

# TABLE OF CONTENTS

---

TABLE OF CONTENTS .....	1
LIST OF TABLES .....	3
LIST OF FIGURES .....	6
ACKNOWLEDGEMENTS .....	7
<b>1. INTRODUCTION .....</b>	<b>10</b>
1.1. BACKGROUND .....	10
1.2. RATIONALE FOR THE 2011 MOD-MSA: READING .....	11
1.3. ELIGIBILITY CRITERIA FOR THE MOD-MSA: READING.....	11
1.4. TEST ADMINISTRATION OF THE 2011 MOD-MSA: READING .....	11
<i>Test Materials</i> .....	11
<i>Test Administration Schedule</i> .....	12
<i>Student Participation</i> .....	13
<i>Accommodations for Assessment</i> .....	13
<i>Test Format</i> .....	14
<i>Security of Test Materials</i> .....	15
1.5. QUALITY CONTROL PROCEDURES .....	15
1.6. ITEM BANK CONSTRUCTION .....	17
<b>2. TEST DESIGN, DEVELOPMENT, AND SCORING OF THE 2011 MOD-MSA: READING .....</b>	<b>19</b>
2.1. TEST DESIGN AND STRUCTURE OF THE 2011 MOD-MSA: READING.....	19
2.2. DEVELOPMENT AND REVIEW OF THE 2011 MOD-MSA: READING .....	21
<i>National Psychometric Council</i> .....	21
<i>Content Review Committee</i> .....	21
<i>Bias Review Committee</i> .....	21
<i>Vision Review Committee</i> .....	22
2.3. THE 2011 MOD-MSA: READING: OPERATIONAL FORM TEST STRUCTURE AND ITEM TYPES .....	22
2.4. ITEM ANALYSES UNDERTAKEN PRIOR TO THE CREATION OF THE OPERATIONAL FORMS .....	23
<i>Classical Item Analyses</i> .....	23
<i>Differential Item Functioning Analyses</i> .....	24
<i>Item Response Theory (IRT) Analyses</i> .....	26
2.5. ITEMS FLAGGED FOR INSPECTION PRIOR TO THE CREATION OF THE OPERATIONAL FORMS.....	26
2.6. ITEMS SELECTED FOR THE 2011 OPERATIONAL TESTS.....	27
2.7. SCORING PROCEDURES OF THE 2011 MOD-MSA: READING .....	28
<b>3. CENTRAL TENDENCY MEASURES AND CLASSICAL ITEM ANALYSES OF THE 2011 MOD-MSA: READING.....</b>	<b>29</b>
3.1. MEASURES OF CENTRAL TENDENCY.....	29
<i>Standard Error of the Mean</i> .....	29
3.2. ITEM-LEVEL DESCRIPTIVE STATISTICS .....	30
<b>4. SCALE CREATION, EQUATING, AND RAW SCORES TO SCALE SCORES CONVERSION VIA ITEM RESPONSE THEORY PROCEDURES.....</b>	<b>44</b>
THE RASCH MODEL.....	44
4.1. EQUATING, CALIBRATION AND SCALING PROCEDURES FOR THE MOD-MSA: READING.....	46
4.2. SPECIFICS OF LINKING AND EQUATING THE 2011 MOD-MSA GRADES 3-8: READING .....	46

4.2.1. Generalized Robust Z Procedure .....	46
4.2.2. Guidelines for Selecting Linking Items .....	47
4.2.3. Step-by-step Procedure for Selecting Linking Items .....	47
4.3. REPORTING SCALE SCORES FOR THE 2011 MOD-MSA: READING .....	54
Raw Score to Scale Score Conversion Tables for the Total Score .....	56
Raw Score to Scale Score Conversion Tables for the Subscales .....	62
4.4. SCORE INTERPRETATION .....	68
Scale Scores .....	68
Performance Level Descriptors .....	68
4.5. FINAL PERFORMANCE LEVEL CUT POINTS FOR THE MOD-MSA: READING .....	68
<b>5. COMPARATIVE CLASSICAL AND IRT STATS ACROSS YEARS FOR THE COMMON ITEMS USED IN GRADES 3-8 .....</b>	<b>70</b>
<b>6. IRT ITEM LEVEL STATISTICS FOR THE 2011 MOD-MSA: READING .....</b>	<b>82</b>
6.1 RATIONALE FOR THE USE OF THE RASCH MODEL .....	82
6.2 FIT STATISTICS FOR THE RASCH MODEL .....	82
6.3. RASCH ITEM LEVEL STATISTICS .....	83
<b>7. TEST RELIABILITY .....</b>	<b>96</b>
7.1. PRECISION AND RELIABILITY (CLASSICAL METHODS) .....	96
Standard Error of Measurement (SEM) of the Test .....	96
Cronbach's Alpha ( $KR_{20}$ ) .....	97
7.2. IRT METHOD IN MEASURING PRECISION OF THE TEST .....	97
7.3. DECISION ACCURACY AND CONSISTENCY AT THE CUT SCORES .....	98
<b>8. TEST VALIDITY .....</b>	<b>100</b>
8.1. TEST VALIDITY FOR THE 2011 MOD-MSA: READING .....	100
Content-Related Evidence .....	100
Evidence from Item Development Methods .....	100
Evidence Based on Excluding Bias Items Before and After DIF Analysis .....	101
Evidence from Internal Structure of the Tests .....	102
8.2. UNIDIMENSIONALITY ANALYSIS FOR THE 2011 MOD-MSA: READING .....	103
<b>9. SUMMARY OF OPERATIONAL TEST RESULTS OF THE 2011 MOD-MSA: READING .....</b>	<b>105</b>
9.1 CLASSICAL DESCRIPTIVE STATISTICS OF THE OPERATIONAL TESTS, USING RAW SCORES .....	105
9.2 SCALE SCORE DESCRIPTIVE STATISTICS .....	105
9.3 FREQUENCY OF STUDENTS AT EACH PROFICIENCY LEVEL BY GRADE .....	106
9.4. TEST CHARACTERISTIC CURVES, TEST INFORMATION FUNCTIONS, AND CONDITIONAL STANDARD ERRORS OF MEASUREMENT .....	107
Conditional Standard Errors of Measurement (CSEM) of the Tests .....	108
<b>10. REFERENCES .....</b>	<b>121</b>
<b>APPENDIX A: OPERATIONAL ITEMS SELECTED AFTER DATA REVIEW .....</b>	<b>124</b>
<b>APPENDIX B: ITEM DISTRACTOR ANALYSES .....</b>	<b>128</b>
<b>APPENDIX C: THE 2011 MOD-MSA, DIFFERENTIAL ITEM FUNCTIONING ANALYSIS (DIF) .....</b>	<b>141</b>
<b>APPENDIX D: FREQUENCY DISTRIBUTION HISTOGRAMS OF SCALE SCORES .....</b>	<b>154</b>
<b>APPENDIX E: ALIGNMENT OF THE 2011 MOD-MSA: READING ITEMS TO THE STATE STANDARDS .</b>	<b>161</b>
<b>APPENDIX F: IDENTIFYING STUDENTS FOR PARTICIPATION IN THE MOD-MSA: READING PROGRAM .....</b>	<b>168</b>