

## **OVERVIEW OF THIS MANUAL**

---

The *Modified Maryland School Assessment Program (Mod-MSA) Technical Manual* for the 2010 Mathematics Administration is organized around ten major sections, plus a section for the appendices. An overview of this manual is provided below.

### **Section 1: Introduction**

This section presents the Mod\_MSA Mathematics test's background, rationale, eligibility criteria, test administration (test materials, test administration schedule, student participation, accommodation used for assessment, test format, and security of test materials), quality control procedures, and item bank construction.

### **Section 2: Test Design, Development, and Scoring of the 2010 Mod-MSA: Mathematics**

This section describes the test design, development, and the scoring process of the Mod-MSA: Mathematics test. It provides the test design and structure, development and review of the test (National Psychometric Council, Content Review Committee, Bias Review Committee, and Vision Review Committee), structure of the operational test, item analyses undertaken prior to the creation of the operational test (classical item analyses, differential item functioning analyses, and item response theory analyses), items flagged for inspection prior to the creation of the operational form, the items selected for the 2010 operational tests, and the scoring procedure for the Mod-MSA: Mathematics.

### **Section 3: Central Tendency Measures and Other Classical Item Analyses of the 2010 Mod-MSA: Mathematics**

This section provides central tendency measures and classical item analyses. It includes explanation of the standard error of the mean, the measure of central tendencies, and item level descriptive statistics.

### **Section 4: Scale Creation, Equating and Raw Scores to Scale Scores Conversion via Item Response Theory (IRT) Procedures**

This section explains the item response theory (IRT) procedures (the Rasch model and the conditional standard error of measurement); calibration and scaling procedure for Grades 3-5; specifics of creating the base scale for the Mod-MSA: Mathematics (Grades 3 -5); calibration and equating for Grades 6 - 8; specifics of scaling and equating for Grades 6-8 (generalized Robust-Z procedure; reporting scale scores (raw score to scale score conversion for the total test score, and by subscales, guide for selecting linking items and step-by-step procedure for selecting linking items); tables of unequated linking item difficulties; reporting of the scale scores (tables of raw scores to scale scores by total test and also by sub-scales/strands); score interpretation (scale scores and performance level descriptors); and the final scale score cutoffs at each performance level.

### **Section 5: Comparative Classical and IRT Stats Across Years for the Common Items Used in Grades 6-8**

This section compares common items used for linking Grades 6 to 8 across the two years of operation. Both p-values from classical statistical analysis and the Rasch difficulties from IRT

analysis are compared for each grade across 2009 and 2010. Graphical displays are also provided for these stats.

## **Section 6: IRT Item Statistics for the 2010 Mod-MSA: Mathematics**

This section begins with a description for the rationale for the use of the Rasch model. It then provides an explanation of the fit indices and ends with the provision of IRT item level analyses.

## **Section 7: Test Reliability**

This section describes the precision and reliability undertaken by classical methods (standard error of measurement of the test and the Cronbach's Alpha). It also includes a brief statement relating to IRT methods in measuring the precision of the test and explains the decision accuracy and consistency at the cut scores. Tables relating to accuracy and consistency are also provided in this section.

## **Section 8: Test Validity**

To investigate the validity evidence of the 2009 Mod-MSA: Mathematics, content-related evidence, evidence from item development methods, differential item functioning (DIF) analysis on gender and ethnicity (white and black students), and evidence from internal structure, were collected. Also, a study comparing the mode of administration was undertaken by Pearson to validate the online administration of the test. Evidence with respect to the unidimensionality of the test was undertaken by a principal component analysis.

## **Section 9: Summary of Operational Test Results for the 2010 Mod-MSA: Mathematics**

This section presents the summary of the operational test results. It includes the classical descriptive test statistics of the raw scores, scale score descriptive statistics for the test, frequency of students in each proficiency level by grades, and an explanation and depiction of the test characteristic curve (TCC), the test information curve, and the conditional standard error of measurement (CSEM) of the tests.

## **Section 10: References**

This section presents the references used in producing this report.

## **Section 11: Appendices**

This section presents Appendix A to Appendix G. Appendix A provides a list of operational items selected after data review, while Appendix B and Appendix C provide the item distractor analyses, and the frequency distribution and histograms of the scale scores, respectively. The standard setting report is provided in Appendix D, and the alignment of the test items with the state standards is provided in Appendix E. Appendix F provides the criterion used in identifying students with disabilities for participation in the Mod-MSA: Mathematics program. Finally, Appendix G provides the comparability study of the paper-and-pencil and online modes of Mod-MSA administration for Grades 4-5.