

ITEM ANALYSIS

Item analysis refers to a statistical technique that helps instructors identify the effectiveness of their test items and plays an important role in contributing to the fairness of the test.

Generally, the process of item analysis works best when class sizes exceed 50 students. In such cases, item analysis can help in identifying potential mistakes in scoring, ambiguous items and alternatives (distractors) that don't work. When performing item analysis, we are analyzing the following important statistical information.

Proportion Answering Correctly (Item Difficulty)

This indicates the proportion of students who got the item right. A high percentage indicates an easy item/question and a low percentage indicates a difficult one. In general, items should have values of difficulty no less than 20 per cent correct and no greater than 80 per cent.

Discrimination Index

This is the difference between the proportion of the top scorers who got an item correct and the proportion of the bottom scorers who got the item right (each of these groups consists of 27 per cent of the total group of students who took the test and is based on the students' total score for the test). The discrimination index range is between -1 and +1. The closer the index is to +1, the more effectively the item distinguishes between the two groups of students. Sometimes an item will discriminate negatively. These items should be revised and eliminated from scoring as they indicate that the lower performing students actually selected the correct response more frequently than the top performers.

Point Biserial

This is the correlation between an individual student's performance on an item and their total score on the test. The values range from -1 to +1. The high positive values are desirable for the correct answer because they indicate that a student who did well on the exam also did well on this question. Negative values are desirable for the alternatives or distractors that were not the correct answer. A score of 0 or less for the correct alternative indicates the question has difficulty distinguishing between those students who know the material and those who do not. The question should be examined and revised and potentially eliminated from scoring.

Distractor Evaluation

All of the incorrect options, or distractors, should actually be distracting. Preferably, each distracter should be selected by a greater proportion of the lower scorers than of the top group. A distractor should attract at least one candidate. If no one selects a distractor, it is important to revise the option and attempt to make the distractor a more plausible choice.

Reliability

The summary statistics found at the beginning of your item analysis include an estimate of the test's reliability. KR 21 is a measure of how well the items in your exam work together to obtain a measure of student achievement. Generally, values of .60 or greater are acceptable for the purposes of classroom tests; however, an accurate interpretation of this value requires the consideration of a number of factors impacting student performance.

Factors that lower the reliability of a test include items which are poorly written, too many items which are very easy or very hard, and too few items overall on a test to obtain a stable estimate of the student's ability.

Adapted from Michigan State University & Barbara Gross Davis Tools for Teaching