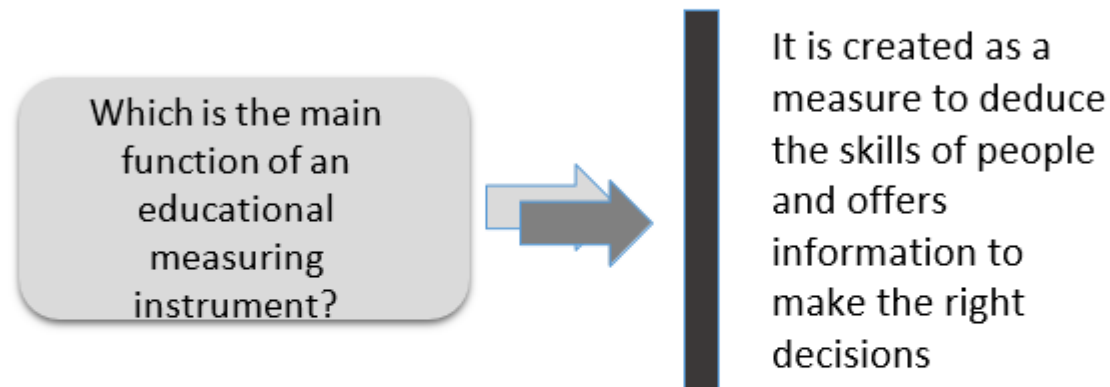


MarkQual<sup>®</sup>

Assessment and qualification system

# **Indexes of Discrimination and Difficulty**

# We will now explain the concepts related to the level of difficulty and the power of discrimination of evaluation instruments.



It is very useful to know the technical indicators which define the quality of the educational instrument being used.

There are three indicators used to describe the psychometric features of the items of an objective test: their level of **difficulty**, their power of **discrimination** and the functioning of their **distractors**.

First, to achieve our purpose, we will describe what in specialized literature is technically understood as the **difficulty** and **discrimination** of an item, and how to calculate them.

# DIFFICULTY

An item difficulty is understood as the proportion of people responding correctly to a test item.



**The higher the proportion is, the lower the difficulty will be.**



This means it is an inverse relationship: the higher the item difficulty is, the lower its index will be. (Wood, 1960).

## Symbol

Usually, this proportion is identified with a  $p$ , and indicates an item difficulty (Crocker & Algina, 1986).

## Calculation

To calculate an item difficulty, you should simply divide the number of people who correctly responded the item with the total number of people responding the item (either correctly or incorrectly).

It is calculated with the following formula:

$$p_i = \frac{A_i}{N_i}$$


$p_i$  = Difficulty index of the item

$A_i$  = Number of correct answers to item  $i$

$N_i$  = Number of correct answers plus the number of incorrect answers to item  $i$

The medium level of difficulty on a test must vary between **0.5 and 0.6**

The desired values of  $p$  should be as follows:



<b>Easy</b>	<b>5%</b> of easy items <b>20%</b> of relatively easy items
<b>Moderate</b>	<b>50%</b> of items with moderate difficulty <b>20%</b> of relatively difficult items
<b>Difficult</b>	<b>5%</b> of difficult items

# Discrimination

If the test and an item measure the same ability or skill, we can expect that the individual who obtained a high score during the entire test should have high probabilities of responding correctly to the item.

We can also expect the opposite, that is to say, the person who obtained low scores on the test should have low chances of responding correctly to the item.

Hence, a good item should **discriminate** those individuals obtaining good scores on the test from those who obtained low scores.

Even though there are many equivalent ways to calculate the discrimination index, we will use the following formula:

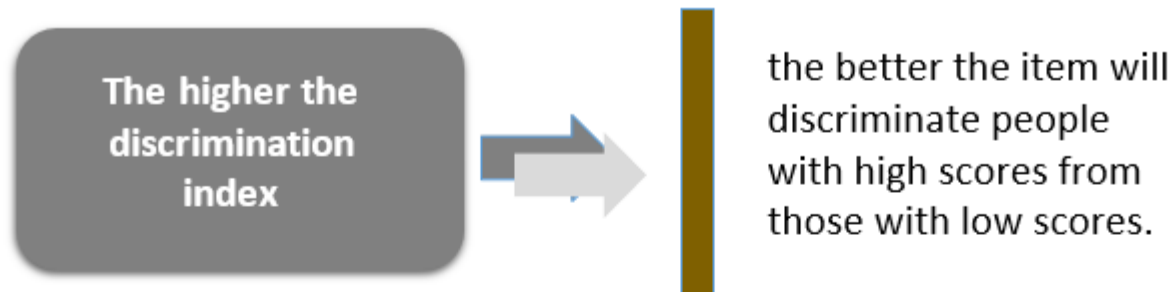
$$D_i = \frac{GA - GB}{N_{\text{largest group}}}$$

$D_i$  = Discrimination index of item  $i$ .

$GA_{\text{correct answers}}$  = Number of correct answers to item  $i$  among the 27% of people with the highest scores on the test.

$GB_{\text{correct answers}}$  = Number of correct answers to item  $i$  among the 27% de people with the lowest scores on the test.

$N_{\text{largest group}}$  = Number of people in the largest group (GA or GB).



**D = 1**  
(maximum value of this indicator)

All the individuals of the **GA** answer correctly to an item and all the individuals of the **GB** answer incorrectly.

**D = -1**  
(maximum negative value)

All the individuals of the **GA** answer incorrectly to an item and all the individuals of the **GB** answer correctly.

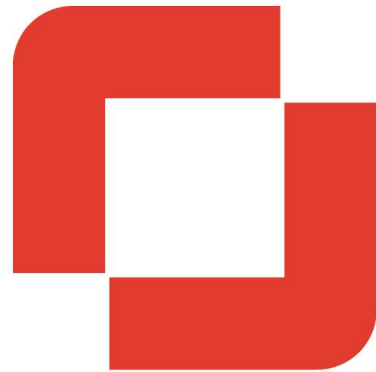
**D = 0**  
(minimum discrimination value)

Both groups answer equally.

Ebel and Frisbie (1986) provided us with a rule to determine the quality of items, in terms of the discrimination index. Let's analyze the table showing us values **D** and their corresponding meanings. Further, on the table, the recommendations for each of these values are reflected.

The table allows us to see the power of differentiation of the reagents according to their D value.

<b>D =</b>	<b>Quality</b>	<b>Recommendations</b>
<b>&gt; 0,39</b>	<b>Excellent</b>	<b>Retain it</b>
<b>0,30 - 0,39</b>	<b>Good</b>	<b>Possibilities to improve it</b>
<b>0,20 - 0,29</b>	<b>Mediocre</b>	<b>Need to review it</b>
<b>0,00 - 0,20</b>	<b>Poor</b>	<b>Dismiss or thoroughly review it</b>
<b>&lt; -0,01</b>	<b>Worst</b>	<b>Definitely dismiss it</b>



MarkQual<sup>®</sup>

Assessment and qualification system