

September 21, 2000

To: Nancy Bene', Janet Bagsby, Mark Moody  
Cc: Rhonda Clymer, Margie Tully  
From: Gary Schaeffer and Dianne Henderson

Subject: Shortening the HSA

One of the issues discussed at a recent conference call was the possible reduction in the number of items in the tests. Reductions are being considered to allow for more fieldtesting during the testing session, to reduce the need for operational items, and to respond to educators requests for shorter tests. Two important considerations in determining the number of items are content coverage (content validity) and test reliability.

Table 1 shows the numbers and percents of each item type in the May 2000 field test. The internal-consistency reliability estimates of the field test forms were .93 - .94. These may be underestimates, given that field test students may not have been motivated and therefore may have guessed on items they could have otherwise answered correctly. A common rule-of-thumb for sufficient test reliability is .90 for high-stakes tests. As such, the current blueprints may contain more items than are needed for sufficient reliability.

For comparison purposes, Table 2 shows the current revised blueprints for Algebra and Geometry. These blueprints were not used in the May 2000 field test. Using the Spearman-Brown prophecy formula and the results from May, the reliabilities of these blueprints were estimated to be .92 - .93.

The Spearman-Brown prophecy formula was used to estimate the reduced numbers of items needed to attain a lower reliability level. A target reliability of .91 was used, to try to ensure a reliability of at least .90. In addition, to address content validity concerns, the proportions of items types were kept approximately the same in the shorter versions of the test as in the 2000 field test versions. Table 3 shows the estimated numbers of items in the blueprints needed to obtain a reliability = .91. The last four columns of the table show the differences in numbers of items between the blueprints with reliability = .91 and (a) the May 2000 field test blueprints for Biology, English, and Government, and (b) the current revised blueprints for Algebra and Geometry. This reduction in reliability generally led to a large reduction in the numbers of items, particularly in SR items. For example, Biology, English, and Government could be reduced by 17-18 SR items. Decreasing the tests by these large numbers of item may lead to concerns about breadth of content coverage and test validity.

Table 4 shows alternative blueprints that reflect an approximately 10-minute reduction in testing time from the May 2000 blueprints for Biology, English, and Government, and from the current revised blueprints for Algebra and Geometry. The numbers of each item type are in similar proportions to the earlier blueprints. These are based on the following approximate time estimates: SR = 1 min; SPR = 3 min; BCR = 5 min; ECR = 15 min.

A 10-minute reduction in operational test time, as outlined in Table 4, appears to be a reasonable approach to consider to shorten the tests and satisfy content validity and reliability concerns.

Table 1. May 2000 Field Test

|            | Total | SR | SR % | SPR | SPR % | BCR | BCR % | ECR | ECR % | May Rel |
|------------|-------|----|------|-----|-------|-----|-------|-----|-------|---------|
| Algebra    | 54    | 37 | 69   | 7   | 13    | 6   | 11    | 4   | 7     | .93     |
| Biology    | 55    | 48 | 87   | 0   | 0     | 7   | 13    | 0   | 0     | .94     |
| English    | 53    | 50 | 94   | 0   | 0     | 2   | 4     | 1   | 2     | .94     |
| Geometry   | 59    | 40 | 68   | 10  | 17    | 5   | 8     | 4   | 7     | .94     |
| Government | 60    | 52 | 87   | 0   | 0     | 7   | 12    | 1   | 2     | .94     |

Table 2. Current Revised Algebra and Geometry Blueprints

|          | Total | SR | SR % | SPR | SPR % | BCR | BCR % | ECR | ECR % | Est. Rel. |
|----------|-------|----|------|-----|-------|-----|-------|-----|-------|-----------|
| Algebra  | 49    | 34 | 69   | 7   | 14    | 5   | 10    | 3   | 6     | .92       |
| Geometry | 49    | 33 | 67   | 9   | 18    | 4   | 8     | 3   | 6     | .93       |

Table 3. Estimates of Numbers of Items for Reliability of .91

|            | Total | SR | SPR | BCR | ECR | SR dif | SPR dif | BCR dif | ECR dif |
|------------|-------|----|-----|-----|-----|--------|---------|---------|---------|
| Algebra    | 42    | 29 | 5   | 5   | 3   | 8      | 2       | 0       | 0       |
| Biology    | 36    | 31 | 0   | 5   | 0   | 17     | 0       | 2       | 0       |
| English    | 35    | 33 | 0   | 1   | 1   | 17     | 0       | 1       | 0       |
| Geometry   | 39    | 26 | 7   | 3   | 3   | 7      | 2       | 1       | 0       |
| Government | 40    | 34 | 0   | 5   | 1   | 18     | 0       | 2       | 0       |

Table 4. Estimates of Numbers of Items to Reduce Testing Time by Approximately 10 Minutes

|            | Total | SR | SPR | BCR | ECR | SR dif | SPR dif | BCR dif | ECR dif | Est. Rel. |
|------------|-------|----|-----|-----|-----|--------|---------|---------|---------|-----------|
| Algebra    | 42    | 29 | 5   | 5   | 3   | 5      | 2       | 0       | 0       | .91       |
| Biology    | 50    | 44 | 0   | 6   | 0   | 4      | 0       | 1       | 0       | .93       |
| English    | 43    | 40 | 0   | 2   | 1   | 10     | 0       | 0       | 0       | .93       |
| Geometry   | 42    | 28 | 7   | 4   | 3   | 5      | 2       | 0       | 0       | .92       |
| Government | 53    | 46 | 0   | 6   | 1   | 6      | 0       | 1       | 0       | .93       |