What Item Analysis Can Tell Us About Item Quality

Definitions:

PBis stands for "point-biserial". The PBis statistic for a particular response option (either the correct response or a distractor) correlates the performance of candidates who did well on the test with whether or not the candidate chose that response option.

P-Value is the percent correct (e.g., the percent of people who answered an item correctly).

Mean Score is the average score for the candidates that selected each response.

Are these items valid?

When the more able candidates are the more likely to answer an item incorrectly, something is clearly wrong! Remember that the PBis statistic for each answer choice correlates candidates' overall scores on the examination with whether or not they choose that answer choice. Accordingly, The PBis for the correct answer choice should not be negative or near zero (.2 is okay, and .1 is marginally okay, depending on the PBis for the distractors). Conversely, the PBis for the incorrect answers (distractors) should be negative or near zero, and never higher than or equal to the PBis for the correct answer.

Note that the p-value and the PBis statistics are quite different in this regard. In good items, the PBis statistic must be higher for the right answer choice than it is for the distractors. However, this does not have to be true of the p-value. If the item is difficult and the PBis statistic for the correct answer is much higher than the PBis statistic for any of the distractors, then regardless of the p-values for the distractors and the right answer, the item may actually be an acceptable item: It successfully differentiates among candidates in the higher ability ranges.

Examine the PBis statistics for each answer option. Make sure that the correct answer has a PBis statistic that is distinctly higher than the PBis statistics for the incorrect answers, and that the PBis statistic for the correct answer is not negative or near zero (i.e., is not below 0.1).

Sample Item Statistics:

```
_____
Item Number  1: GOOD ITEM                  Correct answer = 1

    Accession Number = 1363
   PVALUE =   0.72    PT BISERIAL =        +0.22
        DISTRACTOR     1        2        3        4        5     TOTAL PEOPLE
        N OF PEOPLE    241       9        3       11       70        334
        MEAN SCORE    219.63   215.11   196.67   210.73   210.03
        PVALUE          0.72     0.03     0.01     0.03     0.21
        PT BISERIAL    +0.22    -0.02    -0.10    -0.06    -0.19
```

This item is good because the PBis (PT BISERIAL) statistic for the correct answer is above 0.2 and is substantially higher than the PBis statistics for the distractors.

_____
_____

```
Item Number  2: MIDDLING ITEM                Correct answer = 4

   Accession Number = 1375
   PVALUE =   0.39   PT BISERIAL =        +0.12
      DISTRACTOR      1       2       3       4       5      TOTAL PEOPLE
      N OF PEOPLE    13      87      40     130      64        334
      MEAN SCORE    199.54  215.97  212.03  219.98  219.00
      PVALUE         0.04    0.26    0.12    0.39    0.19
      PT BISERIAL   -0.18   -0.03   -0.10   +0.12   +0.05
```

We can tell that this item sorts candidates appropriately because the PBis for the correct answer is above the minimum of 0.1 and is noticeably higher than the PBis statistics for the distractors. With these PBis values, we can accept this item even though candidates choosing a distractor (#5) and the candidates choosing the correct answer had the same mean score. In fact, these candidates had the same mean score only because the number of people who chose the distractor is much smaller than the number of people who chose the correct answer (see "N of people").

_____

_____

```
Item Number  3: BAD ITEM                  Correct answer = 3

   Accession Number =  255
   PVALUE =   0.34   PT BISERIAL =        -0.07
       DISTRACTOR      1       2       3       4       5      TOTAL PEOPLE
       N OF PEOPLE     2       2     113      23     194        334
       MEAN SCORE    233.00  177.00  215.08  215.48  218.54
       PVALUE         0.01    0.01    0.34    0.07    0.58
       PT BISERIAL   +0.06   -0.16   -0.07   -0.02   +0.09
```

This is a bad item because PBis for correct answer is negative. The better the candidate, the less likely s/he is to choose the correct answer. Equivalently, this item is bad because the PBis for the correct answer is lower than the PBis for some of the distractors: Better candidates are more likely than worse candidates to choose a distractor.

_____