ABSTRACT
The item response curve (IRC) is foundational to selecting quality items for high stakes exams by describing the quality of an item, difficulty, and probability that a student will get the item correct.

BACKGROUND
Producing quality test items in mathematics can be very difficult and labor-intensive. It is not uncommon to inadvertently interject hints, opinions, or prejudices into a test-item so as to sway the approach of the student.

Test items should meet the following criteria: (1) be congruent with key objectives (or psychological constructs); (2) have clearly defined key objectives (or psychological constructs); (3) contribute minimally to error in measurement; (4) be presented in a format that is suitable to the test’s goals; (5) meet specified technical assumptions; (6) be well written and follow prescribed editorial standards; and (7) satisfy ethical and legal concerns (Osterlind, 2010).

If you do not have test items with merit and validity, it does you no good to do any statistical analysis on the data it may produce. As Osterlind states, “a test item or assessment exercise is likewise the heart and soul of valid assessment...no amount of item analysis, statistical manipulation, or psychometric procedure can improve the interpretation of cognition beyond the intrinsic characteristics of the item or exercise” (Osterlind, 2010, p. 217).

QUALITY ITEM
If \( r = 1 \) and \( s = 0 \), which of the following must be true?
(A) \( r < 1 \)
(B) \( s < 1 \)
(C) \( r = 1 \)
(D) \( s = 0 \)
(E) \( r = 0 \)

The item shown is direct and clear in its graphical depiction, the abscissa is labeled ability \( (\theta) \), and the ordinate is labeled probability, or \( P(\theta) \) for probability of theta. The x-axis is labeled ability despite our realizing this term’s limiting implication. In some nonachievement contexts, such as with personality appraisal, the term trait level is more accurate and may be more generic regardless.

The three-parameter IRT model measures difficulty, discrimination, and guessing. The lower horizontal asymptote of the curve (as \( \theta \) decreases without bound, the H.A. is the value that \( P(\theta) \) approaches) tells you the probability of guessing on that item. The inflection point is where the odds of an examinee responding correctly to an item change from less than .5 to greater than .5. The probability of a correct response is also a function of ability level \( (\theta) \) on each item. The slope of the tangent line is a good indicator of discrimination, which is the degree to which the item discriminates between persons in different regions on the latent continuum. In approximating the domain for good discriminating power, take the domain of \( \theta \) where \( f'(\theta) \) is significant for the curve (Osterlind, 2010).

ITEM EXAMPLES

ITEM ANALYSIS
Item 1—good discriminatory power for an examinee in ability level of \( \theta \in [0,1] \). The probability of guessing, \( P(\theta) = 0 \). The inflection point (point at which \( P(\theta) = 0.5 \)) is at approximately \( \theta = 1 \).

Item 2—good discriminatory power for an examinee in ability level of \( \theta \in [-1,1] \). The probability of guessing, \( P(\theta) = 0.25 \). The inflection point (point at which \( P(\theta) = 0.5 \)) is at approximately \( \theta = 0.8 \).

Item 3—good discriminatory power for an examinee in ability level of \( \theta \in [0,1] \). The probability of guessing, \( P(\theta) = 0 \). The inflection point (point at which \( P(\theta) = 0.5 \)) is at approximately \( \theta = 1 \).

Item 4—good discriminatory power for an examinee in ability level of \( \theta \in [-2, -1] \). The probability of guessing, \( P(\theta) = 0 \). The inflection point (point at which \( P(\theta) = 0.5 \)) is at approximately \( \theta = -1.5 \).

Item 5—good discriminatory power for an examinee in ability level of \( \theta \in [-1.5, 0.5] \). The probability of guessing, \( P(\theta) = 0.1 \). The inflection point (point at which \( P(\theta) = 0.5 \)) is at approximately \( \theta = -0.5 \).

Item 6—moderate to low discriminatory power for an examinee in ability level of \( \theta \in [-3, 3] \). The probability of guessing, \( P(\theta) = 0.19 \). The inflection point (point at which \( P(\theta) = 0.5 \)) is at approximately \( \theta = -0.25 \).

CONCLUSION
Item response theory and the item response curve give us additional information on the quality of items in assessments. One of the main facets is the ability to predict the probability of getting an item correct based on the examinee’s ability level as well as the degree of discrimination. Having quality items is essential in assessment and education as it determines where or not our goals of education are being met. Assessment affects decisions about grades, placement, advancement, instructional needs, curriculum, and, in some cases, funding. Armed with information about the difficulty, discrimination, and guessing of items lends us to create better assessments that answer whether students are effectively learning what they are supposed to be learning and where we can improve as instructors.

BIBLIOGRAPHY