

# Section 1. Test Construction and Administration

## Test Development

### *Planning*

Planning for the test development process began with the creation of item development plans for each content area. ETS content leaders collaborated with their content counterparts at MSDE to create these plans. The item bank was reviewed to determine how well the available item pool matched the test form requirements set forth in the test form blueprint. Areas, as defined by the Core Learning Goals, that contained low item counts were given priority when determining which indicators were to be addressed by the item writers. After these critical need areas were defined and addressed, the remaining numbers of items to be developed (which is determined by the requirements set forth in the RFP) were distributed among the remaining indicators in a fashion that would best ensure that future administrations have a sufficient depth of items from which to construct operational forms.

### *Test Specifications and Design*

The basic test design was pre-determined by MSDE and provided to ETS in the form of the content specific “Test Specs – Test Form Matrix” document presented in Tables 1.2 to 1.6. This basic test design document provided direction to session length, item number and type by session, and other form requirements. How the specific items were placed throughout the forms was left to the collaborative efforts of the ETS and MSDE content specialists. Construction of the operational forms was based on test blueprints as approved by MSDE.

### *Item Type*

There were four item types that were utilized by the Maryland HSA exam. These item types were selected response (SR), student produced response (SPR), brief constructed response (BCR), and extended constructed response (ECR). The following table shows how these item types were used on operational forms.

Table 1.1 Number of Items on Operational HSA Forms by Item Type

| Content Area | SR | SPR | BCR | ECR |
|--------------|----|-----|-----|-----|
| Algebra      | 26 | 6   | 3   | 3   |
| Biology      | 48 | -   | 7   | -   |
| English      | 50 | -   | 2   | 1   |
| Geometry     | 26 | 6   | 2   | 3   |
| Government   | 50 | -   | 7   | 1   |

### *Item Writing*

Item writers, at least 50 percent of which were Maryland educators, were contracted to develop quality test items that were aligned with Core Learning Goals. Item writers were selected based on their depth of content knowledge and familiarity with HSA testing program. The item writers were trained on general item writing techniques as well as writing parameters that were specific to the Maryland HSA program. Approximately one month after the initial item writer training, writers were provided a follow-up training session geared to evaluate their writing skills developed up to that point and provide constructive feedback to guide the rest of their writing assignment. Upon completion of their writing assignment, item writers submitted their items to ETS. The items that were accepted started item review and revision process. Many specific requirements of writing for Maryland HSA program can be found in “Guidelines for Item Writers” document.

### *Item Review and Revision*

All items developed for this program underwent a series of editorial reviews in accordance with the following procedures:

- Items edited according to standard rules developed in conjunction with MSDE.
- Items reviewed for accuracy, organization and comprehension, style, usage, consistency and sensitivity.
- Item content reviewed so that each item measures intended Goal-Expectation-Indicator.
- Copyright and/or trademark permission has been obtained for any required materials.
- Internal reviews conducted and historical records will be maintained for all version changes.

After ETS performed required internal reviews, items were submitted to MSDE for their review. If the MSDE content specialist requested a copy, an original version of the item as submitted by the item writer was provided. Any associated stimulus material, graphic, and/or art was provided as well as information regarding the Goal-Expectation-Indicator that each question addressed.

MSDE performed a review of the items and provided feedback to ETS content specialists. These edits were incorporated into the items, then MSDE and ETS content specialists met and conducted a side-by-side review of the items. Any final edits to the items were made. The items were then prepared for Content Review Committee review. All constructed response items were also submitted to Measurement Incorporated (MI) for review.

The final round of reviews involved the Content Review Committee and Bias/Fairness Review Committee. These committees were diverse groups of Maryland educators who reviewed each item and ensured that content in each item accurately reflected what was taught in Maryland schools and that no individual or group would be unfairly favored or disadvantaged due to the content of the items.

Upon the completion of this final round of review, MSDE and ETS content specialists again conducted a side-by-side meeting to evaluate reviews by MI, Content Review Committee, and Bias/Fairness Review Committee. The ETS content specialist then made any necessary edits to the items. The items that survived this process were ready to be placed in field test sections of operational forms.

### **Test Specifications**

All the 2004 operational test forms were constructed from items from the Maryland item bank. The pool of items available for use in the construction of the 2004 forms included all items that had been administered, calibrated and linked to the operational scale. For HSA operational scale was defined in 2002 and included items administered in 2002 and 2003. Items administered prior to 2002 were not eligible for selection of the 2004 forms<sup>2</sup>. In addition, items flagged for poor fit and items that had been flagged for severe differential item functioning (DIF) against one of the focal groups were excluded from the available item pool (see also Section 5 for more details about these analyses and flagging criteria).

Each test included a mixture of selected-response (SR), as well as brief and/or extended constructed-response (BCR, ECR) items. Algebra/Data Analysis and Geometry also included student produced response (SPR) items. Each test form consisted of two sections administered within a single sitting (the two sections were separated by a short break). SR and SPR items were worth one score point and were scored against specific keys. BCR and ECR items varied in number of score points by content area. In Algebra and Geometry BCR items were worth three points and ECR items were worth four points. English I BCR items were worth four points and ECR items were worth six points. The BCR and ECR items for Government were both worth four points and Biology had only BCR items, which were worth four points. Rubrics for items can be found at the following locations:

|                       |   |
|-----------------------|---|
| Algebra and Geometry: | <a href="http://mdk12.org/rubrics/mathematics">http://mdk12.org/rubrics/mathematics</a> .   |
| Biology               | <a href="http://mdk12.org/rubrics/science">http://mdk12.org/rubrics/science</a>             |
| English I             | <a href="http://mdk12.org/rubrics/english">http://mdk12.org/rubrics/english</a>             |
| Government            | <a href="http://mdk12.org/rubrics/socialstudies">http://mdk12.org/rubrics/socialstudies</a> |

In addition, each test form was constructed to meet specific test blueprints. Tables 1.2 to 1.6 indicate distribution of items within each reporting category by item type.

---

<sup>2</sup> Subsequent to the selection for the 2004 forms, a linking study was conducted to place some additional items onto the operational scale. The results of this study are located in Appendix 1.A.

Table 1.2 Algebra/Data Analysis Blueprint

| ALGEBRA/DATA ANALYSIS   |                  |                    |                    |                    |                   |
|---|------------------|--------------------|--------------------|--------------------|-------------------|
| Reporting Category  | Item Type        |                    |                    |                    | Percent of Points |
|   | SR<br>(4pts/ECR) | SPR<br>(3 pts/BCR) | BCR<br>(3 pts/BCR) | ECR<br>(4 pts/ECR) |                   |
| Totals  | 26               | 6                  | 3                  | 3                  |                   |
| Expectation 1.1<br>The student will analyze a wide variety of patterns and functional relationships using the language of mathematics and appropriate technology. |                  |                    |                    |                    | 25%               |
| Expectation 1.2<br>The student will analyze a wide variety of patterns and functional relationships using the language of mathematics and appropriate technology. |                  |                    |                    |                    | 32%               |
| Expectation 3.1<br>The student will collect, organize, analyze, and present data.   |                  |                    |                    |                    | 22%               |
| Expectation 3.2<br>The student will apply the basic concepts of statistics and probability to predict possible outcomes of real-world situations.                 |                  |                    |                    |                    | 21%               |

Table 1.3 Biology Blueprint

| BIOLOGY  |                 |                  |                   |
|--|-----------------|------------------|-------------------|
| Reporting Category   | ITEM TYPE       |                  | Percent of Points |
|  | SR<br>(1 pt/SR) | CR<br>(4 pts/CR) |                   |
| Totals   | 48              | 7                |                   |
| Goal 1<br>Skills and Processes of<br>Biology                         |                 |                  | 21%               |
| Expectation 3.1<br>Structure and Function of<br>Biological Molecules |                 |                  | 16%               |
| Expectation 3.2<br>Structure and Function of<br>Cells and Organisms  |                 |                  | 17%               |
| Expectation 3.3<br>Inheritance of Traits                             |                 |                  | 17%               |
| Expectation 3.4<br>Mechanism of Evolutionary<br>Change               |                 |                  | 12%               |
| Expectation 3.5<br>Interdependence of Organisms<br>in the Biosphere  |                 |                  | 17%               |

Table 1.4 English I Blueprint

| ENGLISH  |           |           |           |                   |
|--|-----------|-----------|-----------|-------------------|
| Reporting Category   | ITEM TYPE |           |           | Percent of Points |
|  | SR        | BCR       | ECR       |                   |
|  | (1pt/SR)  | (3pt/BCR) | (4pt/ECR) |                   |
| TOTALS   | 50        | 2         | 1         |                   |
| Goal 1<br>The student will demonstrate the ability to respond to a text by employing personal experiences and critical analysis.   |           |           |           | 35%               |
| Goal 2<br>The student will demonstrate the ability to compose in a variety of modes by developing content, employing specific forms, and selecting language appropriate for a particular audience and purpose. |           |           |           | 31%               |
| Goal 3<br>The student will demonstrate the ability to control language by applying the conventions of standard English in writing and speaking.  |           |           |           | 20%               |
| Goal 4<br>The student will demonstrate the ability to evaluate the content, organization, and language of texts.   |           |           |           | 14%               |

Table 1.5 Geometry Blueprint

| GEOMETRY  |           |            |            |            |                   |
|---|-----------|------------|------------|------------|-------------------|
| Reporting Category  | ITEM TYPE |            |            |            | Percent of Points |
|   | SR        | SPR        | BCR        | ECR        |                   |
|   | (1pt/SR)  | (1 pt/SPR) | (3 pt/BCR) | (4 pt/ECR) |                   |
| Totals  | 26        | 6          | 2          | 3          |                   |
| Expectation 2.1<br>The student will represent and analyze two and three dimensional figures using tools and technology when appropriate.        |           |            |            |            | 32%               |
| Expectation 2.2<br>The student will apply geometric properties and relationships to solve problems using tools and technology when appropriate. |           |            |            |            | 34%               |
| Expectation 2.3<br>The student will apply concepts of measurement using tools and technology when appropriate.                                  |           |            |            |            | 34%               |

Table 1.6 Government Blueprint

| GOVERNMENT   |           |            |            |                   |
|--|-----------|------------|------------|-------------------|
| Reporting Category   | ITEM TYPE |            |            | Percent of Points |
|  | SR        | BCR        | ECR        |                   |
|  | (1 pt/SR) | (4 pt/BCR) | (4 pt/ECR) |                   |
| Totals   | 50        | 7          | 1          |                   |
| Expectation 1.1<br>The student will demonstrate understanding of the structure and functions of government and politics in the United States   |           |            |            | 26-31%            |
| Expectation 1.2<br>The student will evaluate how the United States government has maintained a balance between protecting rights and maintaining order.  |           |            |            | 23-28%            |
| Goal 2<br>The student will demonstrate an understanding of the history, diversity, and commonality of the peoples of the nation and world, the reality of human interdependence, and the need for global cooperation, through a perspective that is both historical and multicultural. |           |            |            | 15%               |
| Goal 3<br>The student will demonstrate an understanding of geographic concepts and processes to examine the role of culture, technology, and the environment in the location and distribution of human activities throughout history.  |           |            |            | 13%               |
| Goal 4<br>The student will demonstrate an understanding of the historical development and current status of economic principles, institutions, and processes needed to be effective citizens, consumers, and workers.  |           |            |            | 18%               |



## Item Selection and Form Design

In order to conserve the item pool, the operational set of items consisted of both a common set of items shared across forms within an administration and also a unique set of items. Approximately 30% of the total form was common across each of the operational test sections within each of the January and May forms. The balance of the forms consisted of different mixtures of items depending on the form. The guidelines used to construct the forms were listed in Tables 1.7 and 1.8. The exact composition of the forms varied slightly based on available items in the pool.

Table 1.7 January Administration

| Primary Week                    | Make-Up #1   | Make-Up #2 <sup>1</sup>  |
|---------------------------------|--|--|
| January common set - 30%        | January common set - 30%   | January common set - 30%   |
| Unique Items from the pool -70% | Items from January Operational - 35% <sup>2</sup>                | Items from January Operational - 35% <sup>2</sup>                |
|                                 | Unique Items from the pool - 35%                                 | Unique Items from the pool - 35%                                 |
| Field Test Section - 2 versions | Field Test Section - same as 1 <sup>st</sup> operational version | Field Test Section - same as 1 <sup>st</sup> operational version |

Notes. <sup>1</sup>For Government and Biology, the same make-up form was administered for both administrations.

<sup>2</sup>Items from the January Operational administration included in Make-up 1 and 2 must be different.

Table 1.8 May Administration

| Primary Week                    | Make-Up #1   | Make-Up #2   |
|---------------------------------|--|--|
| May Common Set - 30%            | May Common Set - 30%   | May Common Set - 30%   |
| Unique Items from the pool -70% | Items from May Operational - 35% <sup>1</sup>                    | Items from May Operational - 35% <sup>1</sup>                    |
|                                 | Unique Items from the pool - 35%                                 | Unique Items from the pool - 35%                                 |
| Field Test Section - 8 versions | Field Test Section - same as 1 <sup>st</sup> operational version | Field Test Section - same as 1 <sup>st</sup> operational version |

Notes. <sup>1</sup>Items from the May Operational administration included in Make-up 1 and 2 must be different.

In addition to the operational items, an embedded field test section was included with each version of the test form, resulting in several versions of the operational form that

differed only by the set of field test items. These items consisted of either newly written items or previously administered items that had poor item statistics and/or had been revised. Items eligible for re-field testing included items from the 2000-2001 administration years. These items were judged to be acceptable from a content perspective, but had p-values less than 0.25, item-total correlations of less than 0.15, collapsed score levels for constructed response items (i.e., very few responses in the top score levels), very high omit rates or SR items with one best answer, but with positive point-biserials on one or more distracters. For the administration, different versions of the forms were spiraled at the student level.

Forms were constructed using the test construction software associated with the customer item bank. The goal was to match the conditional standard error curve (CSEM) and test characteristic curves (TCC) with the “target” form defined as the base form used to set the operational scale in 2002. The information function, standard error curve, and test characteristic curve were graphical displays based on the item parameters associated with the items selected and were inter-related – that is, changes to the set of items selected will result in changes in all three displays.

The following were general steps completed during the test construction process.

1. For each administration, all forms were constructed simultaneously; in order provide the best opportunity to construct parallel forms.
2. First the common set of items was selected. Then items that matched the test blueprint were selected to match the target test characteristic and standard error curves.
3. During the test construction procedure test developers were careful to ensure that the item selections met all content specifications, including matching items to the test blueprint, distribution of keys, removal of clueing, etc.
4. After the operational forms were selected, the field test sections were constructed. Field test sections did not need to meet any psychometric criteria, but were selected such that the items could be completed within a 30-minute time frame. Field test sections consisted of a set of multiple choice items, a combination of brief constructed response items and multiple choice items, or an extended constructed response item. The field test section was included at the end of Session 2.

In each content area, TCCs and CSEMs for each of the test forms are plotted in figures 1.1 to 1.10. In general the TCCs and CSEMs closely matched the target. Where forms varied in difficulty, differences were minimized in the scale score region of the cut-scores and, in all cases the difference was less than 5% of the total raw score, i.e. the passing raw score difference of the two forms is less than 5%.

Figure 1.1 Test Characteristic Curve: Algebra

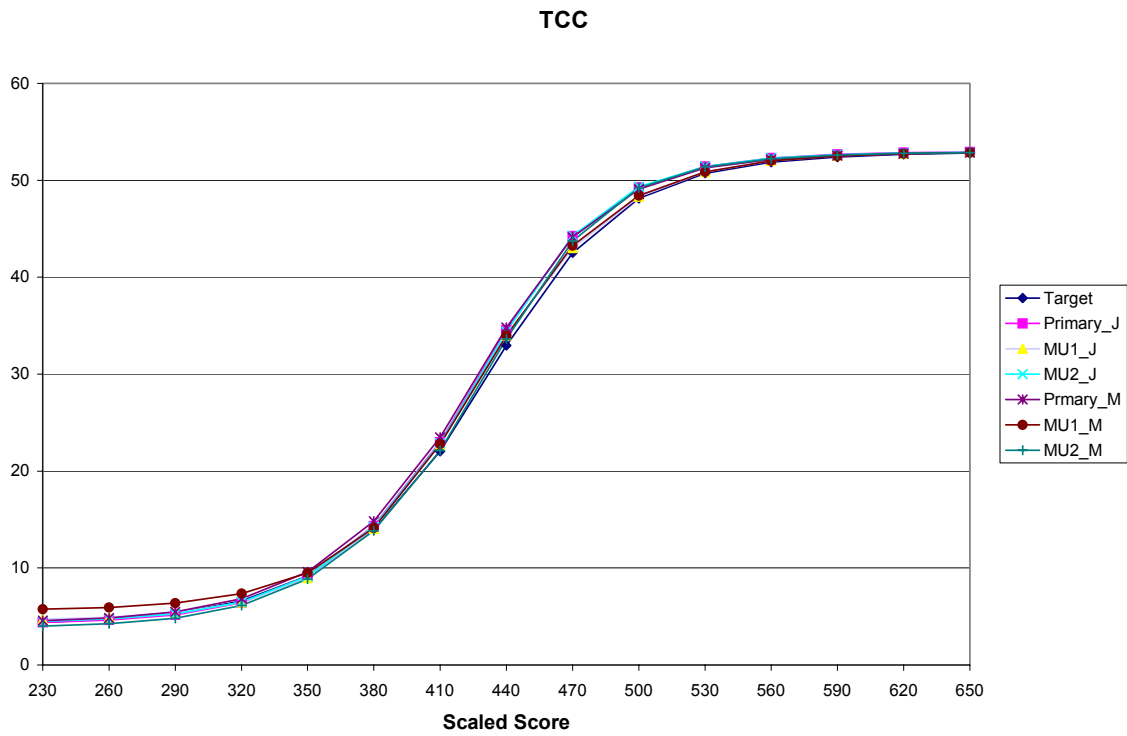


Figure 1.2. Conditional Standard Error of Measurement: Algebra

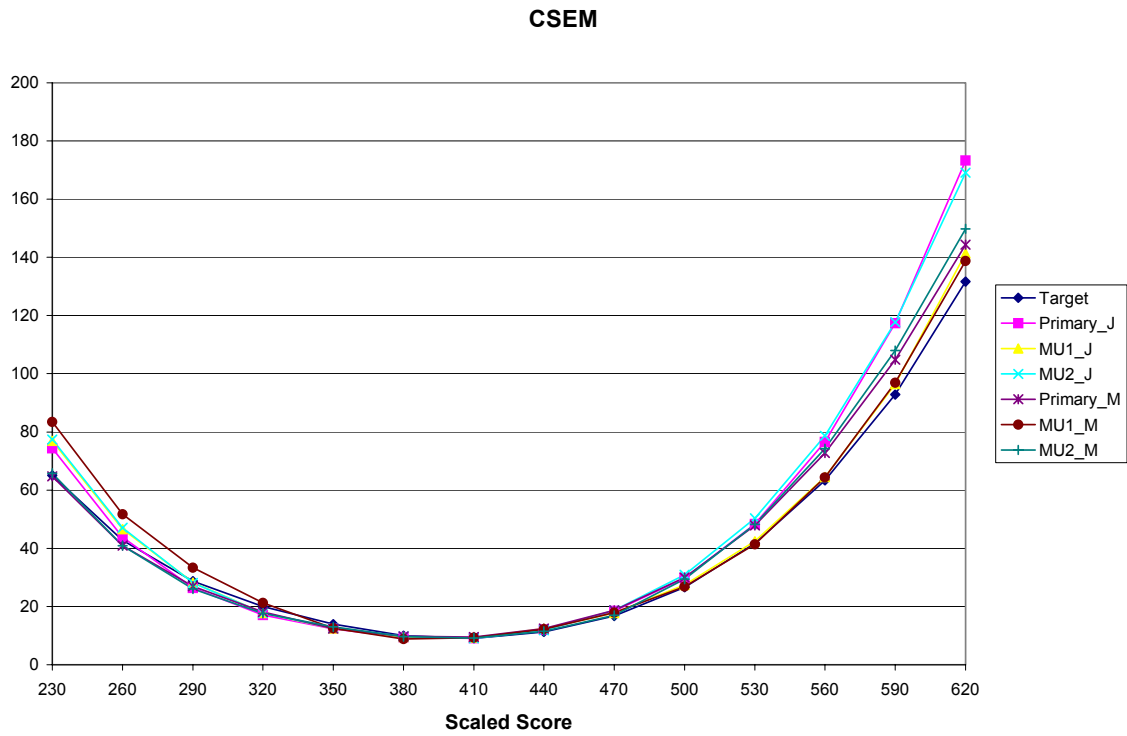


Figure 1.3 Test Characteristic Curve: Biology

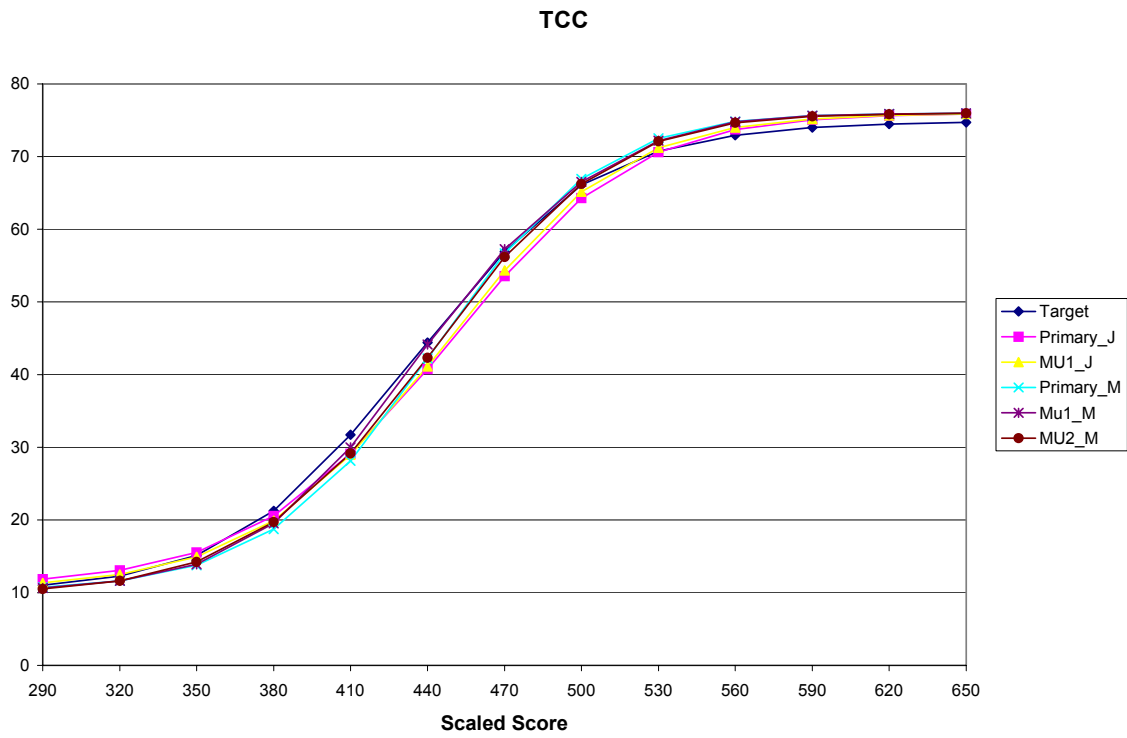


Figure 1.4 Conditional Standard Error of Measurement: Biology

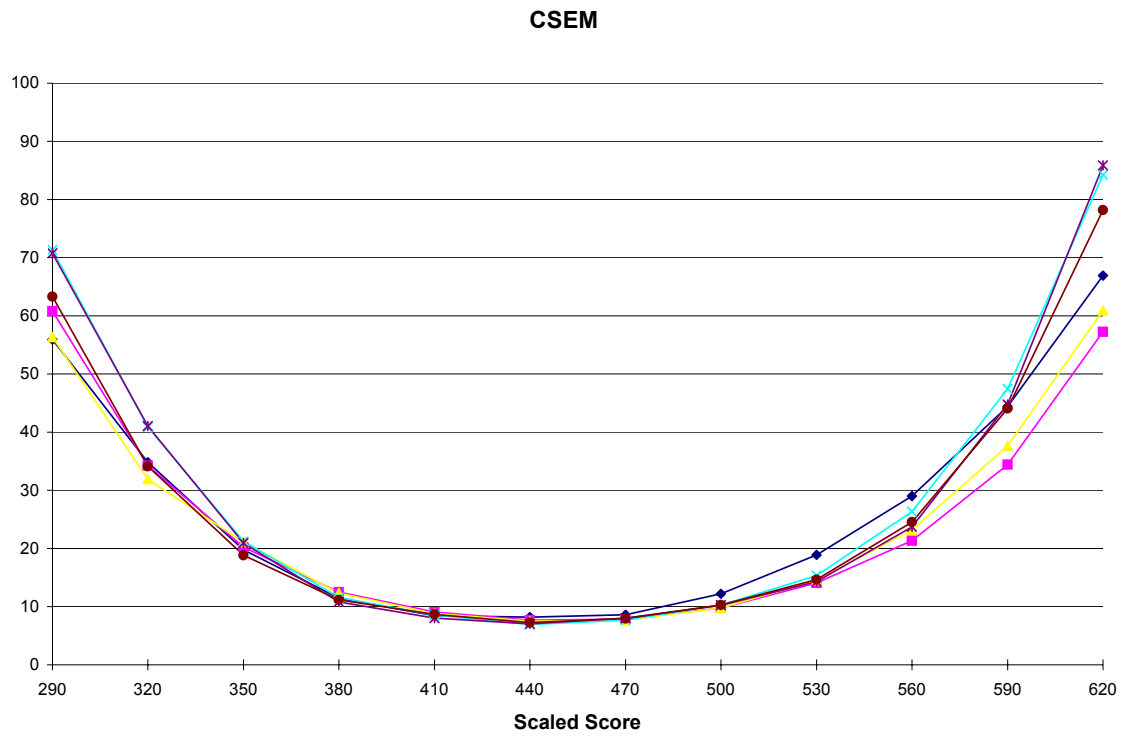


Figure 1.5. Test Characteristic Curve: English I

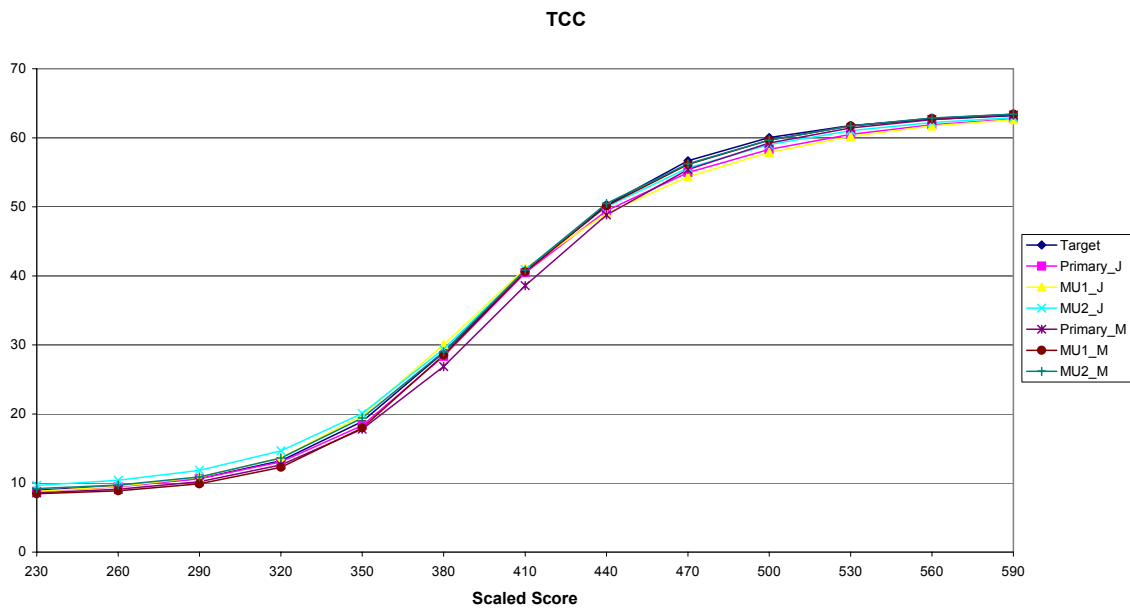


Figure 1.6. Conditional Standard Error of Measurement: English I

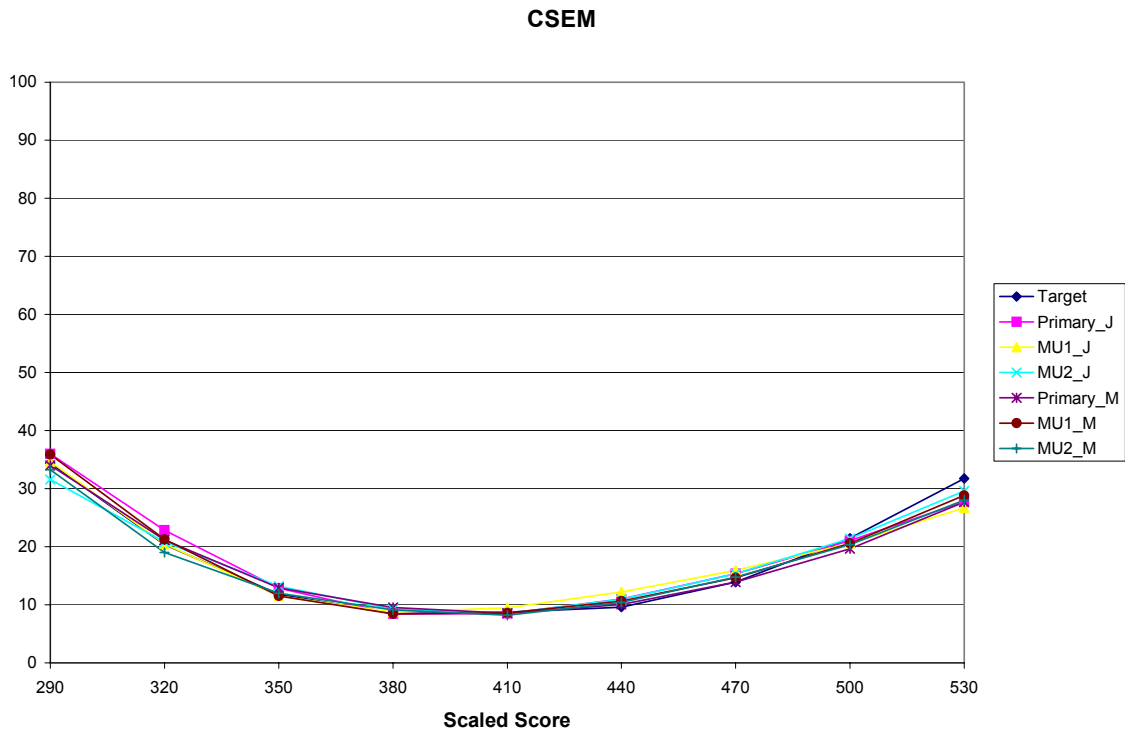


Figure 1.7. Test Characteristic Curve: Geometry

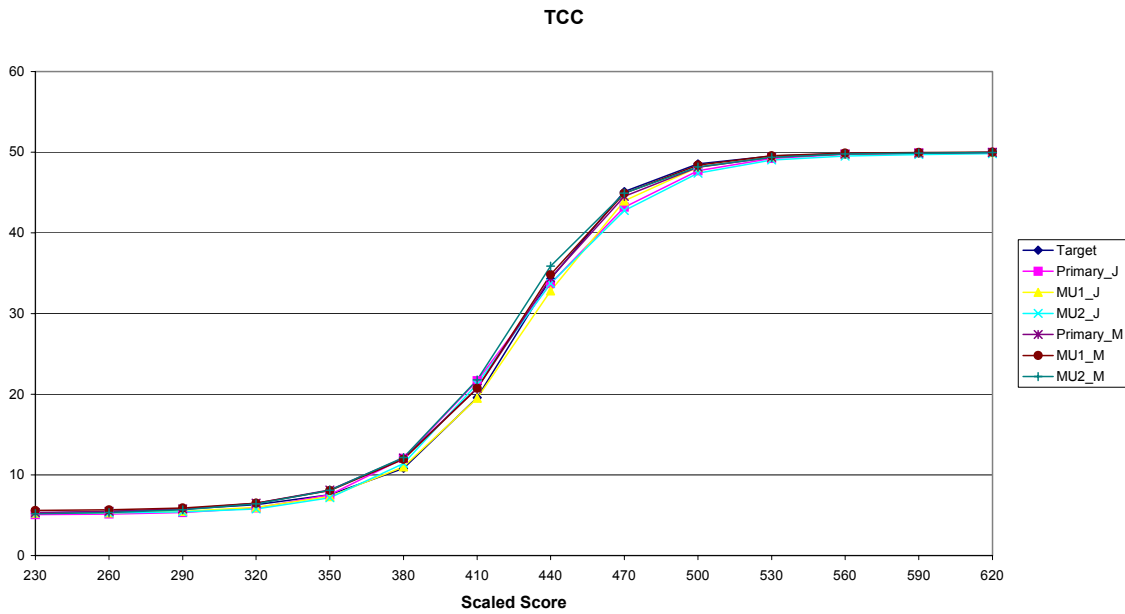


Figure 1.8. Conditional Standard Error of Measurement: Geometry

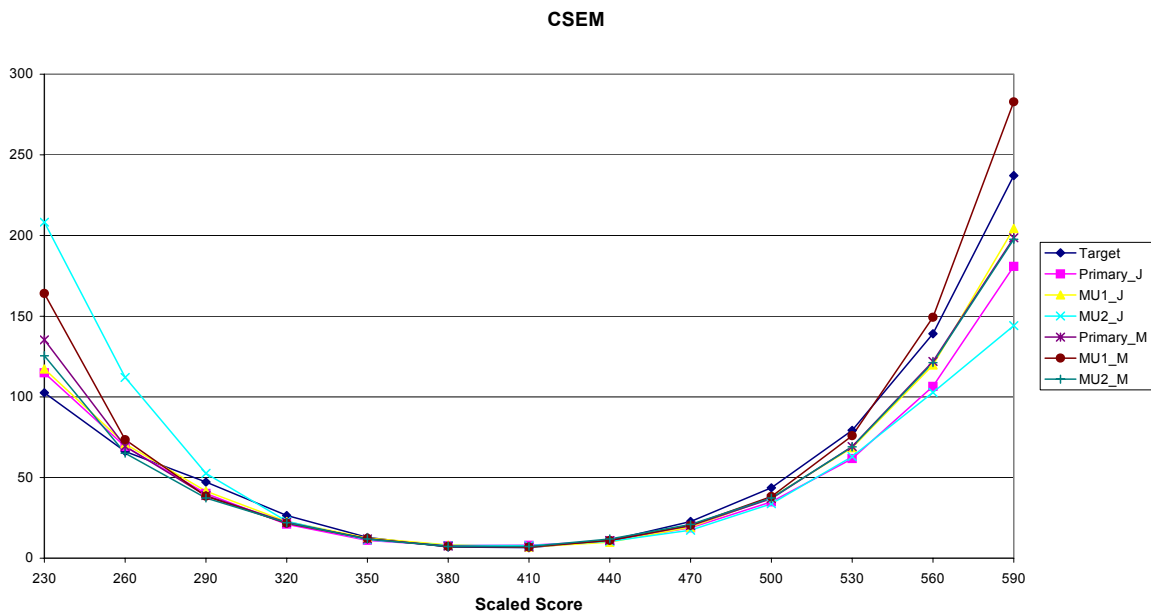


Figure 1.9. Test Characteristic Curve: Government

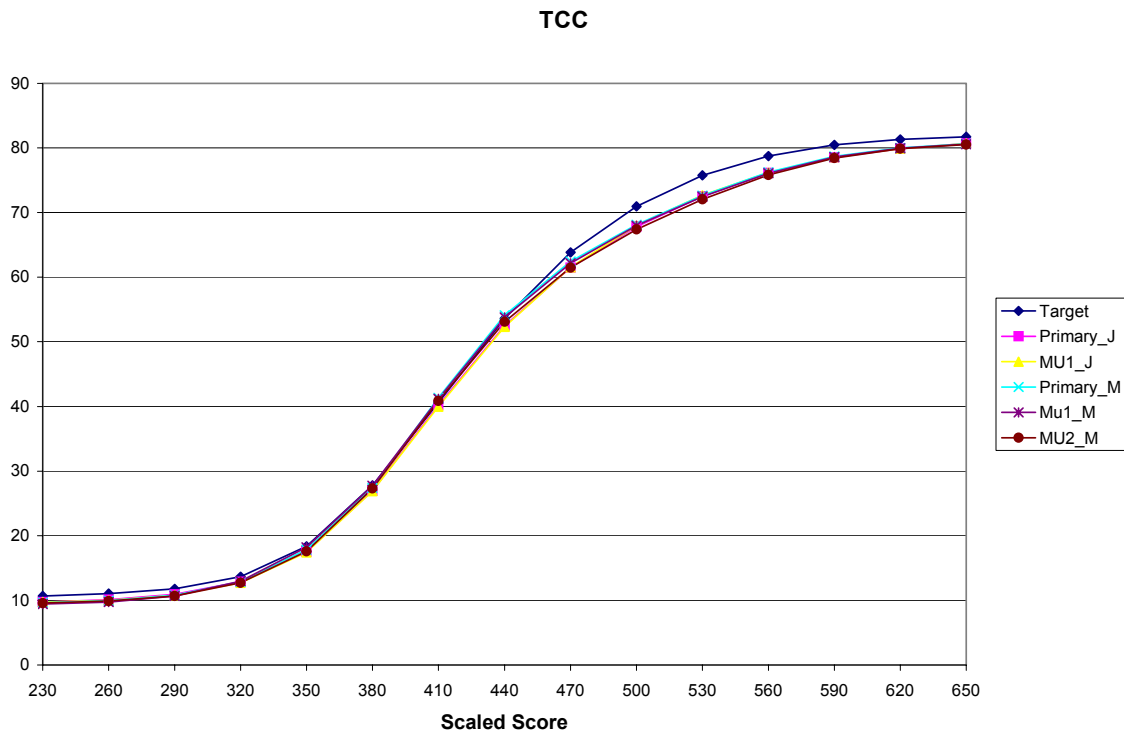


Figure 1.10. Conditional Standard Error of Measurement: Government

