0.48	10.92	21.45	58.47	8.00
			Parameters			-5.40	-1.80	-1.73	3.66
EQRE31735.154	2	23	% of Students	1.45	1.65	37.96	47.88	7.38	3.69
			Parameters			-4.53	-0.49	3.02	1.93

Note. The total number of students is 119 249.

Table 7.1.3 Item Statistics: Junior Reading (English)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQRE31868.162	1	1	R1.0	E	4	0.86	0.36	-1.95	0.67
EQRE31873.162	1	2	R3.0	I	2	0.91	0.26	-2.58	0.60
EQRE31870.162	1	3	R2.0	C	2	0.81	0.37	-1.46	0.70
EQRE31869.162	1	4	R1.0	I	1	0.90	0.31	-2.34	0.67
EQRE31874.162	1	5	R1.0	C	4*	0.62 (2.46)	0.43	-1.80	0.36
EQRE31875.162	1	6	R2.0	I	4*	0.57 (2.26)	0.48	-1.18	0.41
EQRE31913.167	1	7	R1.0	E	2	0.78	0.34	-1.22	0.63
EQRE31915.167	1	8	R2.0	C	3	0.60	0.40	-0.12	0.81
EQRE31918.167	1	9	R3.0	C	1	0.78	0.31	-1.46	0.51
EQRE31914.167	1	10	R1.0	I	2	0.78	0.34	-1.44	0.55
EQRE31919.167	1	11	R1.0	C	4*	0.59 (2.36)	0.53	-0.86	0.61
EQRE31920.167	1	12	R1.0	C	4*	0.50 (1.98)	0.54	-0.33	0.58
EQRE31950.171	1	14	R1.0	C	3	0.79	0.36	-1.35	0.65
EQRE31952.171	1	15	R2.0	I	1	0.81	0.42	-1.33	0.85
EQRE31949.171	1	16	R1.0	I	2	0.58	0.42	-0.01	0.86
EQRE31953.171	1	17	R3.0	I	2	0.53	0.33	0.26	0.57
EQRE31957.171	1	18	R1.0	C	4*	0.56 (2.24)	0.40	-0.96	0.38
EQRE31955.171	1	19	R1.0	C	4*	0.57 (2.26)	0.42	-0.88	0.42
EQRE32023.177	2	1	R2.0	I	2	0.55	0.32	0.17	0.58
EQRE32018.177	2	2	R1.0	I	3	0.73	0.48	-0.74	1.08
EQRE32014.177	2	3	R1.0	E	4	0.97	0.27	-2.94	1.00
EQRE32024.177	2	4	R2.0	I	1	0.88	0.37	-1.84	0.81
EQRE32021.177	2	5	R1.0	I	3	0.62	0.36	-0.24	0.66
EQRE32015.177	2	6	R1.0	E	4	0.63	0.30	-0.34	0.49
EQRE32027.177	2	7	R3.0	C	4	0.58	0.30	-0.09	0.45
EQRE32026.177	2	8	R3.0	E	2	0.83	0.38	-1.53	0.74
EQRE32020.177	2	9	R1.0	I	2	0.72	0.39	-0.84	0.69
EQRE32019.177	2	10	R1.0	I	2	0.82	0.30	-1.69	0.54
EQRE32031.177	2	11	R2.0	I	4*	0.41 (1.64)	0.47	-0.01	0.48
EQRE32029.177	2	12	R1.0	I	4*	0.47 (1.89)	0.43	-0.09	0.40
EQRE31967.173	2	18	R1.0	E	3	0.90	0.29	-2.80	0.53
EQRE31969.173	2	19	R1.0	I	3	0.84	0.40	-1.52	0.82
EQRE31970.173	2	20	R2.0	I	4	0.73	0.34	-1.02	0.56
EQRE31972.173	2	21	R3.0	C	4	0.51	0.28	0.49	0.53
EQRE31974.173	2	22	R2.0	I	4*	0.55 (2.21)	0.46	-0.82	0.50
EQRE31973.173	2	23	R1.0	C	4*	0.48 (1.91)	0.46	-0.40	0.50

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. C = connections; E = explicit; I = implicit; R = reading. () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.4 Distribution of Score Points and Step Parameters for Open-Response Items: Junior Reading (English)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
EQRE31874.162	1	5	% of Students	0.24	0.09	15.08	30.46	46.94	7.20
			Parameters			-8.09	-1.87	-0.81	3.56
EQRE31875.162	1	6	% of Students	0.46	0.24	20.65	36.82	35.94	5.89
			Parameters			-6.50	-1.48	0.07	3.17
EQRE31919.167	1	11	% of Students	0.45	0.20	9.79	45.90	40.19	3.48
			Parameters			-4.28	-2.41	0.11	3.15
EQRE31920.167	1	12	% of Students	1.11	0.49	30.43	39.37	25.15	3.44
			Parameters			-4.21	-0.73	0.69	2.93
EQRE31957.171	1	18	% of Students	0.42	0.20	13.70	49.73	32.58	3.37
			Parameters			-6.07	-2.59	0.74	4.08
EQRE31955.171	1	19	% of Students	0.62	0.15	11.44	52.33	31.99	3.47
			Parameters			-5.24	-2.79	0.76	3.75
EQRE32031.177	2	11	% of Students	1.92	0.88	54.10	23.57	15.58	3.95
			Parameters			-4.53	0.88	0.93	2.67
EQRE32029.177	2	12	% of Students	2.48	0.84	31.29	42.71	18.74	3.93
			Parameters			-4.14	-0.75	1.47	3.08
EQRE31974.173	2	22	% of Students	0.63	0.18	12.04	56.52	26.70	3.93
			Parameters			-4.74	-2.55	0.99	3.04
EQRE31973.173	2	23	% of Students	0.76	0.16	29.08	49.96	17.70	2.34
			Parameters			-5.45	-1.03	1.54	3.32

Note. The total number of students is 130 171.

Table 7.1.5 Item Statistics: Primary Reading (French)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQLF31050.103	1	1	C	L	4	0.58	0.31	0.01	0.57
OQLF31048.103	1	2	C	I	1	0.54	0.35	0.24	0.78
OQLF31046.103	1	3	A	L	3	0.80	0.23	-1.86	0.39
OQLF31047.103	1	4	B	I	4	0.89	0.31	-2.13	0.66
OQLF31051.103	1	5	A	L	4*	0.54 (2.14)	0.33	-1.13	0.30
OQLF31052.103	1	6	A	L	4*	0.59 (2.37)	0.40	-1.15	0.38
OQLF31065.105	1	7	B	L	2	0.86	0.29	-1.96	0.57
OQLF31066.105	1	8	C	E	2	0.85	0.38	-1.48	0.85
OQLF31068.105	1	9	C	I	4	0.74	0.37	-0.80	0.76
OQLF31064.105	1	10	A	I	1	0.69	0.32	-0.67	0.57
OQLF31069.105	1	11	A	I	4*	0.60 (2.40)	0.51	-1.17	0.48
OQLF31071.105	1	12	A	L	4*	0.62 (2.49)	0.49	-0.95	0.51
OQLF31074.106	1	14	B	L	3	0.85	0.23	-2.25	0.43
OQLF31073.106	1	15	A	I	4	0.72	0.36	-0.76	0.70
OQLF31075.106	1	16	C	E	2	0.83	0.27	-1.98	0.47
OQLF31076.106	1	17	C	L	3	0.46	0.29	0.77	0.64
OQLF31078.106	1	18	A	I	4*	0.57 (2.26)	0.41	-1.29	0.41
OQLF31080.106	1	19	A	L	4*	0.62 (2.46)	0.37	-1.51	0.33
OQLF20433.060	2	1	B	I	3	0.77	0.39	-0.99	0.81
OQLF20430.060	2	2	A	E	1	0.96	0.31	-2.56	1.03
OQLF20443.060	2	3	C	L	4	0.57	0.28	0.16	0.54
OQLF20441.060	2	4	C	L	3	0.84	0.32	-1.66	0.64
OQLF20432.060	2	5	A	I	4	0.84	0.31	-1.61	0.63
OQLF20435.060	2	6	B	L	4	0.50	0.35	0.46	0.81
OQLF20439.060	2	7	C	I	3	0.79	0.34	-1.19	0.66
OQLF20444.060	2	8	C	L	1	0.61	0.26	-0.11	0.49
OQLF20440.060	2	9	C	I	3	0.80	0.39	-1.11	0.84
OQLF20437.060	2	10	C	I	2	0.86	0.38	-1.56	0.90
OQLF20445.060	2	11	A	I	4*	0.53 (2.10)	0.41	-0.90	0.41
OQLF20448.060	2	12	A	L	4*	0.52 (2.08)	0.53	-0.53	0.67
OQLF20546.069	2	18	C	I	4	0.41	0.33	0.85	0.91
OQLF20541.069	2	19	A	I	1	0.35	0.28	1.20	0.94
OQLF20543.069	2	20	A	I	1	0.73	0.37	-0.76	0.71
OQLF20545.069	2	21	B	I	4	0.72	0.36	-0.74	0.69
OQLF20549.069	2	22	A	I	4*	0.57 (2.27)	0.36	-1.12	0.29
OQLF20550.069	2	23	B	L	4*	0.54 (2.14)	0.45	-0.20	0.48

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. A = comprehension; B = organization; C = vocabulary and linguistic conventions; L = connections; E = explicit; I = implicit; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.6 Distribution of Score Points and Step Parameters for Open-Response Items: Primary Reading (French)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
OQLF31051.103	1	5	% of Students	0.38	0.29	15.02	58.59	21.03	4.69
			Parameters			-7.04	-3.11	2.12	3.50
OQLF31052.103	1	6	% of Students	0.68	0.33	7.67	53.54	28.61	9.16
			Parameters			-4.24	-3.68	1.03	2.31
OQLF31069.105	1	11	% of Students	0.46	0.44	16.02	39.14	29.93	14.00
			Parameters			-4.86	-1.69	0.37	1.51
OQLF31071.105	1	12	% of Students	0.90	0.43	8.13	39.40	42.38	8.77
			Parameters			-3.64	-2.51	-0.09	2.45
OQLF31078.106	1	18	% of Students	0.27	0.18	9.40	61.18	21.26	7.70
			Parameters			-5.70	-3.19	1.62	2.12
OQLF31080.106	1	19	% of Students	0.50	0.14	8.86	44.71	35.87	9.92
			Parameters			-5.83	-3.32	0.38	2.72
OQLF20445.060	2	11	% of Students	0.26	0.43	20.58	51.54	22.49	4.70
			Parameters			-6.26	-1.71	1.39	2.98
OQLF20448.060	2	12	% of Students	0.38	0.58	16.18	61.1	17.81	3.96
			Parameters			-4.08	-1.71	1.38	2.30
OQLF20549.069	2	22	% of Students	0.78	1.00	16.36	45.96	25.16	10.74
			Parameters			-5.59	-2.51	1.33	2.27
OQLF20550.069	2	23	% of Students	1.16	1.14	12.98	55.64	26.66	2.42
			Parameters			-3.45	-2.29	1.13	3.80

Note. The total number of students is 6588.

Table 7.1.7 Item Statistics: Junior Reading (French)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQLF31248.116	1	1	B	L	4	0.70	0.35	-0.73	0.64
OQLF31245.116	1	2	A	I	3	0.95	0.28	-2.62	0.86
OQLF31249.116	1	3	C	I	2	0.78	0.42	-1.02	0.93
OQLF31250.116	1	4	C	I	1	0.71	0.28	-1.03	0.44
OQLF31251.116	1	5	A	I	4*	0.59 (2.37)	0.47	-1.66	0.44
OQLF31252.116(A)	1	6	A	L	4*	0.55 (2.20)	0.48	-0.78	0.49
OQLF31254.117	1	7	A	E	4	0.93	0.33	-2.22	0.90
OQLF31257.117	1	8	B	L	1	0.75	0.34	-1.04	0.60
OQLF31259.117	1	9	C	L	1	0.68	0.23	-0.76	0.41
OQLF31256.117	1	10	A	I	1	0.86	0.37	-1.70	0.77
OQLF31260.117	1	11	A	L	4*	0.65 (2.59)	0.47	-1.49	0.52
OQLF31262.117	1	12	A	L	4*	0.57 (2.28)	0.48	-0.87	0.43
OQLF31281.120	1	14	A	E	1	0.97	0.16	-3.67	0.60
OQLF31283.120	1	15	B	I	4	0.78	0.37	-1.14	0.73
OQLF31286.120	1	16	C	I	4	0.67	0.39	-0.48	0.72
OQLF31282.120	1	17	A	I	2	0.57	0.35	0.09	0.67
OQLF31309.120	1	18	A	I	4*	0.70 (2.80)	0.44	-2.20	0.42
OQLF31288.120	1	19	A	L	4*	0.65 (2.59)	0.47	-1.70	0.47
OQLF31210.113	2	1	C	I	4	0.33	0.25	1.60	0.68
OQLF31200.113	2	2	A	E	3	0.93	0.31	-2.44	0.81
OQLF31213.113	2	3	C	L	4	0.78	0.41	-1.09	0.79
OQLF31209.113	2	4	B	I	3	0.92	0.28	-2.63	0.63
OQLF31202.113	2	5	A	I	4	0.82	0.42	-1.23	0.90
OQLF31211.113	2	6	C	I	1	0.36	0.18	1.76	0.51
OQLF31204.113	2	7	A	I	3	0.74	0.33	-0.93	0.60
OQLF31205.113	2	8	A	I	4	0.79	0.32	-1.34	0.60
OQLF31201.113	2	9	A	E	2	0.89	0.36	-1.77	0.88
OQLF31206.113	2	10	A	I	2	0.73	0.25	-1.18	0.40
OQLF31215.113	2	11	A	I	4*	0.56 (2.23)	0.43	-0.79	0.42
OQLF31216.113	2	12	A	L	4*	0.65 (2.59)	0.47	-1.25	0.48
OQLF31300.122	2	18	A	L	2	0.96	0.19	-4.03	0.52
OQLF31299.122	2	19	A	I	3	0.66	0.36	-0.46	0.65
OQLF31302.122	2	20	B	I	4	0.61	0.28	-0.19	0.44
OQLF31303.122	2	21	C	L	3	0.37	0.28	1.17	0.81
OQLF31305.122	2	22	A	I	4*	0.55 (2.20)	0.34	-1.14	0.29
OQLF31307.122	2	23	B	L	4*	0.55 (2.19)	0.46	-0.66	0.42

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. A = comprehension; B = organization; C = vocabulary and linguistic conventions; L = connections; E = explicit; I = implicit; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.8 Distribution of Score Points and Step Parameters for Open-Response Items: Junior Reading (French)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
OQLF31251.116	1	5	% of Students	0.15	0.03	15.35	38.85	38.32	7.31
			Parameters			-7.59	-1.89	0.02	2.81
OQLF31252.116(A)	1	6	% of Students	0.88	0.36	16.31	51.57	22.91	7.97
			Parameters			-4.25	-1.94	1.10	1.99
OQLF31260.117	1	11	% of Students	0.21	0.05	5.68	37.35	48.43	8.28
			Parameters			-5.03	-3.05	-0.43	2.57
OQLF31262.117	1	12	% of Students	0.96	0.21	20.17	36.15	34.72	7.78
			Parameters			-4.87	-1.36	0.11	2.66
OQLF31309.120	1	18	% of Students	0.11	0.07	5.54	26.14	49.90	18.23
			Parameters			-6.38	-2.99	-1.17	1.74
OQLF31288.120	1	19	% of Students	0.26	0.02	6.52	42.86	34.41	15.93
			Parameters			-5.27	-3.12	0.15	1.45
OQLF31215.113	2	11	% of Students	0.93	0.52	11.53	57.11	22.24	7.66
			Parameters			-3.95	-2.77	1.42	2.15
OQLF31216.113	2	12	% of Students	0.36	0.49	5.97	36.35	46.90	9.93
			Parameters			-3.95	-3.00	-0.44	2.40
OQLF31305.122	2	22	% of Students	0.23	0.52	17.32	48.39	28.51	5.04
			Parameters			-7.24	-2.43	1.11	3.99
OQLF31307.122	2	23	% of Students	0.98	1.06	19.90	43.79	26.06	8.22
			Parameters			-4.20	-1.56	0.85	2.29

Note. The total number of students is 6145.

Table 7.1.9 Item Statistics: Primary Writing (English)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQWE20246_T	1	13	W2.0	T	4*	0.59 (2.36)	0.55	-0.49	0.69
EQWE20246_V	1	13	W3.0	V	3*	0.75 (2.25)	0.58	-1.41	0.92
EQWE30415	2	13	W2.0	T	4	0.49	0.25	0.74	0.51
EQWE30407	2	14	W1.0	T	3	0.89	0.32	-1.99	0.76
EQWE30408	2	15	W1.0	T	2	0.63	0.38	-0.32	0.71
EQWE30414	2	16	W3.0	V	3	0.81	0.26	-2.09	0.42
EQWE30427_T	2	17	W2.0	T	4*	0.59 (2.36)	0.56	-0.49	0.80
EQWE30427_V	2	17	W3.0	V	3*	0.76 (2.29)	0.56	-1.53	0.99
11042_T	2	24	W2.0	T	4*	0.56 (2.22)	0.55	-0.58	0.72
11042_V	2	24	W3.0	V	3*	0.68 (2.04)	0.59	-1.24	1.07
EQWE30406	2	25	W2.0	T	2	0.77	0.38	-1.12	0.70
EQWE30409	2	26	W3.0	V	4	0.84	0.30	-1.88	0.55
EQWE30419	2	27	W3.0	V	1	0.62	0.36	-0.27	0.73
EQWE30412	2	28	W2.0	T	2	0.75	0.35	-1.08	0.61

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. T = content; V = conventions;

W = writing; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.10 Distribution of Score Points and Step Parameters for Open-Response Items: Primary Writing (English)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
EQWE20246_T	1	13	% of Students	0.93	2.13	9.19	44.48	35.71	7.56
			Parameters			-2.28	-2.05	0.24	2.13
EQWE20246_V	1	13	% of Students	0.93	0.43	15.65	39.82	43.18	
			Parameters			-3.06	-1.30	0.13	
11042_T	2	24	% of Students	0.48	0.76	16.57	47.09	29.03	6.08
			Parameters			-3.55	-1.53	0.55	2.20
11042_V	2	24	% of Students	0.48	0.29	20.29	53.12	25.82	
			Parameters			-3.43	-1.12	0.83	
EQWE30427_T	2	17	% of Students	0.51	2.11	5.78	50.73	34.69	6.18
			Parameters			-2.08	-2.39	0.34	2.17
EQWE30427_V	2	17	% of Students	0.51	0.39	9.91	48.45	40.76	
			Parameters			-3.13	-1.74	0.28	

Note. The total number of students is 119 966.

Table 7.1.11 Item Statistics: Junior Writing (English)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQWE30498_T	1	13	W2.0	T	4*	0.65 (2.60)	0.59	-1.04	0.78
EQWE30498_V	1	13	W3.0	V	3*	0.71 (2.12)	0.59	-1.41	0.97
EQWE30486	2	13	W3.0	V	1	0.62	0.35	-0.25	0.65
EQWE30471	2	14	W1.0	T	4	0.79	0.28	-1.81	0.43
EQWE30474	2	15	W1.0	T	3	0.82	0.26	-2.32	0.37
EQWE30490	2	16	W2.0	V	4	0.50	0.29	0.52	0.52
EQWE30504_T	2	17	W2.0	T	4*	0.60 (2.41)	0.61	-0.84	0.84
EQWE30504_V	2	17	W3.0	V	3*	0.68 (2.05)	0.60	-1.31	1.05
10715_T	2	24	W2.0	T	4*	0.54 (2.17)	0.58	-0.56	0.78
10715_V	2	24	W3.0	V	3*	0.71 (2.12)	0.59	-1.46	0.93
EQWE30472	2	25	W1.0	T	4	0.84	0.29	-2.11	0.48
EQWE30484	2	26	W2.0	V	3	0.75	0.40	-1.09	0.69
EQWE30488	2	27	W3.0	V	4	0.95	0.28	-2.87	0.74
EQWE30478	2	28	W2.0	T	1	0.48	0.23	0.87	0.41

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. T = content; V = conventions; W = writing; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.12 Distribution of Score Points and Step Parameters for Open-Response Items: Junior Writing (English)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
EQWE30498_T	1	13	% of Students	0.76	0.53	7.42	34.23	43.73	13.33
			Parameters			-2.97	-2.29	-0.41	1.50
EQWE30498_V	1	13	% of Students	0.76	0.26	14.63	55.27	29.08	
			Parameters			-3.29	-1.60	0.66	
10715_T	2	24	% of Students	0.70	0.55	17.75	49.47	25.98	5.55
			Parameters			-3.51	-1.51	0.66	2.12
10715_V	2	24	% of Students	0.70	0.27	17.10	50.70	31.22	
			Parameters			-3.49	-1.45	0.56	
EQWE30504_T	2	17	% of Students	0.55	0.49	10.58	43.29	36.83	8.26
			Parameters			-3.26	-2.01	0.06	1.84
EQWE30504_V	2	17	% of Students	0.55	0.37	17.19	57.82	24.07	
			Parameters			-3.37	-1.42	0.86	

Note. The total number of students is 130 465.

Table 7.1.13 Item Statistics: Primary Writing (French)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQEF30521_T	1	13	A	T	4*	0.48 (1.93)	0.53	-0.10	0.63
OQEF30521_V	1	13	C	V	3*	0.64 (1.91)	0.59	-1.15	1.09
OQEF30502	2	13	A	T	2	0.61	0.30	-0.11	0.53
OQEF30511	2	14	C	V	3	0.70	0.27	-0.88	0.44
OQEF30509	2	15	B	T	3	0.81	0.32	-1.44	0.64
OQEF30514	2	16	C	V	2	0.57	0.24	0.17	0.41
OQEF30523_T	2	17	A	T	4*	0.54 (2.14)	0.53	-0.27	0.70
OQEF30523_V	2	17	C	V	3*	0.69 (2.08)	0.57	-1.44	1.16
11248_T	2	24	A	T	4*	0.48 (1.90)	0.57	0.05	0.78
11248_V	2	24	C	V	3*	0.63 (1.90)	0.58	-1.05	1.10
OQEF30503	2	25	A	T	2	0.61	0.19	-0.17	0.28
12291	2	26	C	V	3	0.73	0.26	-1.28	0.38
OQEF30519	2	27	C	V	4	0.64	0.24	-0.42	0.38
OQEF30507	2	28	B	T	3	0.87	0.34	-1.88	0.69

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. A = comprehension; B = organization; C = vocabulary and linguistic conventions; T = content; V = conventions; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.14 Distribution of Score Points and Step Parameters for Open-Response Items: Primary Writing (French)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
OQEF30521_T	1	13	% of Students	0.33	2.40	31.43	39.83	21.82	4.18
			Parameters			-3.31	-0.56	0.91	2.56
OQEF30521_V	1	13	% of Students	0.33	0.14	29.44	49.09	21.00	
			Parameters			-3.79	-0.68	1.02	
11248_T	2	24	% of Students	0.64	2.71	28.28	46.09	19.88	2.40
			Parameters			-2.90	-0.75	1.06	2.78
11248_V	2	24	% of Students	0.64	0.15	28.04	51.32	19.85	
			Parameters			-3.51	-0.74	1.10	
OQEF30523_T	2	17	% of Students	0.21	1.38	17.61	48.48	29.62	2.69
			Parameters			-3.35	-1.41	0.65	3.03
OQEF30523_V	2	17	% of Students	0.21	0.15	12.19	66.78	20.67	
			Parameters			-3.78	-1.60	1.06	

Note. The total number of students is 6595.

Table 7.1.15 Item Statistics: Junior Writing (French)

Item Code	Booklet	Sequence	Expectation	Cognitive Skill	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
12462_T	1	13	A	T	4*	0.68 (2.70)	0.58	-1.11	0.77
12462_V	1	13	C	V	3*	0.76 (2.29)	0.60	-1.68	1.08
OQEF30597	2	13	C	V	4	0.79	0.37	-1.21	0.77
OQEF30585	2	14	A	T	3	0.86	0.23	-2.61	0.40
OQEF30593	2	15	C	V	4	0.65	0.29	-0.45	0.50
OQEF30586	2	16	B	T	3	0.72	0.28	-1.02	0.45
OQEF30601_T	2	17	A	T	4*	0.63 (2.50)	0.56	-0.91	0.75
OQEF30601_V	2	17	C	V	3*	0.76 (2.28)	0.60	-1.66	1.08
12542_T	2	24	A	T	4*	0.65 (2.58)	0.59	-1.10	0.89
12542_V	2	24	C	V	3*	0.78 (2.33)	0.60	-1.79	1.10
OQEF30587	2	25	B	T	4	0.80	0.32	-1.57	0.56
OQEF30599	2	26	C	V	3	0.66	0.23	-0.70	0.34
OQEF30598	2	27	C	V	2	0.77	0.25	-1.52	0.41
OQEF30581	2	28	A	T	4	0.74	0.32	-1.11	0.53

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. A = comprehension; B = organization; C = vocabulary and linguistic conventions; T = content; V = conventions; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.16 Distribution of Score Points and Step Parameters for Open-Response Items: Junior Writing (French)

Item Code	Booklet	Sequence		Score Points					
				Missing	Illegible	10	20	30	40
12462_T	1	13	% of Students	0.45	0.65	7.18	29.11	45.96	16.65
			Parameters			-3.05	-2.12	-0.60	1.33
12462_V	1	13	% of Students	0.45	0.05	9.92	49.93	39.65	
			Parameters			-3.48	-1.82	0.26	
12542_T	2	24	% of Students	0.29	0.16	6.84	38.43	42.36	11.92
			Parameters			-3.54	-2.19	-0.23	1.56
12542_V	2	24	% of Students	0.29	0.02	9.76	46.94	42.99	
			Parameters			-3.68	-1.83	0.14	
OQEF30601_T	2	17	% of Students	0.42	0.62	7.75	42.24	37.79	11.17
			Parameters			-3.10	-2.25	0.00	1.70
OQEF30601_V	2	17	% of Students	0.42	0.02	8.92	53.06	37.58	
			Parameters			-3.41	-1.91	0.34	

Note. The total number of students is 6169.

Table 7.1.17 Item Statistics: Primary Mathematics (English)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQME10026		DV1	AP	D	4*				
EQME31574	1	NV3	KU	N	3	88.03	0.37	-2.03	0.72
EQME20908	2	PV1	KU	P	4	78.80	0.44	-1.19	0.80
EQME31512	3	NV2	AP	N	3	47.74	0.46	0.48	1.17
EQME31542	4	NV3	AP	N	1	55.86	0.47	0.13	0.96
EQME20910	5	PV2	AP	P	2	30.04	0.39	1.23	1.46
EQME31514	6	NV3	PS	N	2	56.88	0.42	0.07	0.75
EQME20812	7	GV2	PS	G	4*	47.00(1.88)	0.49	-0.82	0.37
EQME31587	8	MV2	PS	M	4*	61.50(2.46)	0.69	-1.34	0.53
EQME31516	9	NV2	PS	N	4*	67.25(2.69)	0.65	-1.59	0.40
EQME31526	10	MV2	AP	M	4	36.56	0.41	0.96	1.24
EQME31543	11	PV2	KU	P	4	62.53	0.46	-0.18	0.86
EQME31601	12	NV3	AP	N	2	65.48	0.48	-0.32	0.91
EQME31518	13	DV3	AP	D	4	35.41	0.49	0.86	1.41
EQME31541	14	DV2	AP	D	3	74.10	0.49	-0.80	0.87
EQME31591	15	MV1	KU	M	1	46.95	0.43	0.56	1.05
EQME31525	16	MV2	KU	M	1	83.46	0.43	-1.42	0.85
EQME31567	17	MV1	KU	M	4	64.14	0.47	-0.34	0.80
EQME31559	18	PV1	KU	P	2	42.15	0.37	0.88	0.80
EQME20852	19	PV2	AP	P	1	65.94	0.43	-0.32	0.78
EQME31540	20	NV1	PS	N	4	57.21	0.38	0.16	0.78
EQME31527	21	GV2	KU	G	4	93.27	0.27	-2.81	0.62
EQME20919	22	GV1	AP	G	1	65.42	0.40	-0.43	0.65
EQME10044	23	MV2	KU	M	4	90.59	0.35	-2.22	0.75
EQME10060	24	MV1	AP	M	2	66.48	0.50	-0.41	0.91
EQME10075	25	GV3	AP	G	4*	70.50(2.82)	0.56	-1.83	0.30
EQME31536	26	NV3	AP	N	4*	73.00(2.92)	0.65	-1.74	0.46
EQME31560	27	PV1	AP	P	4*	62.25(2.49)	0.60	-1.31	0.43
EQME31588	28	DV2	PS	D	4*	53.50(2.14)	0.55	-1.53	0.27
EQME31614	29	DV2	KU	D	1	83.39	0.45	-1.41	0.88
EQME31570	30	MV1	AP	M	4	24.16	0.31	1.60	1.45
EQME20820	31	NV1	KU	N	4	90.62	0.41	-1.85	1.04
EQME31505	32	PV1	PS	P	2	36.34	0.45	0.90	1.26
EQME31510	33	GV3	AP	G	3	45.15	0.43	0.66	0.95
EQME31586	34	MV2	PS	M	2	53.97	0.43	0.20	0.79
EQME31594	35	GV2	KU	G	4	59.97	0.37	-0.01	0.63

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. KU = knowledge and understanding; AP = application; PS = thinking; N = number sense and numeration; M = measurement; G = geometry and spatial sense; P = patterning and algebra; D = data management and probability; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

**Table 7.1.18 Distribution of Score Points and Step Parameters for Open-Response Items:
Primary Mathematics (English)**

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
EQME20812	7	% of Students	0.89	0.36	34.20	47.59	8.89	8.07
		Parameters			-6.37	-0.76	2.96	0.88
EQME31587	8	% of Students	0.64	0.13	32.25	13.47	27.42	26.08
		Parameters			-5.37	0.30	-0.83	0.56
EQME10026		% of Students						
		Parameters						
EQME31516	9	% of Students	1.53	0.29	26.05	14.06	17.05	41.02
		Parameters			-5.17	0.22	-0.43	-0.97
EQME10075	25	% of Students	1.06	1.60	16.18	15.67	27.24	38.24
		Parameters			-4.94	-0.66	-1.34	-0.39
EQME31536	26	% of Students	1.01	0.38	12.27	23.07	19.29	43.97
		Parameters			-4.31	-1.72	-0.12	-0.79
EQME31560	27	% of Students	0.98	0.14	20.48	22.58	39.89	15.92
		Parameters			-5.38	-0.80	-0.84	1.79
EQME31588	28	% of Students	1.71	0.45	45.00	13.59	14.98	24.26
		Parameters			-7.48	2.29	-0.22	-0.71

Note. The total number of students is 123 568.

Table 7.1.19 Item Statistics: Junior Mathematics (English)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQME21023	1	NV1	KU	N	3	75.07	0.40	-1.08	0.64
EQME31699	2	PV1	AP	P	3	58.64	0.37	0.00	0.59
EQME10257	3	MV2	AP	M	2	63.70	0.37	-0.28	0.63
EQME31720	4	MV2	AP	M	1	60.93	0.48	-0.14	0.95
EQME31648	5	NV2	AP	N	3	60.87	0.42	-0.15	0.72
EQME31655	6	GV3	PS	G	1	47.39	0.45	0.49	0.88
EQME31744	7	GV2	AP	G	4*	61.75(2.47)	0.50	-2.84	0.22
EQME21011	8	NV1	AP	N	4*	61.25(2.45)	0.63	-0.98	0.46
EQME21009	9	DV2	PS	D	4*	58.25(2.33)	0.69	-0.79	0.53
EQME31702	10	MV2	AP	M	4*	55.50(2.22)	0.71	-0.80	0.60
EQME31729	11	DV2	AP	D	4	69.33	0.38	-0.70	0.58
EQME31636	12	DV3	AP	D	2	62.55	0.41	-0.22	0.72
EQME31675	13	DV2	AP	D	4	37.66	0.31	1.26	0.70
EQME31714	14	PV2	KU	P	4	72.17	0.44	-0.70	0.84
EQME31646	15	PV1	PS	P	3	56.53	0.40	0.09	0.68
EQME21126	16	GV3	AP	G	2	56.08	0.35	0.16	0.59
3469832	17	GV1	PS	G	3	56.88	0.38	0.07	0.63
EQME31635	18	GV1	KU	G	1	87.71	0.32	-1.95	0.67
EQME31665	19	MV2	KU	M	1	93.18	0.21	-3.66	0.45
EQME31678	20	PV1	AP	P	2	61.13	0.36	-0.21	0.53
EQME31711	21	MV2	KU	M	4	15.04	0.34	1.74	1.30
EQME21050	22	MV2	PS	M	2	66.58	0.42	-0.41	0.79
EQME21004	23	NV2	KU	N	3	59.54	0.36	-0.10	0.55
EQME31748	24	MV2	PS	M	3	47.27	0.35	0.65	0.72
EQME31700	25	PV1	AP	P	2	51.56	0.44	0.31	0.82
EQME21015	26	PV1	AP	P	4*	74.50(2.98)	0.61	-2.09	0.36
EQME21056	27	DV3	PS	D	4*	58.00(2.32)	0.68	-1.20	0.44
EQME31662	28	NV3	PS	N	4*	60.75(2.43)	0.67	-1.29	0.45
EQME10275	29	GV3	PS	G	4*	44.50(1.78)	0.67	-0.13	0.64
EQME21068	30	PV1	KU	P	3	48.99	0.50	0.36	1.09
EQME21052	31	NV3	AP	N	4	40.62	0.53	0.64	1.24
EQME31668	32	NV1	PS	N	2	56.61	0.31	0.14	0.47
EQME31645	33	MV2	AP	M	1	59.63	0.45	-0.10	0.82
EQME31669	34	NV3	PS	N	1	38.67	0.40	0.86	1.33
EQME31685	35	DV3	PS	D	4	60.41	0.47	-0.14	0.88
EQME31690	36	DV2	KU	D	3	79.62	0.37	-1.38	0.63

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. KU = knowledge and understanding; AP = application; PS = thinking; N = number sense and numeration; M = measurement; G = geometry and spatial sense; P = patterning and algebra; D = data management and probability; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

**Table 7.1.20 Distribution of Score Points and Step Parameters for Open-Response Items:
Junior Mathematics (English)**

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
EQME31744	7	% of Students	0.49	0.23	34.55	17.02	12.54	35.16
		Parameters			-11.08	1.40	0.75	-2.44
EQME21011	8	% of Students	2.92	0.59	18.83	31.28	21.78	24.59
		Parameters			-3.29	-1.38	0.41	0.33
EQME21009	9	% of Students	3.40	0.73	25.54	29.25	15.54	25.54
		Parameters			-3.16	-0.71	0.70	-0.01
EQME31702	10	% of Students	2.09	0.45	34.33	25.89	12.55	24.67
		Parameters			-3.72	-0.24	0.81	-0.04
EQME21015	26	% of Students	1.14	0.24	22.74	9.89	8.97	57.02
		Parameters			-5.86	0.54	-0.13	-2.90
EQME21056	27	% of Students	1.79	0.39	39.40	18.31	4.87	35.25
		Parameters			-4.95	0.55	1.77	-2.16
EQME31662	28	% of Students	1.50	0.26	32.58	21.01	10.18	34.46
		Parameters			-4.89	-0.07	0.92	-1.12
EQME10275	29	% of Students	3.26	0.50	46.02	26.29	16.54	7.38
		Parameters			-3.34	0.24	0.84	1.72

Note. The total number of students in Primary Reading (English) is 129 882.

Table 7.1.21 Item Statistics: Primary Mathematics (French)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQMF10054	1	NA2	MA	N	4	92.03	0.34	-2.21	0.80
OQMF31572	2	AA2	CC	A	2	89.93	0.36	-2.10	0.76
OQMF20609	3	AA2	HP	A	3	50.41	0.44	0.38	0.96
OQMF31523	4	NA4	CC	N	4	82.96	0.42	-1.31	0.84
OQMF20615	5	AA3	CC	A	2	83.36	0.40	-1.43	0.76
OQMF31503	6	NA1	MA	N	3	72.91	0.45	-0.72	0.81
OQMF20505	7	MA2	HP	M	4*	58.25(2.33)	0.63	-1.53	0.42
OQMF20508	8	GA1	MA	G	4*	80.25(3.21)	0.48	-2.70	0.33
OQMF31597	9	TA1	MA	T	4*	54.75(2.19)	0.60	-1.22	0.46
OQMF10063	10	NA4	HP	N	4*	71.00(2.84)	0.66	-1.75	0.49
OQMF20584	11	GA1	CC	G	3	59.77	0.41	-0.06	0.70
OQMF20577	12	MA3	CC	M	2	80.81	0.42	-1.25	0.76
OQMF20624	13	TA1	MA	T	3	57.41	0.47	0.06	0.91
OQMF20592	14	GA1	HP	G	3	35.62	0.34	1.12	1.10
OQMF10027	15	MA1	HP	M	4	30.97	0.36	1.36	1.05
OQMF31516	16	NA3	HP	N	3	55.34	0.35	0.22	0.53
OQMF31579	17	AA3	MA	A	3	46.10	0.38	0.66	0.90
OQMF31590	18	TA1	HP	T	2	60.09	0.45	-0.05	0.85
OQMF31571	19	AA1	CC	A	3	86.92	0.33	-2.03	0.59
OQMF20575	20	MA2	MA	M	4	69.33	0.41	-0.55	0.69
OQMF31556	21	GA1	MA	G	4	45.40	0.35	0.81	0.64
OQMF31524	22	NA4	MA	N	2	65.37	0.50	-0.29	1.02
OQMF31500	23	NA1	CC	N	3	75.33	0.44	-0.92	0.75
OQMF20621	24	TA1	CC	T	4	82.11	0.42	-1.29	0.81
OQMF31533	25	MA1	MA	M	2	45.96	0.24	1.22	0.39
OQMF10057	26	NA2	MA	N	4*	73.50(2.94)	0.63	-1.88	0.43
OQMF20517	27	GA1	HP	G	4*	64.50(2.58)	0.57	-1.61	0.42
OQMF31582	28	AA1	HP	A	4*	55.75(2.23)	0.57	-1.11	0.36
OQMF20535	29	TA1	HP	T	4*	62.00(2.48)	0.59	-1.55	0.39
OQMF20596	30	GA2	MA	G	2	82.15	0.39	-1.42	0.68
OQMF20581	31	MA4	CC	M	3	65.94	0.44	-0.33	0.81
OQMF31513	32	NA3	CC	N	2	84.29	0.46	-1.33	0.98
OQMF31511	33	NA3	MA	N	2	61.68	0.48	-0.12	0.95
OQMF31510	34	NA2	HP	N	1	54.99	0.27	0.35	0.40
OQMF31592	35	TA2	MA	T	4	41.74	0.43	0.77	1.10
OQMF31542	36	MA2	CC	M	1	78.24	0.48	-0.95	0.98

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. CC = knowledge and understanding; MA = application; HP = thinking; N = number sense and numeration; M = measurement; G = geometry and spatial sense; A = patterning and algebra; T = data management and probability; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

**Table 7.1.22 Distribution of Score Points and Step Parameters for Open-Response Items:
Primary Mathematics (French)**

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
OQMF20505	7	% of Students	0.59	0.18	34.98	21.10	16.45	26.69
		Parameters			-6.59	0.23	0.40	-0.17
OQMF20508	8	% of Students	0.35	0.06	4.69	15.67	32.44	46.79
		Parameters			-5.48	-3.05	-1.72	-0.56
OQMF31597	9	% of Students	0.71	0.05	30.99	36.28	12.07	19.90
		Parameters			-5.82	-0.60	1.57	-0.03
OQMF10063	10	% of Students	0.62	0.09	17.94	20.40	18.47	42.48
		Parameters			-5.27	-0.93	-0.10	-0.70
OQMF10057	26	% of Students	0.92	0.08	18.06	12.75	22.75	45.45
		Parameters			-5.43	-0.28	-1.09	-0.71
OQMF20517	27	% of Students	0.74	0.03	13.19	35.24	29.19	21.61
		Parameters			-5.41	-2.07	0.18	0.87
OQMF31582	28	% of Students	1.77	0.15	28.37	34.36	15.87	19.48
		Parameters			-5.27	-0.75	1.35	0.22
OQMF20535	29	% of Students	1.06	0.05	20.92	29.84	25.18	22.96
		Parameters			-5.92	-1.12	0.24	0.58

Note. The total number of students is 6612.

Table 7.1.23 Item Statistics: Junior Mathematics (French)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQMF31627	1	NA2	CC	N	4	93.55	0.21	-3.41	0.49
OQMF20806	2	AA2	MA	A	1	70.91	0.34	-0.80	0.52
OQMF31670	3	GA1	HP	G	4	52.85	0.19	0.90	0.23
OQMF31658	4	MA2	HP	M	3	49.07	0.34	0.60	0.67
OQMF31681	5	GA2	CC	G	1	81.90	0.37	-1.38	0.70
OQMF31624	6	NA1	MA	N	3	56.44	0.51	0.08	1.10
OQMF10112	7	AA1	HP	A	4*	78.50(3.14)	0.53	-2.43	0.39
OQMF20720	8	NA3	MA	N	4*	62.75(2.51)	0.63	-1.34	0.40
OQMF20782	9	GA2	HP	G	4*	72.50(2.90)	0.57	-1.72	0.46
OQMF10206	10	TA1	HP	T	4*	50.50(2.02)	0.52	-0.51	0.34
OQMF31621	11	NA1	CC	N	1	76.50	0.37	-1.14	0.60
OQMF20759	12	GA1	MA	G	4	81.86	0.41	-1.30	0.79
OQMF20749	13	MA3	CC	M	3	50.87	0.39	0.44	0.77
OQMF31635	14	NA3	HP	N	2	55.71	0.47	0.14	1.04
OQMF31664	15	MA3	MA	M	1	42.99	0.38	0.87	0.70
OQMF31714	16	TA2	MA	T	4	72.11	0.46	-0.65	0.87
OQMF31702	17	AA2	HP	A	1	60.90	0.44	-0.10	0.84
OQMF20736	18	MA1	MA	M	2	52.95	0.37	0.36	0.69
OQMF10143	19	GA2	CC	G	3	81.46	0.40	-1.31	0.75
OQMF20732	20	MA1	HP	M	2	63.61	0.34	-0.31	0.52
OQMF31717	21	TA2	HP	T	3	52.66	0.46	0.27	1.04
OQMF31695	22	AA1	MA	A	3	58.83	0.42	-0.03	0.68
OQMF20728	23	MA1	CC	M	4	62.29	0.54	-0.20	1.07
OQMF31633	24	NA3	MA	N	1	68.50	0.48	-0.49	0.91
OQMF31712	25	TA1	HP	T	3	54.23	0.33	0.32	0.54
OQMF20756	26	MA3	MA	M	4*	67.50(2.70)	0.65	-1.56	0.49
OQMF20768	27	GA1	MA	G	4*	65.25(2.61)	0.65	-1.27	0.50
OQMF31641	28	NA1	HP	N	4*	45.75(1.83)	0.64	-0.18	0.62
OQMF20812	29	AA2	MA	A	4*	64.75(2.59)	0.62	-1.15	0.54
OQMF31678	30	GA1	CC	G	1	74.89	0.41	-0.90	0.73
OQMF31696	31	AA1	CC	A	2	60.93	0.43	-0.08	0.87
OQMF31628	32	NA2	MA	N	1	71.01	0.36	-0.76	0.58
OQMF10208	33	TA1	MA	T	2	46.03	0.45	0.55	0.95
OQMF31687	34	GA2	HP	G	2	66.70	0.42	-0.42	0.73
OQMF31655	35	MA2	CC	M	4	52.92	0.51	0.23	1.08
OQMF10215	36	TA2	CC	T	2	78.66	0.43	-1.13	0.77

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. CC = knowledge and understanding; MA = application; HP = thinking; N = number sense and numeration; M = measurement; G = geometry and spatial sense; A = patterning and algebra; T = data management and probability; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

**Table 7.1.24 Distribution of Score Points and Step Parameters for Open-Response Items:
Junior Mathematics (French)**

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
OQMF10112	7	% of Students	0.29	0.11	6.48	15.27	34.10	43.74
		Parameters			-5.50	-2.31	-1.66	-0.28
OQMF20720	8	% of Students	1.32	0.41	27.70	18.26	22.23	30.09
		Parameters			-5.06	0.02	-0.30	-0.03
OQMF20782	9	% of Students	0.70	0.16	6.87	23.38	39.64	29.24
		Parameters			-4.09	-2.50	-0.98	0.68
OQMF10206	10	% of Students	2.88	0.41	27.90	42.19	17.19	9.44
		Parameters			-4.39	-1.05	1.71	1.67
OQMF20756	26	% of Students	0.75	0.13	22.88	15.30	26.77	34.17
		Parameters			-5.27	-0.20	-0.81	0.05
OQMF20768	27	% of Students	1.17	0.37	18.91	23.93	27.81	27.80
		Parameters			-4.23	-0.98	-0.32	0.43
OQMF31641	28	% of Students	2.96	0.70	37.25	41.41	8.27	9.41
		Parameters			-3.14	-0.38	1.93	0.88
OQMF20812	29	% of Students	1.10	0.13	13.44	27.86	39.79	17.68
		Parameters			-3.99	-1.51	-0.47	1.38

Note. The total number of students is 6160.

Differential Item Functioning (DIF) Analysis Results

The gender-based and SLL-based DIF results are provided for the primary- and junior-division assessments in Tables 7.1.25–7.1.48. Results are presented for two random samples of 2000 examinees. Each table indicates the value of $\hat{\beta}$ and the significance level for each item. The DIF is also provided for those items that had a significant level of DIF, with at least a B- or C-level effect size. The results for items with B-or C-level DIF in both samples are presented in bold type. For gender-based DIF, negative estimates of $\hat{\beta}$ indicate that the girls outperformed the boys; positive $\hat{\beta}$ estimates indicate that the boys outperformed the girls. For SLL-based DIF, negative estimates of $\hat{\beta}$ indicate that the SLLs outperformed the non-SLLs; positive $\hat{\beta}$ estimates indicate that the non-SLLs outperformed the SLLs.

Table 7.1.25 Gender-Based DIF Statistics: Primary Reading (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQRE31619.141	1	1	0.009	0.557		0.017	0.239	
EQRE31617.141	1	2	-0.018	0.080		-0.011	0.267	
EQRE31621.141	1	3	0.005	0.687		-0.010	0.450	
EQRE31622.141	1	4	-0.002	0.850		-0.024	0.067	
EQRE31615.141	1	5	-0.067	0.006	B-	-0.007	0.788	
EQRE31614.141	1	6	-0.017	0.469		0.036	0.122	
EQRE31666.146	1	7	-0.056	0.000		-0.059	0.000	B-
EQRE31667.146	1	8	0.007	0.623		0.028	0.040	
EQRE31662.146	1	9	0.033	0.021		0.030	0.032	
EQRE31664.146	1	10	0.027	0.057		0.024	0.091	
EQRE31660.146	1	11	-0.021	0.362		0.013	0.572	
EQRE31661.146	1	12	0.001	0.964		-0.029	0.259	
EQRE31702.150	1	14	0.005	0.674		0.000	0.982	
EQRE31698.150	1	15	0.024	0.045		0.041	0.001	
EQRE31700.150	1	16	0.012	0.445		-0.004	0.805	
EQRE31703.150	1	17	0.028	0.034		0.040	0.003	
EQRE31697.150	1	18	-0.025	0.310		0.027	0.265	
EQRE31696.150	1	19	0.140	0.000	C+	0.134	0.000	C+
EQRE31819.159	2	1	0.041	0.005		0.029	0.042	
EQRE31816.159	2	2	-0.008	0.491		-0.012	0.267	
EQRE31824.159	2	3	-0.020	0.038		-0.016	0.100	
EQRE31817.159	2	4	0.012	0.384		0.002	0.910	
EQRE31822.159	2	5	-0.012	0.187		-0.012	0.211	
EQRE31814.159	2	6	-0.019	0.132		-0.007	0.559	
EQRE31811.159	2	7	-0.002	0.908		-0.030	0.032	
EQRE31812.159	2	8	-0.015	0.347		-0.001	0.931	
EQRE31820.159	2	9	-0.065	0.000	B-	-0.060	0.000	B-
EQRE31818.159	2	10	-0.017	0.170		-0.036	0.004	
EQRE31809.159	2	11	-0.015	0.545		-0.072	0.005	B-
EQRE31810.159	2	12	-0.114	0.000	C-	-0.087	0.000	B-
EQRE31736.154	2	18	0.000	0.976		0.007	0.467	
EQRE31737.154	2	19	-0.014	0.124		0.009	0.304	
EQRE31738.154	2	20	0.032	0.000		0.035	0.000	
EQRE31740.154	2	21	-0.002	0.872		-0.006	0.691	
EQRE31733.154	2	22	0.061	0.011	B+	0.038	0.123	
EQRE31735.154	2	23	0.021	0.357		-0.035	0.132	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.26 Gender-Based DIF Statistics: Junior Reading (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQRE31868.162	1	1	-0.027	0.009		-0.014	0.155	
EQRE31873.162	1	2	-0.033	0.000		-0.031	0.000	
EQRE31870.162	1	3	-0.003	0.826		-0.003	0.775	
EQRE31869.162	1	4	0.053	0.000		0.043	0.000	
EQRE31874.162	1	5	0.034	0.176		-0.018	0.446	
EQRE31875.162	1	6	0.065	0.008	B+	0.006	0.800	
EQRE31913.167	1	7	0.028	0.030		0.066	0.000	B+
EQRE31915.167	1	8	-0.040	0.006		-0.024	0.100	
EQRE31918.167	1	9	0.006	0.667		0.006	0.649	
EQRE31914.167	1	10	-0.020	0.115		-0.041	0.001	
EQRE31919.167	1	11	0.000	0.984		0.001	0.945	
EQRE31920.167	1	12	-0.010	0.681		0.053	0.026	
EQRE31950.171	1	14	-0.041	0.001		-0.038	0.002	
EQRE31952.171	1	15	-0.038	0.001		-0.004	0.742	
EQRE31949.171	1	16	-0.013	0.361		0.003	0.814	
EQRE31953.171	1	17	0.044	0.004		0.024	0.115	
EQRE31957.171	1	18	0.006	0.778		-0.013	0.559	
EQRE31955.171	1	19	-0.050	0.017		-0.045	0.030	
EQRE32023.177	2	1	-0.006	0.680		-0.049	0.001	
EQRE32018.177	2	2	-0.066	0.000	B-	-0.050	0.000	
EQRE32014.177	2	3	0.000	0.944		-0.018	0.000	
EQRE32024.177	2	4	-0.008	0.390		-0.013	0.184	
EQRE32021.177	2	5	0.033	0.028		0.011	0.458	
EQRE32015.177	2	6	-0.008	0.601		-0.005	0.738	
EQRE32027.177	2	7	-0.044	0.004		-0.061	0.000	B-
EQRE32026.177	2	8	0.049	0.000		0.021	0.048	
EQRE32020.177	2	9	0.021	0.113		0.044	0.001	
EQRE32019.177	2	10	-0.013	0.261		-0.014	0.231	
EQRE32031.177	2	11	-0.010	0.694		0.032	0.217	
EQRE32029.177	2	12	0.014	0.562		-0.005	0.841	
EQRE31967.173	2	18	0.003	0.709		0.026	0.005	
EQRE31969.173	2	19	0.046	0.000		0.009	0.384	
EQRE31970.173	2	20	-0.004	0.751		-0.026	0.053	
EQRE31972.173	2	21	0.090	0.000	C+	0.106	0.000	C+
EQRE31974.173	2	22	-0.071	0.000	B-	-0.061	0.002	B-
EQRE31973.173	2	23	0.041	0.060		0.086	0.000	B+

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.27 Gender-Based DIF Statistics: Primary Reading (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQLF31050.103	1	1	-0.032	0.035		-0.039	0.010	
OQLF31048.103	1	2	-0.011	0.478		-0.004	0.814	
OQLF31046.103	1	3	0.004	0.735		0.005	0.690	
OQLF31047.103	1	4	0.008	0.407		-0.002	0.826	
OQLF31051.103	1	5	0.020	0.374		0.015	0.500	
OQLF31052.103	1	6	-0.066	0.004	B-	-0.043	0.063	
OQLF31065.105	1	7	0.021	0.055		0.007	0.528	
OQLF31066.105	1	8	-0.003	0.776		0.007	0.513	
OQLF31068.105	1	9	0.017	0.189		0.016	0.211	
OQLF31064.105	1	10	0.043	0.002		0.049	0.001	
OQLF31069.105	1	11	0.085	0.001	B+	0.078	0.003	B+
OQLF31071.105	1	12	0.060	0.008	B+	0.048	0.033	
OQLF31074.106	1	14	-0.016	0.173		0.000	0.996	
OQLF31073.106	1	15	-0.003	0.804		0.015	0.272	
OQLF31075.106	1	16	0.009	0.447		0.010	0.406	
OQLF31076.106	1	17	0.015	0.317		-0.005	0.771	
OQLF31078.106	1	18	0.000	0.982		-0.002	0.921	
OQLF31080.106	1	19	0.015	0.523		0.039	0.101	
OQLF20433.060	2	1	-0.024	0.047		-0.021	0.097	
OQLF20430.060	2	2	-0.005	0.449		-0.005	0.348	
OQLF20443.060	2	3	-0.014	0.374		-0.037	0.017	
OQLF20441.060	2	4	0.005	0.631		0.001	0.904	
OQLF20432.060	2	5	-0.017	0.133		-0.008	0.471	
OQLF20435.060	2	6	-0.006	0.702		-0.030	0.049	
OQLF20439.060	2	7	-0.026	0.039		-0.016	0.209	
OQLF20444.060	2	8	-0.041	0.007		-0.053	0.001	
OQLF20440.060	2	9	0.030	0.010		0.030	0.011	
OQLF20437.060	2	10	-0.001	0.900		0.000	0.989	
OQLF20445.060	2	11	-0.084	0.000	B-	-0.080	0.001	B-
OQLF20448.060	2	12	-0.066	0.001	B-	-0.062	0.002	B-
OQLF20546.069	2	18	0.008	0.601		0.007	0.656	
OQLF20541.069	2	19	0.038	0.010		0.033	0.023	
OQLF20543.069	2	20	0.032	0.016		0.029	0.030	
OQLF20545.069	2	21	-0.012	0.371		-0.017	0.229	
OQLF20549.069	2	22	0.036	0.183		0.007	0.810	
OQLF20550.069	2	23	0.007	0.754		0.019	0.390	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.28 Gender-Based DIF Statistics: Junior Reading (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
OQLF31248.116	1	1	-0.029	0.038		-0.031	0.026	
OQLF31245.116	1	2	0.000	0.962		-0.003	0.697	
OQLF31249.116	1	3	0.034	0.004		0.033	0.005	
OQLF31250.116	1	4	0.055	0.000		0.043	0.002	
OQLF31251.116	1	5	-0.063	0.011	B-	-0.092	0.000	C-
OQLF31252.116(A)	1	6	-0.020	0.429		0.009	0.702	
OQLF31254.117	1	7	-0.005	0.508		-0.016	0.042	
OQLF31257.117	1	8	-0.011	0.388		-0.014	0.290	
OQLF31259.117	1	9	-0.031	0.033		-0.025	0.091	
OQLF31256.117	1	10	-0.004	0.713		0.000	0.972	
OQLF31260.117	1	11	-0.024	0.260		-0.039	0.066	
OQLF31262.117	1	12	0.031	0.247		0.042	0.111	
OQLF31281.120	1	14	0.015	0.007		0.010	0.067	
OQLF31283.120	1	15	0.001	0.916		0.000	0.968	
OQLF31286.120	1	16	0.010	0.473		0.023	0.098	
OQLF31282.120	1	17	0.013	0.383		0.020	0.178	
OQLF31309.120	1	18	0.051	0.033		0.086	0.000	B+
OQLF31288.120	1	19	-0.040	0.091		-0.042	0.076	
OQLF31210.113	2	1	0.013	0.366		0.008	0.591	
OQLF31200.113	2	2	-0.019	0.012		-0.012	0.135	
OQLF31213.113	2	3	-0.039	0.002		-0.035	0.005	
OQLF31209.113	2	4	-0.012	0.149		-0.009	0.276	
OQLF31202.113	2	5	0.010	0.389		0.012	0.304	
OQLF31211.113	2	6	0.024	0.118		0.036	0.020	
OQLF31204.113	2	7	-0.005	0.689		0.005	0.720	
OQLF31205.113	2	8	-0.011	0.367		-0.009	0.462	
OQLF31201.113	2	9	0.013	0.184		0.012	0.206	
OQLF31206.113	2	10	-0.038	0.007		-0.033	0.018	
OQLF31215.113	2	11	0.005	0.840		0.031	0.208	
OQLF31216.113	2	12	0.012	0.608		0.029	0.195	
OQLF31300.122	2	18	-0.004	0.475		-0.012	0.040	
OQLF31299.122	2	19	-0.005	0.711		0.003	0.836	
OQLF31302.122	2	20	0.007	0.659		0.022	0.156	
OQLF31303.122	2	21	-0.009	0.562		-0.030	0.043	
OQLF31305.122	2	22	0.031	0.213		-0.011	0.646	
OQLF31307.122	2	23	0.034	0.207		-0.001	0.962	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.29 SLL-Based DIF Statistics: Primary Reading (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQRE31619.141	1	1	-0.003	0.835		-0.006	0.693	
EQRE31617.141	1	2	-0.031	0.006		-0.043	0.000	
EQRE31621.141	1	3	-0.014	0.300		-0.018	0.179	
EQRE31622.141	1	4	0.137	0.000	C+	0.078	0.000	B+
EQRE31615.141	1	5	0.039	0.105		0.032	0.194	
EQRE31614.141	1	6	0.017	0.467		0.077	0.002	B+
EQRE31666.146	1	7	-0.003	0.830		0.013	0.370	
EQRE31667.146	1	8	0.106	0.000	C+	0.121	0.000	C+
EQRE31662.146	1	9	-0.010	0.491		-0.018	0.225	
EQRE31664.146	1	10	-0.012	0.401		-0.019	0.202	
EQRE31660.146	1	11	-0.009	0.707		0.055	0.018	
EQRE31661.146	1	12	-0.037	0.139		-0.040	0.107	
EQRE31702.150	1	14	0.008	0.542		0.030	0.028	
EQRE31698.150	1	15	0.021	0.108		0.016	0.206	
EQRE31700.150	1	16	-0.014	0.392		-0.022	0.167	
EQRE31703.150	1	17	-0.015	0.289		-0.002	0.871	
EQRE31697.150	1	18	-0.018	0.466		-0.021	0.400	
EQRE31696.150	1	19	0.073	0.028	B+	0.089	0.008	C+
EQRE31819.159	2	1	-0.003	0.838		-0.012	0.427	
EQRE31816.159	2	2	-0.019	0.094		-0.017	0.153	
EQRE31824.159	2	3	0.012	0.253		0.016	0.147	
EQRE31817.159	2	4	-0.063	0.000	B-	-0.053	0.000	
EQRE31822.159	2	5	-0.018	0.096		0.017	0.101	
EQRE31814.159	2	6	0.031	0.021		0.041	0.003	
EQRE31811.159	2	7	-0.045	0.002		-0.032	0.032	
EQRE31812.159	2	8	-0.008	0.596		-0.033	0.033	
EQRE31820.159	2	9	-0.073	0.000	B-	-0.036	0.005	
EQRE31818.159	2	10	-0.025	0.062		-0.021	0.104	
EQRE31809.159	2	11	-0.101	0.000	C-	-0.126	0.000	C-
EQRE31810.159	2	12	0.005	0.844		-0.069	0.006	B-
EQRE31736.154	2	18	-0.005	0.645		-0.013	0.220	
EQRE31737.154	2	19	0.007	0.514		0.032	0.002	
EQRE31738.154	2	20	-0.008	0.395		0.009	0.345	
EQRE31740.154	2	21	-0.006	0.705		-0.002	0.888	
EQRE31733.154	2	22	0.050	0.041		0.041	0.094	
EQRE31735.154	2	23	-0.074	0.002	B-	-0.052	0.023	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.30 SLL-Based DIF Statistics: Junior Reading (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQRE31868.162	1	1	-0.018	0.144		0.030	0.017	
EQRE31873.162	1	2	0.059	0.000	B+	0.061	0.000	B+
EQRE31870.162	1	3	-0.003	0.850		0.007	0.611	
EQRE31869.162	1	4	0.006	0.557		0.009	0.422	
EQRE31874.162	1	5	0.028	0.262		0.037	0.140	
EQRE31875.162	1	6	0.119	0.000	C+	0.153	0.000	C+
EQRE31913.167	1	7	-0.037	0.009		-0.008	0.569	
EQRE31915.167	1	8	-0.007	0.633		-0.020	0.199	
EQRE31918.167	1	9	0.011	0.443		0.009	0.527	
EQRE31914.167	1	10	0.026	0.063		0.024	0.078	
EQRE31919.167	1	11	-0.002	0.922		-0.003	0.894	
EQRE31920.167	1	12	0.058	0.013		0.046	0.052	
EQRE31950.171	1	14	-0.029	0.035		-0.022	0.113	
EQRE31952.171	1	15	-0.004	0.762		-0.005	0.688	
EQRE31949.171	1	16	-0.018	0.239		0.023	0.134	
EQRE31953.171	1	17	-0.044	0.005		-0.019	0.238	
EQRE31957.171	1	18	-0.104	0.000	C-	-0.090	0.000	C-
EQRE31955.171	1	19	-0.098	0.000	C-	-0.106	0.000	C-
EQRE32023.177	2	1	0.003	0.859		0.000	0.986	
EQRE32018.177	2	2	0.030	0.032		-0.005	0.705	
EQRE32014.177	2	3	-0.015	0.013		-0.019	0.001	
EQRE32024.177	2	4	-0.004	0.715		0.007	0.527	
EQRE32021.177	2	5	-0.067	0.000	B-	-0.070	0.000	B-
EQRE32015.177	2	6	-0.028	0.073		-0.035	0.026	
EQRE32027.177	2	7	0.029	0.080		-0.009	0.602	
EQRE32026.177	2	8	0.090	0.000	C+	0.049	0.000	
EQRE32020.177	2	9	-0.014	0.312		-0.019	0.180	
EQRE32019.177	2	10	-0.040	0.002		-0.022	0.086	
EQRE32031.177	2	11	0.025	0.323		0.070	0.005	B+
EQRE32029.177	2	12	-0.020	0.427		-0.056	0.028	
EQRE31967.173	2	18	-0.020	0.056		-0.034	0.001	
EQRE31969.173	2	19	-0.019	0.132		-0.022	0.071	
EQRE31970.173	2	20	0.047	0.002		0.042	0.005	
EQRE31972.173	2	21	0.019	0.242		0.027	0.092	
EQRE31974.173	2	22	-0.028	0.166		-0.042	0.045	
EQRE31973.173	2	23	0.030	0.154		-0.020	0.349	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.31 SLL-Based DIF Statistics: Primary Reading (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQLF31050.103	1	1	-0.020	0.347		0.003	0.886	
OQLF31048.103	1	2	0.030	0.143		0.002	0.922	
OQLF31046.103	1	3	-0.014	0.419		-0.022	0.201	
OQLF31047.103	1	4	-0.001	0.928		-0.017	0.210	
OQLF31051.103	1	5	-0.020	0.508		-0.002	0.943	
OQLF31052.103	1	6	0.041	0.193		0.040	0.199	
OQLF31065.105	1	7	-0.006	0.707		0.006	0.705	
OQLF31066.105	1	8	0.026	0.089		0.019	0.204	
OQLF31068.105	1	9	0.035	0.054		0.002	0.897	
OQLF31064.105	1	10	0.004	0.827		0.000	0.996	
OQLF31069.105	1	11	0.001	0.988		0.031	0.401	
OQLF31071.105	1	12	-0.003	0.920		0.006	0.853	
OQLF31074.106	1	14	0.009	0.560		0.007	0.647	
OQLF31073.106	1	15	0.025	0.192		0.015	0.440	
OQLF31075.106	1	16	0.004	0.797		0.007	0.667	
OQLF31076.106	1	17	0.044	0.040		0.014	0.488	
OQLF31078.106	1	18	0.053	0.074		0.022	0.445	
OQLF31080.106	1	19	0.018	0.580		0.004	0.898	
OQLF20433.060	2	1	0.041	0.022		0.050	0.004	
OQLF20430.060	2	2	-0.003	0.777		-0.011	0.247	
OQLF20443.060	2	3	-0.050	0.017		-0.047	0.027	
OQLF20441.060	2	4	-0.013	0.396		0.004	0.815	
OQLF20432.060	2	5	-0.006	0.723		-0.001	0.964	
OQLF20435.060	2	6	-0.020	0.333		-0.005	0.822	
OQLF20439.060	2	7	-0.045	0.010		-0.025	0.130	
OQLF20444.060	2	8	-0.023	0.276		0.005	0.820	
OQLF20440.060	2	9	0.066	0.000	B+	0.067	0.000	B+
OQLF20437.060	2	10	0.015	0.299		0.017	0.243	
OQLF20445.060	2	11	-0.012	0.699		0.022	0.491	
OQLF20448.060	2	12	0.001	0.970		-0.006	0.834	
OQLF20546.069	2	18	-0.018	0.377		-0.028	0.175	
OQLF20541.069	2	19	-0.005	0.793		-0.015	0.434	
OQLF20543.069	2	20	0.017	0.360		-0.009	0.636	
OQLF20545.069	2	21	0.003	0.859		-0.008	0.683	
OQLF20549.069	2	22	-0.071	0.058		-0.077	0.041	B-
OQLF20550.069	2	23	-0.019	0.521		-0.045	0.123	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.32 SLL-Based DIF Statistics: Junior Reading (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
OQLF31248.116	1	1	-0.041	0.076		-0.068	0.004	B-
OQLF31245.116	1	2	-0.003	0.789		-0.023	0.069	
OQLF31249.116	1	3	-0.019	0.362		-0.025	0.221	
OQLF31250.116	1	4	-0.034	0.164		-0.010	0.667	
OQLF31251.116	1	5	-0.095	0.024	C-	-0.120	0.004	C-
OQLF31252.116(A)	1	6	0.111	0.006	C+	0.068	0.095	
OQLF31254.117	1	7	0.008	0.570		0.010	0.486	
OQLF31257.117	1	8	0.022	0.343		0.039	0.090	
OQLF31259.117	1	9	0.042	0.094		0.025	0.306	
OQLF31256.117	1	10	0.006	0.741		0.012	0.530	
OQLF31260.117	1	11	0.017	0.621		0.013	0.705	
OQLF31262.117	1	12	0.069	0.114		-0.019	0.659	
OQLF31281.120	1	14	-0.003	0.779		0.005	0.639	
OQLF31283.120	1	15	-0.020	0.352		-0.023	0.267	
OQLF31286.120	1	16	0.013	0.594		0.003	0.896	
OQLF31282.120	1	17	0.004	0.859		-0.004	0.882	
OQLF31309.120	1	18	-0.077	0.061		-0.053	0.182	
OQLF31288.120	1	19	-0.021	0.590		0.025	0.534	
OQLF31210.113	2	1	0.014	0.583		0.018	0.450	
OQLF31200.113	2	2	0.001	0.923		0.001	0.933	
OQLF31213.113	2	3	0.023	0.287		-0.003	0.906	
OQLF31209.113	2	4	0.021	0.128		0.028	0.061	
OQLF31202.113	2	5	-0.005	0.792		0.011	0.562	
OQLF31211.113	2	6	-0.006	0.819		-0.014	0.581	
OQLF31204.113	2	7	0.016	0.497		-0.003	0.913	
OQLF31205.113	2	8	-0.009	0.678		0.016	0.459	
OQLF31201.113	2	9	0.031	0.061		0.028	0.095	
OQLF31206.113	2	10	0.015	0.537		0.020	0.410	
OQLF31215.113	2	11	0.043	0.292		0.039	0.330	
OQLF31216.113	2	12	0.013	0.734		0.063	0.101	
OQLF31300.122	2	18	0.007	0.469		0.002	0.888	
OQLF31299.122	2	19	-0.035	0.148		-0.039	0.111	
OQLF31302.122	2	20	-0.012	0.643		-0.012	0.627	
OQLF31303.122	2	21	-0.048	0.054		-0.061	0.014	B-
OQLF31305.122	2	22	-0.028	0.501		-0.026	0.526	
OQLF31307.122	2	23	-0.037	0.420		-0.044	0.329	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.33 Gender-Based DIF Statistics: Primary Writing (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQWE20246_T	1	13	0.012	0.596		0.007	0.763	
EQWE20246_V	1	13	-0.017	0.397		-0.036	0.069	
EQWE30415	2	13	-0.035	0.066		0.007	0.713	
EQWE30407	2	14	-0.020	0.076		-0.007	0.537	
EQWE30408	2	15	-0.007	0.677		-0.022	0.206	
EQWE30414	2	16	0.018	0.191		0.010	0.500	
EQWE30427_T	2	17	-0.007	0.736		-0.021	0.329	
EQWE30427_V	2	17	-0.026	0.152		-0.003	0.853	
11042_T	2	24	0.040	0.082		0.075	0.001	B+
11042_V	2	24	0.007	0.701		0.004	0.833	
EQWE30406	2	25	0.005	0.741		0.017	0.257	
EQWE30409	2	26	-0.002	0.899		0.034	0.014	
EQWE30419	2	27	-0.017	0.347		-0.021	0.237	
EQWE30412	2	28	0.041	0.008		0.036	0.023	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.34 Gender-Based DIF Statistics: Junior Writing (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQWE30498_T	1	13	-0.016	0.471		-0.043	0.054	
EQWE30498_V	1	13	-0.051	0.004		-0.044	0.013	
EQWE30486	2	13	-0.058	0.001		-0.070	0.000	B-
EQWE30471	2	14	-0.039	0.006		-0.019	0.178	
EQWE30474	2	15	0.014	0.320		0.019	0.159	
EQWE30490	2	16	0.059	0.001	B+	0.041	0.020	
EQWE30504_T	2	17	0.013	0.555		0.019	0.371	
EQWE30504_V	2	17	-0.012	0.487		-0.013	0.432	
10715_T	2	24	0.112	0.000	C+	0.126	0.000	C+
10715_V	2	24	-0.003	0.858		-0.001	0.975	
EQWE30472	2	25	-0.011	0.382		0.005	0.674	
EQWE30484	2	26	-0.046	0.002		-0.051	0.001	
EQWE30488	2	27	0.000	0.954		-0.006	0.394	
EQWE30478	2	28	0.099	0.000	C+	0.087	0.000	B+

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.35 Gender-Based DIF Statistics: Primary Writing (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
OQEF30521_T	1	13	-0.021	0.389		0.010	0.662	
OQEF30521_V	1	13	0.055	0.003		0.017	0.368	
OQEF30502	2	13	-0.002	0.924		-0.013	0.455	
OQEF30511	2	14	-0.023	0.151		0.004	0.822	
OQEF30509	2	15	-0.024	0.069		0.003	0.818	
OQEF30514	2	16	-0.004	0.824		0.000	0.983	
OQEF30523_T	2	17	0.050	0.019		0.106	0.000	C+
OQEF30523_V	2	17	-0.034	0.026		-0.052	0.001	
11248_T	2	24	0.015	0.499		-0.001	0.962	
11248_V	2	24	-0.049	0.007		-0.063	0.000	B-
OQEF30503	2	25	-0.004	0.833		-0.017	0.332	
12291	2	26	0.027	0.076		0.049	0.001	
OQEF30519	2	27	0.030	0.077		0.026	0.119	
OQEF30507	2	28	-0.002	0.839		0.005	0.642	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.36 Gender-Based DIF Statistics: Junior Writing (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
12462_T	1	13	0.028	0.224		0.037	0.121	
12462_V	1	13	-0.024	0.172		-0.014	0.436	
OQEF30597	2	13	-0.025	0.102		-0.057	0.000	
OQEF30585	2	14	0.009	0.522		0.012	0.339	
OQEF30593	2	15	0.030	0.097		0.031	0.086	
OQEF30586	2	16	0.021	0.208		0.021	0.200	
OQEF30601_T	2	17	0.024	0.287		0.016	0.490	
OQEF30601_V	2	17	0.002	0.907		-0.007	0.668	
12542_T	2	24	0.049	0.019		0.022	0.288	
12542_V	2	24	-0.020	0.265		-0.009	0.638	
OQEF30587	2	25	0.055	0.000		0.029	0.045	
OQEF30599	2	26	-0.007	0.678		-0.010	0.558	
OQEF30598	2	27	0.092	0.000	C+	0.060	0.000	B+
OQEF30581	2	28	0.008	0.626		0.030	0.059	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.37 SLL-Based DIF Statistics: Primary Writing (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQWE20246_T	1	13	0.007	0.768		0.012	0.602	
EQWE20246_V	1	13	-0.055	0.005		-0.058	0.003	
EQWE30415	2	13	-0.020	0.268		0.018	0.331	
EQWE30407	2	14	0.030	0.015		0.021	0.092	
EQWE30408	2	15	-0.018	0.272		-0.034	0.045	
EQWE30414	2	16	-0.033	0.019		-0.046	0.002	
EQWE30427_T	2	17	0.035	0.098		0.034	0.114	
EQWE30427_V	2	17	0.003	0.870		0.024	0.166	
11042_T	2	24	0.023	0.297		0.047	0.031	
11042_V	2	24	-0.010	0.581		-0.066	0.000	B-
EQWE30406	2	25	0.015	0.318		0.007	0.636	
EQWE30409	2	26	-0.007	0.617		0.025	0.077	
EQWE30419	2	27	-0.025	0.150		-0.048	0.006	
EQWE30412	2	28	-0.034	0.027		-0.011	0.478	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.38 SLL-Based DIF Statistics: Junior Writing (English)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQWE30498_T	1	13	0.031	0.154		0.024	0.279	
EQWE30498_V	1	13	0.022	0.206		0.035	0.044	
EQWE30486	2	13	0.001	0.958		-0.018	0.279	
EQWE30471	2	14	0.034	0.024		0.030	0.044	
EQWE30474	2	15	-0.062	0.000	B-	-0.037	0.008	
EQWE30490	2	16	-0.083	0.000	B-	-0.048	0.005	
EQWE30504_T	2	17	-0.020	0.330		-0.020	0.327	
EQWE30504_V	2	17	-0.031	0.090		0.007	0.677	
10715_T	2	24	-0.006	0.775		-0.025	0.234	
10715_V	2	24	0.002	0.914		-0.016	0.406	
EQWE30472	2	25	0.015	0.253		0.002	0.866	
EQWE30484	2	26	0.008	0.604		0.006	0.711	
EQWE30488	2	27	-0.002	0.794		0.004	0.597	
EQWE30478	2	28	0.060	0.001	B+	0.047	0.006	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.39 SLL-Based DIF Statistics: Primary Writing (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQEF30521_T	1	13	-0.001	0.969		0.012	0.706	
OQEF30521_V	1	13	0.003	0.890		-0.006	0.819	
OQEF30502	2	13	-0.035	0.139		-0.033	0.153	
OQEF30511	2	14	-0.021	0.355		0.005	0.821	
OQEF30509	2	15	-0.032	0.091		-0.052	0.006	
OQEF30514	2	16	-0.010	0.673		-0.011	0.650	
OQEF30523_T	2	17	-0.060	0.042	B-	-0.065	0.025	B-
OQEF30523_V	2	17	0.032	0.124		0.061	0.003	B+
11248_T	2	24	0.002	0.957		0.013	0.642	
11248_V	2	24	0.005	0.853		-0.020	0.410	
OQEF30503	2	25	0.056	0.018		0.056	0.018	
12291	2	26	0.025	0.230		0.002	0.943	
OQEF30519	2	27	-0.039	0.095		-0.069	0.003	B-
OQEF30507	2	28	0.046	0.004		0.034	0.035	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs students; + = favouring non-SLLs.

Table 7.1.40 SLL-Based DIF Statistics: Junior Writing (French)

Item Code	Booklet	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
12462_T	1	13	0.044	0.247		0.064	0.081	
12462_V	1	13	0.044	0.127		0.045	0.105	
OQEF30597	2	13	-0.049	0.050		-0.097	0.000	C-
OQEF30585	2	14	0.050	0.031		0.073	0.001	B+
OQEF30593	2	15	-0.089	0.003	C-	-0.095	0.001	C-
OQEF30586	2	16	-0.018	0.534		0.003	0.911	
OQEF30601_T	2	17	-0.086	0.018	B-	-0.022	0.532	
OQEF30601_V	2	17	-0.074	0.007	B-	-0.092	0.001	C-
12542_T	2	24	0.011	0.756		0.018	0.599	
12542_V	2	24	0.029	0.344		-0.020	0.480	
OQEF30587	2	25	0.008	0.763		0.023	0.359	
OQEF30599	2	26	0.054	0.080		0.052	0.089	
OQEF30598	2	27	-0.013	0.612		-0.013	0.622	
OQEF30581	2	28	0.034	0.210		0.012	0.671	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.41 Gender-Based DIF Statistics: Primary Mathematics (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQME31574	1	-0.024	0.010		-0.023	0.017	
EQME20908	2	-0.056	0.000		-0.039	0.001	
EQME31512	3	0.024	0.101		0.039	0.007	
EQME31542	4	-0.039	0.007		-0.042	0.003	
EQME20910	5	-0.003	0.836		-0.001	0.950	
EQME31514	6	0.011	0.470		0.003	0.849	
EQME20812	7	-0.044	0.082		-0.052	0.045	
EQME31587	8	0.282	0.000	C+	0.246	0.000	C+
EQME31516	9	0.062	0.071		0.116	0.001	C+
EQME31526	10	0.053	0.000		0.053	0.000	
EQME31543	11	-0.069	0.000		-0.067	0.000	
EQME31601	12	0.023	0.089		0.026	0.059	
EQME31518	13	0.042	0.002		0.040	0.003	
EQME31541	14	0.001	0.961		0.000	0.990	
EQME31591	15	0.040	0.006		0.045	0.002	
EQME31525	16	-0.014	0.207		-0.007	0.500	
EQME31567	17	0.037	0.008		0.064	0.000	B+
EQME31559	18	-0.012	0.423		-0.051	0.001	
EQME20852	19	0.012	0.387		-0.031	0.027	
EQME31540	20	0.022	0.136		0.016	0.266	
EQME31527	21	-0.028	0.000		-0.037	0.000	
EQME20919	22	-0.017	0.244		-0.004	0.781	
EQME10044	23	0.031	0.000		0.016	0.067	
EQME10060	24	-0.030	0.024		-0.036	0.007	
EQME10075	25	-0.164	0.000	C-	-0.193	0.000	C-
EQME31536	26	-0.091	0.002	C-	-0.079	0.008	
EQME31560	27	0.074	0.008	B+	0.109	0.000	C+
EQME31588	28	-0.095	0.009	C-	-0.143	0.000	C-
EQME31614	29	-0.005	0.641		-0.007	0.510	
EQME31570	30	0.013	0.316		0.010	0.447	
EQME20820	31	0.022	0.006		0.027	0.000	
EQME31505	32	-0.013	0.344		-0.012	0.375	
EQME31510	33	-0.021	0.161		-0.022	0.141	
EQME31586	34	-0.006	0.671		-0.011	0.455	
EQME31594	35	0.023	0.122		0.046	0.002	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.42 Gender-Based DIF Statistics: Junior Mathematics (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME21023	1	0.034	0.009		0.010	0.435	
EQME31699	2	-0.032	0.032		-0.028	0.062	
EQME10257	3	0.046	0.002		0.025	0.090	
EQME31720	4	-0.018	0.194		-0.029	0.037	
EQME31648	5	-0.011	0.435		-0.008	0.553	
EQME31655	6	0.009	0.532		0.013	0.388	
EQME31744	7	0.268	0.000	C+	0.277	0.000	C+
EQME21011	8	-0.019	0.532		-0.176	0.000	C-
EQME21009	9	-0.181	0.000	C-	-0.146	0.000	C-
EQME31702	10	-0.094	0.002	C-	-0.049	0.102	
EQME31729	11	0.012	0.392		0.017	0.220	
EQME31636	12	0.007	0.649		0.034	0.018	
EQME31675	13	-0.040	0.007		-0.040	0.007	
EQME31714	14	-0.063	0.000		-0.071	0.000	
EQME31646	15	-0.020	0.175		0.000	0.992	
EQME21126	16	-0.002	0.917		-0.008	0.615	
3469832	17	0.021	0.167		-0.005	0.714	
EQME31635	18	0.016	0.115		0.005	0.630	
EQME31665	19	0.023	0.003		0.000	0.995	
EQME31678	20	-0.037	0.014		-0.046	0.002	
EQME31711	21	0.021	0.053		-0.008	0.446	
EQME21050	22	-0.002	0.868		0.011	0.448	
EQME21004	23	-0.031	0.043		-0.040	0.008	
EQME31748	24	0.072	0.000	B+	0.063	0.000	B+
EQME31700	25	0.050	0.001		0.057	0.000	
EQME21015	26	0.165	0.000	C+	0.108	0.002	C+
EQME21056	27	-0.036	0.310		-0.014	0.677	
EQME31662	28	-0.072	0.032		0.006	0.867	
EQME10275	29	-0.038	0.139		-0.012	0.640	
EQME21068	30	0.042	0.004		0.110	0.000	C+
EQME21052	31	-0.008	0.583		-0.010	0.491	
EQME31668	32	-0.013	0.405		0.014	0.354	
EQME31645	33	-0.029	0.046		-0.022	0.128	
EQME31669	34	-0.007	0.612		-0.045	0.002	
EQME31685	35	-0.028	0.049		-0.042	0.004	
EQME31690	36	-0.021	0.083		-0.022	0.078	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.43 Gender-Based DIF Statistics: Primary Mathematics (French)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQMF10054	1	0.029	0.000		0.026	0.001	
OQMF31572	2	0.005	0.538		0.015	0.089	
OQMF20609	3	0.027	0.069		0.034	0.021	
OQMF31523	4	-0.062	0.000		-0.075	0.000	
OQMF20615	5	0.036	0.001		0.029	0.009	
OQMF31503	6	0.064	0.000	B+	0.042	0.001	
OQMF20505	7	0.112	0.001	C+	0.115	0.000	C+
OQMF20508	8	-0.059	0.024		-0.076	0.004	
OQMF31597	9	-0.050	0.102		-0.071	0.018	
OQMF10063	10	0.146	0.000	C+	0.159	0.000	C+
OQMF20584	11	0.015	0.325		-0.002	0.884	
OQMF20577	12	-0.033	0.005		-0.020	0.094	
OQMF20624	13	0.019	0.202		0.017	0.239	
OQMF20592	14	-0.007	0.622		-0.018	0.217	
OQMF10027	15	-0.005	0.739		0.021	0.130	
OQMF31516	16	0.017	0.263		0.009	0.560	
OQMF31579	17	-0.019	0.206		-0.025	0.104	
OQMF31590	18	-0.040	0.006		-0.042	0.004	
OQMF31571	19	-0.058	0.000		-0.050	0.000	
OQMF20575	20	0.018	0.193		0.030	0.030	
OQMF31556	21	-0.049	0.001		-0.028	0.071	
OQMF31524	22	0.007	0.581		0.011	0.427	
OQMF31500	23	0.031	0.014		0.030	0.018	
OQMF20621	24	-0.055	0.000		-0.048	0.000	
OQMF31533	25	-0.013	0.400		-0.013	0.428	
OQMF10057	26	0.125	0.000	C+	0.129	0.000	C+
OQMF20517	27	-0.049	0.082		-0.051	0.064	
OQMF31582	28	-0.099	0.002	C-	-0.048	0.122	
OQMF20535	29	-0.101	0.001	C-	-0.136	0.000	C-
OQMF20596	30	-0.010	0.384		-0.024	0.036	
OQMF20581	31	0.061	0.000	B+	0.052	0.000	
OQMF31513	32	0.027	0.006		0.034	0.001	
OQMF31511	33	-0.027	0.049		-0.018	0.212	
OQMF31510	34	0.033	0.033		0.039	0.012	
OQMF31592	35	-0.026	0.074		-0.027	0.059	
OQMF31542	36	0.009	0.440		0.016	0.177	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.44 Gender-Based DIF Statistics: Junior Mathematics (French)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
OQMF31627	1	-0.009	0.266		-0.006	0.448	
OQMF20806	2	-0.052	0.000		-0.055	0.000	
OQMF31670	3	-0.027	0.096		-0.042	0.009	
OQMF31658	4	-0.013	0.379		0.000	0.983	
OQMF31681	5	-0.039	0.001		-0.042	0.000	
OQMF31624	6	0.044	0.002		0.054	0.000	
OQMF10112	7	-0.032	0.230		-0.083	0.002	
OQMF20720	8	-0.031	0.353		0.039	0.252	
OQMF20782	9	-0.067	0.010		-0.029	0.264	
OQMF10206	10	-0.098	0.000	C-	-0.134	0.000	C-
OQMF31621	11	0.007	0.582		0.013	0.308	
OQMF20759	12	-0.014	0.213		-0.006	0.570	
OQMF20749	13	0.022	0.150		0.026	0.090	
OQMF31635	14	0.051	0.000		0.064	0.000	B+
OQMF31664	15	-0.012	0.412		-0.010	0.512	
OQMF31714	16	0.050	0.000		0.063	0.000	B+
OQMF31702	17	-0.013	0.383		-0.035	0.015	
OQMF20736	18	0.034	0.025		0.015	0.332	
OQMF10143	19	-0.047	0.000		-0.042	0.000	
OQMF20732	20	0.002	0.867		0.016	0.277	
OQMF31717	21	0.020	0.169		0.020	0.166	
OQMF31695	22	-0.002	0.868		-0.002	0.868	
OQMF20728	23	0.019	0.172		0.015	0.274	
OQMF31633	24	0.038	0.005		0.037	0.006	
OQMF31712	25	-0.018	0.235		-0.037	0.015	
OQMF20756	26	0.207	0.000	C+	0.176	0.000	C+
OQMF20768	27	-0.034	0.261		-0.025	0.414	
OQMF31641	28	0.018	0.496		0.017	0.501	
OQMF20812	29	0.022	0.393		0.023	0.381	
OQMF31678	30	-0.003	0.827		-0.012	0.361	
OQMF31696	31	-0.005	0.717		0.002	0.910	
OQMF31628	32	0.007	0.641		0.009	0.529	
OQMF10208	33	-0.060	0.000		-0.051	0.000	
OQMF31687	34	-0.002	0.894		-0.001	0.930	
OQMF31655	35	0.025	0.081		0.017	0.240	
OQMF10215	36	0.014	0.256		-0.006	0.601	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.45 SLL-Based DIF Statistics: Primary Mathematics (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME31574	1	-0.024	0.009		-0.027	0.005	
EQME20908	2	-0.044	0.000		-0.040	0.001	
EQME31512	3	0.021	0.140		0.013	0.352	
EQME31542	4	-0.019	0.181		-0.025	0.083	
EQME20910	5	-0.043	0.001		-0.038	0.005	
EQME31514	6	0.002	0.896		-0.010	0.496	
EQME20812	7	0.078	0.002	B+	0.076	0.002	B+
EQME31587	8	0.054	0.088		0.038	0.228	
EQME31516	9	-0.100	0.004	C-	-0.128	0.000	C-
EQME31526	10	0.007	0.625		0.010	0.494	
EQME31543	11	-0.011	0.455		-0.010	0.501	
EQME31601	12	0.016	0.257		0.018	0.194	
EQME31518	13	0.040	0.002		0.024	0.060	
EQME31541	14	-0.011	0.402		0.032	0.015	
EQME31591	15	0.022	0.126		0.001	0.934	
EQME31525	16	0.009	0.433		0.027	0.018	
EQME31567	17	-0.046	0.001		-0.052	0.000	
EQME31559	18	0.012	0.404		0.019	0.194	
EQME20852	19	0.000	0.982		-0.016	0.241	
EQME31540	20	-0.009	0.558		-0.020	0.178	
EQME31527	21	0.017	0.033		0.005	0.547	
EQME20919	22	-0.023	0.107		0.019	0.192	
EQME10044	23	0.011	0.252		0.009	0.321	
EQME10060	24	-0.007	0.600		0.012	0.359	
EQME10075	25	0.054	0.104		0.063	0.058	
EQME31536	26	-0.125	0.000	C-	-0.064	0.034	
EQME31560	27	-0.012	0.670		-0.010	0.726	
EQME31588	28	0.053	0.138		0.012	0.739	
EQME31614	29	-0.001	0.898		-0.001	0.918	
EQME31570	30	0.012	0.379		0.006	0.659	
EQME20820	31	0.013	0.155		0.004	0.651	
EQME31505	32	0.030	0.031		0.026	0.059	
EQME31510	33	0.034	0.021		-0.010	0.520	
EQME31586	34	0.007	0.634		0.002	0.891	
EQME31594	35	-0.037	0.014		-0.018	0.238	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.46 SLL-Based DIF Statistics: Junior Mathematics (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME21023	1	0.006	0.634		0.008	0.546	
EQME31699	2	-0.015	0.316		-0.006	0.698	
EQME10257	3	0.019	0.183		-0.009	0.531	
EQME31720	4	-0.010	0.486		-0.025	0.082	
EQME31648	5	0.029	0.043		0.012	0.406	
EQME31655	6	0.018	0.218		0.012	0.408	
EQME31744	7	0.047	0.226		0.078	0.039	B+
EQME21011	8	-0.022	0.467		-0.007	0.810	
EQME21009	9	-0.060	0.055		-0.116	0.000	C-
EQME31702	10	-0.151	0.000	C-	-0.082	0.006	
EQME31729	11	0.061	0.000	B+	0.085	0.000	B+
EQME31636	12	0.045	0.002		0.056	0.000	
EQME31675	13	0.011	0.453		-0.010	0.486	
EQME31714	14	-0.030	0.022		-0.029	0.030	
EQME31646	15	0.011	0.480		0.002	0.896	
EQME21126	16	-0.018	0.230		-0.021	0.171	
3469832	17	0.042	0.006		0.036	0.019	
EQME31635	18	-0.010	0.320		0.004	0.690	
EQME31665	19	-0.010	0.233		0.014	0.080	
EQME31678	20	-0.005	0.730		-0.008	0.577	
EQME31711	21	-0.036	0.001		-0.062	0.000	
EQME21050	22	0.040	0.005		0.021	0.141	
EQME21004	23	-0.051	0.001		-0.078	0.000	
EQME31748	24	-0.032	0.034		-0.004	0.791	
EQME31700	25	-0.025	0.093		-0.039	0.008	
EQME21015	26	0.037	0.303		0.060	0.093	
EQME21056	27	0.106	0.002	C+	0.108	0.002	C+
EQME31662	28	-0.051	0.129		-0.120	0.000	C-
EQME10275	29	0.070	0.007	B+	0.054	0.034	
EQME21068	30	0.014	0.336		0.048	0.001	
EQME21052	31	-0.020	0.140		0.028	0.044	
EQME31668	32	0.007	0.664		-0.012	0.434	
EQME31645	33	-0.060	0.000		-0.011	0.441	
EQME31669	34	-0.016	0.263		-0.043	0.003	
EQME31685	35	0.039	0.006		0.018	0.214	
EQME31690	36	0.032	0.011		0.011	0.398	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.47 SLL-Based DIF Statistics: Primary Mathematics (French)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQMF10054	1	-0.007	0.567		-0.001	0.939	
OQMF31572	2	0.007	0.574		-0.003	0.805	
OQMF20609	3	-0.015	0.444		-0.025	0.216	
OQMF31523	4	-0.004	0.774		-0.003	0.832	
OQMF20615	5	0.002	0.882		-0.026	0.093	
OQMF31503	6	0.016	0.360		-0.019	0.296	
OQMF20505	7	0.032	0.474		0.056	0.202	
OQMF20508	8	-0.027	0.466		0.006	0.871	
OQMF31597	9	-0.084	0.047		-0.033	0.433	
OQMF10063	10	0.017	0.688		0.009	0.834	
OQMF20584	11	0.042	0.035		0.050	0.013	
OQMF20577	12	-0.015	0.355		-0.005	0.748	
OQMF20624	13	-0.003	0.879		0.001	0.954	
OQMF20592	14	-0.023	0.244		-0.036	0.067	
OQMF10027	15	-0.033	0.083		-0.025	0.192	
OQMF31516	16	-0.022	0.288		-0.018	0.387	
OQMF31579	17	-0.019	0.361		-0.023	0.272	
OQMF31590	18	-0.033	0.103		-0.028	0.160	
OQMF31571	19	0.000	0.987		0.022	0.132	
OQMF20575	20	0.009	0.644		0.040	0.037	
OQMF31556	21	0.077	0.000	B+	0.052	0.013	
OQMF31524	22	-0.009	0.640		-0.007	0.704	
OQMF31500	23	0.011	0.535		0.009	0.617	
OQMF20621	24	0.011	0.480		-0.001	0.944	
OQMF31533	25	0.021	0.325		0.050	0.019	
OQMF10057	26	-0.068	0.129		-0.079	0.077	
OQMF20517	27	-0.034	0.375		-0.078	0.044	
OQMF31582	28	0.090	0.037	C+	0.117	0.007	C+
OQMF20535	29	0.039	0.351		0.045	0.276	
OQMF20596	30	-0.004	0.794		0.014	0.397	
OQMF20581	31	-0.006	0.752		0.007	0.735	
OQMF31513	32	0.013	0.344		0.019	0.188	
OQMF31511	33	-0.005	0.802		0.007	0.692	
OQMF31510	34	-0.015	0.483		-0.017	0.442	
OQMF31592	35	-0.037	0.060		-0.036	0.073	
OQMF31542	36	0.004	0.784		0.002	0.906	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.48 SLL-Based DIF Statistics: Junior Mathematics (French)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQMF31627	1	-0.025	0.078		0.000	0.985	
OQMF20806	2	0.008	0.740		0.017	0.479	
OQMF31670	3	0.016	0.565		0.010	0.714	
OQMF31658	4	0.012	0.656		-0.016	0.536	
OQMF31681	5	-0.027	0.176		-0.034	0.093	
OQMF31624	6	0.021	0.378		-0.022	0.359	
OQMF10112	7	0.028	0.542		0.042	0.365	
OQMF20720	8	-0.041	0.476		-0.026	0.653	
OQMF20782	9	0.001	0.988		0.024	0.582	
OQMF10206	10	0.037	0.423		0.002	0.967	
OQMF31621	11	0.036	0.110		0.029	0.205	
OQMF20759	12	0.005	0.812		-0.032	0.126	
OQMF20749	13	-0.036	0.166		-0.020	0.442	
OQMF31635	14	-0.034	0.172		-0.025	0.324	
OQMF31664	15	0.036	0.164		0.049	0.058	
OQMF31714	16	-0.008	0.708		-0.033	0.151	
OQMF31702	17	0.013	0.590		0.032	0.199	
OQMF20736	18	-0.042	0.098		0.009	0.723	
OQMF10143	19	0.005	0.810		0.007	0.736	
OQMF20732	20	-0.008	0.759		0.013	0.602	
OQMF31717	21	0.025	0.314		0.024	0.336	
OQMF31695	22	0.003	0.893		0.005	0.848	
OQMF20728	23	0.024	0.313		0.029	0.220	
OQMF31633	24	0.002	0.948		0.025	0.280	
OQMF31712	25	0.010	0.720		-0.018	0.488	
OQMF20756	26	-0.090	0.091		-0.103	0.049	C-
OQMF20768	27	0.096	0.060		0.060	0.235	
OQMF31641	28	-0.009	0.836		0.004	0.934	
OQMF20812	29	0.020	0.650		0.013	0.774	
OQMF31678	30	-0.011	0.619		-0.046	0.041	
OQMF31696	31	-0.027	0.271		-0.019	0.456	
OQMF31628	32	0.060	0.012	B+	0.048	0.045	
OQMF10208	33	0.007	0.793		0.042	0.096	
OQMF31687	34	0.008	0.726		0.000	0.998	
OQMF31655	35	-0.006	0.794		-0.014	0.567	
OQMF10215	36	0.018	0.400		0.044	0.034	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

The Grade 9 Assessment of Mathematics

Classical Item Statistics and IRT Item Parameters

Table 7.1.49 Item Statistics: Grade 9 Applied Mathematics, Winter (English)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQME10515	1	NA1.03	KU	N	2	68.60	0.27	-0.70	0.51
EQME10662	2	NA1.04	AP	N	2	39.22	0.26	1.27	0.64
EQME10617	3	NA1.06	PS	N	1	35.29	0.17	1.93	0.48
EQME21452	4	NA1.06	PS	N	2	50.05	0.37	0.48	0.90
EQME31771	6	NA2.04	KU	N	3	56.54	0.29	0.16	0.56
EQME31812	7	NA2.08	AP	N	3	59.55	0.29	-0.02	0.56
EQME31763	9	NA2.02	AP	N	4*	57.75 (2.31)	0.44	-1.10	0.25
EQME21453	10	LR1.04	AP	R	4	68.05	0.33	-0.60	0.60
EQME10512	12	LR2.02	AP	R	2	59.41	0.20	0.00	0.34
EQME31816	14	LR3.01	KU	R	2	55.80	0.39	0.14	0.86
EQME10549	15	LR3.04	AP	R	1	44.65	0.27	0.85	0.61
EQME10520	16	LR3.05	PS	R	4	53.75	0.25	0.52	0.51
3434098	18	LR4.04	AP	R	4	69.67	0.29	-0.72	0.54
EQME31839	19	LR4.05	PS	R	4	56.15	0.29	0.18	0.58
EQME31826	23	LR3.03	AP	R	4*	56.00 (2.24)	0.46	-0.74	0.30
EQME10505	24	MG1.03	KU	M	1	61.31	0.36	-0.27	0.69
EQME10538	25	MG2.03	PS	M	3	49.70	0.37	0.36	0.91
EQME10579	26	MG2.05	AP	M	4	51.97	0.39	0.20	0.89
EQME10581	27	MG2.05	PS	M	3	58.55	0.26	0.06	0.51
EQME10522	28	MG3.01	KU	M	1	85.83	0.24	-1.88	0.57
EQME21454	30	MG2.03	PS	M	4*	48.25 (1.93)	0.55	-0.29	0.46
EQME31774	NR	LR2.01	KU	R	3	82.31	0.30	-1.42	0.68
EQME31858	NR	LR4.04	PS	R	3	68.48	0.32	-0.56	0.61
EQME10513	NR	LR2.03	AP	R	1	52.24	0.23	0.53	0.44
EQME31761	NR	MG3.02	AP	M	3	40.45	0.18	1.46	0.48
EQME31859	NR	LR4.05	KU	R	3	87.27	0.27	-1.83	0.68
EQME31750	NR	NA2.02	KU	N	4	60.39	0.35	-0.09	0.68
EQME31784	NR	LR2.02	AP	R	4*	74.75 (2.99)	0.39	-2.89	0.27
EQME31847	NR	LR4.02	PS	R	4*	46.25 (1.85)	0.48	0.27	0.39
EQME31869	NR	MG3.02	AP	M	4*	47.50 (1.90)	0.52	-0.23	0.34
EQME31802	NR	NA1.05	PS	N	4*	34.50 (1.38)	0.44	0.28	0.33

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. KU = knowledge and understanding; AP = application; PS = thinking; N = number sense and algebra; M = measurement and geometry; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.50 Distribution of Score Points and Step Parameters for Open-Response Items: Grade 9 Applied Mathematics, Winter (English)

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
EQME21454	20	% of Students	4.88	1.80	40.35	21.68	15.64	15.64
		Parameters			-3.07	0.54	0.65	0.71
EQME31763	9	% of Students	5.06	2.96	37.70	9.21	5.79	39.28
		Parameters			-4.31	3.05	1.11	-4.23
EQME31784	NR	% of Students	0.41	0.14	15.25	18.57	15.38	50.24
		Parameters			-8.23	-1.03	0.17	-2.47
EQME31802	NR	% of Students	10.60	3.00	63.95	4.57	6.50	11.41
		Parameters			-3.19	4.71	-0.21	-0.20
EQME31826	23	% of Students	6.50	1.23	33.28	12.21	20.84	25.95
		Parameters			-3.57	1.64	-1.01	-0.02
EQME31847	NR	% of Students	9.36	1.15	30.42	27.53	26.26	5.29
		Parameters			-2.31	-0.07	0.32	3.15
EQME31869	NR	% of Students	13.40	1.78	33.33	22.70	3.94	24.90
		Parameters			-1.96	0.45	3.18	-2.58

Note. The total number of students is 18 148. NR = not released.

Table 7.1.51 Item Statistics: Grade 9 Applied Mathematics, Spring (English)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQME10515	1	NA1.03	KU	N	2	71.59	0.29	-0.70	0.51
EQME10662	2	NA1.04	AP	N	2	39.56	0.27	1.27	0.64
EQME10617	3	NA1.06	PS	N	1	36.45	0.19	1.93	0.48
EQME21452	4	NA1.06	PS	N	2	49.34	0.40	0.48	0.90
EQME31790	5	NA2.03	KU	N	2	65.15	0.17	-0.52	0.28
EQME31842	8	NA1.05	PS	N	4*	48.50 (1.94)	0.49	-0.15	0.33
EQME21453	10	LR1.04	AP	R	4	71.46	0.30	-0.60	0.60
EQME10514	11	LR2.02	KU	R	3	85.09	0.20	-2.31	0.40
EQME31815	13	LR2.02	AP	R	1	26.91	0.30	1.62	1.23
EQME10549	15	LR3.04	AP	R	1	47.39	0.28	0.85	0.61
EQME10520	16	LR3.05	PS	R	4	51.53	0.28	0.52	0.51
EQME10501	17	LR4.01	KU	R	2	93.18	0.20	-2.67	0.60
3434098	18	LR4.04	AP	R	4	71.92	0.29	-0.72	0.54
EQME31819	20	LR4.06	PS	R	3	66.13	0.39	-0.25	0.82
EQME31865	21	LR2.01	AP	R	4*	72.75 (2.91)	0.46	-1.25	0.40
EQME21441	22	LR3.01	AP	R	4*	67.75 (2.71)	0.45	-1.43	0.28
EQME10505	24	MG1.03	KU	M	1	67.67	0.35	-0.27	0.69
EQME10538	25	MG2.03	PS	M	3	54.36	0.40	0.36	0.91
EQME10579	26	MG2.05	AP	M	4	58.03	0.43	0.20	0.89
EQME10522	28	MG3.01	KU	M	1	86.15	0.26	-1.88	0.57
EQME31801	29	MG3.02	AP	M	1	51.12	0.39	0.47	0.92
EQME31769	31	MG3.01	AP	M	4*	39.75 (1.59)	0.50	0.16	0.37
EQME31791	NR	NA2.04	AP	N	2	64.33	0.38	-0.17	0.79
EQME10596	NR	MG2.03	PS	M	2	51.09	0.40	0.46	0.99
EQME31830	NR	NA2.01	KU	N	3	41.05	0.26	1.27	0.59
EQME10659	NR	LR3.04	KU	R	4	85.93	0.20	-2.36	0.42
EQME31778	NR	LR4.02	PS	R	1	36.81	0.25	1.50	0.65
EQME10532	NR	LR2.02	AP	R	4	50.97	0.33	0.55	0.66
EQME31863	NR	NA2.02	AP	N	4*	44.75 (1.79)	0.50	-0.10	0.36
EQME31786	NR	LR4.06	PS	R	4*	47.75 (1.91)	0.54	-0.22	0.44
EQME31828	NR	MG2.02	PS	M	4*	42.75 (1.71)	0.45	0.22	0.40

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. KU = knowledge and understanding; AP = application; PS = thinking; N = number sense and algebra; M = measurement and geometry; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.52 Distribution of Score Points and Step Parameters for Open-Response Items: Grade 9 Applied Mathematics, Spring (English)

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
EQME21441	22	% of Students	3.98	0.60	21.95	17.24	10.31	45.93
		Parameters			-4.08	0.12	1.06	-2.84
EQME31769	31	% of Students	11.81	2.49	49.90	15.43	3.57	16.81
		Parameters			-2.42	1.90	2.79	-1.61
EQME31786	NR	% of Students	5.61	0.72	42.19	21.03	14.68	15.77
		Parameters			-3.19	0.79	0.83	0.71
EQME31828	NR	% of Students	7.58	0.86	32.62	46.40	4.72	7.82
		Parameters			-2.58	-0.57	3.83	0.19
EQME31842	8	% of Students	10.57	2.09	31.28	26.58	8.65	20.83
		Parameters			-2.15	0.17	2.28	-0.89
EQME31863	NR	% of Students	6.48	3.44	45.15	18.30	8.84	17.79
		Parameters			-2.98	1.40	1.55	-0.36
EQME31865	21	% of Students	2.98	0.40	6.53	18.47	38.75	32.87
		Parameters			-2.11	-2.20	-1.26	0.58

Note. The total number of students is 22 182. NR = not released.

Table 7.1.53 Item Statistics: Grade 9 Academic Mathematics, Winter (English)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQME10421	1	NA1.01	KU	N	2	60.87	0.31	-0.11	0.62
EQME10346	4	NA2.05	KU	N	4	76.18	0.32	-1.10	0.64
EQME31969	5	NA2.08	PS	N	2	62.07	0.32	-0.16	0.61
EQME21251	7	LR1.01	PS	R	3	43.32	0.30	0.85	0.66
EQME10413	8	LR2.01	AP	R	4	85.62	0.23	-2.10	0.50
EQME31970	9	LR2.04	KU	R	2	76.32	0.36	-0.94	0.73
EQME10384	10	LR2.05	PS	R	2	61.27	0.26	-0.31	0.47
EQME31932	11	LR3.02	AP	R	1	44.58	0.22	1.08	0.50
EQME31906	14	LR3.04	PS	R	4*	53.25 (2.13)	0.53	-0.44	0.42
EQME10320	15	AG1.02	KU	G	4	70.75	0.29	-0.76	0.55
EQME21264	16	AG1.03	AP	G	1	61.64	0.49	-0.13	1.17
EQME31921	17	AG2.01	PS	G	4	43.05	0.25	1.05	0.60
EQME31943	20	AG3.02	AP	G	4	64.67	0.42	-0.27	0.88
EQME10367	21	AG3.03	PS	G	1	64.38	0.27	-0.29	0.51
EQME21275	22	AG1.03	AP	G	4*	67.50 (2.70)	0.63	-0.82	0.61
EQME21230	24	MG1.01	KU	M	3	56.86	0.31	0.15	0.56
EQME31935	26	MG2.03	AP	M	1	53.28	0.22	0.51	0.41
EQME10438	27	MG2.06	PS	M	2	59.75	0.35	-0.03	0.71
EQME31960	31	MG3.02	PS	M	4*	73.50 (2.94)	0.52	-1.70	0.36
EQME31971	NR	LR3.03	KU	R	3	82.50	0.38	-1.22	0.88
EQME21274	NR	NA2.07	AP	N	2	79.65	0.24	-1.68	0.43
EQME31948	NR	MG3.01	KU	M	1	66.85	0.31	-0.45	0.59
EQME31912	NR	AG2.03	AP	G	2	52.05	0.28	0.47	0.55
EQME10392	NR	NA2.07	PS	N	4	39.32	0.41	0.85	1.09
EQME10361	NR	MG1.05	PS	M	1	62.07	0.26	-0.20	0.45
EQME10325	NR	MG3.01	AP	M	4	67.61	0.37	-0.45	0.72
EQME31995	NR	AG3.01	KU	G	1	45.78	0.41	0.56	1.06
EQME31964	NR	MG2.06	AP	M	4*	87.25 (3.49)	0.42	-2.15	0.39
EQME31894	NR	LR2.05	AP	R	4*	60.75 (2.43)	0.55	-1.20	0.46
EQME31988	NR	AG3.04	PS	G	4*	53.75 (2.15)	0.53	-0.29	0.43
EQME31957	NR	NA2.08	AP	N	4*	54.00 (2.16)	0.53	-0.41	0.36

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. KU = knowledge and understanding; AP = application; PS = thinking; G = analytic geometry; M = measurement and geometry; N = number sense and algebra; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.54 Distribution of Score Points and Step Parameters for Open-Response Items: Grade 9 Academic Mathematics, Winter (English)

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
EQME21275	22	% of Students	6.88	0.86	18.82	13.05	16.67	43.72
		Parameters			-1.94	-0.29	-0.44	-0.61
EQME31894	NR	% of Students	1.00	0.34	28.73	20.37	24.24	25.32
		Parameters			-4.91	-0.07	-0.25	0.43
EQME31906	14	% of Students	5.86	1.44	21.79	40.88	10.59	19.44
		Parameters			-2.34	-1.23	2.01	-0.22
EQME31957	NR	% of Students	10.57	2.08	28.31	16.48	15.67	26.88
		Parameters			-2.00	0.55	0.16	-0.36
EQME31960	31	% of Students	2.01	0.55	20.90	13.02	7.16	56.37
		Parameters			-4.44	0.14	0.73	-3.21
EQME31964	NR	% of Students	1.27	0.34	6.08	6.22	14.09	72.00
		Parameters			-3.34	-1.00	-1.77	-2.50
EQME31988	NR	% of Students	7.01	0.91	20.94	33.00	24.65	13.50
		Parameters			-2.17	-1.00	0.52	1.47

Note. The total number of students is 43 408. NR = not released.

Table 7.1.55 Item Statistics: Grade 9 Academic Mathematics, Spring (English)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
EQME10421	1	NA1.01	KU	N	2	62.82	0.31	-0.11	0.62
EQME31982	2	NA2.03	PS	N	4	49.48	0.42	0.48	0.99
10279	3	NA2.04	AP	N	2	51.38	0.31	0.54	0.61
EQME10346	4	NA2.05	KU	N	4	79.47	0.31	-1.10	0.64
EQME31980	6	NA2.02	AP	N	4*	68.00 (2.72)	0.52	-1.25	0.36
EQME21251	7	LR1.01	PS	R	3	46.59	0.31	0.85	0.66
EQME10413	8	LR2.01	AP	R	4	86.77	0.25	-2.10	0.50
EQME10384	10	LR2.05	PS	R	2	66.76	0.26	-0.31	0.47
EQME31952	12	LR3.03	KU	R	1	74.05	0.42	-0.66	0.89
EQME31939	13	LR2.04	AP	R	4*	55.50 (2.22)	0.56	-0.62	0.43
EQME10320	15	AG1.02	KU	G	4	72.29	0.30	-0.76	0.55
EQME21264	16	AG1.03	AP	G	1	63.90	0.45	-0.13	1.17
EQME31983	18	AG2.02	AP	G	1	67.50	0.38	-0.36	0.75
EQME10368	19	AG3.02	KU	G	4	50.98	0.44	0.40	1.07
EQME10367	21	AG3.03	PS	G	1	64.29	0.31	-0.29	0.51
EQME21239	23	AG3.05	PS	G	4*	77.75 (3.11)	0.52	-2.09	0.40
EQME10366	25	MG2.02	PS	M	2	64.75	0.50	-0.19	1.17
EQME31968	28	MG3.01	KU	M	2	88.07	0.41	-1.32	1.22
EQME31922	29	MG3.01	AP	M	2	69.02	0.37	-0.45	0.74
EQME31963	30	MG2.02	AP	M	4*	75.50 (3.02)	0.56	-1.60	0.46
EQME31972	NR	NA1.03	AP	N	2	69.89	0.37	-0.49	0.74
EQME10426	NR	MG2.06	PS	M	1	67.49	0.35	-0.40	0.66
EQME31941	NR	LR2.04	KU	R	3	63.04	0.40	-0.12	0.82
EQME10334	NR	AG2.01	PS	G	2	44.51	0.38	0.74	1.00
EQME31883	NR	AG3.02	AP	G	1	55.73	0.41	0.21	0.96
EQME31973	NR	LR3.02	AP	R	3	71.11	0.29	-0.74	0.50
10267	NR	MG1.02	KU	M	4	83.22	0.31	-1.44	0.65
EQME31958	NR	MG2.02	AP	M	1	59.35	0.40	0.05	0.83
EQME31927	NR	LR3.04	PS	R	4*	50.75 (2.03)	0.57	-0.41	0.49
EQME31917	NR	MG3.01	PS	M	4*	77.25 (3.09)	0.50	-1.44	0.38
EQME31946	NR	AG1.01	AP	G	4*	59.25 (2.37)	0.45	-1.23	0.37

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. KU = knowledge and understanding; AP = application; PS = thinking; G = analytic geometry; M = measurement and geometry; N = number sense and algebra; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.56 Distribution of Score Points and Step Parameters for Open-Response Items: Grade 9 Academic Mathematics, Spring (English)

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
EQME21239	23	% of Students	0.63	0.13	12.74	17.87	12.15	56.49
		Parameters			-5.35	-1.22	0.30	-2.08
EQME31917	NR	% of Students	4.25	0.30	8.42	16.90	13.36	56.77
		Parameters			-2.03	-1.76	0.11	-2.07
EQME31927	NR	% of Students	3.02	0.71	38.25	24.37	18.61	15.03
		Parameters			-3.59	0.30	0.61	1.05
EQME31939	13	% of Students	3.61	0.80	34.97	16.80	22.09	21.73
		Parameters			-3.60	0.69	-0.21	0.65
EQME31946	NR	% of Students	0.74	0.10	22.05	28.51	35.97	12.62
		Parameters			-6.08	-0.80	-0.28	2.24
EQME31963	30	% of Students	1.29	0.23	20.58	7.22	15.58	55.10
		Parameters			-4.43	0.63	-1.22	-1.36
EQME31980	6	% of Students	3.30	0.83	19.80	19.32	13.16	43.59
		Parameters			-3.47	-0.44	0.56	-1.63

Note. The total number of students is 55 535. NR = not released.

Table 7.1.57 Item Statistics: Grade 9 Applied Mathematics, Winter (French)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQMF21302	1	PNA1	CC	N	3	46.83	0.19	1.15	0.38
OQMF21306	2	PNA1	MA	N	1	47.66	0.33	0.57	0.83
OQMF21320	4	PNA2	MA	N	1	44.63	0.41	0.68	0.93
OQMF21312	5	PNA2	CC	N	4	55.65	0.34	0.64	0.60
OQMF10281	6	PNA3	HP	N	2	49.59	0.19	0.79	0.43
OQMF21308	8	PNA1	MA	N	4*	66.75 (2.67)	0.44	-1.41	0.26
OQMF10295	10	PRA1	CC	R	3	70.25	0.34	-0.74	0.67
OQMF21340	11	PRA1	MA	R	3	74.93	0.30	-1.20	0.61
OQMF31775	13	PRA2	CC	R	1	63.36	0.21	-0.27	0.38
OQMF21370	15	PRA3	CC	R	2	62.53	0.28	-0.16	0.55
OQMF10312	16	PRA3	MA	R	2	68.32	0.42	-0.41	0.92
OQMF21374	17	PRA3	HP	R	4	46.28	0.35	0.52	0.60
OQMF10313	19	PRA3	HP	R	4*	59.50 (2.38)	0.26	-1.16	0.13
OQMF21403	21	PMA2	CC	M	2	60.33	0.32	-0.01	0.58
OQMF21406	22	PMA2	HP	M	3	36.36	0.17	2.20	0.36
OQMF21397	23	PMA2	MA	M	3	63.64	0.35	-0.18	0.58
OQMF21409	24	PMA3	MA	M	4	68.60	0.34	-0.41	0.69
OQMF31830	25	PMA3	HP	M	2	45.73	0.31	0.78	0.74
OQMF31844	26	PMA4	CC	M	3	50.14	0.35	0.75	0.83
OQMF21430	28	PMA4	HP	M	2	64.46	0.28	-0.34	0.47
OQMF21420	30	PMA3	HP	M	4*	59.50 (2.38)	0.50	-0.97	0.37
13475	NR	PRA2	MA	R	4	33.88	0.38	1.15	1.13
OQMF31759	NR	PNA3	MA	N	4	44.63	0.28	0.93	0.61
OQMF21427	NR	PMA4	MA	M	2	76.31	0.24	-1.29	0.45
OQMF21346	NR	PRA1	HP	R	2	57.02	0.31	0.18	0.50
OQMF21386	NR	PMA1	HP	M	2	34.99	0.27	1.27	1.11
OQMF21318	NR	PNA2	CC	N	3	47.93	0.25	0.82	0.57
OQMF31838	NR	PMA1	MA	M	4*	54.25 (2.17)	0.59	-0.38	0.45
OQMF21363	NR	PRA2	MA	R	4*	40.50 (1.62)	0.44	0.56	0.26
OQMF31765	NR	PNA3	HP	N	4*	59.75 (2.39)	0.51	-0.97	0.38
OQMF31793	NR	PRA1	MA	R	4*	47.25 (1.89)	0.44	-0.05	0.41

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. CC = knowledge and understanding; MA = application; HP = thinking; N = number sense and algebra; M = measurement and geometry; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.58 Distribution of Score Points and Step Parameters for Open-Response Items: Grade 9 Applied Mathematics, Winter (French)

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
OQMF10313	19	% of Students	8.26	2.20	16.25	25.62	19.83	27.82
		Parameters			-2.35	-2.20	1.15	-1.25
OQMF21308	8	% of Students	3.31	1.93	21.21	15.98	16.80	40.77
		Parameters			-3.94	0.21	-0.19	-1.72
OQMF21363	NR	% of Students	19.83	3.31	30.03	19.83	15.98	11.02
		Parameters			-1.00	0.89	0.81	1.52
OQMF21420	30	% of Students	3.31	0.83	30.58	19.56	14.88	30.85
		Parameters			-3.99	0.31	0.48	-0.66
OQMF31765	NR	% of Students	2.75	0.83	25.62	27.00	16.25	27.55
		Parameters			-3.92	-0.50	0.83	-0.31
OQMF31793	NR	% of Students	5.23	1.10	20.94	58.68	5.51	8.54
		Parameters			-2.46	-1.69	3.70	0.26
OQMF31838	NR	% of Students	6.06	1.93	23.97	31.13	17.36	19.56
		Parameters			-2.28	-0.69	0.93	0.52

Note. The total number of students is 363. NR = not released.

Table 7.1.59 Item Statistics: Grade 9 Applied Mathematics, Spring (French)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQMF21306	2	PNA1	MA	N	1	48.84	0.35	0.57	0.83
OQRE13414	3	PNA2	CC	N	1	48.51	0.26	0.74	0.56
OQMF21312	5	PNA2	CC	N	4	46.85	0.29	0.64	0.60
OQMF10281	6	PNA3	HP	N	2	49.83	0.25	0.79	0.43
OQMF21319	7	PNA3	MA	N	2	29.17	0.18	1.99	0.78
OQMF21334	9	PNA3	HP	N	4*	71.75 (2.87)	0.35	-1.74	0.23
OQMF10295	10	PRA1	CC	R	3	74.03	0.30	-0.74	0.67
OQMF21340	11	PRA1	MA	R	3	80.55	0.28	-1.20	0.61
OQMF21040	12	PRA1	HP	R	2	53.37	0.25	0.43	0.58
OQMF21357	14	PRA2	MA	R	2	41.77	0.32	0.93	0.82
OQMF21370	15	PRA3	CC	R	2	62.32	0.30	-0.16	0.55
OQMF21374	17	PRA3	HP	R	4	52.82	0.29	0.52	0.60
OQMF21364	18	PRA2	MA	R	4*	48.00 (1.92)	0.37	-0.16	0.28
OQMF31842	20	PMA1	HP	M	3	47.51	0.26	0.82	0.55
OQMF21403	21	PMA2	CC	M	2	59.78	0.32	-0.01	0.58
OQMF21406	22	PMA2	HP	M	3	38.34	0.16	2.20	0.36
OQMF21409	24	PMA3	MA	M	4	66.74	0.34	-0.41	0.69
OQMF31844	26	PMA4	CC	M	3	42.65	0.36	0.75	0.83
13486	27	PMA4	MA	M	3	66.08	0.13	-0.90	0.21
OQMF21392	29	PMA1	MA	M	4*	42.50 (1.70)	0.40	-0.08	0.22
OQMF21371	NR	PRA3	MA	R	3	55.14	0.42	0.20	1.18
OQMF21433	NR	PMA4	HP	M	1	51.05	0.30	0.51	0.64
OQMF31778	NR	PRA2	CC	R	4	47.18	0.38	0.57	1.07
OQMF21416	NR	PMA3	HP	M	3	45.41	0.25	1.01	0.51
OQMF31746	NR	PNA2	MA	N	1	65.52	0.30	-0.32	0.63
OQMF21301	NR	PNA1	CC	N	2	44.86	0.24	1.08	0.49
OQMF21379	NR	PRA3	HP	R	4*	48.50 (1.94)	0.52	-0.15	0.37
OQMF31841	NR	PMA3	HP	M	4*	63.50 (2.54)	0.43	-1.32	0.28
OQMF21311	NR	PNA1	MA	N	4*	55.75 (2.23)	0.45	-0.58	0.28
OQMF21348	NR	PRA1	MA	R	4*	53.75 (2.15)	0.48	-0.33	0.41

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. CC = knowledge and understanding; MA = application; HP = thinking; N = number sense and algebra; M = measurement and geometry; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

Table 7.1.60 Distribution of Score Points and Step Parameters for Open-Response Items: Grade 9 Applied Mathematics, Spring (French)

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
OQMF21311	NR	% of Students	7.40	4.31	23.87	24.53	9.94	29.94
		Parameters			-2.10	-0.32	1.99	-1.89
OQMF21334	9	% of Students	2.65	1.33	9.72	21.22	25.75	39.34
		Parameters			-3.02	-2.43	-0.62	-0.88
OQMF21348	NR	% of Students	5.41	1.55	15.47	47.96	15.25	14.36
		Parameters			-1.93	-1.93	1.79	0.76
OQMF21364	18	% of Students	5.30	1.66	28.29	40.44	14.14	10.17
		Parameters			-3.47	-0.91	2.41	1.31
OQMF21379	NR	% of Students	11.05	3.20	33.59	17.90	12.27	21.99
		Parameters			-1.94	0.81	0.81	-0.28
OQMF21392	29	% of Students	14.70	3.54	47.51	3.20	7.73	23.31
		Parameters			-2.95	7.26	-2.17	-2.47
OQMF31841	NR	% of Students	4.09	0.66	27.07	17.90	9.50	40.77
		Parameters			-4.37	0.50	1.30	-2.74

Note. The total number of students is 905. NR = not released.

Table 7.1.61 Item Statistics: Grade 9 Academic Mathematics, Winter (French)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQMF20976	1	DNA1	MA	N	4	71.77	0.37	-0.68	0.70
OQMF20941	3	DNA3	CC	N	4	71.77	0.40	-0.62	0.86
OQMF10394	4	DNA3	MA	N	1	58.09	0.43	-0.03	0.75
OQMF20951	5	DNA4	HP	N	2	50.82	0.25	0.63	0.54
OQMF20912	8	DNA4	MA	N	4*	65.50 (2.62)	0.51	-1.11	0.31
13433	9	DRA1	HP	R	1	64.70	0.39	-0.24	0.76
OQMF31906	10	DRA1	MA	R	3	55.34	0.33	0.24	0.68
13429	11	DRA2	CC	R	2	71.12	0.35	-0.69	0.66
OQMF20968	12	DRA2	MA	R	2	52.65	0.34	0.38	0.66
OQMF20967	14	DRA3	HP	R	3	53.05	0.26	0.36	0.47
OQMF31921	15	DRA1	MA	R	4*	51.50 (2.06)	0.51	-0.67	0.35
OQMF10332	17	DGA1	MA	G	3	61.76	0.38	-0.22	0.76
OQMF10337	18	DGA1	CC	G	2	50.56	0.26	0.75	0.44
OQMF10338	19	DGA1	HP	G	1	37.79	0.37	1.04	0.90
OQMF31872	21	DGA2	HP	G	4	46.37	0.27	0.83	0.62
OQMF31881	22	DGA1	MA	G	4*	57.75 (2.31)	0.54	-0.60	0.47
OQMF21001	23	DMA1	HP	M	2	48.00	0.33	0.51	0.65
OQMF31873	24	DMA2	MA	M	3	48.98	0.27	0.71	0.53
OQMF10351	25	DMA2	CC	M	2	61.49	0.35	-0.06	0.72
OQMF10373	28	DMA4	CC	M	2	72.69	0.30	-0.86	0.62
OQMF21026	29	DMA4	MA	M	3	46.56	0.42	0.56	0.98
OQMF20926	30	DMA3	HP	M	4*	67.25 (2.69)	0.56	-0.93	0.41
OQMF31905	0	DNA4	MA	N	4	54.55	0.32	0.29	0.60
OQMF21018	0	DMA3	CC	M	1	58.81	0.39	0.02	0.74
OQMF31944	0	DNA2	CC	N	1	62.21	0.31	-0.17	0.57
OQMF31949	0	DRA3	MA	R	2	78.00	0.23	-1.52	0.42
OQMF31911	0	DGA2	CC	G	3	52.91	0.33	0.38	0.66
OQMF31957	0	DMA3	HP	M	3	46.95	0.32	0.70	0.69
OQMF20904	0	DNA3	MA	N	4*	56.75 (2.27)	0.52	-0.80	0.47
OQMF31940	0	DRA3	HP	R	4*	67.75 (2.71)	0.54	-0.99	0.36
OQMF31902	0	DMA4	HP	M	4*	64.25 (2.57)	0.54	-0.78	0.46

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. CC = knowledge and understanding; MA = application; HP = thinking; G = analytic geometry; M = measurement and geometry; N = number sense and algebra; R = linear relations; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

**Table 7.1.62 Distribution of Score Points and Step Parameters for Open-Response Items:
Grade 9 Academic Mathematics, Winter (French)**

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
OQMF20904	NR	% of Students	1.11	0.65	22.46	36.15	26.39	13.23
		Parameters			-4.18	-1.06	0.49	1.55
OQMF20912	8	% of Students	4.39	3.01	27.11	4.65	17.35	43.48
		Parameters			-3.27	2.88	-2.60	-1.45
OQMF20926	30	% of Students	6.42	2.10	13.10	23.71	10.74	43.94
		Parameters			-1.57	-1.42	0.99	-1.71
OQMF31881	22	% of Students	3.67	1.24	21.22	33.01	20.04	20.83
		Parameters			-2.86	-1.01	0.70	0.76
OQMF31902	NR	% of Students	4.32	0.59	18.27	12.18	43.68	20.96
		Parameters			-2.71	-0.08	-1.73	1.39
OQMF31921	15	% of Students	3.41	1.90	37.72	24.49	10.15	22.33
		Parameters			-4.03	0.46	1.66	-0.75
OQMF31940	NR	% of Students	8.06	1.64	16.44	16.70	7.33	49.84
		Parameters			-1.75	-0.56	1.19	-2.84

Note. The total number of students is 1527. NR = not released.

Table 7.1.63 Item Statistics: Grade 9 Academic Mathematics, Spring (French)

Item Code	Sequence	Expectation	Cognitive Skill	Strand	Answer Key/ Max. Score	CTT Item Statistics		IRT Item Parameters	
						Difficulty	Item-Total Correlation	Location	Slope
OQMF20976	1	DNA1	MA	N	4	72.42	0.35	0.70	-0.68
OQMF20936	2	DNA2	CC	N	3	53.12	0.23	0.42	0.55
OQMF10394	4	DNA3	MA	N	1	61.81	0.34	0.75	-0.03
OQMF20951	5	DNA4	HP	N	2	50.52	0.27	0.54	0.63
OQRE13442	6	DNA4	MA	N	1	66.73	0.32	0.59	-0.40
OQMF20905	7	DNA2	MA	N	4*	41.75 (1.67)	0.39	0.28	-0.35
13433	9	DRA1	HP	R	1	64.49	0.38	0.76	-0.24
13429	11	DRA2	CC	R	2	72.62	0.34	0.66	-0.69
OQMF31888	13	DRA3	MA	R	2	77.98	0.33	0.69	-1.02
OQMF20967	14	DRA3	HP	R	3	56.20	0.25	0.47	0.36
OQMF20916	16	DRA3	HP	R	4*	73.00 (2.92)	0.45	0.32	-1.45
OQMF10332	17	DGA1	MA	G	3	65.69	0.37	0.76	-0.22
OQMF10337	18	DGA1	CC	G	2	49.76	0.23	0.44	0.75
OQMF10338	19	DGA1	HP	G	1	40.11	0.30	0.90	1.04
OQMF20994	20	DGA2	CC	G	4	71.98	0.26	0.46	-0.88
OQMF31881	22	DGA1	MA	G	4*	57.00 (2.28)	0.55	0.47	-0.60
OQMF21001	23	DMA1	HP	M	2	53.00	0.29	0.65	0.51
OQMF10351	25	DMA2	CC	M	2	60.89	0.37	0.72	-0.06
OQMF31896	26	DMA3	HP	M	3	46.32	0.27	0.60	0.86
OQMF21019	27	DMA3	CC	M	1	75.74	0.23	0.41	-1.30
OQMF10373	28	DMA4	CC	M	2	75.06	0.34	0.62	-0.86
OQMF10375	31	DMA4	HP	M	4*	56.75 (2.27)	0.52	0.33	-0.57
OQMF10395	NR	DNA3	CC	N	3	83.59	0.24	0.48	-1.87
OQMF31868	NR	DRA2	MA	R	1	66.09	0.35	0.68	-0.32
OQMF21011	NR	DMA2	MA	M	4	69.94	0.33	0.60	-0.59
OQMF31886	NR	DRA1	MA	R	4	90.71	0.35	1.10	-1.62
OQMF31932	NR	DGA2	HP	G	4	30.62	0.30	1.03	1.48
OQMF31917	NR	DMA4	MA	M	1	38.55	0.26	0.65	1.33
OQMF31898	NR	DNA4	MA	N	4*	35.75 (1.43)	0.44	0.35	0.46
OQMF31874	NR	DMA2	HP	M	4*	69.50 (2.78)	0.51	0.40	-1.52
OQMF31899	NR	DRA1	MA	R	4*	62.50 (2.50)	0.61	0.55	-0.74

Note. The guessing parameter was set at a constant of 0.2 for multiple-choice items. CC = knowledge and understanding; MA = application; HP = thinking; G = analytic geometry; M = measurement and geometry; N = number sense and algebra; R = linear relations; NR = not released; () = Mean score for OR items.

*Maximum score code for open-response (OR) items.

**Table 7.1.64 Distribution of Score Points and Step Parameters for Open-Response Items:
Grade 9 Academic Mathematics, Spring (French)**

Item Code	Sequence		Score Points					
			Missing	Illegible	10	20	30	40
OQMF10375	31	% of Students	6.20	5.88	30.10	12.57	9.49	35.75
		Parameters			-2.28	1.29	0.61	-1.90
OQMF20905	7	% of Students	3.44	2.28	55.00	20.62	4.12	14.53
		Parameters			-5.21	2.03	3.74	-1.96
OQMF20916	16	% of Students	3.08	1.24	12.29	16.45	20.70	46.24
		Parameters			-2.87	-1.09	-0.59	-1.24
OQMF31874	NR	% of Students	1.24	0.28	16.89	21.66	22.22	37.71
		Parameters			-4.54	-0.94	-0.16	-0.42
OQMF31881	22	% of Students	3.28	0.96	21.14	34.83	21.86	17.93
		Parameters			-2.86	-1.01	0.70	0.76
OQMF31898	NR	% of Students	7.85	5.76	51.04	23.06	3.64	8.65
		Parameters			-2.62	1.39	3.60	-0.52
OQMF31899	NR	% of Students	3.64	1.32	21.14	24.46	17.45	31.99
		Parameters			-2.55	-0.65	0.36	-0.12

Note. The total number of students is 2498. NR = not released.

Differential Item Functioning (DIF) Analysis Results

The gender- and SLL-based DIF results for the Grade 9 Assessment of Mathematics are provided in Tables 7.1.65–7.1.76. Only for the English-language test are the results from two random samples of 2000 examinees. For the French-language test, gender-based DIF analysis was conducted for one sample of students, and SLL-based DIF analysis was not conducted. In both cases, DIF analysis was limited by the relatively small population of students who wrote the French-language test. Each table indicates the value of $\hat{\beta}$, the significance level and the category of the effect size for each item for B- and C-level DIF items. For gender-based DIF, negative estimates of $\hat{\beta}$ indicate that the girls outperformed the boys; positive $\hat{\beta}$ estimates indicate that the boys outperformed the girls. For SLL-based DIF, negative estimates of $\hat{\beta}$ indicate that the SLLs outperformed the non-SLLs; positive $\hat{\beta}$ estimates indicate that the non-SLLs outperformed the SLLs.

Table 7.1.65 Gender-Based DIF Statistics: Grade 9 Applied Mathematics, Winter (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10515	1	-0.015	0.306		-0.018	0.218	
EQME10662	2	-0.009	0.543		0.001	0.971	
EQME10617	3	-0.013	0.395		-0.015	0.314	
EQME21452	4	0.077	0.000	B+	0.060	0.000	B+
EQME31771	6	-0.033	0.031		-0.031	0.044	
EQME31812	7	-0.013	0.396		0.013	0.405	
EQME31763	9	-0.138	0.002	C-	-0.166	0.000	C-
EQME21453	10	0.056	0.000		0.056	0.000	
EQME10512	12	0.040	0.011		0.019	0.225	
EQME31816	14	0.151	0.000	C+	0.122	0.000	C+
EQME10549	15	0.061	0.000	B+	0.030	0.053	
EQME10520	16	-0.001	0.940		0.008	0.632	
3434098	18	0.043	0.003		-0.012	0.405	
EQME31839	19	-0.025	0.109		-0.008	0.588	
EQME31826	23	-0.216	0.000	C-	-0.162	0.000	C-
EQME10505	24	-0.016	0.280		-0.028	0.063	
EQME10538	25	-0.044	0.004		-0.039	0.010	
EQME10579	26	-0.068	0.000	B-	-0.063	0.000	B-
EQME10581	27	-0.048	0.002		-0.034	0.027	
EQME10522	28	-0.008	0.468		-0.018	0.105	
EQME21454	30	0.090	0.005	C+	0.102	0.002	C+
EQME31774	NR	0.023	0.052		0.044	0.000	
EQME31784	NR	0.134	0.000	C+	0.150	0.000	C+
EQME31847	NR	0.066	0.031	B+	0.096	0.002	C+
EQME31869	NR	0.010	0.796		-0.108	0.006	C-
EQME31858	NR	0.045	0.002		0.045	0.002	
EQME10513	NR	-0.032	0.043		-0.047	0.003	
EQME31761	NR	0.008	0.611		0.020	0.192	
EQME31859	NR	-0.012	0.236		-0.020	0.065	
EQME31750	NR	-0.086	0.000	B-	-0.061	0.000	B-
EQME31802	NR	0.083	0.011	B+	0.104	0.001	C+

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.66 Gender-Based DIF Statistics: Grade 9 Applied Mathematics, Spring (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10515	1	-0.021	0.139		-0.006	0.674	
EQME10662	2	-0.024	0.118		-0.014	0.374	
EQME10617	3	0.003	0.821		-0.014	0.356	
EQME21452	4	0.066	0.000	B+	0.063	0.000	B+
EQME31790	5	-0.014	0.365		0.022	0.162	
EQME31842	8	0.083	0.020	B+	0.014	0.704	
EQME21453	10	0.030	0.036		0.028	0.046	
EQME10514	11	-0.018	0.135		-0.036	0.001	
EQME31815	13	-0.020	0.139		0.003	0.842	
EQME10549	15	-0.012	0.453		0.038	0.015	
EQME10520	16	0.019	0.222		0.022	0.173	
EQME10501	17	-0.011	0.178		0.001	0.938	
3434098	18	0.000	0.975		-0.010	0.471	
EQME31819	20	0.029	0.040		0.033	0.021	
EQME31865	21	0.065	0.032	B+	0.077	0.013	B+
EQME21441	22	0.060	0.125		0.045	0.249	
EQME10505	24	-0.010	0.511		-0.027	0.059	
EQME10538	25	-0.046	0.002		-0.067	0.000	B-
EQME10579	26	-0.064	0.000	B-	-0.072	0.000	B-
EQME10522	28	-0.001	0.938		-0.010	0.364	
EQME31801	29	0.060	0.000	B+	0.027	0.074	
EQME31769	31	-0.047	0.191		-0.019	0.602	
EQME31791	NR	-0.060	0.000	B-	-0.051	0.000	
EQME10596	NR	0.030	0.047		0.015	0.329	
EQME31830	NR	-0.025	0.101		-0.046	0.003	
EQME31863	NR	-0.107	0.002	C-	-0.055	0.117	
EQME31786	NR	0.152	0.000	C+	0.141	0.000	C+
EQME10659	NR	-0.009	0.412		0.004	0.751	
EQME31778	NR	0.093	0.000	C+	0.072	0.000	B+
EQME10532	NR	-0.022	0.145		-0.027	0.075	
EQME31828	NR	-0.194	0.000	C-	-0.199	0.000	C-

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.67 Gender-Based DIF Statistics: Grade 9 Academic Mathematics, Winter (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10421	1	-0.018	0.245		-0.006	0.690	
EQME10346	4	-0.062	0.000	B-	-0.037	0.005	
EQME31969	5	-0.013	0.378		-0.017	0.266	
EQME21251	7	0.004	0.788		0.028	0.072	
EQME10413	8	-0.008	0.464		-0.011	0.294	
EQME31970	9	0.007	0.582		0.017	0.186	
EQME10384	10	0.043	0.004		0.074	0.000	B+
EQME31932	11	0.050	0.001		0.034	0.028	
EQME31906	14	0.236	0.000	C+	0.216	0.000	C+
EQME10320	15	-0.044	0.002		-0.069	0.000	B-
EQME21264	16	-0.076	0.000	B-	-0.058	0.000	
EQME31921	17	0.036	0.020		0.056	0.000	
EQME31943	20	-0.053	0.000		-0.061	0.000	B-
EQME10367	21	-0.004	0.794		-0.027	0.073	
EQME21275	22	-0.087	0.013	B-	-0.045	0.195	
EQME21230	24	0.027	0.079		0.030	0.048	
EQME31935	26	-0.028	0.078		-0.001	0.958	
EQME10438	27	-0.020	0.180		-0.062	0.000	B-
EQME31960	31	0.087	0.020	B+	0.097	0.010	C+
EQME31971	NR	0.003	0.790		-0.009	0.424	
EQME21274	NR	0.014	0.273		0.010	0.432	
EQME31948	NR	-0.027	0.069		-0.008	0.576	
EQME31912	NR	0.072	0.000	B+	0.072	0.000	B+
EQME31964	NR	0.015	0.609		0.005	0.856	
EQME31894	NR	-0.116	0.000	C-	-0.028	0.372	
EQME31988	NR	0.095	0.002	C+	0.050	0.103	
EQME10392	NR	-0.004	0.806		0.018	0.218	
EQME10361	NR	0.021	0.162		0.036	0.016	
EQME10325	NR	0.025	0.080		0.017	0.239	
EQME31995	NR	-0.030	0.043		0.001	0.934	
EQME31957	NR	-0.236	0.000	C-	-0.262	0.000	C-

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.68 Gender-Based DIF Statistics: Grade 9 Academic Mathematics, Spring (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10421	1	-0.007	0.624		0.016	0.283	
EQME31982	2	0.018	0.225		0.032	0.032	
10279	3	0.009	0.563		0.068	0.000	B+
EQME10346	4	-0.016	0.182		-0.036	0.004	
EQME31980	6	0.186	0.000	C+	0.154	0.000	C+
EQME21251	7	0.024	0.135		0.046	0.003	
EQME10413	8	-0.011	0.321		0.000	0.970	
EQME10384	10	0.052	0.000		0.059	0.000	B+
EQME31952	12	-0.002	0.855		-0.018	0.175	
EQME31939	13	0.018	0.601		0.019	0.575	
EQME10320	15	-0.055	0.000		-0.090	0.000	C-
EQME21264	16	-0.106	0.000	C-	-0.084	0.000	B-
EQME31983	18	0.002	0.879		0.008	0.603	
EQME10368	19	0.037	0.012		0.009	0.522	
EQME10367	21	-0.031	0.041		-0.045	0.003	
EQME21239	23	0.150	0.000	C+	0.163	0.000	C+
EQME10366	25	0.020	0.142		0.016	0.237	
EQME31968	28	0.033	0.001		0.013	0.188	
EQME31922	29	-0.038	0.009		-0.044	0.002	
EQME31963	30	-0.026	0.458		0.017	0.633	
EQME31972	NR	-0.067	0.000	B-	-0.082	0.000	B-
EQME31927	NR	-0.058	0.059		-0.105	0.001	C-
EQME31917	NR	-0.079	0.018	B-	-0.101	0.003	C-
EQME10426	NR	-0.021	0.155		0.003	0.862	
EQME31941	NR	0.037	0.012		0.051	0.000	
EQME10334	NR	0.067	0.000	B+	0.073	0.000	B+
EQME31883	NR	-0.076	0.000	B-	-0.083	0.000	B-
EQME31973	NR	0.071	0.000	B+	0.075	0.000	B+
10267	NR	0.012	0.330		-0.015	0.203	
EQME31958	NR	-0.019	0.201		0.001	0.957	
EQME31946	NR	-0.189	0.000	C-	-0.191	0.000	C-

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.69 Gender-Based DIF Statistics: Grade 9 Applied Mathematics, Winter (French)

Item Code	Sequence	Sample 1		
		$\hat{\beta}$	p-value	Level of DIF
OQMF21302	1	-0.082	0.094	
OQMF21306	2	0.072	0.135	
OQMF21320	4	-0.074	0.118	
OQMF21312	5	-0.120	0.013	C-
OQMF10281	6	0.038	0.480	
OQMF21308	8	-0.245	0.026	C-
OQMF10295	10	-0.022	0.645	
OQMF21340	11	0.189	0.000	C+
OQMF31775	13	-0.052	0.291	
OQMF21370	15	-0.034	0.481	
OQMF10312	16	0.067	0.140	
OQMF21374	17	0.104	0.030	C+
OQMF10313	19	0.098	0.417	
OQMF21403	21	0.022	0.639	
OQMF21406	22	0.001	0.977	
OQMF 21397	23	-0.013	0.785	
OQMF21409	24	-0.076	0.098	
OQMF31830	25	0.042	0.403	
OQMF31844	26	-0.007	0.885	
OQMF21430	28	0.044	0.353	
OQMF21420	30	0.151	0.201	
13475	NR	0.015	0.750	
OQMF31759	NR	0.071	0.152	
OQMF21427	NR	-0.004	0.923	
OQMF31838	NR	-0.257	0.009	C-
OQMF21363	NR	-0.067	0.572	
OQMF21346	NR	-0.078	0.096	
OQMF21386	NR	-0.058	0.248	
OQMF21318	NR	-0.169	0.000	C-
OQMF31765	NR	0.171	0.106	
OQMF31793	NR	0.035	0.678	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.70 Gender-Based DIF Statistics: Grade 9 Applied Mathematics, Spring (French)

Item Code	Sequence	Sample 1		
		$\hat{\beta}$	p-value	Level of DIF
OQMF21306	2	0.051	0.100	
OQRE13414	3	-0.053	0.095	
OQMF21312	5	-0.120	0.000	C-
OQMF10281	6	0.078	0.014	B+
OQMF21319	7	-0.065	0.028	B-
OQMF21334	9	0.035	0.620	
OQMF10295	10	0.050	0.065	
OQMF21340	11	0.059	0.019	B+
OQMF21040	12	0.006	0.846	
OQMF21357	14	0.028	0.354	
OQMF21370	15	-0.029	0.344	
OQMF21374	17	0.145	0.000	C+
OQMF21364	18	-0.064	0.298	
OQMF31842	20	0.049	0.121	
OQMF21403	21	-0.057	0.059	
OQMF21406	22	-0.031	0.329	
OQMF 21397	23	-0.027	0.374	
OQMF21409	24	-0.032	0.267	
OQMF31844	26	0.053	0.078	
13486	27	0.001	0.980	
OQMF21392	29	0.076	0.349	
OQMF21371	NR	-0.087	0.004	B-
OQMF21433	NR	0.003	0.925	
OQMF31778	NR	-0.017	0.568	
OQMF21416	NR	-0.051	0.106	
OQMF21379	NR	-0.025	0.741	
OQMF31746	NR	-0.011	0.698	
OQMF21301	NR	0.079	0.012	B+
OQMF31841	NR	-0.041	0.598	
OQMF21311	NR	0.076	0.325	
OQMF21348	NR	-0.091	0.123	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.71 Gender-Based DIF Statistics: Grade 9 Academic Mathematics, Winter (French)

Item Code	Sequence	Sample 1		
		$\hat{\beta}$	p-value	Level of DIF
OQMF20976	1	0.045	0.037	
OQMF20941	3	-0.021	0.331	
OQMF10394	4	-0.057	0.017	
OQMF20951	5	0.060	0.018	B+
OQMF20912	8	0.264	0.000	C+
13433	9	-0.020	0.395	
OQMF31906	10	-0.040	0.104	
13429	11	0.002	0.924	
OQMF20968	12	0.095	0.000	C+
OQMF20967	14	0.013	0.598	
OQMF31921	15	0.158	0.004	C+
OQMF10332	17	-0.024	0.307	
OQMF10337	18	0.033	0.196	
OQMF10338	19	-0.044	0.064	
OQMF31872	21	-0.008	0.742	
OQMF31881	22	-0.176	0.001	C-
OQMF21001	23	0.013	0.610	
OQMF31873	24	0.064	0.012	B+
OQMF10351	25	0.006	0.807	
OQMF10373	28	-0.009	0.698	
OQMF21026	29	-0.013	0.587	
OQMF20926	30	-0.080	0.171	
OQMF31905	NR	-0.003	0.905	
OQMF21018	NR	0.035	0.150	
OQMF31944	NR	-0.024	0.332	
OQMF20904	NR	-0.045	0.329	
OQMF31940	NR	-0.118	0.057	
OQMF31949	NR	0.037	0.079	
OQMF31911	NR	-0.060	0.014	B-
OQMF31957	NR	0.068	0.006	B+
OQMF31902	NR	-0.026	0.616	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.72 Gender-Based DIF Statistics: Grade 9 Academic Mathematics, Spring (French)

Item Code	Sequence	Sample 1		
		$\hat{\beta}$	p-value	Level of DIF
OQMF20976	1	0.053	0.002	
OQMF20936	2	-0.043	0.034	
OQMF10394	4	-0.029	0.122	
OQMF20951	5	0.047	0.019	
OQRE13442	6	0.011	0.560	
OQMF20905	7	-0.069	0.106	
13433	9	-0.034	0.065	
13429	11	-0.021	0.215	
OQMF31888	13	-0.009	0.578	
OQMF20967	14	0.076	0.000	B+
OQMF20916	16	0.425	0.000	C+
OQMF10332	17	-0.042	0.023	
OQMF10337	18	-0.002	0.938	
OQMF10338	19	-0.040	0.036	
OQMF20994	20	-0.005	0.772	
OQMF31881	22	-0.090	0.020	C-
OQMF21001	23	0.031	0.112	
OQMF10351	25	0.006	0.738	
OQMF31896	26	0.013	0.509	
OQMF21019	27	0.023	0.170	
OQMF10373	28	0.023	0.160	
OQMF10375	31	-0.188	0.000	C-
OQMF10395	0	-0.025	0.086	
OQMF31868	0	-0.036	0.049	
OQMF21011	0	-0.026	0.137	
OQMF31898	0	-0.006	0.885	
OQMF31886	0	-0.003	0.797	
OQMF31932	0	-0.036	0.048	
OQMF31917	0	0.067	0.001	B+
OQMF31874	0	-0.127	0.002	C-
OQMF31899	0	0.020	0.632	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.73 SLL-Based DIF Statistics: Grade 9 Applied Mathematics, Winter (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10515	1	-0.046	0.018		-0.049	0.011	
EQME10662	2	-0.019	0.350		-0.001	0.958	
EQME10617	3	-0.025	0.213		-0.052	0.008	
EQME21452	4	0.028	0.156		0.054	0.006	
EQME31771	6	-0.075	0.000	B-	-0.076	0.000	B-
EQME31812	7	-0.041	0.046		0.004	0.860	
EQME31763	9	0.189	0.000	C+	0.105	0.053	
EQME21453	10	-0.006	0.752		0.031	0.115	
EQME10512	12	0.039	0.066		0.018	0.400	
EQME31816	14	0.132	0.000	C+	0.117	0.000	C+
EQME10549	15	0.000	0.996		0.011	0.599	
EQME10520	16	-0.043	0.043		-0.027	0.198	
3434098	18	0.079	0.000	B+	0.086	0.000	B+
EQME31839	19	0.083	0.000	B+	0.063	0.002	B+
EQME31826	23	0.182	0.000	C+	0.124	0.013	C+
EQME10505	24	0.049	0.015		0.031	0.125	
EQME10538	25	-0.057	0.005		-0.059	0.004	B-
EQME10579	26	-0.133	0.000	C-	-0.129	0.000	C-
EQME10581	27	-0.024	0.243		0.009	0.642	
EQME10522	28	0.032	0.040		0.032	0.042	
EQME21454	30	-0.034	0.413		-0.007	0.866	
EQME31774	NR	0.085	0.000	B+	0.098	0.000	C+
EQME31784	NR	0.087	0.067		0.159	0.001	C+
EQME31847	NR	0.031	0.428		0.068	0.085	
EQME31869	NR	-0.275	0.000	C-	-0.266	0.000	C-
EQME31858	NR	0.075	0.000	B+	0.065	0.001	B+
EQME10513	NR	0.001	0.978		-0.028	0.175	
EQME31761	NR	-0.050	0.016		-0.031	0.131	
EQME31859	NR	0.006	0.697		-0.004	0.814	
EQME31750	NR	-0.030	0.138		-0.038	0.060	
EQME31802	NR	-0.227	0.000	C-	-0.238	0.000	C-

Note. B = moderate DIF; C = large DIF; - = favouring SLL students; + = favouring non-SLL students.

Table 7.1.74 SLL-Based DIF Statistics: Grade 9 Applied Mathematics, Spring (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10515	1	-0.023	0.156		-0.027	0.099	
EQME10662	2	-0.044	0.012		-0.021	0.249	
EQME10617	3	-0.035	0.053		-0.026	0.149	
EQME21452	4	0.030	0.083		0.034	0.056	
EQME31790	5	0.007	0.699		-0.001	0.963	
EQME31842	8	-0.010	0.814		-0.019	0.650	
EQME21453	10	-0.002	0.903		0.005	0.786	
EQME10514	11	0.041	0.005		0.027	0.066	
EQME31815	13	-0.028	0.074		-0.027	0.090	
EQME10549	15	-0.014	0.433		-0.039	0.036	
EQME10520	16	-0.079	0.000	B-	-0.058	0.002	
EQME10501	17	0.033	0.004		0.053	0.000	
3434098	18	0.117	0.000	C+	0.114	0.000	C+
EQME31819	20	0.150	0.000	C+	0.131	0.000	C+
EQME31865	21	-0.080	0.031	B-	-0.079	0.031	B-
EQME21441	22	0.121	0.009	C+	0.137	0.003	C+
EQME10505	24	0.064	0.000	B+	0.054	0.002	
EQME10538	25	-0.015	0.377		-0.004	0.834	
EQME10579	26	-0.075	0.000	B-	-0.089	0.000	C-
EQME10522	28	-0.009	0.514		-0.006	0.653	
EQME31801	29	0.050	0.004		0.060	0.001	B+
EQME31769	31	-0.061	0.130		-0.140	0.000	C-
EQME31791	NR	-0.049	0.004		-0.007	0.678	
EQME10596	NR	-0.010	0.555		-0.007	0.673	
EQME31830	NR	-0.179	0.000	C-	-0.199	0.000	C-
EQME31863	NR	-0.120	0.003	C-	-0.104	0.008	C-
EQME31786	NR	0.262	0.000	C+	0.274	0.000	C+
EQME10659	NR	-0.014	0.306		-0.020	0.136	
EQME31778	NR	0.053	0.002		0.062	0.000	B+
EQME10532	NR	-0.004	0.819		-0.054	0.003	
EQME31828	NR	-0.119	0.000	C-	-0.093	0.004	C-

Note. B = moderate DIF; C = large DIF; - = favouring SLL students; + = favouring non-SLL students.

Table 7.1.75 SLL-Based DIF Statistics: Grade 9 Academic Mathematics, Winter (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10421	1	-0.073	0.000	B-	-0.094	0.000	C-
EQME10346	4	-0.122	0.000	C-	-0.133	0.000	C-
EQME31969	5	-0.009	0.586		-0.031	0.069	
EQME21251	7	0.017	0.337		0.027	0.126	
EQME10413	8	0.029	0.031		0.046	0.000	
EQME31970	9	0.103	0.000	C+	0.112	0.000	C+
EQME10384	10	0.001	0.954		0.018	0.303	
EQME31932	11	0.001	0.936		-0.006	0.738	
EQME31906	14	0.155	0.000	C+	0.166	0.000	C+
EQME10320	15	-0.072	0.000	B-	-0.038	0.014	
EQME21264	16	-0.102	0.000	C-	-0.125	0.000	C-
EQME31921	17	-0.013	0.483		-0.021	0.237	
EQME31943	20	-0.075	0.000	B-	-0.085	0.000	B-
EQME10367	21	0.155	0.000	C+	0.153	0.000	C+
EQME21275	22	-0.140	0.000	C-	-0.137	0.000	C-
EQME21230	24	0.080	0.000	B+	0.055	0.002	
EQME31935	26	0.037	0.043		0.053	0.004	
EQME10438	27	-0.021	0.202		-0.014	0.398	
EQME31960	31	-0.041	0.329		-0.007	0.876	
EQME31971	NR	0.017	0.209		0.029	0.034	
EQME21274	NR	0.015	0.306		0.003	0.836	
EQME31948	NR	-0.056	0.000		-0.080	0.000	B-
EQME31912	NR	0.087	0.000	B+	0.099	0.000	C+
EQME31964	NR	0.210	0.000	C+	0.194	0.000	C+
EQME31894	NR	0.016	0.653		-0.012	0.748	
EQME31988	NR	0.199	0.000	C+	0.179	0.000	C+
EQME10392	NR	-0.023	0.162		-0.001	0.931	
EQME10361	NR	0.120	0.000	C+	0.091	0.000	C+
EQME10325	NR	0.003	0.866		-0.003	0.840	
EQME31995	NR	-0.074	0.000	B-	-0.058	0.001	
EQME31957	NR	-0.438	0.000	C-	-0.431	0.000	C-

Note. B = moderate DIF; C = large DIF; - = favouring SLL students; + = favouring non-SLL students.

Table 7.1.76 SLL-Based DIF Statistics: Grade 9 Academic Mathematics, Spring (English)

Item Code	Sequence	Sample 1			Sample 2		
		$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQME10421	1	-0.117	0.000	C-	-0.096	0.000	C-
EQME31982	2	-0.020	0.155		-0.016	0.245	
10279	3	0.008	0.602		0.000	0.976	
EQME10346	4	-0.093	0.000	C-	-0.104	0.000	C-
EQME31980	6	0.085	0.013	B+	0.079	0.019	B+
EQME21251	7	0.009	0.533		0.015	0.312	
EQME10413	8	0.013	0.206		0.014	0.167	
EQME10384	10	0.056	0.000		0.046	0.001	
EQME31952	12	-0.005	0.692		-0.002	0.837	
EQME31939	13	-0.118	0.000	C-	-0.119	0.000	C-
EQME10320	15	-0.040	0.002		-0.022	0.077	
EQME21264	16	-0.105	0.000	C-	-0.111	0.000	C-
EQME31983	18	0.119	0.000	C+	0.107	0.000	C+
EQME10368	19	0.031	0.022		0.027	0.046	
EQME10367	21	0.128	0.000	C+	0.133	0.000	C+
EQME21239	23	0.303	0.000	C+	0.283	0.000	C+
EQME10366	25	-0.018	0.153		-0.022	0.090	
EQME31968	28	0.043	0.000		0.030	0.001	
EQME31922	29	0.005	0.681		-0.004	0.742	
EQME31963	30	-0.236	0.000	C-	-0.174	0.000	C-
EQME31972	NR	-0.060	0.000	B-	-0.083	0.000	B-
EQME31927	NR	0.180	0.000	C+	0.203	0.000	C+
EQME31917	NR	-0.248	0.000	C-	-0.289	0.000	C-
EQME10426	NR	-0.036	0.007		-0.019	0.155	
EQME31941	NR	0.056	0.000		0.062	0.000	B+
EQME10334	NR	0.081	0.000	B+	0.062	0.000	B+
EQME31883	NR	-0.087	0.000	B-	-0.076	0.000	B-
EQME31973	NR	0.096	0.000	C+	0.096	0.000	C+
10267	NR	0.051	0.000		0.048	0.000	
EQME31958	NR	-0.042	0.002		-0.047	0.001	
EQME31946	NR	0.000	0.986		-0.013	0.625	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

The Ontario Secondary School Literacy Test (OSSLT)

Classical Item Statistics and IRT Item Parameters

Table 7.1.77 Item Statistics: OSSLT (English)

Item Code	Section	Sequence	Cognitive Skill	Answer Key/Max. Score Code ¹	CTT Item Statistics		IRT Item Parameters
					Difficulty	Item-Total Correlation	Location
EQRE32141.185	I	1	R1	2	0.70	0.43	-0.56
EQRE32144.185	I	2	R2	3	0.78	0.25	-0.93
EQRE32147.185	I	3	R2	2	0.56	0.41	0.36
EQRE32143.185	I	4	R3	2	0.63	0.40	-0.05
EQRE32145.185	I	5	R1	3	0.47	0.35	0.93
EQRE32149.185	I	6	R3	3*	0.80 (2.41)	0.46	-1.50
EQWE30522	II	1	W1	3	0.91	0.27	-2.23
EQWE30531	II	2	W2	3	0.77	0.31	-0.93
EQWE30553	II	3	W3	4	0.88	0.33	-1.92
EQWE30554	II	4	W3	1	0.73	0.41	-0.72
EQWE30565T	III	1	W4	3*	0.92 (2.77)	0.27	-2.29
EQWE30565V	III	1	W3	2*	0.99 (1.97)	0.23	-3.09
1W3_066T	IV	1	W4	6*	0.69 (4.12)	0.55	-0.64
1W3_066V	IV	1	W3	4*	0.83 (3.30)	0.52	-1.60
EQRE32090.180	V	1	R1	2	0.90	0.26	-2.03
EQRE32095.180	V	2	R1	4	0.54	0.33	0.55
EQRE32098.180	V	3	R2	2	0.57	0.32	0.35
EQRE32091.180	V	4	R3	1	0.67	0.29	-0.22
EQRE32093.180	V	5	R2	2	0.79	0.32	-0.95
EQRE32097.180	V	6	R2	4	0.87	0.33	-1.80
EQRE32099.180	V	7	R2	3*	0.86 (2.58)	0.36	-1.31
1W2_059T	VI	1	W4	6*	0.76 (4.53)	0.54	-1.22
1W2_059V	VI	1	W3	4*	0.81 (3.24)	0.51	-1.72
EQWE30530	VII	1	W2	1	0.70	0.45	-0.52
EQWE30541	VII	2	W3	4	0.75	0.37	-0.80
EQWE30527	VII	3	W1	1	0.71	0.33	-0.51
EQWE30546	VII	4	W3	3	0.77	0.35	-0.91
EQRE32234.194	VIII	1	R1	2	0.78	0.18	-0.87
EQRE32236.194	VIII	2	R2	2	0.67	0.40	-0.28
EQRE32239.194	VIII	3	R2	2	0.86	0.26	-1.64
EQRE32237.194	VIII	4	R2	3	0.82	0.28	-1.24
EQRE32238.194	VIII	5	R3	3	0.88	0.40	-1.92

Note. The slope parameter was set at 0.588 for all items, and the guessing parameter was set at 0.20 for all multiple-choice items. R1 = understanding explicitly; R2 = understanding implicitly; R3 = making connections; W1 = developing main idea; W2 = organizing information; W3 = using conventions; W4 = topic development; () = Mean score for open-response items.

¹Answer key for multiple-choice items and maximum score category for open-response items.

*Open-response items (OR, SW or LW).

Table 7.1.77 Item Statistics: OSSLT (English) (continued)

Item Code	Section	Sequence	Cognitive Skill	Answer Key/Max. Score Code ¹	CTT Item Statistics		IRT Item Parameters
					Difficulty	Item-Total Correlation	Location
EQRE32241.194	VIII	6	R3	4	0.76	0.39	-0.90
EQRE32235.194	VIII	7	R1	2	0.77	0.36	-0.89
EQRE32243.194	VIII	8	R2	2	0.56	0.30	0.45
EQRE32245.194	VIII	9	R2	4	0.77	0.40	-0.91
EQRE32169.188	IX	1	R2	2	0.71	0.37	-0.51
EQRE32174.188	IX	2	R2	4	0.89	0.25	-1.96
EQRE32173.188	IX	3	R2	4	0.83	0.29	-1.39
EQRE32175.188	IX	4	R2	2	0.73	0.31	-0.59
EQRE32172.188	IX	5	R3	4	0.84	0.33	-1.50
EQRE32177.188	IX	6	R2	3*	0.77 (2.31)	0.45	-1.61
EQRE32178.188	IX	7	R3	3*	0.87 (2.62)	0.32	-2.01
EQWE30574T	X	1	W4	3*	0.89 (2.66)	0.31	-1.99
EQWE30574V	X	1	W3	2*	0.99 (1.97)	0.22	-2.83
EQRE21594.129	XI	1	R3	3	0.67	0.37	-0.27
EQRE21593.129	XI	2	R1	4	0.94	0.27	-2.62
EQRE21599.129	XI	3	R2	2	0.83	0.38	-1.39
EQRE21590.129	XI	4	R2	3	0.95	0.27	-3.00
EQRE21598.129	XI	5	R2	4	0.86	0.26	-1.69
EQRE21596.129	XI	6	R2	4	0.38	0.20	1.71

Note. The slope parameter was set at 0.588 for all items, and the guessing parameter was set at 0.20 for all multiple-choice items. R1 = understanding explicitly; R2 = understanding implicitly; R3 = making connections; W1 = developing main idea; W2 = organizing information; W3 = using conventions; W4 = topic development; () = Mean score for open-response items.

¹Answer key for multiple-choice items and maximum score category for open-response items.

*Open-response items (OR, SW or LW).

Table 7.1.78 Distribution of Score Points and Step Parameters for Open-Response Reading and Short-Writing Tasks: OSSLT (English)

Item Code	Section	Sequence		Off Topic	Missing	Illegible	Score 1	Score 2	Score 3
EQRE32149.185	I	6	% of Students	1.67	0.51	0.02	6.43	39.21	52.16
			Parameters				-2.18	-2.30	-0.02
EQWE30565T	III	1	% of Students	0.13	0.55	0.06	1.56	17.49	80.20
			Parameters				-2.20	-3.13	-1.53
EQWE30565V	III	1	% of Students	0.13	0.55	0.06	1.22	98.03	
			Parameters				-1.67	-4.52	
EQRE32099.180	V	7	% of Students	1.35	4.10	0.08	5.33	14.76	74.38
			Parameters				-0.92	-1.52	-1.49
EQRE32177.188	IX	6	% of Students	0.38	0.56	0.01	12.10	42.32	44.63
			Parameters				-3.45	-1.67	0.28
EQRE32178.188	IX	7	% of Students	0.12	0.67	0.01	1.64	32.42	65.13
			Parameters				-1.86	-3.57	-0.58
EQWE30574T	X	1	% of Students	0.25	0.89	0.02	0.51	29.35	68.99
			Parameters				-0.59	-4.59	-0.78
EQWE30574V	X	1	% of Students	0.25	0.89	0.02	0.93	97.91	
			Parameters				-0.93	-4.74	

Note. The total number of students is 142 915.

Table 7.1.79 Distribution of Score Points and Step Parameters for Long-Writing Tasks: OSSLT (English)

Item Code	Section	Sequence		Off Topic	Missing	Illegible	Score 1.0	Score 1.5	Score 2.0	Score 2.5
1W3_066T	IV	1	% of Students	0.03	0.63	0.03	1.16	1.08	1.34	2.94
			Parameters				-3.50*		-0.22	-1.49
1W3_066V	IV	1	% of Students		0.63	0.03	0.21	0.30	2.02	7.58
			Parameters				-1.51*		-2.47	-2.38
1W2_059T	VI	1	% of Students	0.01	0.37	0.01	0.08	0.08	0.68	1.43
			Parameters				-1.69*		-2.64	-1.95
1W2_059V	VI	1	% of Students		0.37	0.01	0.14	0.16	1.81	7.20
			Parameters				-1.70*		-3.06	-2.48
Item Code	Section	Sequence		Score 3.0	Score 3.5	Score 4.0	Score 4.5	Score 5.0	Score 5.5	Score 6.0
1W3_066T	IV	1	% of Students	7.02	14.79	23.70	23.47	12.93	7.96	2.91
			Parameters	-1.46	-1.27	-0.82	-0.18	0.57	0.66	1.35
1W3_066V	IV	1	% of Students	33.48	34.06	21.70				
			Parameters	-2.43	-0.75	-0.08				
1W2_059T	VI	1	% of Students	4.50	8.81	19.52	22.95	20.10	14.51	6.95
			Parameters	-2.29	-1.64	-1.62	-0.80	-0.32	0.06	0.65
1W2_059V	VI	1	% of Students	44.16	30.87	15.27				
			Parameters	-2.81	-0.43	0.16				

Note. The total number of students is 142 915.

*Scores 1.0 and 1.5 have only one step parameter, since the two categories were collapsed.

Table 7.1.80 Item Statistics: OSSLT (French)

Item Code	Section	Sequence	Cognitive Skill	Answer Key/Max. Score Code ¹	CTT Item Statistics		IRT Item Parameters
					Difficulty	Item-Total Correlation	Location
OQLF31476.128	I	1	L1	3	0.93	0.27	-2.92
OQLF31477.128	I	2	L1	1	0.65	0.39	-0.59
OQLF31470.128	I	3	L3	2	0.81	0.29	-1.60
OQLF31475.128	I	4	L2	4	0.71	0.16	-0.78
OQLF31472.128	I	5	L2	3	0.94	0.26	-3.06
OQLF31468.128	I	6	L3	3*	0.73 (2.20)	0.15	-2.20
OQEF30688	II	1	É1	4	0.59	0.35	-0.24
OQEF30670	II	2	É2	1	0.69	0.36	-0.84
OQEF30663	II	3	É3	3	0.21	0.18	3.03
OQEF30680	II	4	É3	4	0.82	0.34	-1.79
OQEF30690T	III	1	É4	3*	0.87 (2.60)	0.35	-2.31
OQEF30690V	III	1	É3	2*	0.98 (1.95)	0.23	-3.23
2É3 029T	IV	1	É4	6*	0.73 (4.37)	0.60	-1.47
2É3 029V	IV	1	É3	4*	0.84 (3.37)	0.59	-2.31
OQLF31500.125	V	1	L2	4	0.86	0.33	-2.03
OQLF31499.125	V	2	L2	3	0.58	0.27	-0.09
OQLF31502.125	V	3	L1	4	0.77	0.38	-1.33
OQLF31504.125	V	4	L2	2	0.81	0.25	-1.57
OQLF31503.125	V	5	L1	3	0.83	0.34	-1.77
OQLF31498.125	V	6	L3	2	0.74	0.35	-1.13
OQLF31506.125	V	7	L2	3*	0.78 (2.33)	0.34	-1.51
2É2 028T	VI	1	É4	6*	0.73 (4.40)	0.58	-1.57
2É2 028V	VI	1	É3	4*	0.83 (3.33)	0.56	-2.33
OQEF30683	VII	1	É3	2	0.94	0.28	-3.30
OQEF30673	VII	2	É2	1	0.64	0.36	-0.51
OQEF30689	VII	3	É1	3	0.72	0.32	-1.02
OQEF30666	VII	4	É3	4	0.96	0.13	-3.53
OQLF31440.132	VIII	1	L3	3	0.89	0.31	-2.42
OQLF31449.132	VIII	2	L1	1	0.97	0.19	-3.92
OQLF31444.132	VIII	3	L2	1	0.46	0.30	0.62
OQLF31451.132	VIII	4	L3	3	0.70	0.38	-0.85
OQLF31450.132	VIII	5	L1	4	0.92	0.28	-2.87
OQLF31448.132	VIII	6	L2	3	0.91	0.29	-2.72
OQLF31446.132	VIII	7	L2	4	0.42	0.25	0.88
OQLF31453.132	VIII	8	L2	4	0.85	0.21	-1.99
OQLF31442.132	VIII	9	L2	2	0.82	0.38	-1.73

Note. The slope parameter was set at 0.588 for all items, and the guessing parameter was set at 0.20 for all multiple-choice items. L1 = understanding explicitly; L2 = understanding implicitly; L3 = making connections; É1 = developing main idea; É2 = organizing information; É3 = using conventions; É4 = topic development; () = Mean score for open-response items.

¹Answer key for multiple-choice items and maximum score category for open-response items.

*Open-response items (OR, SW or LW).

Table 7.1.80 Item Statistics: OSSLT (French) (continued)

Item Code	Section	Sequence	Cognitive Skill	Answer Key/Max. Score Code ¹	CTT Item Statistics		IRT Item Parameters
					Difficulty	Item-Total Correlation	Location
OQLF31408.130	IX	1	L2	1	0.97	0.16	-4.05
OQLF31410.130	IX	2	L2	2	0.56	0.40	-0.09
OQLF31409.130	IX	3	L2	2	0.86	0.35	-2.11
OQLF31405.130	IX	4	L3	1	0.90	0.24	-2.55
OQLF31406.130	IX	5	L2	4	0.56	0.28	-0.01
OQLF31402.130	IX	6	L2	3*	0.75 (2.25)	0.37	-1.74
OQLF31401.130	IX	7	L3	3*	0.83 (2.50)	0.47	-2.18
OQEF30684T	X	1	É4	3*	0.82 (2.45)	0.53	-1.85
OQEF30684V	X	1	É3	2*	0.96 (1.91)	0.35	-2.73
OQLF31424.134	XI	1	L2	3	0.77	0.29	-1.30
OQLF31427.134	XI	2	L2	1	0.90	0.25	-2.43
OQLF31420.134	XI	3	L3	1	0.91	0.29	-2.65
OQLF31428.134	XI	4	L1	2	0.93	0.24	-3.05
OQLF31422.134	XI	5	L2	1	0.81	0.32	-1.64
OQLF31425.134	XI	6	L2	4	0.65	0.33	-0.52

Note. The slope parameter was set at 0.588 for all items, and the guessing parameter was set at 0.20 for all multiple-choice items. R1 = understanding explicitly; R2 = understanding implicitly; R3 = making connections; W1 = developing main idea; W2 = organizing information; W3 = using conventions; W4 = topic development; () = Mean score for open-response items.

¹Answer key for multiple-choice items and maximum score category for open-response items.

*Open-response items (OR, SW or LW).

Table 7.1.81 Distribution of Score Points and Step Parameters for Open-Response Reading and Short-Writing Tasks: OSSLT (French)

Item Code	Section	Sequence		Off Topic	Missing	Illegible	Score 1.0	Score 2.0	Score 3.0
OQLF31468.128	I	6	% of Students	0.11	0.21	0.00	5.94	67.36	26.38
			Parameters				-4.49	-3.14	1.03
OQEF30690T	III	1	% of Students	0.13	0.82	0.02	3.10	30.87	65.06
			Parameters				-2.62	-3.28	-1.02
OQEF30690V	III	1	% of Students	0.13	0.82	0.02	3.31	95.72	
			Parameters				-2.37	-4.09	
OQLF31506.125	V	7	% of Students	0.49	4.09	0.04	19.08	15.43	60.87
			Parameters				-2.57	-0.48	-1.48
OQLF31402.130	IX	6	% of Students	1.36	0.84	0.02	10.39	47.66	39.72
			Parameters				-3.09	-2.28	0.16
OQLF31401.130	IX	7	% of Students	0.71	1.18	0.00	7.44	29.22	61.45
			Parameters				-3.21	-2.37	-0.97
OQEF30684T	X	1	% of Students	1.07	1.92	0.02	5.87	34.66	56.46
			Parameters				-2.21	-2.66	-0.68
OQEF30684V	X	1	% of Students	1.07	1.92	0.02	2.95	94.04	
			Parameters				-1.34	-4.11	

Note. The total number of students is 5352.

Table 7.1.82 Distribution of Score Points and Step Parameters for Long-Writing Tasks: OSSLT (French)

Item Code	Section	Sequence		Off Topic	Missing	Illegible	Score 1.0	Score 1.5	Score 2.0	Score 2.5
2É3_029T	IV	1	% of Students	0.00	0.36	0.00	0.82	0.65	0.99	1.48
			Parameters				-4.04*		-1.08	-1.70
2É3_029V	IV	1	% of Students	0.00	0.36	0.00	0.34	0.22	1.76	7.46
			Parameters				-2.33*		-2.77	-3.26
2É2_028T	VI	1	% of Students	0.00	0.67	0.00	0.06	0.15	1.01	1.79
			Parameters				-2.03*		-3.56	-2.03
2É2_028V	VI	1	% of Students	0.00	0.67	0.00	0.21	0.30	2.99	8.03
			Parameters				-2.23*		-3.77	-2.77
Item Code	Section	Sequence		Score 3.0	Score 3.5	Score 4.0	Score 4.5	Score 5.0	Score 5.5	Score 6.0
2É3_029T	IV	1	% of Students	5.23	11.42	23.80	22.50	12.31	13.23	7.21
			Parameters	-2.64	-1.98	-1.81	-0.85	-0.15	-0.65	0.21
2É3_029V	IV	1	% of Students	30.23	29.28	30.36				
			Parameters	-3.02	-1.30	-1.17				
2É2_028T	VI	1	% of Students	7.29	11.90	19.73	20.98	14.78	14.74	6.89
			Parameters	-2.80	-1.76	-1.68	-1.00	-0.44	-0.63	0.28
2É2_028V	VI	1	% of Students	28.27	32.88	26.64				
			Parameters	-2.81	-1.48	-0.90				

Note. The total number of students is 5352.

*Scores 1.0 and 1.5 have only one step parameter, since the two categories were collapsed.

Differential Item Functioning (DIF) Analysis Results

The gender- and SLL-based DIF results for the OSSLT are provided in Tables 7.1.83–7.1.85. Only for the English-language test are the results from two random samples of 2000 students. For the French-language test, gender-based DIF analysis was conducted for one sample of students, and SLL-based DIF analysis was not conducted. In both cases, DIF analysis was limited by the relatively small population of students who wrote the French-language test. Each table indicates the value of $\hat{\beta}$, the significance level and the category of the effect size for each item for B- and C-level DIF items. For gender-based DIF, negative estimates of $\hat{\beta}$ indicate that the girls outperformed the boys; positive $\hat{\beta}$ estimates indicate that the boys outperformed the girls. For SLL-based DIF, negative estimates of $\hat{\beta}$ indicate that the SLLs outperformed the non-SLLs; positive $\hat{\beta}$ estimates indicate that the non-SLLs outperformed the SLLs.

Table 7.1.83 Gender-Based DIF Statistics: OSSLT (English)

Item Code	Section	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQRE32141.185	I	1	0.033	0.014		0.028	0.027	
EQRE32144.185	I	2	0.031	0.014		0.036	0.004	
EQRE32147.185	I	3	0.030	0.034		0.028	0.049	
EQRE32143.185	I	4	0.047	0.001		0.047	0.001	
EQRE32145.185	I	5	0.080	0.000	B+	0.081	0.000	B+
EQRE32149.185	I	6	0.121	0.000	C+	0.090	0.000	B+
EQWE30522	II	1	-0.013	0.130		-0.013	0.137	
EQWE30531	II	2	-0.022	0.096		-0.045	0.000	
EQWE30553	II	3	-0.064	0.000	B-	-0.050	0.000	
EQWE30554	II	4	-0.072	0.000	B-	-0.073	0.000	B-
EQWE30565T	III	1	-0.004	0.820		-0.010	0.517	
EQWE30565V	III	1	0.006	0.393		0.000	0.949	
1W3_066T	IV	1	-0.037	0.184		-0.030	0.284	
1W3_066V	IV	1	-0.030	0.076		-0.035	0.038	
EQRE32169.188	IX	1	-0.036	0.008		-0.041	0.002	
EQRE32174.188	IX	2	-0.024	0.016		-0.020	0.040	
EQRE32173.188	IX	3	-0.028	0.016		-0.003	0.804	
EQRE32175.188	IX	4	-0.042	0.002		-0.029	0.034	
EQRE32172.188	IX	5	-0.044	0.000		-0.039	0.000	
EQRE32177.188	IX	6	0.045	0.034		0.025	0.238	
EQRE32178.188	IX	7	-0.044	0.008		-0.009	0.595	
EQRE32090.180	V	1	0.022	0.024		0.003	0.741	
EQRE32095.180	V	2	0.073	0.000	B+	0.080	0.000	B+
EQRE32098.180	V	3	0.038	0.012		0.058	0.000	
EQRE32091.180	V	4	0.069	0.000	B+	0.047	0.001	
EQRE32093.180	V	5	0.071	0.000	B+	0.047	0.000	
EQRE32097.180	V	6	0.016	0.122		0.014	0.141	
EQRE32099.180	V	7	0.072	0.005	B+	0.098	0.000	B+
1W2_059T	VI	1	0.042	0.082		0.014	0.574	
1W2_059V	VI	1	-0.054	0.001		-0.002	0.917	
EQWE30530	VII	1	-0.021	0.109		-0.037	0.004	
EQWE30541	VII	2	-0.044	0.001		-0.040	0.002	
EQWE30527	VII	3	0.005	0.744		-0.008	0.569	
EQWE30546	VII	4	-0.017	0.179		-0.049	0.000	
EQRE32234.194	VIII	1	0.020	0.126		0.035	0.007	
EQRE32236.194	VIII	2	-0.070	0.000	B-	-0.048	0.000	
EQRE32239.194	VIII	3	0.068	0.000	B+	0.025	0.016	
EQRE32237.194	VIII	4	-0.013	0.279		-0.005	0.644	
EQRE32238.194	VIII	5	-0.034	0.000		-0.024	0.009	
EQRE32241.194	VIII	6	-0.098	0.000	C-	-0.064	0.000	B-
EQRE32235.194	VIII	7	-0.008	0.535		-0.023	0.064	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.83 Gender-Based DIF Statistics: OSSLT (English) (continued)

Item Code	Section	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQRE32243.194	VIII	8	0.019	0.201		0.007	0.626	
EQRE32245.194	VIII	9	0.011	0.366		0.025	0.045	
EQWE30574T	X	1	0.040	0.015		0.034	0.039	
EQWE30574V	X	1	0.010	0.189		0.007	0.303	
EQRE21594.129	XI	1	-0.055	0.000		-0.035	0.015	
EQRE21593.129	XI	2	0.011	0.165		0.026	0.001	
EQRE21599.129	XI	3	0.019	0.091		0.028	0.011	
EQRE21590.129	XI	4	0.001	0.871		0.014	0.031	
EQRE21598.129	XI	5	0.036	0.001		0.029	0.005	
EQRE21596.129	XI	6	0.019	0.218		0.003	0.837	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.84 Gender-Based DIF Statistics: OSSLT (French)

Item Code	Section	Sequence	Sample 1		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQLF31476.128	I	1	0.036	0.000	
OQLF31477.128	I	2	-0.027	0.056	
OQLF31470.128	I	3	0.010	0.426	
OQLF31475.128	I	4	0.025	0.086	
OQLF31472.128	I	5	-0.002	0.741	
OQLF31468.128	I	6	0.014	0.433	
OQEF30688	II	1	0.042	0.004	
OQEF30670	II	2	0.037	0.006	
OQEF30663	II	3	-0.019	0.136	
OQEF30680	II	4	-0.022	0.051	
OQEF30690T	III	1	0.050	0.010	
OQEF30690V	III	1	0.007	0.421	
2É3_029T	IV	1	0.069	0.018	B+
2É3_029V	IV	1	-0.018	0.348	
OQLF31408.130	IX	1	-0.005	0.345	
OQLF31410.130	IX	2	-0.029	0.050	
OQLF31409.130	IX	3	-0.040	0.000	
OQLF31405.130	IX	4	-0.030	0.001	
OQLF31406.130	IX	5	0.004	0.806	
OQLF31402.130	IX	6	-0.068	0.002	B-
OQLF31401.130	IX	7	0.032	0.131	
OQLF31500.125	V	1	-0.009	0.408	
OQLF31499.125	V	2	-0.080	0.000	B-
OQLF31502.125	V	3	0.009	0.460	
OQLF31504.125	V	4	0.068	0.000	B+

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.84 Gender-Based DIF Statistics: OSSLT (French) (continued)

Item Code	Section	Sequence	Sample 1		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF
OQLF31503.125	V	5	-0.010	0.383	
OQLF31498.125	V	6	0.054	0.000	
OQLF31506.125	V	7	0.052	0.081	
2É2 028T	VI	1	0.018	0.519	
2É2 028V	VI	1	-0.040	0.027	
OQEF30683	VII	1	-0.012	0.094	
OQEF30673	VII	2	-0.024	0.095	
OQEF30689	VII	3	-0.026	0.055	
OQEF30666	VII	4	-0.011	0.094	
OQLF31440.132	VIII	1	0.007	0.482	
OQLF31449.132	VIII	2	0.007	0.207	
OQLF31444.132	VIII	3	-0.018	0.230	
OQLF31451.132	VIII	4	0.015	0.271	
OQLF31450.132	VIII	5	0.046	0.000	
OQLF31448.132	VIII	6	0.033	0.000	
OQLF31446.132	VIII	7	-0.034	0.025	
OQLF31453.132	VIII	8	-0.055	0.000	
OQLF31442.132	VIII	9	-0.011	0.344	
OQEF30684T	X	1	-0.013	0.556	
OQEF30684V	X	1	0.035	0.007	
OQLF31424.134	XI	1	-0.004	0.741	
OQLF31427.134	XI	2	0.016	0.099	
OQLF31420.134	XI	3	0.008	0.357	
OQLF31428.134	XI	4	0.002	0.757	
OQLF31422.134	XI	5	-0.024	0.045	
OQLF31425.134	XI	6	0.018	0.217	

Note. B = moderate DIF; C = large DIF; - = favouring female students; + = favouring male students.

Table 7.1.85 SLL-Based DIF Statistics: OSSLT (English)

Item Code	Section	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	<i>p</i> -value	Level of DIF	$\hat{\beta}$	<i>p</i> -value	Level of DIF
EQRE32141.185	I	1	-0.137	0.000	C-	-0.109	0.000	C-
EQRE32144.185	I	2	0.078	0.000	B+	0.038	0.009	
EQRE32147.185	I	3	-0.092	0.000	C-	-0.070	0.000	B-
EQRE32143.185	I	4	0.049	0.002		0.002	0.919	
EQRE32145.185	I	5	-0.152	0.000	C-	-0.087	0.000	B-
EQRE32149.185	I	6	-0.009	0.692		0.018	0.417	
EQWE30522	II	1	0.021	0.074		0.000	0.969	
EQWE30531	II	2	-0.076	0.000	B-	-0.039	0.004	
EQWE30553	II	3	-0.029	0.021		-0.042	0.000	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

Table 7.1.85 SLL-Based DIF Statistics: OSSLT (English) (continued)

Item Code	Section	Sequence	Sample 1			Sample 2		
			$\hat{\beta}$	p-value	Level of DIF	$\hat{\beta}$	p-value	Level of DIF
EQWE30554	II	4	-0.105	0.000	C-	-0.100	0.000	C-
EQWE30565T	III	1	-0.007	0.701		-0.021	0.187	
EQWE30565V	III	1	0.040	0.000		0.027	0.000	
1W3_066T	IV	1	0.344	0.000	C+	0.293	0.000	C+
EQRE32169.188	IX	1	0.015	0.333		0.014	0.333	
EQRE32174.188	IX	2	-0.012	0.288		-0.017	0.135	
EQRE32173.188	IX	3	-0.059	0.000	B-	-0.023	0.078	
EQRE32175.188	IX	4	0.064	0.000	B+	0.090	0.000	C+
EQRE32172.188	IX	5	0.146	0.000	C+	0.137	0.000	C+
EQRE32177.188	IX	6	0.014	0.536		0.000	0.988	
EQRE32178.188	IX	7	-0.017	0.354		-0.033	0.054	
EQRE32090.180	V	1	0.027	0.020		0.001	0.955	
EQRE32095.180	V	2	0.041	0.010		0.077	0.000	B+
EQRE32098.180	V	3	-0.144	0.000	C-	-0.106	0.000	C-
EQRE32091.180	V	4	-0.083	0.000	B-	-0.063	0.000	B-
EQRE32093.180	V	5	0.241	0.000	C+	0.181	0.000	C+
EQRE32097.180	V	6	-0.065	0.000	B-	-0.056	0.000	
EQRE32099.180	V	7	-0.189	0.000	C-	-0.115	0.000	C-
1W2_059T	VI	1	-0.344	0.000	C-	-0.273	0.000	C-
1W2_059V	VI	1	0.043	0.017		0.012	0.487	
EQWE30530	VII	1	-0.036	0.014		-0.027	0.055	
EQWE30541	VII	2	0.026	0.085		0.012	0.411	
EQWE30527	VII	3	-0.061	0.000	B-	-0.075	0.000	B-
EQWE30546	VII	4	0.013	0.371		0.035	0.011	
EQRE32234.194	VIII	1	0.049	0.002		0.036	0.016	
EQRE32236.194	VIII	2	-0.146	0.000	C-	-0.139	0.000	C-
EQRE32239.194	VIII	3	0.006	0.672		0.006	0.648	
EQRE32237.194	VIII	4	-0.012	0.407		-0.017	0.179	
EQRE32238.194	VIII	5	0.149	0.000	C+	0.168	0.000	C+
EQRE32241.194	VIII	6	0.022	0.148		0.020	0.157	
EQRE32235.194	VIII	7	0.102	0.000	C+	0.105	0.000	C+
EQRE32243.194	VIII	8	-0.052	0.001		-0.048	0.003	
EQRE32245.194	VIII	9	0.015	0.316		0.027	0.052	
EQWE30574T	X	1	-0.123	0.000	C-	-0.084	0.000	B-
EQWE30574V	X	1	0.016	0.049		0.018	0.009	
EQRE21594.129	XI	1	-0.016	0.309		-0.016	0.300	
EQRE21593.129	XI	2	0.123	0.000	C+	0.100	0.000	C+
EQRE21599.129	XI	3	0.032	0.021		0.007	0.591	
EQRE21590.129	XI	4	0.025	0.004		0.021	0.010	
EQRE21598.129	XI	5	0.048	0.000		0.031	0.012	
EQRE21596.129	XI	6	-0.078	0.000	B-	-0.035	0.032	

Note. B = moderate DIF; C = large DIF; - = favouring SLLs; + = favouring non-SLLs.

**Education Quality and
Accountability Office**



2 Carlton Street, Suite 1200, Toronto ON M5B 2M9

Telephone: 1-888-327-7377 Web site: www.eqao.com

© 2011 Queen's Printer for Ontario