

CHAPTER I: HIGH SCHOOL ASSESSMENTS

Section 1. Introduction

The Maryland High School Assessments (MD HSAs) consist of end-of-course tests in Algebra/Data Analysis, Biology, English, and Government. The MD HSAs are referred to as “end-of-course” tests because students take each test as they complete the appropriate coursework. In addition, results from the Algebra/Data Analysis, Biology, and English administrations are used in the Maryland State Department of Education (MSDE) Adequate Yearly Progress (AYP) reports, required under the No Child Left Behind (NCLB) Act for the 2009 school year.

Students entering ninth grade in the 2005–2006 school year and in subsequent school years are required to earn satisfactory scores on all four MD HSAs in order to earn a Maryland High School Diploma.¹ Information on the interpretation of scores is provided to students, parents, schools, and other stakeholders via the MSDE website.

Historically, the MD HSAs have consisted of paper-and-pencil tests that contain selected-response (SR) items, which require students to choose between four short response options; brief constructed response (BCR) items, which require students to write a short response; and extended constructed response (ECR) items, which require students to write a longer response. In addition, Algebra/Data Analysis tests include items that require a student-produced response (SPR); students must grid in their responses on the answer document. The SR and SPR items are machine-scored; the BCR and ECR items are scored by raters.

The administration options and test design of the MD HSAs changed beginning with the May 2009 administration. As of May 2009, the MD HSAs are administered online as well as in paper-and-pencil format. Also, the BCR and ECR items that appeared in previous test administrations have been replaced by SR and SPR items. Consequently, the October, January, and April administrations were conducted in paper-and-pencil format only and included BCR and ECR items; the May and Summer (July/August) administrations were given both online and in paper-and-pencil format and contained only SR items and, in the case of Algebra/Data Analysis, SPR items. A study of the comparability of online and paper forms of the May 2009 MD HSAs was conducted and the resulting report is provided in Appendix 1C. All MD HSA items, regardless of test administration date, are based on content outlined in Maryland’s Core Learning Goals (CLGs).²

This report provides information about the October 2008 administration and the January, April, May, and Summer 2009 administrations. For the October 2008 administration, one form was administered. For the January 2009 administration, two forms were administered. One form was

¹ More information on the testing requirement for graduation is available on the Maryland State Department of Education website at http://mdk12.org/assessments/high_school/index.html.

² The HSA Core Learning Goals documents can be found on the Maryland School Improvement website at <http://www.mdk12.org/assessments/standards/9-12.html>.

used for the primary administration, and one form was used for the makeup administration. Administration of the MD HSAs in April occurred for the first time in 2009; one form was administered for this session. For the May 2009 administration, thirteen test forms were administered: eleven for the primary testing window and one for each of two makeup testing windows. For the Summer 2009 administration, two primary forms were administered: one for the first week of testing and one for the second week of testing.

Each test form consisted of operational and field test items. The operational items were used to produce student scores. Field test items were scored along with the operational items, but examinees' scores on these items were not included in the computation of their total test scores. Performance on the field test items was analyzed, and all flagged items were reviewed. Field test items that were approved by ETS content specialists and calibrated were then marked as available for use in the item bank. Items that were deemed unacceptable were marked as "Unavailable" and may be revised and field tested again in the future. With the exception of items selected for public release, the operational items that are returned to the item bank must remain unused for at least one year to minimize item exposure.

The item response model used to calibrate the items in the MD HSAs is the three-parameter logistic (3PL) model. This model is used to generate both total test scores and subscores. In the past, the total test scores were generated using item-pattern (IP) scoring, and the subscores were created using raw score (RS) to scale score (SS) conversion tables. Beginning with the May 2009 administration, subscores are calculated using IP scoring instead of RS to SS tables. This change was implemented to provide consistency in scoring between total test scores and subscores and to mitigate possible confusion due to the use of different scoring methods within the HSAs. Total test results in the scale score metric are reported to students. Subscores are not reported to students but are aggregated at the classroom level to provide teachers and administrators with additional information about student performance in each of the subscore categories.

Beginning with the 2004 administration, pre-equated item parameters typically have been used to generate student scores. Prior to 2004, students' scores were based on item parameters estimated after each the administration.³ When pre-equated item parameters are used, the parameters are not estimated following an administration; instead, existing bank parameters are used to produce student scores. Using this approach, scores can be calculated and assigned to students immediately after their answer documents have been processed.

An exception to the practice of using pre-equated item parameters to generate student scores occurred for the May 2009 administration. Given the removal of BCR and ECR items, Maryland's technical advisory committee, the National Psychometrics Council (NPC), recommended that the operational item parameters should be estimated and equated using data collected in the May 2009 administration. Comparison of post-equated item parameters with pre-equated item parameters would enable MSDE and the NPC to assess whether the change in test design affected item parameters. MSDE and the NPC determined that differences between the two sets of parameters were negligible, and they decided to report scores based on the post-test calibration and equating of the May 2009 operational items. They also decided that in the future,

³In the post-test equating, anchor items representative of the content and difficulty of the test forms were used to equate the test forms using a Stocking and Lord procedure (CTB/McGraw-Hill, December 2003).

student scores would continue to be generated using the pre-equated item parameters residing in the item bank.

All technical support and analyses were carried out in accordance with both the *ETS Standards for Quality and Fairness* (2002) and the *Standards for Educational and Psychological Testing*, issued jointly by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999).

The MD HSA chapter of this technical report consists of six other sections and three appendices. Section 2 describes the procedures used for test construction, and administration. Section 3 discusses the validity of the MD HSAs. Section 4 delineates the scoring procedures and score types. Section 5 describes the results of analyses of test reliability and decision consistency and decision accuracy. Section 6 provides summary statistics and descriptive information about student characteristics. Section 7 gives the results of the analysis of the field test data, including classical item analysis, differential item functioning, and item calibration and scaling. Appendix 1A provides classical item statistics for each operational test item by form administered. Appendix 1B provides classical item statistics for each field test item administered. Appendix 1C presents the design, analyses, results and conclusions of a study conducted to examine the comparability of the online and paper forms of the May 2009 MD HSAs.