Grade 4 Writing........................................................................................................................................... 27
  Monitoring Comments: Grade 4 Writing........................................................................................................ 28
Grade 7 Writing................................................................................................................................................ 29
  Monitoring Comments: Grade 7 Writing........................................................................................................ 30
Summary of Marking Reliability ....................................................................................................................... 31
Considerations for Marking .............................................................................................................................. 32
Some noteworthy comments from participants ............................................................................................... 33
Appendix 1: Monitoring session Scorer demographics .................................................................................... 34
  Summary of responses to Likert scale statements (sample): ........................................................................ 34
Foundation Skills Assessment: Marking Monitoring

Large scale assessments in British Columbia reflect evidence-based best practices as recommended by national and international assessment experts in “Standards for Educational and Psychological Testing” (published collaboratively by the American Educational Research Association, National Council of Measurement in Education, and the International Test Commission (AERA, NCME, & ITC, 2014)). This document, referred to as the “Standards”, provides guidance on all aspects of the assessment process from development, administration, to scoring and reporting. According to Standard 10.2: Monitoring Accuracy and Reliability of Scoring, test developers have a responsibility to “monitor and document the accuracy and reliability of scoring, and correct sources of scoring errors” (AERA et al., 2014; p. 118). In adherence to this standard, the Ministry of Education monitors district and school-based scoring of the written-response sections of the Foundation Skills Assessment (FSA) by re-scoring a sample of student response booklets from each district and a select number of independent schools.

This report describes the degree of consistency between the score local markers assigned to their students’ responses in district/school-based scoring sessions and scores assigned by the markers participating in the provincial monitoring session. The report also provides general comments regarding student performance based on the participants’ observations during the monitoring session.

The scoring of the written-response questions for the FSA is based on scoring rubrics which were developed under guidance from Standard 10.1: Developing Procedures for Human Scoring (AERA et al., 2014), and adapted from the BC Performance Standards. Scoring rubrics and related scoring materials are available at www.bced.gov.bc.ca/assessment/fsa/training.htm.

Scorers

In total, 53 scorers participated in the marking monitoring session. Each school district was invited to send one educator, familiar with FSA marking in their district, to participate in the FSA monitoring session. In addition, four independent school educators and one educator from the Yukon were invited to participate. (Appendix 1: Monitor Session Scorer Demographics)

Scoring Process

Using the FSA Scoring Guide, scorers were familiarized with the FSA Scoring Rubrics and trained in holistic marking, using the ‘best line of fit”. The training they received is similar to the training markers should receive at their local marking sites.

Scorers used professional judgment, guided by the scoring materials, to give each student response a fair and reliable reading and score. Differences of one scale point between the score assigned at the monitoring session and the score assigned at the district/school-based scoring session are considered to be reasonable. An 80% or higher correlation (scores are within one scale point) between locally assigned scores and the scores assigned at the provincial monitoring session indicates that local scores are highly aligned with the provincial scoring standards specific to each assessment.

Note: Percentage differences may not add up to 100% due to rounding.
Over 3,100 student response booklets from grades 4 and 7 were remarked during the monitoring session.

**Scoring Feedback**

During the scoring session, scorers’ observations were compiled, based on student responses for each of the FSA items. A summary of observations for Reading, Writing and Numeracy can be found at the end of each section of the report.

Scorers also provided general feedback on the marking support materials and the session. This feedback is found at the end of the report.
Grade 4 Numeracy

Written-Response Question 1
Curriculum Theme – Number, Computational Fluency

Grade 4
Percentage of score differences for Numeracy Question 1

97.5% of the local scores were within 1 point of the scores assigned at the monitoring session.
Grade 4 Numeracy (continued)

Written-Response Question 2
Curriculum Theme – Patterns

Grade 4

Percentage of score differences for Numeracy Question 2

<table>
<thead>
<tr>
<th>Difference</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>63.0%</td>
</tr>
<tr>
<td>1</td>
<td>31.2%</td>
</tr>
<tr>
<td>2</td>
<td>5.3%</td>
</tr>
<tr>
<td>3</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

94.2% of the local scores were within 1 point of the scores assigned at the monitoring session.
Grade 4 Numeracy (continued)

Written-Response Question 3
Curriculum Theme – Number, Computational Fluency, Geometry and Measurement

Grade 4

Percentage of score differences for Numeracy Question 3

92.7% of the local scores were within 1 point of the scores assigned at the monitoring session.
Monitoring Comments: Grade 4 Numeracy
General comments regarding student performance on numeracy written response questions, according to the participants of the July 2018 monitoring session

Student Strengths
- Students demonstrated a strong ability to show data visually.
- Students exhibited strong analysis skills; tabulating survey data and giving answers based on the data.

Areas Requiring Improvement
- Students need to read instructions/questions more thoroughly.
- Students need to show all their work on the paper.
- Students need to check that they have completely answered the question.

Implications for Instruction
The marking team suggests that teachers could:
- encourage students to clearly show all their work/steps/thinking and to identify the final answer by circling or answering the questions asked in a sentence.
- encourage students to reflect on their answers to see if they make sense or are logical.
- provide direct instruction representing, comparing and interpreting data using concrete graphs.
- continue to encourage students to express their thinking in concrete ways.
Grade 7 Numeracy

Written-Response Question 1
Curriculum Theme – Number, Computational Fluency

Grade 7
Percentage of score differences for Numeracy Question 1

<table>
<thead>
<tr>
<th>Difference between the local score and the score assigned at monitoring</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>63.3%</td>
</tr>
<tr>
<td>1</td>
<td>33.5%</td>
</tr>
<tr>
<td>2</td>
<td>2.9%</td>
</tr>
<tr>
<td>3</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

96.8% of the local scores were within 1 point of the scores assigned at the monitoring session.
Grade 7 Numeracy (continued)

Written-Response Question 2
Curriculum Theme – Number, Computational Fluency, Patterns

Grade 7

Percentage of score differences for Numeracy Question 2

96.5% of the local scores were within 1 point of the scores assigned at the monitoring session.
Grade 7 Numeracy (continued)

**Written-Response Question 3**

Curriculum Theme – Number, Computational Fluency, Geometry and Measurement

Grade 7

**Percentage of score differences for Numeracy Question 3**

- 62.9% within 0 points
- 31.9% within 1 point
- 4.5% within 2 points
- 0.7% within 3 points

94.8% of the local scores were within 1 point of the scores assigned at the monitoring session.
Monitoring Comments: Grade 7 Numeracy
General comments regarding student performance on numeracy written response questions, according to the participants of the July 2018 monitoring session

Student Strengths
▪ Students attempted the task and were able to come up with at least a partial solution.
▪ Students work was clear and organized.
▪ Students demonstrated creative approaches to solving numeracy problems.

Areas Requiring Improvement
▪ Students need to read instructions/questions more thoroughly.
▪ Students need to improve their understanding of mathematical vocabulary (for example, least probability and greatest probability).
▪ Students need to show all their work on the paper.
▪ Students need to check that they have completely answered the question.

Implications for Instruction
The marking team suggests that teachers could:
• encourage students to clearly show all their work/steps/thinking and to identify the final answer by circling or answering the questions asked in a sentence.
• provide direct instruction on mathematical vocabulary and the difference between discrete and continuous data and how this data is represented.
• provide opportunities for students to draw and label various types of graphs.
• encourage students to reflect on their answers to see if they make sense or are logical.
Grade 4 Reading

Written-Response: Theme 1, Question 1

96.1% of the local scores were within 1 point of the scores assigned at the monitoring session.
Grade 4 Reading (continued)

Written-Response: Theme 1, Question 2

Grade 4

Percentage of score differences for Reading Question 2

97.3% of the local scores were within 1 point of the scores assigned at the monitoring session.
Written-Response: Theme 1, Question 3

Grade 4

Percentage of score differences for Reading Question 3

95.9% of the local scores were within 1 point of the scores assigned at the monitoring session.

Grade 4 Reading (continued)
Grade 4 Reading (continued)

Written-Response: Theme 2, Question 1

Grade 4

Percentage of score differences for Reading Question 1

93.8% of the local scores were within 1 point of the scores assigned at the monitoring session.

54.5%

39.3%

6.2%
Written-Response: Theme 2, Question 2

Grade 4

Percentage of score differences for Reading Question 2

96.0% of the local scores were within 1 point of the scores assigned at the monitoring session.
Written-Response: Theme 2, Question 3

Grade 4

Percentage of score differences for Reading Question 3

96.4% of the local scores were within 1 point of the scores assigned at the monitoring session.
Monitoring Comments: Grade 4 Reading
General comments regarding student performance on reading written response questions according to the participants of the July 2018 monitoring session

Student Strengths
- Students demonstrated good comprehension of the texts and task.
- Students were able to provide details and specific information to demonstrate their understanding.
- Some students demonstrated a personal connection to the texts.
- Some students were able to make insightful inferences.

Areas Requiring Improvement
- Students need to go beyond literal comprehension to more inferential thinking.
- Students need to support ideas with information from both texts.
- Students need to base inferences on evidence from texts.

Implications for Instruction
The marking team suggests that teachers could:
- focus on guiding students on ways to support their thinking using evidence from the texts.
- work with students to compare two texts.
- work with students to make connections beyond text-to-self.
- use student exemplars to help model student responses.
Grade 7 Reading

Written-Response: Theme 1, Question 1

95.6% of the local scores were within 1 point of the scores assigned at the monitoring session.
Written-Response: Theme 1, Question 2

Grade 7

Percentage of score differences for Reading Question 2

96.0% of the local scores were within 1 point of the scores assigned at the monitoring session.
Written-Response: Theme 1, Question 3

Grade 7

Percentage of score differences for Reading Question 3

95.5% of the local scores were within 1 point of the scores assigned at the monitoring session.
Grade 7 Reading (continued)

Written-Response: Theme 2, Question 1

93.9% of the local scores were within 1 point of the scores assigned at the monitoring session.
Written-Response: Theme 2, Question 2

Grade 7

Percentage of score differences for Reading Question 2

96.0% of the local scores were within 1 point of the scores assigned at the monitoring session.
Written-Response: Theme 2, Question 3

Grade 7

Percentage of score differences for Reading Question 3

94.6% of the local scores were within 1 point of the scores assigned at the monitoring session.

Grade 7 Reading (continued)
Monitoring Comments: Grade 7 Reading
General comments regarding student performance on reading written response questions according to the participants of the July 2018 monitoring session

Student Strengths
- Students were able to summarize the texts.
- Students have a basic understanding of the texts and the task.
- Students were able to make good inferences and include details.
- Many students referenced both texts to make comparisons.

Areas Requiring Improvement
- Students need to make specific references to the texts.
- Students need to make meaningful inferences and provide insight.
- Students need to compare texts of different genres.
- Students need to cite textual evidence to demonstrate their understanding.

Implications for Instruction
The marking team suggests that teachers could:
- provide instruction making inferences that are more meaningful and that draw insights to deepen comprehension.
- provide instruction on how to reference texts to locate main ideas and supporting details.
- consider using the criteria from the scoring rubrics to guide student learning.
- choose to use texts based on current issues to allow students to use background knowledge to support their answers.
- use graphic organizers to provide students with more opportunities to compare, contrast, and make connections.
Grade 4 Writing

85.5% of the local scores were within 1 point of the scores assigned at the monitoring session.
Monitoring Comments: Grade 4 Writing

General comments regarding student performance on the writing written response question according to the participants of the July 2018 monitoring session

Student Strengths
- Students were able to relate to the writing topic.
- Students were able to use good detail.
- Students were able to generate many ideas.
- Students were able to structure their writing using a beginning, middle and ending.

Areas Requiring Improvement
- Students need to pay attention to transitions.
- Students need to work on sentence variety.
- Students need to use engaging details to support reactions.
- Students need to pay attention to sentence structure, paragraphing, and use of conventions.

Implications for Instruction
The marking team suggests that teachers could:
- encourage students to make personal connections.
- use student exemplars to help model writing.
- familiarize students with the scoring rubrics and performance standards.
Grade 7 Writing

Percentage of score differences for Writing Question 1

83.0% of the local scores were within 1 point of the scores assigned at the monitoring session.
Monitoring Comments: Grade 7 Writing

General comments regarding student performance on the writing written response question according to the participants of the July 2018 monitoring session.

Student Strengths

▪ Students were able to develop a personal response.
▪ Students were able to use background knowledge to make relevant personal connections.

Areas Requiring Improvement

▪ Students need to use engaging details to support reactions.
▪ Students need to use engaging introductions and endings that provide closure.
▪ Students need to use transition words between paragraphs.
▪ Students need to improve sentence structure, varied language, and conventions.

Implications for Instruction

The marking team suggests that teachers could:

▪ instruct students on how to engage the reader.
▪ provide instruction on developing beginning/conclusions with impact that goes beyond summary.
▪ focus on vocabulary development, effective use of transition words, and writing conventions.
▪ provide students with student exemplars of effective writing along with scoring rubrics and performance standards for self and peer assessment along with descriptive feedback on how to improve their writing.
Summary of Marking Reliability

Maintaining consistent FSA marking standards across the province helps to ensure that no matter where a student takes the FSA, his or her responses will be marked in a consistent manner. The purpose of the FSA marking monitoring session was to gather evidence for the reliability, validity and fairness of the FSA marking process in accordance to the Standards (AERA et al., 2014). Overall results show a high level of consistency; between 83.0 - 97.5% of the scores assigned at the monitoring session were within 1 point of the local scores. No significant difference in scoring consistency was found between the grades, although marking reliability for numeracy was slightly higher than for reading and writing for both grades.

The results from the FSA marking monitoring session provide evidence for the reliability, validity, and fairness of the FSA marking process at the provincial level. The results also allow the Ministry to assess the reliability of the marks assigned at district levels, provide useful information to FSA markers about the FSA marking process in terms of how to effectively use the assessment rubric and knowing what to look for when marking student responses, etc. A district level and an independent school summary level monitoring report are also generated to help inform local scoring.

At the end of the session, feedback was collected from the participants. This information is summarized in Appendix 1. Overall, the feedback was extremely positive – all participants enjoyed the session, agreed that it was a valuable professional development experience, and suggested they learned useful information to improve their local scoring activities.
Considerations for Marking
General comments regarding district/school marking according to the participants of the July 2018 monitoring session.

Score holistically. Holistic scoring gives an overall score for the student response. When scoring holistically, the focus is on what the student CAN do and the strengths demonstrated in the student’s work. No one aspect of the student’s work carries more importance or weight than another when scoring holistically.

Give the student response a complete and careful read for a general impression. Use professional experience, and based on the rubric, give each student paper a fair score. Be aware of personal biases when scoring and refer to the rubric and exemplars, especially when in doubt.

Based on the balance of strengths and weaknesses according to the rubric, arrive at an overall score for the response.

Refer to the rubrics regularly to stay focused. The scoring rubrics form the basis for assessment; refer to them often, especially if unsure what score to give the student’s work. Markers should review the rubrics at the beginning of each scoring session, and continue to refer to them during scoring, adding notes and highlighting aspects of importance.

Score for the ‘best line of fit’. Markers should give an overall score for the student response, based on the score of the rubric where most of their response falls. Often students will score across two or three score points on the rubric. The final score should reflect where most of a student’s response lies.

When being scored, this student’s work received the greatest number of checkmarks for descriptors in Score 4, therefore, ‘best line of fit’ would be a score of 4. (even though there is a calculation error in this numeracy example)
Some noteworthy comments from participants

This was a great opportunity to review and look over how the marking process takes place. It was beneficial to see the 'holistic marking' process and to see how different perspectives on papers are taken into account. Overall a very useful process to take part in.

line of best fit was interesting and a useful tool I will take back to my district; I would have liked to start marking more quickly and do fewer exemplars; I felt marking in teams and the experts in the room was wonderful support.

I can now go back to my district with more clarity about the score allocation, particularly those for the math exemplars

I really enjoy coming to this; I enjoy the training, the marking; I also really enjoy working with other educators because I live in the north and am isolated from the rest of the province. It is important to work together as a group. I thought this was very well organized. I have attended 2-3 previous marking sessions. I thought it was really well organized; time frame was great. Well done!

One thing that I will take back was the importance of implications for instructions and how to support our students. Also, the proper application of the holistic rubric is something else I will share.

Being exposed to a variety of papers from schools across the province was enlightening. Discussions with colleagues at the table were valuable. Discussion around exemplars for numeracy rubrics went too long. Sometimes the discussion of exemplars took longer than scoring.

Learning line of best fit important! I have a much better understanding of what is being looked for in math now. There were a number of questions that were not in alignment with my scoring. I appreciate that FSAs can provide very valuable information to teachers to use, but not all teachers are on board and not all PVPs provide the info to teachers in a timely manner

Enjoyed discussion with other colleagues in the profession. Will take back lots of info about marking especially the line of best fit in math. Facilitators were great, upbeat and very supportive.
Appendix 1: Monitoring session Scorer demographics

48 participants completed the feedback form.

Who were they?
- 21% classroom educators
- 69% administrators
- 17% other

Years of experience:
- 6% had less than 10 years of experience
- 79% had more than 16 years of experience

Where did they teach?
- 85% taught at English language schools
- 6% taught at French immersion schools

Experience with the FSA:
- 46% had experience with field testing the FSA
- 33% participated in the district-based marking of the FSA
- 12% participated in the FSA marking refresher
- 4% participated in the FSA standard setting sessions

Summary of responses to Likert scale statements (sample):

100% agreed or strongly agreed that the training provided an opportunity to develop a common understanding of the scoring criteria

100% agreed or strongly agreed that the exemplar papers were useful in explaining and supporting score allocation

92% agreed or strongly agreed that the marking leaders were effective in guiding and coordinating the monitored marking process

98% agreed or strongly agreed that the monitored marking process was efficient and well-coordinated

100% agreed or strongly agreed that the monitored marking was a valuable professional development experience