



ITEM ANALYSIS

Multiple-Choice Questions: Difficulty and Discrimination Indices

Good multiple-choice questions (items) follow the Goldilocks rule: they should not be too difficult or too easy, but just right. Furthermore, they should appropriately discriminate between students who really know the answer and students who do not. In other words, they reflect a student's true score. Unfortunately, the true score is theoretical; we cannot really know what it is supposed to be. However, we can create good test items that mitigate error so that a student's observed score is as close to their true score as possible. So how do we know if our test questions are good?

An item analysis allows one to determine whether a multiple-choice question discriminates between students who know the material from those who do not and consists of calculating two indices for each question: a difficulty index and a discrimination index (Salkind, 2017).

Difficulty Index (D) – describes the total number of students who got an item correct.

$$D = (S_H + S_L)/T$$

Where

D = difficulty index

S_H = number of students in the high group (see below) who answered the question correctly

S_L = number of students in the low group (see below) who answered the question correctly

T = the total number of responses for the item

Interpreting the difficulty index requires students to be divided into high and low groups.

High and low groups – defined by the top and bottom 50% of total exam scores. To calculate the high and low groups:

1. For each student, tally the total number of multiple-choice questions that they got correct.
2. Rank and sort the score totals (still linked to each student) from highest to lowest.
3. Identify the top 50%* of scores as the high group, and the bottom 50% of scores as the low group. If you have an odd number of students who took the exam, use your judgement of whether the median score best fits into the high or low group. (*Much of the literature will say to use the top 27%, but using the top 50% is OK for smaller class sizes.)

Ideally, if 50% of the students are the high group then 50% of the students are the low group.

