

Test Overview and Design

Introduction

The Maryland School Assessment (MSA) tests are measures of students' knowledge relative to the Maryland Voluntary State Curriculum. The MSA Science test was added to established assessments in reading and mathematics to form part of the MSA program. Administered annually in the spring, the MSA program was established to meet the requirements of the No Child Left Behind Act (NCLB) of 2001. In 2006, Pearson was contracted by Maryland State Department of Education (MSDE) to develop, administer and maintain the MSA Science test. This report provides technical details of work accomplished during the 2007-2008 test administration cycle.

Purpose

The purpose of this MSA Technical Manual is to provide objective information regarding technical aspects of the 2008 MSA Science operational test. This volume is intended to be one source of information to Maryland K-12 educational stakeholders (including testing coordinators, educators, parents and other interested citizens) about the development, implementation, scoring and technical attributes of the MSA Science. Other sources of information regarding the MSA Science test, provided in paper or online format, include the MSA Science administration manual, implementation material and training materials.

The information provided here fulfills professional and scientific guidelines (AERA, APA, NCME, 1999) for field test technical reports of large scale educational assessments and is intended for use by qualified users within schools who use the MSA Science and interpret the results. Specifically, information was selected for inclusion in this report based on NCLB requirements and the following *Standards for Educational and Psychological Testing*:

- Standards 6.1 – 6.15 Supporting Documentation for Tests
- Standards 10.1—10.12 Testing Individuals with Disabilities
- Standards 13.1—13.19 Educational Testing and Assessment

This manual provides information about the MSA Science test regarding:

1. Content of the tests;
2. Test form design;
3. Identification of ineffective items;
4. Reliability of the tests;
5. Difficulty of the test questions;
6. Equating of test forms;
7. Detection of item bias;
8. Scoring and reporting the results of the tests;

Each of these facets in the MSA Science test development and use cycle is critical to validity of test scores and interpretation of results. This technical manual covers all of these topics for the 2007-2008 testing year.

Test Overview

In 2002, the Maryland State Department of Education adopted the testing program known as the Maryland School Assessment (MSA). The first two subjects to be established under this new testing program were Reading and Mathematics. The Science test was added and the first field administration was conducted in the spring of 2007, followed by the first operational test in 2008. The MSA Science test is currently targeted at grade 5 and grade 8 students to assess achievement in Science.

Purpose and Use

By assessing student achievement against the Science academic standards, the MSA Science test serves two important uses. First, it provides information about what students learned in school to parents, teachers, and educators to inform the improvement of individualized instructional programs, classroom education and school performance. Second, the MSA Science test provides an accountability tool to measure performance levels of students, schools and districts against the Science academic standards.

Test Content, Specifications and Design

The MSA Science test was designed to align to the Voluntary State Curriculum (VSC) that specifies curricular indicators and objectives that contributed directly to measuring content standards. According to MSDE's website, the VSC defines what students should know and be able to do and "is the document that aligns the Maryland Content Standards and the Maryland Assessment Program". The VSC is formatted so that it delineates the content standards, which are broad, measurable statements about what students should know and be able to do. Next, the VSC has indicator statements that provide the next level of specificity and begin to narrow the focus for teachers. Finally, the objectives provide teachers with very clear information about what specific learning should occur. The VSC is widely disseminated to Maryland educational stakeholders, to include teachers, central office staff, students, parents and other stakeholders.

In order to ensure that MSDE is in accordance with the federal law that requires states to align their tests to their content standards, the VSC serves as the guiding document for test development and design. Developing the items for testing was a collaborative effort between MSDE, educators and Pearson. Teachers, administrators and content specialists were recruited from all over Maryland for different test development committees. These committees reviewed items developed for MSA Science test.

The basic test specifications were established by MSDE and provided to Pearson to guide the test administration. Since the inception of the Science test, there have been two test administrations-a census field test in 2007 and an operational test in 2008. Both administrations were conducted under the same testing conditions. Accordingly, the field test was designed so that it would meet the requirements of the operational administration test blueprint. The major difference is that there would be fewer scored items on the operational form, but the same number of overall items. For the 2008 operational test, two base forms (i.e., two forms of scored operational items) were used. Each form had a total of 77 items on the grade 5 form and 75 items on the grade 8 form. Grade 5 tests had 66 core (operational) items and 11 field test items for grade 5. The grade 8 test had 64 core items with 11 field test items. For both grade tests, only core items contributed

to student scores. The two base forms share a set of 20 common items. These common items are discrete (i.e., non-passage based) selected response (SR) items.

There are two MSA Science test blueprints available, one for grade 5 and one for grade 8 and there are six standards assessed across each grade with 66 items in the grade 5 test and 64 items in the grade 8 test, as presented in Table 1 and 2.

Table 1. Grade 5 MSA Science Standards Assessed

Standard	
1.0	Skills and Processes
2.0	Earth/Space Science
3.0	Life Science
4.0	Chemistry
5.0	Physics
6.0	Environmental
Total Number of items: 66	
Total number of points:72	

Table 2. Grade 8 MSA Science Standards Assessed

Standard	
1.0	Skills and Processes
2.0	Earth/Space Science
3.0	Life Science
4.0	Chemistry
5.0	Physics
6.0	Environmental
Total Number of items: 64	
Total number of points: 72	

MSA Science 2008 Field test Design

Field test forms were composed of selected response (SR) items and brief constructed response (BCR). Items were either stand-alone (not linked to other items), linked to a lab set stimulus (e.g., technical graph or figure), or linked to a technical passage stimulus. Each field test form included 20 stand-alone SR items that were common across all 10 forms. Each lab set and its linked SR items appeared on two separate field test forms with the same SR items appeared on each form. The 10 field test item sets were combined with the core items. Field test item sets 1-5 were embedded in Base Form A and 6-10 in Base Form B. In other words, operational forms 1 through 5 share the same core items and are differentiated by a unique field test item set within each form.

MSDE and Pearson worked together to finalize the structure of the 2008 field test forms. At each grade, 10 field test forms were produced. The intent of the test build process was to have each form be parallel in terms of number of SR items, BCR items and stimulus materials. In addition, the field test forms were designed to be equivalent to the operational base forms plus embedded field test in terms of total numbers of SR and BCR items. All 10 forms per grade had the same number of SR and BCR items. In addition, a goal of item selection was to balance, to the extent possible, coverage of the standards across the 10 field test forms per grade. On a per form basis, initial item selections were performed by Pearson and then shared with MSDE for review and

approval. Since Form 1 at each grade was the Braille/large print form, items were selected for Form 1 on the basis of feedback provided by the low-vision panel in October.

The 2008 forms (and all subsequent operational assessments) would be spiraled at the student-level. Spiraling at the student-level supports the assumption that examinee groups responding to each test form are randomly equivalent; an assumption that will further strengthen the link across forms.