ORGANIZATION OF EASTER CARIBBEAN STATES (OECS)

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

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

ITEM ANALYSIS

25 November 2025

I. Introduction

This report presents the psychometric properties of the OECS Early Reading Assessment (OERA) and the OECS Early Mathematics Assessment (OEMA), administered across all English-speaking OECS Member States in 2025. The analyses were conducted to evaluate item quality, test reliability, and alignment with UIS technical standards for SDG 4.1.1a reporting. Key indicators examined include internal consistency, item difficulty, item discrimination, point-biserials, distractor functioning, and overall student performance. The report also incorporates post-administration review by subject matter experts to ensure interpretive validity and alignment with regional curriculum expectations.

II. Summary of Findings

The 2025 item analysis indicates that the OERA demonstrated strong technical performance, while the OEMA—though psychometrically acceptable—exhibited greater variability across items and domains, highlighting areas for targeted refinement. Both assessments remain appropriate for OECS SDG reporting and for guiding instructional decision-making at the classroom and system levels.

Domain	Reliability	Discrimination	Difficulty	Interpretation
OERA	Excellent ($\approx$0.90)	Strong (0.29 / 0.56)	Moderate (0.63)	Highly cohesive, high-quality item set
OEMA	Acceptable (0.79–0.86)	Mixed (0.20 / 0.39)	Variable	More heterogeneous; several items need revision

OERA

- Reliability is high, reflecting strong internal cohesion and robust alignment with SDG 4.1.1a reporting requirements.
- Item functioning is generally strong across all psychometric indicators.
- Only minor adjustments, such as rebalancing the difficulty of a small number of items, may be considered for future cycles.

OEMA

- Reliability is moderate, suggesting potential multidimensionality and variability in difficulty across mathematics strands.
- Three items exhibited discrimination and point-biseral values below the acceptable threshold of 0.2 and were excluded from the data set (Q6, Q10, Q4(a))
- Several items exhibited lower discrimination and will be targeted for revision in subsequent cycles to enhance measurement precision.

Overall, the reading assessment demonstrated a more unified construct with strong item cohesion, whereas the mathematics assessment reflected expected multidimensionality and greater variability in item quality. These findings support the continued use of both assessments for regional SDG reporting while guiding targeted improvements in mathematics item development.

III.   Methodology

Data were derived from all Grade 2 public-school students participating in the 2025 cycle of the OECS Early Grade Assessments. Both multiple-choice and constructed response items were included where applicable. Psychometric analyses followed standard UIS and ACER procedures:

- Classical Test Theory (CTT) indices calculated for all dichotomous and polytomous items (difficulty, discrimination, point-biserials).
- Internal consistency reliability estimated using Cronbach's α for mixed formats.
- Test-level performance statistics calculated to evaluate overall mastery and variation.
- Fleiss' Kappa (κ) was computed to assess inter-rater reliability, appropriate for items scored on a fixed rating scale.

Following live administration, subject matter experts reviewed item-level statistics, including difficulty, discrimination, point-biserials, distractor functioning, and gender DIF. Each item was evaluated for clarity, alignment with Grade 2 learning objectives, cognitive demand, and scoring consistency. Items flagged for low discrimination or distractor issues were discussed, with recommendations to retain, revise, or remove based on consensus. All decisions were documented to ensure transparency and alignment with UIS SDG 4.1.1a reporting standards.

OERA Results Summary

Test-Level Performance

- Mean score: 14.23 / 21 (~67.8%)
- Standard deviation: 5.49

Performance across the student population shows expected variation, indicating that most students achieved foundational mastery of early reading skills. The observed distribution reflects typical differences in ability within the Grade 2 cohort, without extreme clustering at either high or low ends.

Test Reliability

- Cronbach's α ≈ 0.90

This indicates high internal consistency, reflecting that items coherently measure the underlying reading construct. The high reliability supports the use of OERA scores for both population-level reporting and instructional decision-making.

Item-Level Performance

- Difficulty: 0.49–0.74 (mean = 0.63)

Items are moderately challenging, appropriate for Grade 2 students, and balanced to differentiate students across ability levels.

- Discrimination: 0.23–0.36 (mean = 0.29)

Most items effectively distinguish higher-performing from lower-performing students, supporting accurate ranking of student performance.

- Point-biserials: 0.42–0.66 (mean = 0.56)

Positive correlations between item scores and total scores confirm that items contribute meaningfully to the overall test measurement.

Gender DIF

- Differences between males and females were small and stable across performance bands.
- Female advantage modest (F: 55%, M: 50%).
- Item-level DIF mostly within ±3%; no evidence of systematic gender bias.

Overall, the OERA demonstrated strong psychometric properties: excellent reliability, effective discrimination, functional distractors, and minimal gender DIF. The assessment is well-suited for UIS population reporting and provides a reliable measure of early reading skills for Grade 2 students.

OEMA Results Summary

Test-Level Performance

- Mean score: 24.53 (weighted scale)

Performance reflected the integration of multiple-choice and constructed-response items across diverse mathematics domains. Scores indicated moderate mastery, with expected variation due to the multidimensional nature of mathematics skills assessed at Grade 2.

Test Reliability

- Cronbach's α = 0.85

The assessment demonstrated acceptable internal consistency. Slightly lower reliability than OERA reflects the heterogeneity of mathematics skills and the presence of multiple content strands within the test.

Item-Level Performance

- Difficulty: 0.40–0.82

The wide range of item difficulty ensured coverage of easier and more challenging content, supporting differentiation across student ability levels.

- Discrimination: 0.15–0.33 (mean $\approx$ 0.20)

Some items provided strong discrimination (e.g., Q24), while others (e.g., Q14, Q25) demonstrated weaker, though still functional, discrimination. Items with shallow discrimination have been flagged for refinement.

- Point-biserials: 0.26–0.51 (mean ≈ 0.39)

Moderate correlations suggest most items contributed adequately to the measurement of overall mathematics ability, though variability across items reflects multidimensionality.

Gender DIF

- Male and female performance patterns were similar across items.
- DIF values were within acceptable thresholds (<10 percentage points), with no items flagged for structural gender bias.

Inter-Rater Reliability

- Fleiss' Kappa: $\kappa = 0.784$ (SE = 0.0073, $p < .001$)
- Agreement strongest for full-credit responses ($\kappa = 0.918$) and mid-level responses ($\kappa = 0.807$); substantial for zero-score responses ($\kappa = 0.621$).
- Items with highly skewed distributions showed perfect observed agreement but limited κ contribution.
- Results confirm consistent application of scoring procedures and reliable differentiation of student performance.

Overall Interpretation

Overall, the OEMA demonstrated adequate reliability and generally functional item performance but greater variability across items and content strands than OERA. Several items require refinement, particularly those with shallow discrimination or weak distractors. Moderate reliability suggests possible multidimensionality or uneven strand difficulty.

Appendix A: 2025 OERA Item Analysis

TEST STATISTICS

Statistic	Value
# of students	6596
# of questions	21
min value	0
max value	21
mode	20
median	16
mean	14.39
std. dev.	5.21
variance	30.18
α	0.88

ITEM STATISTICS

Item ID	DIFF	DIS	RPBIS
Q1	0.70	0.26	0.50
Q2	0.60	0.32	0.50
Q3	0.67	0.29	0.53
Q4	0.65	0.31	0.53
Q5	0.70	0.28	0.52
Q6	0.60	0.33	0.53
Q7	0.63	0.25	0.43
Q8	0.74	0.23	0.46
Q9	0.69	0.30	0.60
Q10	0.69	0.31	0.62
Q11	0.62	0.32	0.50
Q12	0.64	0.35	0.65
Q13	0.63	0.34	0.61
Q14	0.49	0.25	0.36
Q15	0.66	0.33	0.60
Q16	0.62	0.27	0.42
Q17	0.65	0.33	0.62
Q18	0.65	0.34	0.65
Q19	0.63	0.36	0.66
Q20	0.62	0.34	0.60
Q21	0.65	0.29	0.48

DISTRACTOR ANALYSIS

Item ID	Key	Difficulty Index			Discrimination Index			Point Biserial Correlation		
		A	B	C	A	B	C	A	B	C
Q1	C	0.19	0.10	**0.70**	-0.17	-0.09	**0.26**	-0.17	-0.06	**0.23**
Q2	A	**0.58**	0.13	0.27	**0.32**	-0.12	-0.19	**0.12**	-0.11	0.00
Q3	C	0.15	0.17	**0.66**	-0.15	-0.14	**0.29**	-0.10	-0.08	**0.18**
Q4	A	**0.64**	0.15	0.19	**0.30**	-0.15	-0.16	**0.12**	-0.01	-0.10
Q5	C	0.14	0.16	**0.69**	-0.13	-0.14	**0.26**	-0.20	-0.01	**0.19**
Q6	A	**0.59**	0.24	0.16	**0.31**	-0.21	-0.11	**0.07**	-0.03	-0.02
Q7	B	0.21	**0.62**	0.15	-0.11	**0.24**	-0.14	-0.08	**0.15**	-0.08
Q8	C	0.13	0.13	**0.71**	-0.11	-0.12	**0.20**	-0.11	-0.11	**0.19**
Q9	B	0.19	**0.68**	0.12	-0.19	**0.30**	-0.11	-0.17	**0.23**	-0.06
Q10	A	**0.68**	0.17	0.13	**0.30**	-0.17	-0.13	**0.19**	-0.11	-0.07
Q11	C	0.21	0.16	**0.61**	-0.15	-0.16	**0.31**	-0.03	-0.05	**0.11**
Q12	C	0.13	0.22	**0.63**	-0.13	-0.21	**0.34**	-0.06	-0.08	**0.17**
Q13	B	0.20	**0.63**	0.16	-0.18	**0.32**	-0.15	-0.06	**0.18**	-0.11
Q14	C	0.30	0.19	**0.49**	-0.07	-0.17	**0.23**	0.05	-0.12	**0.09**
Q15	B	0.15	**0.64**	0.18	-0.15	**0.31**	-0.18	-0.07	**0.13**	-0.05
Q16	B	0.18	**0.61**	0.18	-0.12	**0.24**	-0.14	-0.09	**0.11**	-0.01
Q17	C	0.19	0.15	**0.64**	-0.19	-0.14	**0.32**	-0.21	0.00	**0.23**
Q18	C	0.13	0.21	**0.64**	-0.13	-0.21	**0.32**	-0.03	-0.15	**0.19**
Q19	B	0.13	**0.62**	0.23	-0.13	**0.34**	-0.22	-0.04	**0.19**	-0.12
Q20	C	0.21	0.17	**0.60**	-0.19	-0.15	**0.32**	-0.05	-0.08	**0.15**
Q21	B	0.16	**0.64**	0.17	-0.13	**0.26**	-0.15	-0.09	**0.16**	-0.06

Appendix B: 2025 OEMA Item Analysis (Post-Exclusions)

TEST STATISTICS

Statistic	Value
# of students	6645
# of questions	25
min value	2
max value	41
mode	22
median	24
mean	24.42
std. dev.	7.45
variance	55.49
Cronbach's α	0.85

ITEM STATISTICS

Item ID	DIFF	DIS	RPBIS
Q1	0.82	0.16	0.38
Q2	0.67	0.20	0.35
Q3	0.70	0.18	0.32
Q4	0.71	0.23	0.41
Q5	0.52	0.28	0.44
Q7	0.70	0.25	0.45
Q8	0.81	0.16	0.38
Q9	0.57	0.23	0.36
Q11	0.61	0.27	0.44
Q12	0.68	0.23	0.41
Q13	0.66	0.26	0.43
Q14	0.40	0.18	0.29
Q15	0.71	0.24	0.44
Q16	0.76	0.19	0.37
Q17	0.59	0.21	0.34
Q18	0.54	0.22	0.36
Q19	0.42	0.19	0.31
Q20	0.69	0.24	0.43
Q21	0.48	0.30	0.48
Q22	0.53	0.27	0.43
Q23	0.56	0.24	0.39
Q24	0.48	0.33	0.51

DISTRACTOR ANALYSIS

Item ID	KEY	Difficulty Index			Discrimination Index			Point Biserial Correlation		
		A	B	C	A	B	C	A	B	C
Q1	B	0.09	**0.82**	0.08	-0.08	**0.16**	-0.07	-0.19	**0.21**	-0.04
Q2	A	**0.67**	0.18	0.14	**0.20**	-0.11	-0.07	**0.17**	-0.11	-0.06
Q3	C	0.10	0.18	**0.70**	-0.06	-0.10	**0.18**	-0.06	-0.06	**0.11**
Q4	B	0.13	**0.71**	0.13	-0.10	**0.23**	-0.11	-0.08	**0.13**	-0.07
Q5	C	0.24	0.22	**0.52**	-0.14	-0.12	**0.28**	-0.01	-0.04	**0.07**
Q7	A	**0.70**	0.14	0.14	**0.25**	-0.11	-0.13	**0.13**	-0.08	-0.07
Q8	C	0.09	0.08	**0.81**	-0.08	-0.07	**0.16**	-0.16	-0.10	**0.21**
Q9	A	**0.57**	0.21	0.20	**0.23**	-0.06	-0.15	**0.14**	-0.07	-0.07
Q11	C	0.15	0.23	**0.61**	-0.13	-0.13	**0.27**	-0.17	0.02	**0.14**
Q12	A	**0.68**	0.18	0.12	**0.23**	-0.14	-0.08	**0.12**	-0.08	-0.04
Q13	B	0.13	**0.66**	0.19	-0.11	**0.26**	-0.14	-0.09	**0.12**	-0.04
Q14	A	**0.40**	0.23	0.35	**0.18**	-0.02	-0.14	**-0.02**	-0.05	0.06
Q15	C	0.12	0.15	**0.71**	-0.11	-0.12	**0.24**	-0.12	-0.09	**0.18**
Q16	A	**0.76**	0.12	0.10	**0.19**	-0.08	-0.09	**0.16**	-0.10	-0.10
Q17	C	0.10	0.30	**0.59**	-0.08	-0.11	**0.21**	-0.16	-0.02	**0.15**
Q18	B	0.22	**0.54**	0.22	-0.13	**0.22**	-0.07	-0.05	**0.05**	0.01
Q19	B	0.32	**0.42**	0.24	-0.07	**0.19**	-0.10	-0.02	**0.02**	0.00
Q20	C	0.18	0.11	**0.69**	-0.14	-0.09	**0.24**	-0.14	-0.12	**0.22**
Q21	B	0.23	**0.48**	0.27	-0.13	**0.30**	-0.15	-0.05	**0.07**	0.00
Q22	C	0.22	0.23	**0.53**	-0.12	-0.13	**0.27**	-0.07	-0.01	**0.09**
Q23	C	0.19	0.23	**0.56**	-0.10	-0.13	**0.24**	-0.04	-0.02	**0.05**
Q24	A	**0.48**	0.22	0.28	**0.33**	-0.13	-0.18	**-0.04**	0.02	0.01

Q25	0.46	0.16	0.26
Q1a	0.40	0.36	0.59
Q1b	0.66	0.19	0.38
Q2a	0.57	0.38	0.60
Q2b	0.33	0.27	0.54
Q3a	0.62	0.32	0.51
Q3b	0.76	0.20	0.42
Q4a	0.53	0.15	0.23
Q5a	0.52	0.28	0.42
Q5b	0.82	0.16	0.35
Q6a	0.56	0.33	0.52
Q6a	0.73	0.23	0.45

Q25	**C**	0.26	0.25	**0.46**	-0.06	-0.09	**0.16**	-0.04	-0.01	**0.05**

Appendix C: 2025 OERA Gender DIF

Sex		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	AVG.
Female	P1	20%	15%	21%	16%	20%	14%	17%	22%	18%	19%	16%	13%	13%	18%	19%	17%	18%	18%	16%	12%	15%	**17%**
	P2	49%	36%	37%	35%	46%	31%	37%	52%	42%	41%	33%	34%	35%	34%	36%	35%	36%	36%	36%	32%	35%	**37%**
	P3	64%	45%	53%	44%	61%	37%	54%	68%	64%	66%	40%	52%	45%	35%	47%	44%	61%	56%	51%	45%	50%	**51%**
	P4	78%	55%	70%	63%	76%	51%	62%	80%	88%	88%	53%	80%	69%	40%	74%	53%	84%	85%	82%	70%	61%	**69%**
	P5	90%	77%	89%	86%	91%	80%	76%	90%	97%	98%	77%	96%	92%	58%	95%	75%	96%	98%	97%	92%	80%	**87%**
F Total		62%	47%	56%	51%	61%	45%	51%	65%	65%	65%	46%	58%	54%	38%	57%	47%	62%	62%	59%	53%	51%	**55%**
Male	P1	20%	18%	19%	16%	20%	18%	17%	24%	18%	17%	15%	15%	16%	14%	14%	18%	15%	14%	14%	13%	19%	**16%**
	P2	43%	35%	38%	35%	42%	33%	41%	50%	43%	40%	32%	31%	33%	32%	37%	36%	36%	36%	35%	33%	36%	**37%**
	P3	62%	46%	52%	45%	61%	38%	51%	65%	62%	66%	40%	55%	50%	36%	45%	42%	60%	55%	50%	46%	50%	**51%**
	P4	77%	56%	71%	61%	74%	48%	67%	79%	86%	87%	53%	80%	69%	40%	75%	49%	86%	87%	79%	68%	65%	**69%**
	P5	89%	79%	89%	86%	92%	77%	83%	93%	97%	97%	78%	96%	93%	57%	95%	71%	96%	97%	95%	89%	84%	**87%**
M Total		56%	45%	52%	47%	56%	41%	50%	60%	59%	59%	42%	53%	50%	35%	51%	42%	56%	55%	52%	48%	49%	**50%**
Grand Total		**59%**	**46%**	**54%**	**49%**	**59%**	**43%**	**51%**	**62%**	**62%**	**62%**	**44%**	**55%**	**52%**	**36%**	**54%**	**44%**	**59%**	**58%**	**56%**	**50%**	**50%**	**52%**

PERCENTILE		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	AVG.
0-20th	P1	0%	-3%	2%	0%	0%	-4%	0%	-1%	0%	2%	1%	-1%	-3%	3%	5%	-1%	3%	4%	1%	-1%	-4%	**0%**
21-40th	P2	6%	0%	-1%	0%	4%	-3%	-3%	2%	-1%	1%	2%	3%	2%	1%	-1%	-1%	0%	0%	1%	0%	-1%	**1%**
41-60th	P3	1%	-1%	2%	-1%	0%	-1%	4%	2%	2%	0%	0%	-3%	-4%	-2%	2%	3%	1%	1%	1%	-1%	1%	**0%**
61-80th	P4	1%	0%	-1%	2%	2%	3%	-5%	1%	2%	1%	0%	0%	0%	1%	-1%	3%	-2%	-2%	3%	2%	-4%	**0%**
81-100th	P5	0%	-2%	0%	0%	-1%	3%	-7%	-3%	0%	1%	0%	-1%	0%	1%	0%	3%	0%	1%	1%	3%	-3%	**0%**
Grand Total		**6%**	**2%**	**5%**	**4%**	**6%**	**3%**	**1%**	**4%**	**6%**	**6%**	**4%**	**5%**	**4%**	**3%**	**6%**	**5%**	**6%**	**6%**	**7%**	**5%**	**2%**	**5%**

Appendix D: 2025 OEMA Gender DIF

Sex		Q1	Q2	Q3	Q4	Q5	Q7	Q8	Q9	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25
Female	P1	68%	45%	51%	43%	24%	45%	66%	30%	31%	42%	39%	21%	44%	53%	33%	28%	24%	39%	17%	22%	31%	15%	30%
	P2	86%	64%	67%	66%	32%	66%	88%	56%	52%	62%	57%	27%	69%	73%	54%	40%	26%	61%	22%	34%	37%	14%	30%
	P3	93%	76%	72%	78%	43%	81%	93%	71%	66%	71%	70%	30%	83%	86%	64%	50%	28%	77%	38%	42%	50%	26%	43%
	P4	97%	81%	83%	86%	57%	92%	95%	74%	76%	82%	84%	38%	90%	91%	69%	62%	42%	85%	59%	58%	60%	52%	47%
	P5	99%	88%	90%	94%	84%	97%	98%	85%	90%	93%	94%	61%	97%	97%	84%	80%	63%	95%	78%	84%	83%	86%	67%
F Total		89%	71%	73%	74%	48%	77%	89%	64%	64%	71%	70%	35%	78%	81%	61%	53%	37%	72%	43%	48%	53%	39%	44%
Male	P1	59%	44%	49%	45%	22%	42%	60%	30%	29%	43%	36%	24%	42%	55%	35%	31%	23%	45%	17%	25%	32%	17%	30%
	P2	84%	62%	63%	68%	35%	59%	81%	52%	49%	63%	54%	25%	66%	76%	52%	39%	26%	66%	27%	41%	38%	19%	35%
	P3	91%	71%	72%	79%	46%	74%	90%	62%	61%	75%	65%	30%	82%	82%	64%	47%	33%	76%	45%	50%	51%	34%	48%
	P4	96%	77%	79%	87%	67%	90%	95%	71%	79%	83%	81%	37%	89%	90%	69%	62%	45%	87%	66%	69%	68%	56%	52%
	P5	99%	87%	90%	97%	85%	96%	98%	77%	90%	94%	94%	69%	96%	95%	83%	81%	67%	96%	86%	87%	89%	88%	73%
M Total		85%	68%	70%	75%	50%	72%	84%	58%	61%	71%	65%	37%	74%	79%	60%	52%	39%	73%	48%	54%	55%	43%	48%
Grand Total		87%	69%	72%	74%	49%	74%	86%	61%	62%	71%	68%	36%	76%	80%	61%	52%	38%	73%	46%	51%	54%	41%	46%

PERCENTILE		Q1	Q2	Q3	Q4	Q5	Q7	Q8	Q9	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25
0-20th	P1	10%	1%	2%	-2%	2%	3%	6%	0%	2%	-1%	3%	-3%	3%	-2%	-1%	-3%	0%	-6%	0%	-4%	-1%	-2%	0%
21-40th	P2	1%	2%	4%	-2%	-2%	7%	7%	4%	3%	-1%	3%	2%	2%	-3%	2%	1%	0%	-5%	-5%	-8%	-1%	-5%	-5%
41-60th	P3	1%	5%	0%	-2%	-3%	7%	4%	9%	4%	-4%	5%	0%	1%	3%	0%	3%	-5%	1%	-7%	-8%	-1%	-8%	-5%
61-80th	P4	1%	4%	4%	-1%	-10%	2%	0%	2%	-3%	-1%	3%	1%	1%	1%	0%	0%	-3%	-2%	-7%	-11%	-8%	-4%	-4%
81-100th	P5	0%	2%	0%	-3%	-1%	1%	0%	8%	1%	-2%	1%	-8%	1%	1%	1%	0%	-4%	-1%	-8%	-3%	-6%	-2%	-6%
Grand Total		4%	4%	3%	-1%	-2%	5%	4%	6%	3%	-1%	4%	-2%	3%	1%	1%	1%	-2%	-1%	-4%	-6%	-3%	-4%	-4%

Sex		1(a)	1(b)	2(a)	2(b)	3(a)	3(b)	4(a)	5(a)	5(b)	6(a)	6(b)	AVG.
Female	P1	3%	46%	19%	4%	28%	52%	37%	23%	62%	20%	48%	**34%**
	P2	8%	60%	42%	11%	49%	79%	45%	37%	83%	45%	73%	**50%**
	P3	18%	65%	64%	21%	69%	88%	50%	49%	91%	60%	86%	**61%**
	P4	43%	77%	84%	36%	86%	93%	60%	65%	96%	81%	93%	**73%**
	P5	81%	86%	98%	67%	94%	97%	72%	81%	99%	93%	98%	**87%**
F Total		31%	67%	62%	28%	66%	83%	53%	52%	87%	61%	80%	**61%**
Male	P1	2%	44%	15%	4%	25%	54%	38%	21%	64%	18%	43%	**33%**
	P2	14%	57%	34%	11%	49%	79%	48%	42%	86%	42%	68%	**50%**
	P3	23%	61%	57%	19%	69%	86%	49%	52%	91%	62%	85%	**61%**
	P4	50%	72%	80%	32%	86%	92%	56%	66%	93%	78%	92%	**73%**
	P5	84%	89%	97%	65%	96%	98%	69%	86%	98%	89%	97%	**88%**
M Total		34%	64%	56%	26%	64%	81%	52%	53%	86%	57%	76%	**60%**
Grand Total		**33%**	**66%**	**59%**	**27%**	**65%**	**82%**	**52%**	**52%**	**86%**	**59%**	**78%**	**61%**

PERCENTILE		1(a)	1(b)	2(a)	2(b)	3(a)	3(b)	4(a)	5(a)	5(b)	6(a)	6(b)	AVG.
0-20th	P1	1%	2%	4%	0%	3%	-2%	-1%	2%	-2%	3%	5%	**1%**
21-40th	P2	-5%	2%	8%	0%	0%	0%	-3%	-5%	-4%	2%	5%	**0%**
41-60th	P3	-6%	4%	7%	2%	0%	2%	1%	-4%	0%	-2%	1%	**0%**
61-80th	P4	-8%	5%	3%	4%	1%	1%	4%	-1%	2%	2%	1%	**-1%**
81-100th	P5	-3%	-3%	1%	2%	-2%	0%	3%	-5%	0%	5%	1%	**-1%**
Grand Total		**-3%**	**3%**	**6%**	**2%**	**2%**	**1%**	**1%**	**-2%**	**0%**	**4%**	**4%**	**1%**

Appendix E: 2025 OEMA Constructed Response IRR

	Score (0)	Score (1)	Score (2)
Q1	2	40	0
Q2	2	40	0
Q3	5	0	37
Q4	39	2	1
Q5	2	40	0
Q6	0	0	42
Q7	30	12	0
Q8	0	42	0
Q9	3	39	0
Q10	0	42	0
Q11	2	40	0
Q12	0	42	0

m	42
n	12
pa	0.893437863
pe	0.5060547367
kappa	0.7842632677
s.e	0.007343653333
z	106.7947018
p-value	0

alpha	0.05
lower	0.7698699716
upper	0.7986565637

q	0.1686507937	0.6726190476	0.1587301587
b	0.1402077035	0.2202026644	0.1335348954
κ	0.6206750376	0.8072640936	0.9180971008
s.e	0.009838022268	0.009838022268	0.009838022268
z	63.08941174	82.05552616	93.32130745
p	0	0	0