Evaluating your distractors

By Dr. Michaela Geddes   **November 2015**

In addition to looking at an item's p-value and discrimination index to determine how well an item is functioning, it is also important to analyze the distractor choice.

The study of distractors is important for SMEs to better understand the performance of an item. Accordingly, distractor analyses can be used in an item's revision process. Distractor evaluation is also helpful during key validation as it can help determine whether an item has a key error or more than one correct answer.

This backgrounder discusses two methods for distractor evaluation: **tabular** and **graphical**.
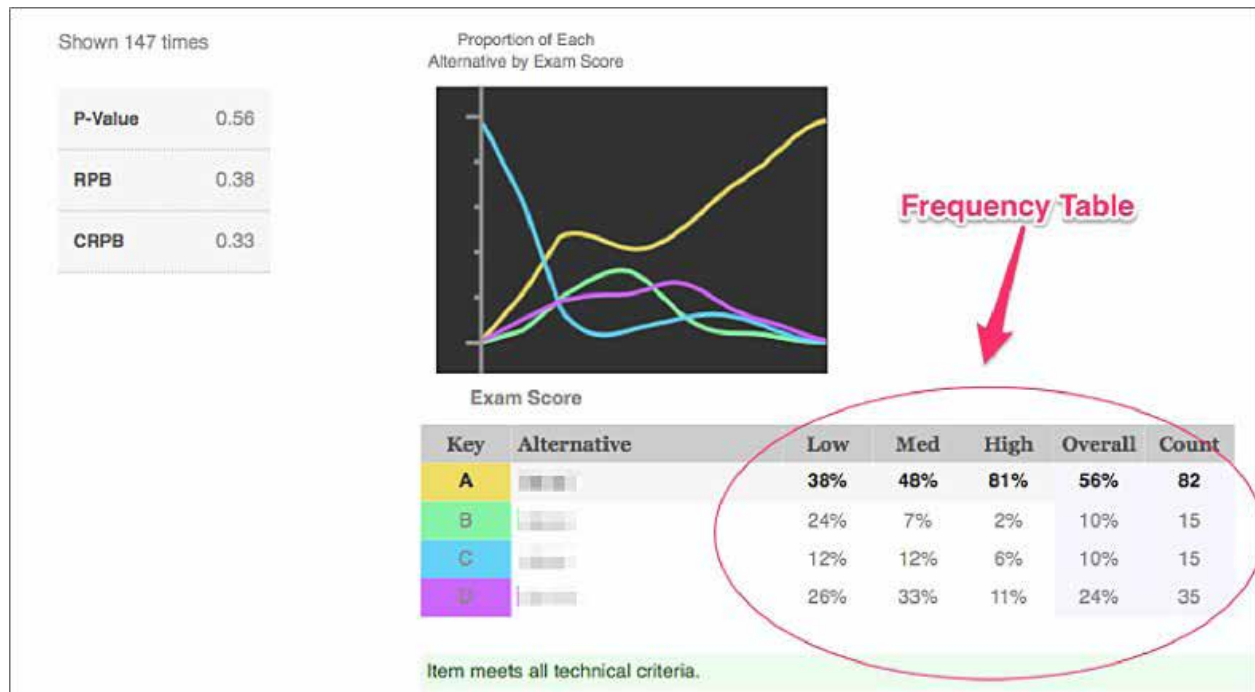
Tabular Method For Studying Distractor Performance

The tabular method comprises a frequency table for all options of a multiple-choice item according to ability groups. Ability groups may be defined as low, medium, and high performers, as shown in the examples below.

In Figure 1, the correct answer, option A, was chosen 56% of the time; the correct answer was chosen by 38% of the low ability group, 48% of the medium ability group, and 81% of the high ability group. This monotonically increasing pattern is desirable of a response for a correct answer.

Option B, a distractor, has a low response rate (2%) for the high ability group, and a higher response rate (24%) for the low ability group. This monotonically decreasing pattern is desirable for a well-performing distractor.
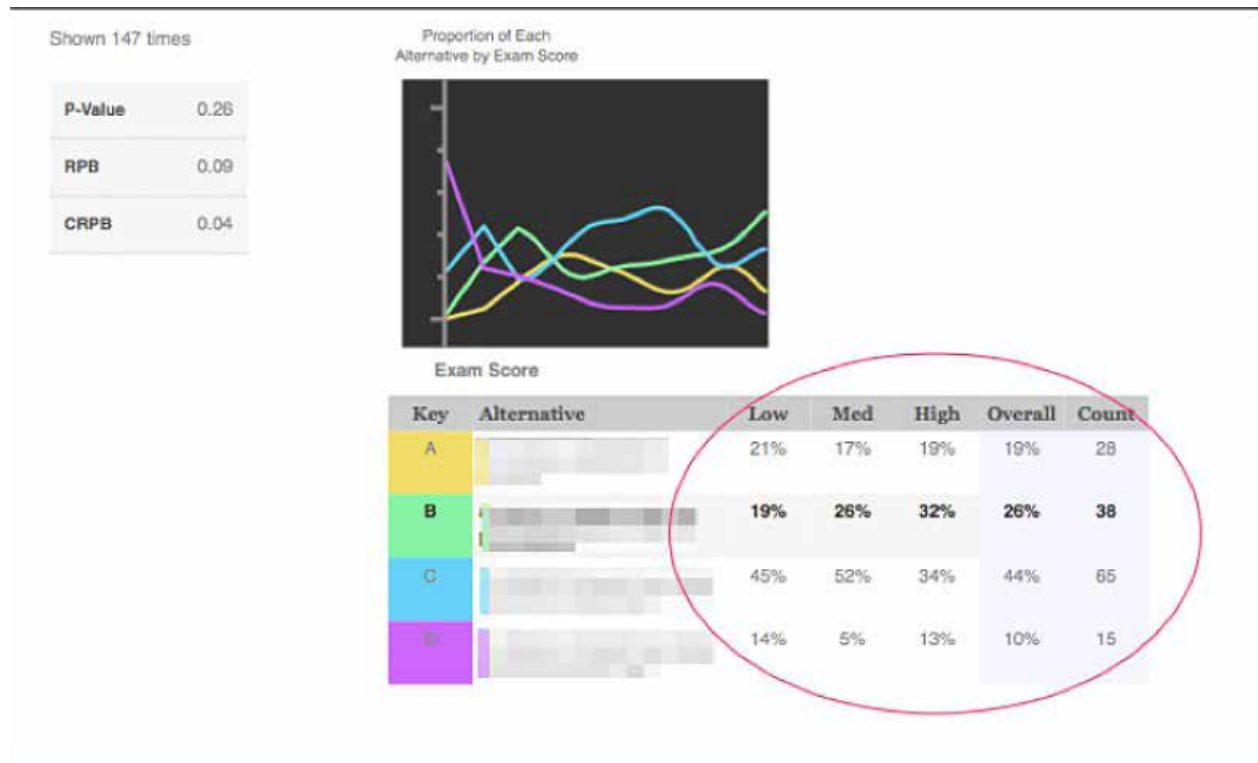
Figure 1. Frequency table for a 4-option multiple-choice item: Item A.

The second item, shown in Figure 2, exhibits a distractor pattern that presents a problem. First, the correct answer, option B, is selected less frequently (26%) than a distractor, Option C (44%), overall. In addition, Option C is more often chosen by the middle group and less often by the high or low ability groups. Accordingly, the information presented in Option C is more attractive to medium ability examinees and less attractive to high or low ability examinees. The opposite pattern in seen in Options A and D, where the options are more often chosen by the low or high ability groups and less often by the middle ability groups. Such nonmonotonic patterns are undesirable for distractors, and these distractors should ideally be rejected and replaced with distractors that present a monotonically decreasing pattern.

While some people like to see the numbers in a frequency table, a visual display of how the options are selected is also helpful.

Figure 2. Frequency table for a 4-option multiple-choice item: Item B.

			Shown 147 times				
P-Value	0.26						
RPB	0.09						
CRPB	0.04						

Proportion of Each Alternative by Exam Score

Exam Score

Key	Alternative	Low	Med	High	Overall	Count
A		21%	17%	19%	19%	28
B		19%	26%	32%	26%	38
C		45%	52%	34%	44%	65
D		14%	5%	13%	10%	15

Graphical Method For Studying Distractor Performance

Graphical methods use trace lines for the correct answer and for the distractors. Figure 3 shows four trace lines, one for each option. The correct choice is shown in green and is monotonically increasing as a function of total test score – a desirable pattern.

The blue trace line, a distractor, monotonically decreases – a desirable pattern for a well-functioning distractor. Ideally, each distractor should exhibit this pattern; an examinee's tendency to select a wrong option decreases with the examinee's ability.

However, in this example, the yellow trace line shows a flat performance across all ability examinees. This option does not discriminate across ability groups and is thus making a limited contribution to the measurement capabilities of the test. Since this option is also infrequently selected, it is not a plausible distractor to most candidates. As such, this distractor should be revised or dropped from the test and replaced with a new distractor.

The purple trace line shows a low response rate overall. It is generally a distractor that is implausible and therefore not typically selected by examinees. It too should be revised or dropped and replaced with a new distractor.

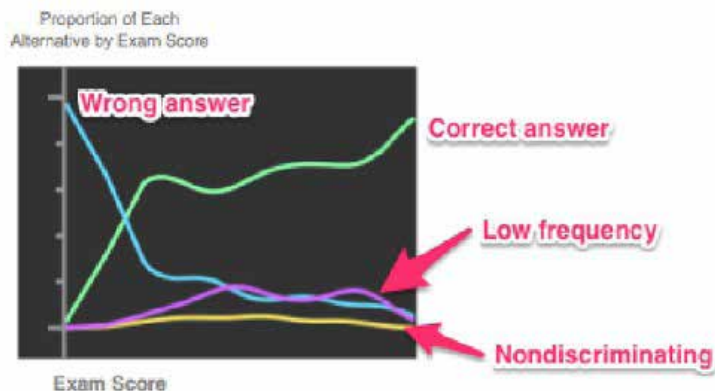## Graphical Method For Studying Distractor Performance

Graphical methods use trace lines for the correct answer and for the distractors. Figure 3 shows four trace lines, one for each option. The correct choice is shown in green and is monotonically increasing as a function of total test score – a desirable pattern.

The blue trace line, a distractor, monotonically decreases – a desirable pattern for a well-functioning distractor. Ideally, each distractor should exhibit this pattern; an examinee's tendency to select a wrong option decreases with the examinee's ability.

However, in this example, the yellow trace line shows a flat performance across all ability examinees. This option does not discriminate across ability groups and is thus making a limited contribution to the measurement capabilities of the test. Since this option is also infrequently selected, it is not a plausible distractor to most candidates. As such, this distractor should be revised or dropped from the test and replaced with a new distractor.

The purple trace line shows a low response rate overall. It is generally a distractor that is implausible and therefore not typically selected by examinees. It too should be revised or dropped and replaced with a new distractor.

Figure 3. Four types of trace lines.

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