

Reading your Item Analysis Report: Item is making a limited contribution to the measurement capacity of the test

By Dr. Michaela Geddes

September 2015

Following an exam administration, Yardstick often produces an Item Analysis report from their proprietary software called COGs. The item analysis report provides information about each item in terms of its difficulty, discrimination, and the distribution of responses across alternatives. This backgrounder focuses on a comment from the COGs item analysis report that says **“Item is making a limited contribution to the measurement capability of the test.”**

This comment is specifically linked to the corrected item discrimination index (CRPB). Item discrimination values range from +1.00 to -1.00 and tell you how strong the relationship is between performance on a particular item and performance on the test as a whole:

- Positive values indicate that higher performing examinees answer the item correctly while lower performing examinees select an incorrect option
- Positive values close to zero (or zero) indicate that high and low scorers have an extremely similar chance of answering the question correctly
- Negative values indicate that examinees who perform poorly overall are more likely to answer the question correctly than high performing examinees

What does it mean?

Items are automatically assigned the comment "Item is making a limited contribution to the measurement capability of the test," based on a predetermined range for the corrected item discrimination index (CRPB). As a default, the CRPB range is between 0.0 and 0.2, exclusive, though one can adjust this range based on the testing program requirements. The default range is based on best practices in which the rule of thumb is that discrimination values greater than or equal to 0.2 are deemed acceptable. Accordingly, if the discrimination index is less than 0.2 the item should be reviewed.

Based on this "rule of thumb," one should review items with CRPB's less than 0.2 as such items may have poor predictive ability of total test performance.

What should you do?

An example of an item that is "making a limited contribution to the measurement capability of the test" is shown in Figure 1. As you can see on the left hand side, the CRPB is 0.1074, and should therefore be reviewed by content experts during key validation. From the table below the graph, one can see that nearly half (49%) of low performers select the correct answer and 66% of high performers select the correct answer. From a review of this item, one may conclude that this question is acceptable for scoring, though in the future it may be worth developing a new distractor for option B.

Figure 1. Example of item with "limited contribution to the measurement capabilities of the test"

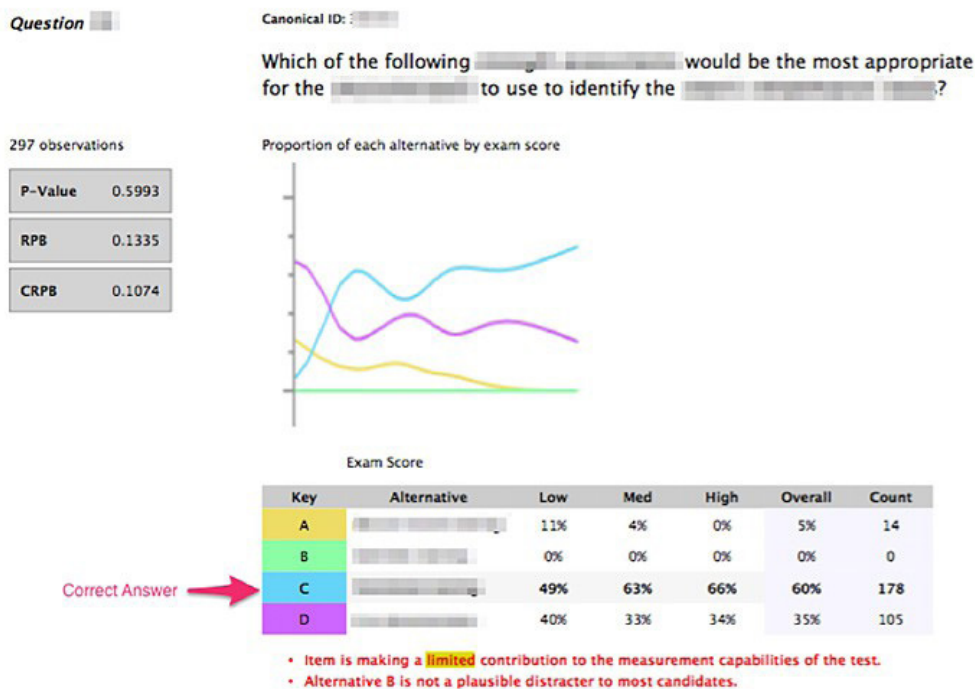
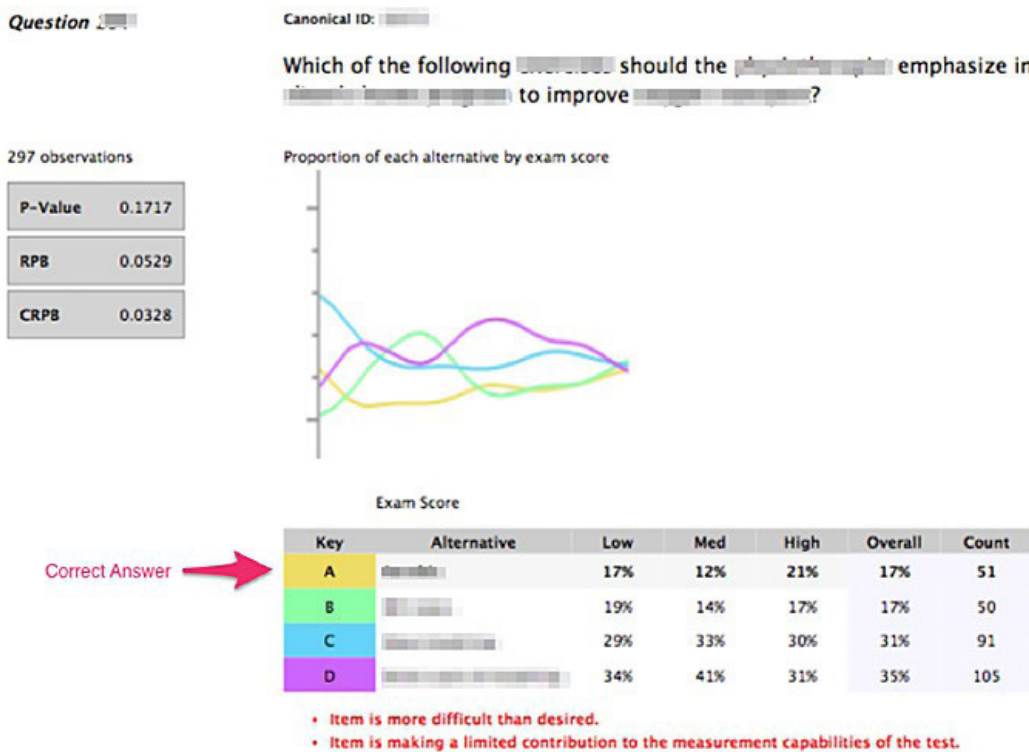


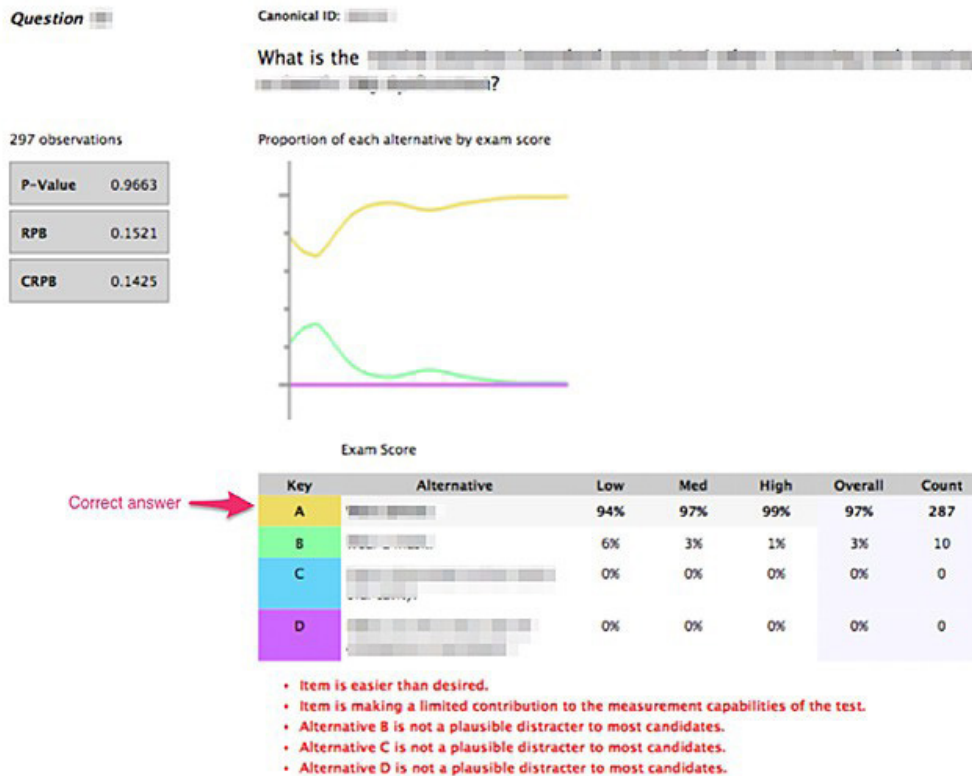
Figure 2 shows a different example of an item with a low CRPB. In this example, option A is keyed as “correct,” yet only 17% of examinees select the correct answer (a percentage less than guessing); a very difficult question. Another 17% of examinees selected option B, 31% selected option C, and 35% selected option D. You can also see that 21% of high performers and 17% of low performers selected the correct answer; this item is poorly discriminating between high and low performers. Taking all the information into account, it is clear that this item is functioning poorly and thus the “item is making a limited contribution to the measurement capabilities of the test.” Based on a review of this item, one may conclude that although option A is the only correct answer, the item should be removed from an examinees total score.

Figure 2. Example of poor item with “limited contribution to the measurement capabilities of the test”



A final example is shown in Figure 3 below. This item exhibits a low CRPB (0.1425), likely because nearly all (97%) examinees selected the correct answer; an easy item. Accordingly, this item poorly discriminates between high and low examinees because nearly all examinees got the correct answer. Following a key validation review, one may conclude to score the question on the test, but to retire the question from future operational administrations, as the question is too easy and does not contribute to the purpose of the test (e.g., differentiate competent from not-yet-competent).

Figure 3. Example of easy item with “limited contribution to the measurement capabilities of the test”



To read more articles related to eLearning, examination, and instructional design go to www.getyardstick.com and check out our blog.