ORGANIZATION OF EASTER CARIBBEAN STATES (OECS)

**PROGRAMME FOR EDUCATIONAL ADVANCEMENT
AND RELEVANT LEARNING (PEARL)**

**EARLY READING ASSESSMENT (OERA)
AND
EARLY MATHEMATICS ASSESSMENT (OEMA)**

TEST DEVELOPMENT REPORT

31 July 2025

I.   Introduction

The importance of high-quality learning data in driving effective and equitable education reform cannot be overstated. Understanding who is not learning, what foundational skills are lacking, and where support is most urgently needed is essential to ensuring that *every learner succeeds* (UIS, 2019; World Bank, 2019). In the Eastern Caribbean, this vision has materialized through the collaborative development of the OECS Early Grade Reading Assessment (OERA) and the Early Grade Mathematics Assessment (OEMA)—two diagnostic, regionally harmonized tools designed to evaluate literacy and numeracy proficiency among Grade 2 students.

After years of planning and consultation, the OERA and OEMA were created under the Programme for Educational Advancement and Relevant Learning (PEARL) with the support and leadership of assessment and education officers from eight OECS Member States. These assessments aim to provide reliable, actionable data that can inform the refinement of the OECS Harmonized Curriculum, guide targeted teacher professional development, and support broader data-driven decisions in education policy and planning (OECS Commission, 2022; UIS, 2019).

The instruments were designed through an iterative, evidence-based process that emphasized alignment with both regional learning standards and international benchmarks, particularly the Global Proficiency Frameworks and Minimum Proficiency Levels (UIS, 2022; OECS Commission, 2021). Content was drawn directly from the Grade 2 OECS Harmonized Primary Curriculum for Reading and Mathematics, ensuring coherence with what is expected to be taught across the region (OECS Curriculum and Assessment Framework, 2022).

Special attention was paid to the cultural and linguistic relevance of all items: reading passages reflect OECS contexts, and mathematics problems include concrete, pictorial, and abstract representations rooted in real-world Caribbean applications (USAID & RTI, 2017; ACER, 2018). Bias and fairness reviews were conducted to ensure all students—regardless of background—had an equal opportunity to demonstrate their skills. The development process incorporated global best practices including the EGRA Toolkit, the UIS Global Framework for Measuring Learning Outcomes, and the OECS Curriculum and Assessment Framework (UIS, 2020; OECS, 2022).

In 2024, the OERA and OEMA were piloted across eight OECS Member States as part of PEARL. The pilot gathered both quantitative and qualitative data to evaluate the technical quality of the assessments and the feasibility of full-scale regional administration. The pilot analysis was conducted using internationally recognized psychometric standards and methodologies (ACER, 2018; UIS, 2022). This included analyses of item difficulty, discrimination, and internal consistency to assess validity and reliability.

In May 2025, a three-day Assessment Revision Workshop convened curriculum specialists, assessment officers, and technical experts from across the region. The workshop aimed to:

- Analyze pilot results and identify technical and content-related issues;
- Revise items based on performance, clarity, length, and cultural relevance;

- Finalize assessment instruments for the 2025 roll-out; and
- Confirm an implementation plan grounded in standardized and equitable administration procedures.

The finalized versions of the OERA and OEMA are now aligned with the OECS Harmonized Curriculum and meet many of the eligibility requirements for reporting on Sustainable Development Goal (SDG) 4.1.1a, which monitors the proportion of children achieving minimum proficiency in reading and mathematics by the end of primary education (UIS, 2019). These assessments not only facilitate valid cross-country comparisons but also support the global education monitoring agenda. Most importantly, they empower OECS Member States to make informed, data-driven decisions that improve learning outcomes and ensure that no learner is left behind.

II.   Development and Expertise

The development of the OECS Grade 2 Assessments in Reading and Mathematics was heavily informed by the OECS Assessment Framework (2018) and the OECS Curriculum and Assessment Framework (2023) which reflect current trends in assessment practices in the region. The goal was to develop regional assessments that reflect the recommendations received as part of the consultations as well as represent trends such as:

1. Balanced systems of assessment that have an appropriate emphasis on diagnostic, formative and summative assessment.
2. Reflection on learning outcomes in the new OHPC so that strategies may be developed to enhance teaching through professional development and/or the further enhancement of certain elements of the OHPC.
3. Assessment design and administration in a manner that results in the generation of reliable and valid data.
4. Information for educators that creates opportunities to discuss assessment results with learners in a manner that results in learners reflecting on what they have learnt, what they need to know and developing, in conjunction with their teachers' strategies to advance their learning.
5. Go beyond measuring the level of learning to identify specific gaps in learning and provide learners, teachers and parents with evidence that may be used to plan learning.
6. Generate results which are shared with learners, parents and educators on a timely basis to measure progress, inform teaching and support system enhancement.
7. Are administered in a positive, affirming manner that builds on learner confidence.
8. Are conducted in a manner that protects the rights of learners.
9. Provide data that are stored and shared responsibly.

The regional assessments have been developed and the pilot designed to support decision making and it is anticipated that assessment data will be shared and used as follows:

Level	Data Use
School	• Students receive results and feedback that, guided by their teachers, they can understand and use to improve their learning. • Teachers use individual results, along with other evidence to guide responsive and differentiated learning. • Principals use school results to determine priorities for Continuous School Improvement which brings a focus to specific professional development and support necessary to achieve goals.
District	• Education Officers use District results to work with schools as they develop their Continuous School Improvement Plans. These are amalgamated into District plans for which professional development and resource support is sought.
Education Ministry	• Chief Education Officers/Directors use National results to guide the development of national goals for education and to determine budgetary resources which will be made available to support Continuous School and System Improvement
Government	• To show evidence of the achievement associated with the investment in Education as part of its public accountability
OECS	• To enhance education in the region and inform priorities for research and collaboration.

Reporting templates for the dissemination of the data will be drafted to ensure that the different bodies receive the information from the assessments in a form that facilitates evidence based decision making. The development of such reporting templates and the circulation of results in this form will increase the utility of regional assessments and permit maximum use of assessment data by all educators. This coupled with a greater emphasis on formative assessment that is embedded in the new OHPC, will lead to informed decisions about teaching and learning enhancements which will result in bringing the OECS closer to having "every learner succeed".

Consultation and Professional Development

The OERA and OEMA development was an iterative process led by Assessment Officers and Curriculum Officers from Member States, and advised by PICT Assessment Specialists. Subject matter experts developed test items based on learning outcomes in the OHPC including reading passages selected to reflect culturally relevant contexts from the OECS region and mathematics problems included real-world applications to engage students meaningfully.

Consultations with subject matter experts and policymakers to review test content for alignment with curriculum objectives ensured regional assessments met instructional expectations. Bias and fairness

reviews were conducted to ensure cultural and linguistic appropriateness, making certain that all students, regardless of background, had an equal opportunity to demonstrate their skills.

In order to align the OERA and OEMA with international and regional best practices in early literacy and numeracy assessment, key references included the UIS Global Framework for Measuring Learning Outcomes and the EGRA Toolkit for early reading assessment. Details of the GPF alignment of the OERA and OEMA are located in the OECS Regional Assessment Framework.

The professional development for members of the assessment team was embedded in the assessment development process. The following members of the assessment team, detailed in Table 1, have received Certificates of Completion from St. Francis Xavier University as part of the Comprehensive Professional Development Programme. These individuals participated in numerous capacity building activities and completed self-reflections in order to qualify for receipt of their Certificates.

Table 1: OERA/OEMA Assessment and Item Writing Team

Country	Name	Role
Antigua and Barbuda	Ayisha Bellot	Testing and Measurement Officer
	Deavene Guishard	Education Officer Assessment, Testing & Examinations; CXC Local Registrar
British Virgin Islands	Khaliah Callwood	Curriculum Officer - Math
	Janice George-Harris	Education Officer - Language
Dominica	Shermine Bardouille	Curriculum Officer – Language Arts
	Candia Robinson	Senior Education Officer- Curriculum
	Krishna Robin	Measurement Officer
	Meloticia Mills	Item Writer - Mathematics
Grenada	Valjean Alexander	Assessment Officer
	Jason Douglas	Curriculum Development Officer
	Henva Simeon	Primary School Principal
	Gillian Thomas	Teacher
St. Kitts and Nevis	Ornella Bacchus	Assessment Coordinator
	Dalton Esdaille	Assessment Officer
Saint Lucia	Janeta Joseph	Teacher
	Patterson Abraham	Assessment Officer
St. Vincent & the Grenadines	Gaylene Wickham	Education Officer/Literacy
	Fitzroy Stapleton	Education Officer Curriculum

Country	Name	Role
	Bernadette Greaves	SEO Exams
	Kenneth Holder	Principal

OECS Harmonised Primary Reading Curriculum Alignment

The OERA, is designed to gather information about student progress at a time when expectations for success shift from the Early to the Transitional stage of reading. During this shift, students engage with increasingly complex text where their reliance on oral processing of text fades and the more efficient processing of text that occurs with silent reading emerges. The information provided by the OERA is intended to provide a clearer path toward school planning, decisions about classroom resources, and differentiated instruction that is responsive to each learner within the diversity of learners in today's classrooms.

The purpose of reading and viewing instruction is to develop readers who enjoy and interact meaningfully with a wide range of genres and text forms. Meaningful interaction with texts is developed by learning to access and build on background knowledge and use various textual sources of meaning, vocabulary and language structures, and meaningful application of graphophonic elements of the text. The revised OECS Harmonised Primary Curriculum (OHPC) provides educators with Primary School and Grade Level pedagogical and instructional support that fulfils this purpose.

Grade K-6 Essential Learning Outcomes (ELOs) for Reading and Viewing

- **ELO 2:** Students will demonstrate a variety of ways to use background knowledge and interests to select and engage critically with a range of culturally diverse paper based, visual, and digital texts for pleasure and personal growth.

- **ELO 3**: Students will interact with understanding and critical thought to a variety of genres and text forms using vocabulary, comprehension strategies, and graphophonic cues.

- **ELO 4:** Students will develop their understanding of how an author's choice of vocabulary, language, genre, text form, text features and style influence the meaning of text and define the author's craft.

OECS Harmonised Primary Mathematics Curriculum

All questions on the OECS EARLY Mathematics Assessment (OEMA) are linked to learning outcomes set out in the Grade 2 Mathematics OHPC. The tasks presented aim to allow students to demonstrate factual knowledge (knowledge of basic mathematical facts and information); conceptual knowledge (understanding of mathematics concepts and principles) and procedural knowledge (ability to perform mathematical procedures and follow rules for solving problems).

The knowledge, skills, attitudes, and mathematical processes that students should demonstrate by the end of Grade 2 are highlighted under six content strands, namely, Number Sense (N); Operations with Numbers (O); Patterns and Relationships (PR); Geometric Thinking (G); Measurement (M), and Data Handling and Probability (D).

In Grade 2, it is expected that instructional time should focus on six areas: (1) extending understanding of base-ten notation; (2) building fluency with addition and subtraction; (3) identifying, describing, extending, creating, and making predictions about a variety of patterns, including those found in real-life contexts; (4) analysing and classifying two-dimensional shapes as polygons or non-polygons; (5) using standard units of measure; and (6) collecting, organising, visualising and analysing data, and using language such as impossible, possible and certain, and making predictions.

III.   Pilot Administration

The pilot OERA included six passages spanning four genres, with questions structured to evaluate students' performance across three cognitive levels (literal, inferential, evaluative) and three levels of difficulty. These items were developed by the OERA Development Team and aligned with the OECS Learning Standards to reflect the authentic cultural and social experiences of students in the region. To simulate a familiar classroom environment, the layout of the student test booklets resembled a typical Grade 2 reading comprehension lesson, supported by illustrations purchased from image banks or generated via AI tools, in consultation with the PICT.

Students participated in the assessment across two sittings. Trained Assessors and Invigilators facilitated the sessions using a standardized script that introduced the assessment, explained its purpose, offered emotional support, and shared relevant context for each passage—such as character names, nonfiction topics, and a brief story overview—to ensure students felt confident and prepared. Table 2 outlines the specifications of the 2024 Pilot OERA.

Table 2: *Outline of Pilot OERA*

Curriculum Alignment	Mode of Administration	Time Allotment	Strategic Focus	Cognitive Levels	Question Formats	Number of Items
Genres: two fiction, two nonfiction, one visual text, and one poetry	**Day 1:** Individual Administration by the Assessor; Paper based for the students. Tablet based for Assessor (Oral Reading and Comprehension)	**Day 1:** Oral Reading (15–20 minutes per student).	Comprehension Vocabulary Grammar Word Work	Literal Analytic Analytic (usage)	Selected Response	32
					Constructed Response	16
	Day 2: Group Administration: Group Administration by the Invigilator Paper based for students (Silent Reading Comprehension)	**Day 2:** Silent Reading and Comprehension (15–20 minutes per passage; 45–60 minutes total).				

In Grade 2, mathematics instruction is expected to emphasize six key domains: (1) extending understanding of base-ten notation; (2) building fluency with addition and subtraction; (3) identifying, describing, extending, creating, and predicting patterns in both mathematical and real-world contexts; (4) analyzing and classifying two-dimensional shapes as polygons or non-polygons; (5) applying standard units of measurement; and (6) collecting, organizing, and interpreting data using appropriate language for probability (e.g., "impossible," "possible," "certain") (OECS Commission, 2022).

Each item on the pilot OECS Early Grade Mathematics Assessment (OEMA) was aligned to the Grade 2 Mathematics learning outcomes from the draft OECS Harmonized Primary Curriculum (OHPC). Assessment tasks were designed to measure factual knowledge (basic facts and procedures), conceptual understanding (relationships between ideas), and procedural knowledge (application of methods to solve problems) as described in global mathematics assessment frameworks (UIS, 2019; OECS Learning Standards, 2018).

The Constructed Response Assessment was administered in a group setting. Students received printed booklets with space for open responses, and the Invigilator read each question aloud, providing time for written answers. Similarly, the Selected Response Assessment was group-administered, using printed test booklets and bubble-style answer sheets. The dual-format approach was intended to capture a comprehensive view of student thinking, offering opportunities for students to show both reasoning and calculation processes. Table 3 outlines the specifications of the 2024 Pilot OEMA.

Table 3: *Outline of OEMA*

Item Formats	Number of Items	Mode of Administration	Time Allotment	Curriculum Strands	Cognitive Levels
Selected Response (Multiple choice)	35	Group administration; Pencil and paper (answers shaded on answer sheets)	2 minutes per item, with the total testing time estimated between 60–80 minutes.	Number Sense; Operations with Numbers; Patterns and Relationships; Geometrical Thinking; Measurement; Data Handling and Probability	Factual knowledge Conceptual knowledge Procedural knowledge
Constructed Response	6	Group administration: Pencil and paper (answers written in question booklet)			

Field Trial Sample Selection

In 2022, Member States identified 84 schools that would participate in the pilot of the new OHPC based on an estimated total population of 341 public primary schools in the OECS region. As it is important that the schools included be as representative of the population as possible, it was proposed that the national numbers be based on the ratio of schools in each Member State to the overall number of schools in the Region.

With the intention of a full-scale rollout of the OEMA and OERA to all grade 2 schools, field tests were conducted across OECS Member States on a portion of pilot school students to validate test items and identify areas for improvement. While the assessment tools are designed for Grade 2 students at the end of their academic year, after discussion with leadership of Member States on logistical constraints, it was decided that field testing would be delayed to the next academic year.

Over 600 Grade 3 pilot school students (approx. 30% of the pilot student population) were identified for the field test given that this population of students closely demonstrate the competencies of students at the end of Grade 2 and would be representative of the target population. Students were selected using a stratified simple random sampling methodology to ensure equal representation of gender and ability across the population. Countries were responsible for ensuring that the sample of students selected met the agreed-upon criteria, ensuring that gender distribution and abilities are represented. Table 4 shows the disaggregation of student data by school and gender.

Table 4: Disaggregation Of Student Population and Sample Count By School And Gender

Country	Population				Sample			
	Schools	Males	Females	TOTAL	Schools	Males	Females	TOTAL
Anguilla	6	79	80	159	2	13	6	19
Antigua and Barbuda	33	365	311	676	7	37	36	73
British Virgin Islands	11	114	121	235	3	19	20	39
Dominica	57	263	228	491	6	37	29	66
Grenada	56	684	634	1318	7	41	49	90
St. Kitts and Nevis	31	260	248	508	1	26	28	54
St. Lucia	77	1118	1051	2169	7	83	66	149
St. Vincent and the Grenadines	67	721	665	1386	14	74	75	149
TOTAL	**341**	**3604**	**3338**	**6942**	**47**	**330**	**309**	**639**

Assessment Feedback

After the pilot administration of the OERA and OEMA, a comprehensive survey gathered feedback from regional assessment teams, invigilators, assessors, pilot school principals and teachers, and national coordinators to guide revisions. Readers and mathematicians on the development teams praised the assessments' alignment with Grade 2 curriculum standards, the collaborative design process, and use of Tangerine software. They expressed concerns about assessment length—particularly for struggling students—and sustainability post-OECS support, and recommended shorter assessments, smaller working groups, enhanced standardization materials, and ongoing professional development, including dedicated post-analysis reviews.

Invigilators and assessors reported that professional development sessions and guides were valuable but could be better organized: they asked for earlier distribution of materials, clearer accommodation guidance, more concise scripts, and improved assessment layouts. Both groups noted student fatigue from assessment length and complexity, suggesting fewer passages and questions. Meanwhile, school leaders, teachers, and National Focal Points flagged logistical and procedural issues—such as inconvenient preparation schedules, underutilized or misleveled sample passages, and the need for more timely communication and better software for automatic error counting. They also recommended conducting assessments in familiar classrooms and offering training on bubble-sheet use. Despite these challenges, overall stakeholder sentiment was positive, with most learners appearing comfortable during the assessments.

While some points have been addressed during the Assessment Revision Workshop, a summary of the recommendations is included in Table 5 below:

Table 5: *Post-Pilot Recommendations*

Content	Administration	Professional Development
Reduce assessment length and complexity, particularly for struggling students.	Shorten and simplify scripts and administration guides.	Enhance collaboration through smaller working groups
Revise the format to make it easier for learners to navigate.	Provide guides and professional development sessions earlier in the process.	Improve scheduling to ensure teachers have enough time to prepare learners effectively.
Ensure sample passages are appropriate for Grade 2 reading levels.	Ensure clearer instructions on accommodations for students with special needs.	Transition online meetings to in-person sessions for increased participation
	Address the sustainability of assessments after OECS support ends	Ongoing training and professional development in

IV.   Quantitative Analysis

A key component of the pilot phase for the OECS Early Grade Reading Assessment (OERA) and OECS Early Grade Mathematics Assessment (OEMA) was the statistical validation of each instrument to ensure that they reliably measure the intended constructs. Internal validity and construct reliability analyses were conducted and shared with curriculum and assessment officers across the OECS to confirm that the assessments accurately reflect reading and mathematical competence in early primary learners (OECS Commission, 2022).

Although the instruments were designed for administration at the end of Grade 2, the pilot was administered to Grade 3 students—a population believed to best represent end-of-Grade 2 competencies due to the timing of the pilot window. Of 639 students assessed, 401 produced complete data sets from both Oral and Silent Reading components. Incomplete data were due to logistical issues such as electronic scoring delays in Antigua and Barbuda, privacy protocols in Dominica, and general absenteeism during field testing. A total of 333 students completed both Constructed and Selected Response sections.

In addition to overall test score reliability, item-level analysis such as distractor analysis was also performed on the Selected Response items for both OERA and OEMA, following international best practices for early grade assessment (UNESCO UIS, 2019; RTI International, 2016). Table 6 below details of the distribution of sample students disaggregated by school and gender:

Table 6: *Disaggregation Of Student Expected and Reported Sample Count By School And Gender*

Table 9: Results of the Pilot Administration

Statistic	OERA	OEMA
# of students	402	333
# of questions	48	45
min value	4	14
max value	42	50
median	33	33
mean	32.75	32.74
std. dev.	9.54	7.41
SEM	0.02	0.02
variance	90.71	54.78
skew	0.79	-0.11
KR-20	0.91	0.86
KR-21	0.90	0.86
α	0.91	0.82

Country	Expected Sample				Reported Sample			
	Schools	Males	Females	Total	Schools	Males	Females	TOTAL
Anguilla	2	13	6					
Antigua and Barbuda	7	37	36					
British Virgin Islands	3	19	20					
Dominica	6	37	29					
Grenada	7	41	49					
St. Kitts and Nevis	1	26	28	54	1	24	25	49
St. Lucia	7	83	66	149	6	63	39	102
St. Vincent and the Grenadines	14	74	75	149	13	55	53	108
TOTAL	47	330	309	639	36	215	186	401

Test Validity and Reliability

Three key reliability coefficients were calculated to assess internal consistency:

- Kuder-Richardson Formula 20 (KR-20) evaluates the consistency of dichotomously scored items and accounts for item variance.
- Kuder-Richardson Formula 21 (KR-21) provides a simplified reliability estimate assuming equal item difficulty.
- Cronbach's Alpha (α) measures the average inter-item correlation to assess the internal coherence of the instrument.

Results for the OERA yielded coefficients above 0.90 for all three metrics, signifying excellent internal consistency (Cortina, 1993). Similarly, reliability coefficients for the OEMA exceeded 0.80, indicating strong construct reliability suitable for cross-national comparison and future longitudinal analysis.

These results affirm that both instruments are robust in their design and measurement capacity, making them appropriate for high-quality, system-level assessment aligned with Sustainable Development Goal 4.1.1(a) (UN SDG Indicators, 2024).

Item Analysis

A detailed item analysis was undertaken to further evaluate the quality and effectiveness of the OECS Early Grade Mathematics Assessment (OEMA) and the Early Grade Reading Assessment (OERA). This analysis focused on three key metrics widely recognized in educational measurement: item difficulty, discrimination index, and point-biserial correlation

Table 8: Item Analysis of OERA

Item ID	Difficulty Index		Discrimination Index		Point Biserial Correlation	
SF1	0.83	easy	0.14	poor	0.35	keep
SF2	0.84	easy	0.15	poor	0.40	keep
SF3	0.67	easy	0.19	poor	0.26	keep
SF4	0.64	easy	0.18	poor	0.12	improve
SF5	0.83	easy	0.14	poor	0.27	keep
SF7	0.79	easy	0.19	poor	0.30	keep
OB1	0.81	easy	0.19	poor	0.40	keep
P5	0.51	easy	0.17	poor	0.11	improve
P6	0.50	easy	0.17	poor	0.05	improve
SD1	0.88	easy	0.12	poor	0.47	keep
SD3	0.90	easy	0.10	poor	0.52	keep
SD4	0.61	easy	0.19	poor	0.33	keep
SD6	0.52	easy	0.19	poor	0.33	keep
SD7	0.82	easy	0.17	poor	0.52	keep
SD8	0.82	easy	0.17	poor	0.52	keep

These indices provide critical insights into how well individual test items function in differentiating student performance and contribute to overall test reliability (Haladyna & Rodriguez, 2013).

Item Difficulty Index

The item difficulty index is calculated as the proportion of students who answered an item correctly, ranging from 0.00 to 1.00. Items with values:

- Below 0.30 are considered difficult,
- Between 0.30 and 0.70 are moderate, and
- Above 0.70 are easy.
- Items exceeding 0.95 are categorized as very easy, possibly indicating limited diagnostic value

Table 8: Item Analysis of OERA

Item ID	Difficulty Index		Discrimination Index		Point Biserial Correlation	
SF1	0.83	easy	0.14	poor	0.35	keep
SF2	0.84	easy	0.15	poor	0.40	keep
SF3	0.67	easy	0.19	poor	0.26	keep
SF4	0.64	easy	0.18	poor	0.12	improve
SF5	0.83	easy	0.14	poor	0.27	keep
SF7	0.79	easy	0.19	poor	0.30	keep
OB1	0.81	easy	0.19	poor	0.40	keep
P5	0.51	easy	0.17	poor	0.11	improve
P6	0.50	easy	0.17	poor	0.05	improve
SD1	0.88	easy	0.12	poor	0.47	keep
SD3	0.90	easy	0.10	poor	0.52	keep
G5	0.67	easy	0.17	poor	0.30	keep
M4	0.18	difficult	0.17	poor	0.26	keep

Item Discrimination Index

The discrimination index determines how well an item distinguishes between high- and low-performing students. This was calculated by comparing the performance of the top 27% and bottom 27% of students on each item. Items with discrimination values below 0.20 were flagged as underperforming and marked for review and possible revision or elimination (Popham, 2008).

Table 9: Item Analysis of OEMA

Item ID	Difficulty Index		Discrimination Index		Point Biserial Correlation	
Q1	0.89	easy	0.09	poor	0.30	keep
Q2	0.62	moderate	0.19	poor	0.33	keep
Q7	0.71	easy	0.18	poor	0.35	keep
Q8	0.43	moderate	0.16	poor	0.24	keep
Q11	0.41	moderate	0.09	poor	0.13	improve
Q12	0.73	easy	0.18	poor	0.34	keep
Q13	0.66	moderate	0.17	poor	0.32	keep
Q14	0.71	easy	0.19	poor	0.38	keep
Q17	0.42	moderate	0.19	poor	0.32	keep
Q18	0.74	easy	0.17	poor	0.36	keep
Q19	0.33	moderate	0.06	poor	0.12	improve
Q20	0.73	easy	0.17	poor	0.34	keep
Q21	0.77	easy	0.19	poor	0.41	keep
Q23	0.71	easy	0.16	poor	0.29	keep
Q26	0.44	moderate	0.11	poor	0.17	improve
Q28	0.86	easy	0.13	poor	0.36	keep
Q31	0.29	difficult	0.07	poor	0.17	improve
Q33	0.53	moderate	0.17	poor	0.28	keep
Q35	0.48	moderate	0.18	poor	0.27	keep
Q1A (1)	0.93	easy	0.04	very poor	0.14	improve
Q1B (2)	0.72	easy	0.16	poor	0.35	keep
Q3B (1)	0.83	easy	0.13	poor	0.29	keep

Point-Biseral Correlation

The point-biserial correlation assesses the strength of the relationship between performance on an individual item and the overall test score. Items with a value below 0.20 were flagged for review, while negative values would indicate a potential issue with item validity, such as miskeying or misinterpretation by students (Downing, 2005).

All flagged items from both the OERA and OEMA were reviewed during a three-day Assessment Revision Workshop convened with regional education officers and subject matter experts. The workshop focused on:

Table 9: Item Analysis of OEMA

Item ID	Difficulty Index		Discrimination Index		Point Biserial Correlation	
Q1	0.89	easy	0.09	poor	0.30	keep
Q2	0.62	moderate	0.19	poor	0.33	keep
Q7	0.71	easy	0.18	poor	0.35	keep
Q8	0.43	moderate	0.16	poor	0.24	keep
Q11	0.41	moderate	0.09	poor	0.13	improve
Q12	0.73	easy	0.18	poor	0.34	keep
Q4A (1)	0.94	easy	0.05	poor	0.16	improve
Q4B (1)	0.97	very easy	0.02	poor	0.12	improve
Q5A (2)	0.58	moderate	0.14	poor	0.35	keep
Q5B (1)	0.86	easy	0.11	poor	0.27	keep
Q6 (2)	0.75	easy	0.19	poor	0.47	keep

subject matter experts. The workshop focused on:

- Conducting distractor analysis for items with weak discrimination and low point-biserial correlations.
- Revising item stems and distractors where needed for clarity or content alignment.
- Improving cognitive rigor and representation across difficulty levels to align better with OECS Learning Standards.

This systematic approach ensures that each item on the assessments not only aligns with curriculum expectations but also functions statistically to support valid and reliable interpretation of student performance (AERA, APA & NCME, 2014).

V.   Workshop Revisions

Data collected from the pilot administration of the OERA and OEMA were not used as a measure of student learning, but rather served as a validation tool for the assessment instruments themselves. The primary objective of the pilot was to identify and address limitations related to test design, item quality, alignment to curriculum standards, and practicality of implementation across OECS Member States. Through critical review of the pilot results, the regional assessment team identified specific elements in need of revision, proposing improvements related to item wording, formatting, content balance, cultural relevance, and test duration.

Workshop participants focused on evaluating the OERA in terms of:

- Curriculum alignment with the OECS Harmonized Primary Curriculum (OHPC)

- Cognitive demand of items and alignment to Bloom's revised taxonomy
- Item structure and quality
- Fairness and cultural relevance
- Assessment length and student fatigue
- Administration logistics and feasibility
- Global Proficiency Framework (GPF) alignment for reporting to the UNESCO Institute for Statistics (UIS)

Special attention was given to ensuring that the assessment met the minimum requirements for global learning reporting, particularly for Sustainable Development Goal 4.1.1a.

As shown in Table 10, significant revisions were made to the OERA. In response to challenges related to the administration and scoring of the Oral Reading component, this portion was eliminated from the final version. The final assessment now consists solely of Silent Reading Comprehension passages, each carefully selected and revised to reflect international Grade 2 reading standards in terms of text complexity, vocabulary load, and thematic relevance.

Table 10: *Summary of OERA Revisions*

OERA	Pilot	Final
Number of Passages	6	3
Genres	4	3
Number of Item	48	21
Time Allotment	>3 hr	1 hr

As reflected in Table 11, the pilot OEMA was revised to:Remove redundant or low-performing items

- Increase the clarity of item stems and distractors
- Improve alignment with Grade 2 essential learning outcomes
- Reduce student fatigue and total test time

Table 11: *Summary of OEMA Revisions*

OERA	Pilot	Final
Number of Papers	2	2
Number of Items	42	27
Time Allotment	>3 hr	1.5 hr

Reductions were based on both statistical analyses and expert consensus during the Assessment Revision Workshop. The revised OEMA maintains two components—Constructed Response and Selected Response—with a sharper focus on assessing factual, conceptual, and procedural knowledge across all six mathematics strands. Detailed item-level revisions are available in Appendices C and D, including item replacements, cognitive level adjustments, and revised illustrations and formatting.

References

ACER. (2018). *International Standards for Early Grade Assessment Development*. Australian Council for Educational Research.

ACER. (2021). *OECS Early Grade Reading Assessment: Technical Report on Pilot Administration*. Australian Council for Educational Research.

OECS Commission. (2021). *OECS Harmonized Primary Curriculum: Mathematics and Language Arts*. Organisation of Eastern Caribbean States.

OECS Commission. (2022). *OECS Curriculum and Assessment Framework*. Organisation of Eastern Caribbean States.

OECS Commission. (2023). *OECS Early Grade Reading and Mathematics Assessments (OERA and OEMA): Pilot and Final Reports*. Organisation of Eastern Caribbean States.

UIS. (2018). *Global Framework for Measuring Learning Outcomes*. UNESCO Institute for Statistics.

UIS. (2019). *Report on SDG 4.1.1a: Global Monitoring of Learning Outcomes*. UNESCO Institute for Statistics.

UIS. (2020). *Global Proficiency Frameworks for Reading and Mathematics*. UNESCO Institute for Statistics.

UIS. (2022). *Global Minimum Proficiency Levels (MPLs) for Early Education Assessment*. UNESCO Institute for Statistics.

USAID & RTI International. (2017). *EGRA Toolkit for Early Grade Reading Assessment*. U.S. Agency for International Development and Research Triangle Institute.

USAID & RTI International. (2021). *Guidance on Culturally Responsive and Fair Assessments*. U.S. Agency for International Development and Research Triangle Institute.

World Bank. (2019). *Learning Outcomes and Education Equity: Advancing Educational Quality through Data and Accountability*. The World Bank Group.

Appendix A: GPF Alignment and Psychometric Analysis of Pilot OERA

Table 12: Pilot OERA Item Statistics

Item ID	GPL Alignment		Difficulty Index		Discrimination Index		Point Biserial Correlation	
SF1	2	R1.2.1_P	0.83	easy	0.14	poor	0.35	keep
SF2	2	R1.2.1_E	0.84	easy	0.15	poor	0.40	keep
SF3	2	R1.2.1_E	0.67	easy	0.19	poor	0.26	keep
SF4	2	Not aligned	0.64	easy	0.18	poor	0.12	improve
SF5	2	R1.2.1_E	0.83	easy	0.14	poor	0.27	keep
SF6	2	R1.1.1_E	0.62	easy	0.29	good	0.38	keep
SF7	2	R1.1.1_E	0.79	easy	0.19	poor	0.30	keep
OB1	4	R1.2.1_P	0.81	easy	0.19	poor	0.40	keep
OB2	4	R1.2.1_M	0.70	easy	0.26	good	0.25	keep
OB3	4	R2.2.1_P	0.64	easy	0.27	good	0.37	keep
OB4	4	Not aligned	0.63	easy	0.22	acceptable	0.28	keep
OB5	4	R1.2.1_M	0.69	easy	0.22	acceptable	0.35	keep
OB6	4	R1.2.1_P	0.72	easy	0.24	acceptable	0.27	keep
OB7	4	R1.3.1_M	0.71	easy	0.27	good	0.41	keep
OB8	4	R2.2.1_E	0.65	easy	0.29	good	0.48	keep
P1	3	R1.2.1_E	0.67	easy	0.22	acceptable	0.38	keep
P2	3	R2.1.1_E	0.70	easy	0.28	good	0.34	keep
P3	3	R1.2.1_M	0.61	easy	0.32	good	0.41	keep
P4	3	R1.3.1_E	0.58	easy	0.28	good	0.26	keep
P5	3	R1.2.1_M	0.51	easy	0.17	poor	0.11	improve
P6	3	Not aligned	0.50	easy	0.17	poor	0.05	improve
P7	3	R2.1.1_E	0.53	easy	0.21	acceptable	0.21	keep
P8	3	R2.3.1_E	0.59	easy	0.30	good	0.32	keep
SD1	3	R1.2.1_M	0.88	easy	0.12	poor	0.47	keep
SD2	3	R1.2.1_M	0.74	easy	0.21	acceptable	0.50	keep
SD3	3	R1.2.1_E	0.90	easy	0.10	poor	0.52	keep
SD4	3	Not aligned	0.61	easy	0.19	poor	0.33	keep
SD5	3	Not aligned	0.61	easy	0.27	good	0.41	keep
SD6	3	Not aligned	0.52	easy	0.19	poor	0.33	keep

SD7	3	R1.1.1_M	0.82	easy	0.17	poor	0.52	keep
SD8	3	R2.3.1_M	0.82	easy	0.17	poor	0.52	keep
G1	3	R2.2.1_M	0.64	easy	0.25	good	0.48	keep
G2	3	R2.2.1_P	0.62	easy	0.37	acceptable	0.66	keep
G3	3	R1.2.1_E	0.53	easy	0.26	good	0.46	keep
G4	3	R1.2.1_E	0.74	easy	0.26	good	0.66	keep
G5	3	R1.1.1_M	0.67	easy	0.17	poor	0.30	keep
G6	3	R2.2.1_E	0.67	easy	0.23	acceptable	0.57	keep
G7	3	Not aligned	0.51	easy	0.30	good	0.47	keep
G8	3	R1.2.1_E	0.51	easy	0.31	good	0.49	keep
G8b	3	R2.2.1_E	0.66	easy	0.28	good	0.49	keep
M1	3	R1.2.1_P	0.67	easy	0.30	good	0.66	keep
M2	3	R1.1.1_M	0.67	easy	0.32	good	0.53	keep
M3	3	R2.1.1_M	0.48	easy	0.37	acceptable	0.49	keep
M4	3	Not aligned	0.18	difficult	0.17	poor	0.26	keep
M5	3	R1.2.1_M	0.73	easy	0.27	good	0.70	keep
M6	3	R1.1.1_M	0.64	easy	0.27	good	0.55	keep
M7	3	R2.2.1_P	0.54	easy	0.35	acceptable	0.51	keep
M8	3	R2.2.1_E	0.58	easy	0.31	good	0.49	keep

Table 13: *Pilot OERA Distractor Analysis*

OERA Silent Reading			Difficulty Index			Discrimination Index			Point Biserial Correlation		
#	Key	GPL Alignment	A	B	C	A	B	C	A	B	C
SF1	C	2 R1.2.1_P	0.099	0.052	**0.830**	-0.085	-0.028	0.165	-0.337	-0.188	0.433
SF2	B	2 R1.2.1_E	0.075	**0.816**	0.075	-0.061	0.189	-0.061	-0.271	**0.467**	-0.266
SF3	C	2 R1.2.1_E	0.094	0.212	**0.670**	-0.061	-0.075	0.193	-0.220	-0.170	**0.339**
SF4	B	2 Not aligned	0.146	**0.613**	0.198	-0.028	**0.241**	-0.137	-0.113	**0.372**	-0.250
SF5	B	2 R1.2.1_E	0.137	**0.797**	0.052	-0.113	0.198	-0.038	-0.310	**0.421**	-0.167
SF6	A	2 R1.1.1_E	**0.571**	0.226	0.189	**0.321**	-0.175	-0.099	**0.554**	-0.394	-0.240
SF7	B	2 R1.1.1_E	0.127	**0.745**	0.108	-0.104	**0.241**	-0.085	-0.342	**0.492**	-0.253
OB1	A	4 R1.2.1_P	**0.807**	0.099	0.061	0.198	-0.085	-0.047	**0.497**	-0.309	-0.253
OB2	A	4 R1.2.1_M	**0.656**	0.175	0.151	**0.321**	-0.160	-0.108	**0.523**	-0.370	-0.239
OB3	B	4 R2.2.1_P	0.099	**0.637**	0.231	-0.066	**0.292**	-0.160	-0.259	**0.505**	-0.320
OB4	B	4 Not aligned	0.236	**0.623**	0.104	-0.146	**0.269**	-0.052	-0.269	**0.448**	-0.215
OB5	C	4 R1.2.1_M	0.170	0.189	**0.623**	-0.146	-0.137	**0.335**	-0.419	-0.290	**0.582**
OB6	B	4 R1.2.1_P	0.132	**0.689**	0.165	-0.118	**0.307**	-0.151	-0.284	**0.507**	-0.344
OB7	C	4 R1.3.1_M	0.151	0.132	**0.698**	-0.127	-0.118	**0.297**	-0.334	-0.319	**0.548**
OB8	C	4 R2.2.1_E	0.118	0.226	**0.618**	-0.094	-0.193	**0.358**	-0.306	-0.415	**0.630**
P1	C	3 R1.2.1_E	0.146	0.217	**0.623**	-0.094	-0.108	**0.250**	-0.257	-0.274	**0.457**
P2	A	3 R2.1.1_E	**0.637**	0.208	0.146	**0.340**	-0.175	-0.123	**0.565**	-0.348	-0.333
P3	A	3 R1.2.1_M	**0.599**	0.165	0.212	**0.377**	-0.142	-0.179	**0.565**	-0.245	-0.378
P4	C	3 R1.3.1_E	0.264	0.137	**0.580**	-0.137	-0.104	**0.283**	-0.257	-0.244	**0.449**
P5	B	3 R1.2.1_M	0.354	**0.467**	0.151	-0.094	**0.255**	-0.099	-0.121	**0.382**	-0.274
P6	C	3 Not aligned	0.274	0.179	**0.519**	-0.052	-0.090	**0.203**	-0.081	-0.251	**0.322**
P7	C	3 R2.1.1_E	0.217	0.255	**0.486**	-0.146	-0.108	**0.321**	-0.253	-0.177	**0.437**
P8	B	3 R2.3.1_E	0.189	**0.590**	0.160	-0.156	**0.321**	-0.090	-0.339	**0.505**	-0.202

Appendix B: GPF Alignment and Psychometric Analysis of Pilot OEMA

Table 14: Pilot OEMA Item Statistics

Item ID	GPL Alignment		Difficulty Index		Discrimination Index		Point Biserial Correlation	
Q1A (1)	2	N1.3.1_E	0.93	easy	0.04	very poor	0.14	improve
Q1B (2)	-	N/A	0.72	easy	0.16	poor	0.35	keep
Q1	2	N1.1.1b_P	0.89	easy	0.09	poor	0.30	keep
Q2	2	N1.3.1_M	0.62	moderate	0.19	poor	0.33	keep
Q3	3	N2.1.1_P	0.63	moderate	0.21	acceptable	0.33	keep
Q4	3	N2.1.1_P	0.74	easy	0.22	acceptable	0.39	keep
Q5	2	N1.1.2_P	0.74	easy	0.22	acceptable	0.46	keep
Q6	2	N1.1.2_M	0.69	moderate	0.21	acceptable	0.35	keep
Q7	2	N1.1.2_M	0.71	easy	0.18	poor	0.35	keep
Q2 (3)	2	N1.3.2_E	0.70	moderate	0.24	acceptable	0.50	keep
Q8	2	N1.4.1_M	0.43	moderate	0.16	poor	0.24	keep
Q9	3	N1.3.3_P	0.75	easy	0.22	acceptable	0.41	keep
Q10	2	N1.3.2_E	0.66	moderate	0.30	good	0.49	keep
Q11	2	N1.3.1_E	0.41	moderate	0.09	poor	0.13	improve
Q12	4	A3.2.3a_M	0.73	easy	0.18	poor	0.34	keep
Q13	2	N1.4.1_E	0.66	moderate	0.17	poor	0.32	keep
Q14	2	N1.3.7_M	0.71	easy	0.19	poor	0.38	keep
Q3A (1)	4	A1.1.2_M	0.66	moderate	0.32	good	0.51	keep
Q3B (1)	4	A1.1.2_M	0.83	easy	0.13	poor	0.29	keep
Q15	2	N1.1.2_M	0.42	moderate	0.25	good	0.41	keep
Q16	2	A1.1.1_M	0.79	easy	0.20	acceptable	0.43	keep
Q17	2	N1.1.2_M	0.42	moderate	0.19	poor	0.32	keep
Q4A (1)	2	G2.1.1_M	0.94	easy	0.05	poor	0.16	improve
Q4B (1)	2	G2.1.1_M	0.97	very easy	0.02	poor	0.12	improve
Q18	2	G1.1.1_E	0.74	easy	0.17	poor	0.36	keep
Q19	4	G1.1.2_E	0.33	moderate	0.06	poor	0.12	improve
Q20	2	G2.1.1_M	0.73	easy	0.17	poor	0.34	keep
Q21	3	G1.1.2_E	0.77	easy	0.19	poor	0.41	keep
Q22	2	G1.1.1_E	0.71	easy	0.24	acceptable	0.45	keep
Q23	4	G1.1.2_E	0.71	easy	0.16	poor	0.29	keep
Q24	3	G3.1.1_M	0.56	moderate	0.18	poor	0.31	keep
Q5A (2)	2	M1.1.1a_M	0.58	moderate	0.14	poor	0.35	keep
Q5B (1)	2	M1.1.2_E	0.86	easy	0.11	poor	0.27	keep
Q25	2	M2.1.2_E	0.44	moderate	0.23	acceptable	0.34	keep

Q26	4	M1.1.2b_M	0.44	moderate	0.11	poor	0.17	improve
Q27	2	M1.1.2_E	0.73	easy	0.20	acceptable	0.41	keep
Q28	2	M1.1.2_E	0.86	easy	0.13	poor	0.36	keep
Q29	4	M1.2.1_P	0.49	moderate	0.31	good	0.52	keep
Q30	2	M3.1.1a_M	0.54	moderate	0.29	good	0.47	keep
Q31	4	M1.2.1_P	0.29	difficult	0.07	poor	0.17	improve
Q6 (2)	2	S1.1.2_M	0.75	easy	0.19	poor	0.47	keep
Q32	2	S1.1.2_M	0.47	moderate	0.37	excellent	0.58	keep
Q33	-	N/A	0.53	moderate	0.17	poor	0.28	keep
Q34	2	S1.1.2_M	0.45	moderate	0.22	acceptable	0.36	keep
Q35	5	S2.1.1_P	0.48	moderate	0.18	poor	0.27	keep

Table 15: *Pilot OEMA Distractor Analysis*

OEMA Constructed Response				Difficulty Index			Discrimination Index			Point Biserial Correlation		
#	Key	GPL Alignment		A	B	C	A	B	C	A	B	C
Q1	B	2	N1.1.1b_P	0.047	**0.896**	0.038	-0.048	**0.115**	-0.038	-0.161	**0.369**	-0.166
Q2	A	2	N1.3.1_M	**0.580**	0.142	0.245	**0.216**	-0.077	-0.106	**0.374**	-0.206	-0.156
Q3	A	3	N2.1.1_P	**0.646**	0.198	0.127	**0.168**	-0.067	-0.072	**0.298**	-0.090	-0.189
Q4	C	3	N2.1.1_P	0.090	0.179	**0.708**	-0.091	-0.135	**0.260**	-0.247	-0.201	**0.415**
Q5	A	2	N1.1.2_P	**0.722**	0.151	0.104	**0.245**	-0.125	-0.087	**0.490**	-0.294	-0.224
Q6	B	2	N1.1.2_M	0.099	**0.712**	0.175	-0.082	**0.255**	-0.149	-0.208	**0.440**	-0.266
Q7	B	2	N1.1.2_M	0.156	**0.717**	0.113	-0.111	**0.202**	-0.067	-0.194	**0.353**	-0.165
Q8	C	2	N1.4.1_M	0.160	0.406	**0.401**	-0.125	0.029	**0.139**	-0.236	0.058	**0.234**
Q9	A	3	N1.3.3_P	**0.703**	0.160	0.123	**0.255**	-0.125	-0.106	**0.463**	-0.260	-0.232
Q10	B	2	N1.3.2_E	0.250	**0.618**	0.118	-0.216	**0.322**	-0.082	-0.347	**0.515**	-0.190
Q11	A	2	N1.3.1_E	**0.392**	0.387	0.198	**0.072**	0.067	-0.106	**0.174**	0.102	-0.198
Q12	B	4	A3.2.3a_M	0.108	**0.670**	0.193	-0.053	**0.250**	-0.159	-0.131	**0.437**	-0.302
Q13	C	2	N1.4.1_E	0.127	0.198	**0.642**	-0.111	-0.106	**0.240**	-0.189	-0.241	**0.429**
Q14	B	2	N1.3.7_M	0.217	**0.627**	0.142	-0.154	**0.255**	-0.077	-0.294	**0.450**	-0.158
Q15	C	2	N1.1.2_M	0.420	0.099	**0.467**	-0.072	-0.101	**0.197**	-0.114	-0.254	**0.348**
Q16	A	2	A1.1.1_M	**0.764**	0.118	0.099	**0.231**	-0.101	-0.101	**0.464**	-0.208	-0.291
Q17	A	2	N1.1.2_M	**0.392**	0.226	0.354	**0.168**	-0.038	-0.091	**0.309**	-0.082	-0.111
Q18	C	2	G1.1.1_E	0.057	0.175	**0.698**	-0.048	-0.091	**0.221**	-0.173	-0.149	**0.427**
Q19	A	4	G1.1.2_E	**0.358**	0.325	0.241	**0.087**	0.043	-0.053	**0.167**	0.098	-0.074
Q20	B	2	G2.1.1_M	0.146	**0.670**	0.099	-0.082	**0.250**	-0.082	-0.133	**0.436**	-0.184
Q21	A	3	G1.1.2_E	**0.708**	0.090	0.123	**0.269**	-0.082	-0.096	**0.532**	-0.233	-0.195
Q22	C	2	G1.1.1_E	0.127	0.137	**0.670**	-0.091	-0.082	**0.250**	-0.222	-0.144	**0.456**
Q23	A	4	G1.1.2_E	**0.755**	0.080	0.080	**0.212**	-0.063	-0.063	**0.446**	-0.122	-0.187
Q24	C	3	G3.1.1_M	0.193	0.241	**0.491**	-0.111	-0.043	**0.240**	-0.164	-0.040	**0.371**
Q25	B	2	M2.1.2_E	0.283	**0.425**	0.222	-0.135	**0.240**	-0.024	-0.138	**0.383**	-0.063
Q26	C	4	M1.1.2b_M	0.198	0.245	**0.481**	-0.048	-0.038	**0.163**	-0.055	-0.027	**0.261**
Q27	C	2	M1.1.2_E	0.146	0.118	**0.646**	-0.130	-0.072	**0.293**	-0.257	-0.135	**0.506**
Q28	A	2	M1.1.2_E	**0.759**	0.071	0.094	**0.236**	-0.063	-0.096	**0.538**	-0.181	-0.255
Q29	B	4	M1.2.1_P	0.212	**0.434**	0.283	-0.063	**0.298**	-0.154	-0.119	**0.504**	-0.216

Q30	C	2	M3.1.1a_M	0.208	0.175	**0.547**	-0.163	-0.072	**0.317**	-0.223	-0.118	**0.476**
Q31	C	4	M1.2.1_P	0.425	0.203	**0.302**	0.048	-0.082	**0.115**	0.093	-0.118	**0.238**
Q32	A	2	S1.1.2_M	**0.472**	0.189	0.264	**0.346**	-0.125	-0.144	**0.549**	-0.191	-0.203
Q33	B	-	N/A	0.212	**0.538**	0.165	-0.053	**0.173**	-0.034	-0.096	**0.306**	-0.041
Q34	A	2	S1.1.2_M	**0.453**	0.090	0.250	**0.269**	-0.082	-0.082	**0.390**	-0.192	-0.092
Q35	B	5	S2.1.1_P	0.094	**0.491**	0.198	-0.087	**0.250**	-0.048	-0.164	**0.367**	-0.054

Appendix C: *Revision Tables for OERA*

Table 16: *OERA Visual Passage*

Pilot Items					Field Test Items					Final
#	DIFF	DIS	RPBIS	Changes	#	DIFF	DIS	RPBIS	Changes	#
1	0.83	0.14	0.35	item choices	1	0.89	0.04	-0.03		1
2	0.84	0.15	0.40	item choices	2	0.89	0.11	0.49		2
3	0.67	0.19	0.26	item choices	3	0.89	0.11	0.49		3
4	0.64	0.18	0.12	image	4	0.86	0.14	0.54		4
5	0.83	0.14	0.27	item choices	5	0.93	0.07	0.56		5
6	0.62	0.29	0.38		6	0.86	0.07	0.12		6
7	0.79	0.19	0.30	item choices	7	0.71	0.14	0.38		7
8	n/A	n/a	n/a	item choices	8	0.89	0.04	0.59		8

Table 17: *OERA Fiction Passage*

Pilot Items					Field Test Items					Final
#	DIFF	DIS	RPBIS	Changes	#	DIFF	DIS	RPBIS	Changes	#
1	0.81	0.19	0.40	*new item*	1	0.96	0.04	0.75		1
2	0.70	0.26	0.25	*new item*	2	0.93	0.07	0.56		2
3	0.64	0.27	0.37	*new item*	3	0.82	0.04	-0.03		3
4	0.63	0.22	0.28	*new item*	4	0.86	0.07	0.62		4
5	0.69	0.22	0.35	*new item*	5	0.93	0.00	0.17		5
6	0.72	0.24	0.27	*new item*	6	0.64	0.14	0.11		6
7	0.71	0.27	0.41	*new item*	7	0.86	0.14	0.79		7
8	0.65	0.29	0.48	*new item*	8	0.71	0.14	0.51		8

Table 18: *OERA Expository Passage*

Pilot Items					Field Test Items					Final
#	DIFF	DIS	RPBIS	Changes	#	DIFF	DIS	RPBIS	Changes	#
1	0.81	0.19	0.40		1	1.00	0.00	#DIV/0!		1
2	0.70	0.26	0.25		4	0.89	0.11	0.68		2
3	0.64	0.27	0.37		3	0.96	0.04	0.27		3
8	0.65	0.29	0.48		2	1.00	0.00	#DIV/0!		8
				new item	5	0.96	0.04	0.35		
4	0.63	0.22	0.28	removed						

5	0.69	0.22	0.35	removed
6	0.72	0.24	0.27	removed
7	0.71	0.27	0.41	removed

Appendix D: *Revision Tables for OEMA*

Table 19: Summary of Revision of OEMA Selected Response Items

	Pilot Items					Field Test Items				Final
#	DIFF	DIS	RPBIS	Changes	#	DIFF	DIS	RPBIS	Changes	#
1	0.89	0.09	0.30	formatting item choices	1	1.00	0.00	#DIV/0!		1
3	0.63	0.21	0.33		2	0.72	0.06	0.30		2
4	0.74	0.22	0.39	question stem item choices	3	0.83	0.17	0.65		3
6	0.69	0.21	0.35		5	0.83	0.06	0.16		4
7	0.71	0.18	0.35	question stem item choices	6	0.56	0.22	0.59		5
8	0.43	0.16	0.24	formatting item choices	7	0.17	0.17	0.37		6
9	0.75	0.22	0.41	image	8	0.83	0.06	0.30		7
10	0.66	0.30	0.49	image	9	0.94	-0.06	-0.08		8
11	0.41	0.09	0.13	diagram	10	0.67	0.22	0.43		9
12	0.73	0.18	0.34	question stem	11	0.22	0.00	-0.10		10
13	0.66	0.17	0.32	question stem	12	0.56	0.22	0.61		11
14	0.71	0.19	0.38	question stem	13	0.61	0.28	0.58		12
16	0.79	0.20	0.43	formatting	14	0.94	0.06	0.44		13
17	0.42	0.19	0.32	diagram	15	0.28	0.06	0.15		14
21	0.77	0.21	0.41	question stem item choices	17	0.67	0.22	0.68		15
23	0.71	0.16	0.29	item choices	19	0.67	0.33	0.85		16
24	0.56	0.18	0.31	diagram	20	0.50	0.39	0.59		17
25	0.44	0.23	0.34		21	0.56	0.00	0.07		18
26	0.44	0.11	0.17	question stem item choices	22	0.44	0.22	0.29		19
27	0.73	0.19	0.41	formatting	23	0.78	0.22	0.63		20
29	0.49	0.31	0.52	formatting	25	0.61	0.28	0.52		21
30	0.54	0.29	0.47	formatting	26	0.61	0.17	0.52	image	22
31	0.29	0.07	0.17	image	27	0.67	0.11	0.22		23
32	0.47	0.37	0.58		28	0.50	0.50	0.87		24
35	0.48	0.18	0.27		30	0.39	0.06	0.32		25
22	0.71	0.24	0.45		18	0.83	0.06	0.43	removed	
28	0.86	0.13	0.36		24	0.83	0.17	0.57	removed	
34	0.45	0.19	0.36	diagram	29	0.33	0.22	0.44	removed	
5	0.74	0.22	0.46		4	0.78	0.22	0.63	removed	
18	0.74	0.17	0.36		16	0.72	0.17	0.56	removed	

2	0.62	0.19	0.33	removed
15	0.42	0.25	0.41	removed
19	0.33	0.06	0.12	removed
20	0.73	0.17	0.34	removed
33	0.53	0.17	0.28	removed

Table 20: *Summary of Revision of OEMA Constructed Response Items*

Pilot Items					Field Test Items					Final
#	DIFF	DIS	RPBIS	Changes	#	DIFF	DIS	RPBIS	Changes	#
CR1A	0.93	0.04	0.14	diagram question stem	CR1A	0.36	0.21	0.63		CR1A
CR1B	0.72	0.16	0.35	diagram question stem	CR1B	0.75	0.36	0.61		CR1B
CR2	0.70	0.24	0.50	diagram question stem	CR2A	0.64	0.43	0.65		CR2A
					CR2B	0.21	0.29	0.64	question stem	CR2B
CR3A	0.66	0.32	0.51	diagram	CR3A	0.57	0.29	0.73	diagram	CR3A
CR3B	0.83	0.13	0.29	question stem	CR3B	0.89	0.21	0.70	diagram	CR3B
CR4A	0.94	0.05	0.16	question stem	CR4A	0.57	0.29	0.61		CR4A
CR4B	0.97	0.02	0.12		CR4B	1.00	0.00	#DIV/0!		CR4B
CR5A	0.58	0.14	0.35	question stem	CR5A	0.64	0.21	0.37	question stem	CR5A
CR5B	0.86	0.11	0.27	diagram	CR5B	1.00	0.00	#DIV/0!	question stem	CR5B
CR6	0.75	0.22	0.47	diagram	CR6	0.32	0.50	0.68	item added	CR6A
										CR6B