Appendix 5.A Maryland High School Assessment Special Study: Directional Statements accompanying the Government Constructed Responses

Maryland High School Assessment Special Study:

Directional Statements Accompanying the Government Constructed Responses

December 2004

Educational Testing Service

Maryland High School Assessment Special Study: Directional Statements accompanying the Government Constructed Responses

Background

The HSA is based on a pre-equated design– that is, items were not recalibrated following administrations, and instead bank parameters were used for scoring. As a result, the items must appear exactly as they did in the administration associated with the bank parameters. Any change to the item can result in change how students interact with the item and the resulting item parameters. Therefore, items cannot be modified: text cannot be edited or revised or graphics altered.

In the evolution of the item writing process, the directional statements associated with the Government brief and extended constructed response items were modified to be more specific, beginning with the May, 2004 administration (see Figure 5.A.1). In reviewing the item bank, there were several items that could be used on future forms, however, these items included the previous directional statements and formatting (see Figure 5.A.2). As a result, available items have two different formats and future test forms could include items with both types of formatting. While changing all of the items to the "new" format would be desirable, MSDE was concerned that this change could impact item performance. To obtain new item parameters, the items would need to be re-field tested, which would decrease the numbers of items available for form construction in the short term, would delay the field testing of newly written items, and would increase the development costs associated with these existing items. In reviewing the change it was hypothesized that item performance would not differ based on the modification. Therefore, a study was completed during the May 2004 administration that involved printing two items in both the old and new formats to help evaluate whether or not the items would need to be field tested if they were reformatted to the new style guidelines.

Figure 5.A.1. Government Brief Constructed Response Item: With Instruction

69. Read the sentences below and use them to complete the BRIEF CONSTRUCTED RESPONSE that follows.

Read the scenario below.
Recently a city ordinance [law] was passed that banned skateboard riding on most city streets and sidewalks. You and your friends believe this is an unjust law.

- **Describe two legal ways you and your friends could try to get this law changed.**
- **Explain why each of your choices would be effective.**
- **Include details and examples to support your answer.**

Write your answer on the lines in your Answer Book.

Figure 5.A.2. Government Brief Constructed Response Item: Without Instruction

69.

Read the scenario below.
Recently a city ordinance [law] was passed that banned skateboard riding on most city streets and sidewalks. You and your friends believe this is an unjust law.

- **Describe two legal ways you and your friends could try to get this law changed.**
- **Explain why each of your choices would be effective.**
- **Include details and examples to support your answer.**

Write your answer on the lines in your Answer Book.

Method and Results

In May 2004, two BCR items were selected and included in the field test sections in both the old and new formats. The classical item statistics in Table 5.A.1 show that the two versions of the items were very similar in terms of p-values and poly-serial correlations. We also compared the IRT parameter estimates of the items in each format, and noted that these values were very similar as well (see Table 5.A.2). Figures 5.A.3 and 5.A.4 show the item characteristic curves for the two different versions of items 1 and 2. Figure 5.A.5 and 5.A.6 show the item characteristic curves for each response option for the two different versions of items 1 and 2. Figure 5.A.7 and 5.A.8 show the item information function for the two versions of item 1 and 2.

Table 5.A.1: Classical Item Statistics

	P value		Poly-serial correlation	
	New	Old	New	Old
Item 1 Reappointment/Political Power	MD52236 N=6813	MD68796 N=6378	MD52236 N=6813	MD68796 N=6378
	0.19	0.20	0.31	0.31
Item 2 Due Process/ Public Safety v Rights	MD52234 N=6378	MD68795 N= 6307	MD52234 N=6378	MD68795 N=6307
	0.78	0.76	0.65	0.65

Appendix 5.A

Table 5.A.2. Frequency Distribution of Score Points

	Percent Score 0		Percent Score 1		Percent Score 2		Percent Score 3		Percent Score 4	
	New	Old	New	Old	New	Old	New	Old	New	Old
Item 1 Reappointment/Political Power	0.50	0.50	0.26	0.24	0.21	0.22	0.03	0.03	0.00	0.00
Item 2 Due Process/ Public Safety v Rights	0.19	0.20	0.42	0.40	0.34	0.35	0.04	0.04	0.00	0.00

Table 5.A.3. IRT Parameter Estimates

	A-Value		B1-Value		B2-Value		B3-Value		B4-Value	
	New	Old	New	Old	New	Old	New	Old	New	Old
Item 1 Reappointment/Political Power	0.03125	0.02926	414.1	420.0	435.3	433.8	497.0	498.2	561.7	553.4
Item 2 Due Process/ Public Safety v Rights	0.02321	0.02237	371.1	373.1	424.3	422.5	500.6	507.1	578.7	570.3

Conclusion and Recommendation

Any time item parameter estimates are obtained with different samples, some differences occur due to sampling error. This study found that there were minimal differences between the two sets of item parameters. Thus, this change in the directions does not appear to have had an important or systematic effect on item performance.

Appendix 5.A

Figure 5.A.3: Item Characteristic Curve for CR item 1.

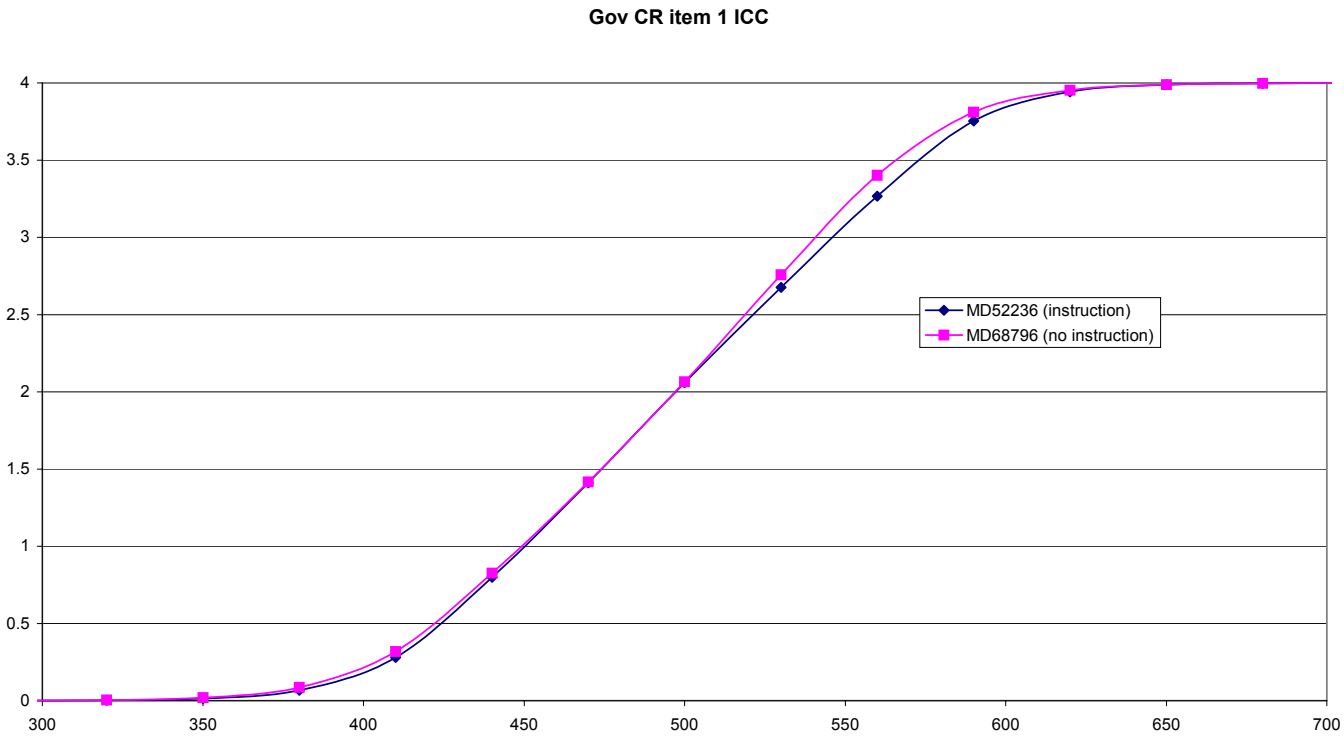
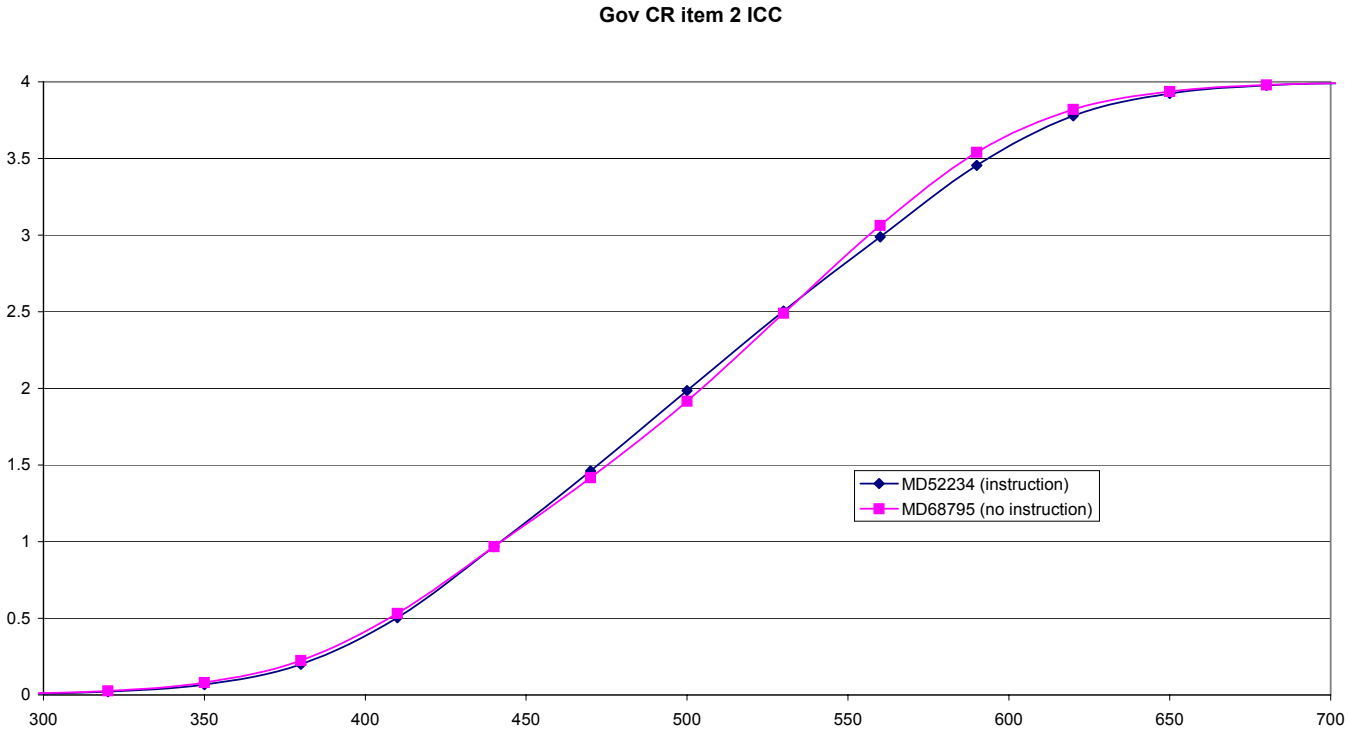


Figure 5.A.4: Item Characteristic Curve for CR item 2



Appendix 5.A

Figure 5.A.5: Item Characteristic Curve for each Response Option of Item 1

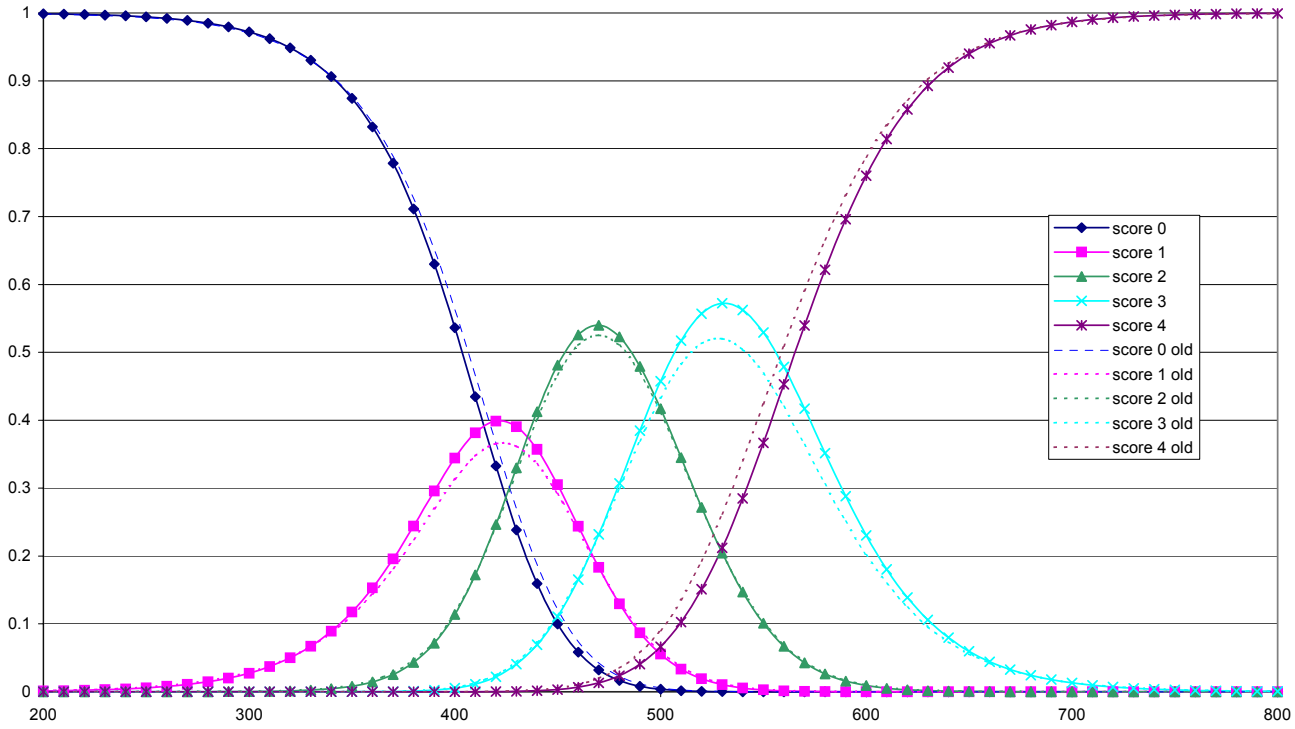
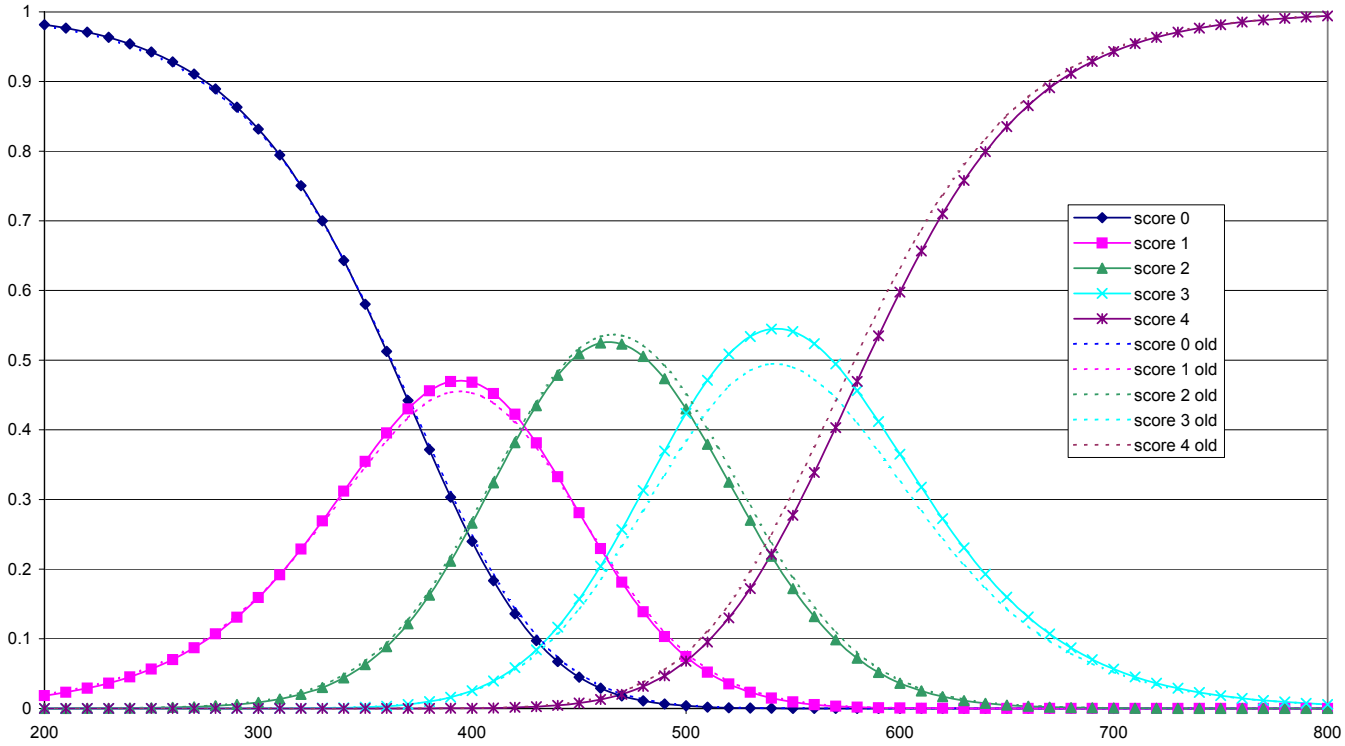


Figure 5.A.6: Item Characteristic Curve for each Response Option of Item 2



Appendix 5.A

Figure 5.A.7: Information function for CR item 1

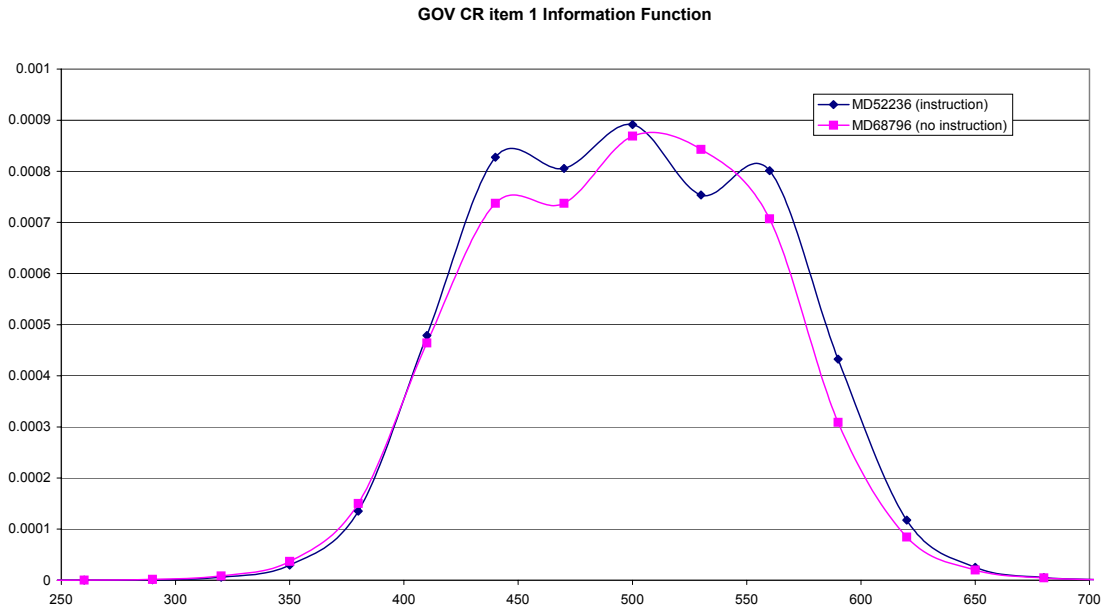


Figure 5.A.8: Information function for CR item 2