

Section 2. Validity

Validity is one of the most important attributes of assessment quality. It refers to the degree to which evidence supports the interpretations of test scores by proposed users of tests and is one of the most fundamental considerations in developing and evaluating tests (AERA, APA, & NCME, 1999). Validity is not based on a single study or type of study, but should be considered an ongoing process of gathering evidence supporting the interpretation of the resulting test scores. This process begins with the test design and continues throughout the entire assessment process, including design, content specifications, item development, psychometric quality and inferences made from the results.

The development of test content for each HSA was overseen by a content expert who has a depth of knowledge and teaching experience related to the course in which the HSA was administered. The appropriate content leads that had similar qualifications reviewed the test development work of these individuals.

The test development process itself provided numerous opportunities for the client to review test content and make changes to ensure that the items, both individually and as collections within forms, were valid measures of the knowledge and skills of Maryland students according to course standards. Every item that was created is referenced to a particular instructional standard (goal, expectation, and indicator). At various points during the internal ETS development process, that specific reference was either confirmed or changed to reflect changes to the item. When the item went to a committee of Maryland educators for a content review, the members of the committee made individual judgments on the match of the item content with the standard it was intended to measure and the appropriateness for the typical age of students being tested. These judgments were tabulated and reviewed by the content experts who use the information to decide which items will advance to the field test stage of development.

The constructs measured by each HSA were described in detail in the Maryland high school curriculum standards (Core Learning Goals). All ETS content staff working on item development had been trained in the Core Learning Goals. The test blueprint documents presented in Section 1 (see Tables 1.2 to 1.6 in Section 1) were created in collaboration with committees of Maryland educators and were directly derived from the Maryland goals, expectations, and indicators. These Learning Goals can be found on the MSDE website at <http://www.mdk12.org>.

Although all eligible students participated in the HSA and information about student performance was provided to students, parents, teachers and other stakeholders, scores for all content areas had no consequences for individual students during this time. Geometry scores were also used for AYP as a component of the Maryland No Child Left Behind (NCLB) Accountability program. Information on the interpretation of scores was provided to students, parents, schools and other stakeholders via the MSDE website.

In addition to the validation documentation gathered and maintained by MSDE, this report contains relevant empirical information in support of the Maryland HSA as follows.

- Section 3 provides detailed information concerning the particular scores that were reported for the Maryland HSA and includes an evaluation of different procedures for reporting subscore performance.
- Section 4 provides demographic information for the population of students who were administered the Maryland HSA as well as summaries of test level statistics. Summary statistics and reliability estimates were reported for the student population and by subgroups. Score distributions as well as the passing rates for all administrations and evidence that the tests were not speeded were also provided in this section.
- Section 5 includes documentation of the analysis procedures as well as distributions of item p-values and item-total correlations from the field testing activities. This section also includes an empirical evaluation of the impact of changing test directions for brief-constructed response items in the Government assessment.
- Appendix 2.A presents the results of factor analyses of the Maryland HSA item responses.