

TABLE OF CONTENTS

Table of Contents	1
List of Tables	3
List of Figures	5
Acknowledgements	6
1. Introduction	10
1.1. Background	10
1.2. Rationale for the 2010 Mod-MSA: Reading	11
1.3. Eligibility Criteria for the Mod-MSA: Reading	11
1.4. Test Administration of the 2010 Mod-MSA: Reading.....	11
Test Materials	11
Test Administration Schedule	12
Student Participation	13
Accommodations for Assessment	13
Test Format	14
Security of Test Materials	15
1.5. Quality Control Procedures	15
1.6. Item Bank Construction.....	17
2. Test Design, Development, and Scoring of the 2010 Mod-MSA: Reading	18
2.1. Test Design and Structure of the 2010 Mod-MSA: Reading	18
2.2. Development and Review of the 2010 Mod-MSA: Reading	19
National Psychometric Council.....	20
Content Review Committee	20
Bias Review Committee.....	20
Vision Review Committee	20
2.3. Operational Form Structure of the 2010 Mod-MSA: Reading.....	20
2.4. Item Analyses Undertaken Prior to the Creation of the Operational Forms.....	22
Classical Item Analyses.....	22
Differential Item Functioning Analyses	22
Item Response Theory (IRT) Analyses	25
2.5. Items Flagged for Inspection Prior to the Creation of the Operational Forms	25
2.6. Items Selected for the 2010 Operational Tests.....	25
2.7. Scoring Procedures of the 2010 Mod-MSA: Reading.....	27
3. Central Tendency Measures and Classical Item Analyses of the 2010 Mod-MSA: Reading.....	28
3.1. Measures of Central Tendency	28
Standard Error of the Mean	28
3.2. Item-Level Descriptive Statistics	29
4. Scale Creation, Equating and Raw Scores to Scale Scores Conversion via Item Response Theory Procedures	43
The Rasch Model	43
4.1. Calibration and Scaling Procedures for Grades 3 to 5.....	45

4.2. Specifics for Creating the Base Scale for the Mod-MSA: Reading Grades 3-5.....	45
4.3. Calibration Equating the 2010 Mod-MSA: Reading Grades 6-8	46
4.4. Specifics of Linking and Equating the 2010 Mod-MSA Grades 6-8: Reading.....	46
4.5. Reporting Scale Scores for the 2010 Mod-MSA: Reading	51
4.6. Score Interpretation	65
4.7. Final Performance Level Cut Points for the Mod-MSA: Reading	65
5. Comparative Classical and IRT Stats Across Years for the Common Items Used in Grades 6-8	67
6. IRT Item Level Statistics for the 2010 Mod-MSA: Reading.....	73
6.1 Rationale for the Use of the Rasch Model.....	73
6.2 Fit Statistics for the Rasch Model	73
Mean-Square	74
6.3. Rasch Item Level Statistics	74
7. Test Reliability	87
7.1. Precision and Reliability (Classical Methods)	87
Standard Error of Measurement (SEM) of the Test	87
Cronbach’s Alpha (KR ₂₀).....	88
7.2. IRT Method in Measuring Precision of the Test.....	88
7.3. Decision Accuracy and Consistency at the Cut Scores	89
8. Test Validity	91
8.1. Test Validity for the 2010 Mod-MSA: Reading.....	91
Evidence from Internal Structure of the Tests.....	93
8.2. Unidimensionality Analysis for the 2010 Mod-MSA: Reading.....	95
9. Summary of Operational Test Results of the 2010 Mod-MSA: Reading.....	96
9.1 Classical Descriptive Test Statistics.....	96
9.2 Scale Score Descriptive Statistics	96
9.3 Frequency of Students at Each Proficiency Level by Grade	97
9.4. Test Characteristic Curves, Test Information Functions, and Conditional Standard Errors of Measurement	98
10. References.....	112
Appendix A: Operational Items Selected After Data Review	115
Appendix B: Item Distractor Analyses	119
Appendix C: Frequency Distribution Histograms of Scale Scores.....	132
Appendix D: Standard Setting Report.....	139
Appendix E: Alignment of the 2010 Mod-MSA: Reading Items to the State Standards	215
Appendix F: Identifying Students for Participation in the Mod-MSA: Reading Program	222
Appendix G: Comparability Study of Paper and Pencil, and Online Administration of the Mod- MSA.....	224