

Item Analysis

Techniques to Improve Test Items and Instruction

OFFICE OF DISTANCE LEARNING
ASSESSMENT & TESTING



DISTANCE@FSU
OFFICE OF DISTANCE LEARNING

Item Analysis in a Nutshell

Check the effectiveness of test items:

1. Score the exam and sort the results by score.
2. Select an equal number of students from each end, e.g. top 25% (upper 1/4) and bottom 25% (lower 1/4).
3. Compare the performance of these two groups on each of the test items.

Item Analysis in a Nutshell

For any well-written item

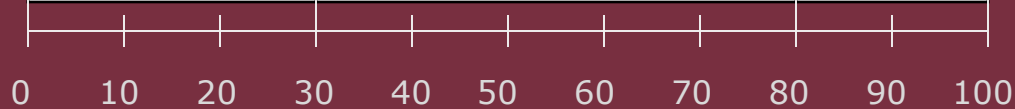
- a greater portion of students in the upper group should have selected the correct answer.
- a greater portion of students in the lower group should have selected each of the distracter (incorrect) answers.



Item Difficulty Level: Definition

The percentage of students who answered the item correctly.

High (Difficult)	Medium (Moderate)	Low (Easy)
$\leq 30\%$	$> 30\% \text{ AND } < 80\%$	$\geq 80\%$



Item Difficulty Level: Examples

Number of students who answered each item = 50

Item No.	No. Correct Answers	% Correct	Difficulty Level
1	15	30	High
2	25	50	Medium
3	35	70	Medium
4	45	90	Low



Item Difficulty Level: Discussion

- Is a test that nobody failed too easy?
- Is a test on which nobody got 100% too difficult?
- Should items that are “too easy” or “too difficult” be thrown out?



What is Item Discrimination?

- Generally, students who did well on the exam should select the correct answer to any given item on the exam.
- The **Discrimination Index** distinguishes for each item between the performance of students who did well on the exam and students who did poorly.

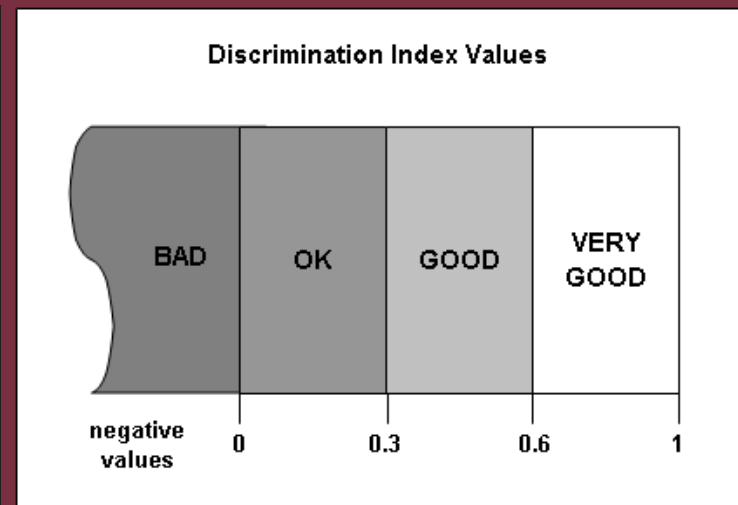
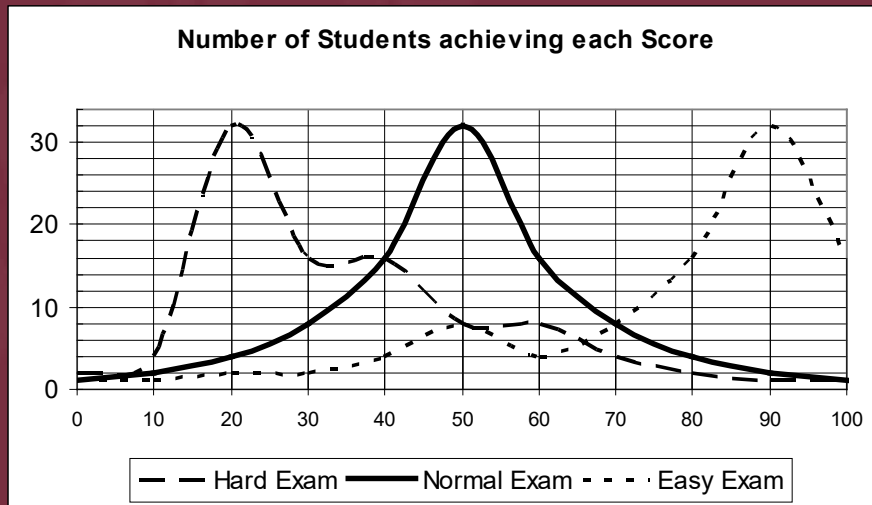


How does it work?

- For each item, subtract the number of students in the lower group who answered correctly from the number of students in the upper group who answered correctly.
- Divide the result by the number of students in one group.
- The Discrimination Index is listed in decimal format and ranges between -1 and 1.



What is a “good” value?



- For exams with a normal distribution, a discrimination of 0.3 and above is good; 0.6 and above is very good.
- Values close to 0 mean that most students performed the same on an item.
- The index should never be negative.

Item Discrimination: Examples

Item No.	Number of Correct Answers in Group		Item Discrimination Index
	Upper 1/4	Lower 1/4	
1	90	20	0.7
2	80	70	0.1
3	100	0	1
4	100	100	0
5	50	50	0
6	20	60	-0.4

Number of students per group = 100

Item Discrimination: Discussion

- What factors could contribute to low item discrimination between the two groups of students?
- What is a likely cause for a negative discrimination index?

Quick Reference

Use the following table as a guideline to determine whether an item (or its corresponding instruction) should be considered for revision.

Item Discrimination (D)	Item Difficulty		
	High	Medium	Low
$D \leq 0\%$	<i>review</i>	<i>review</i>	<i>review</i>
$0\% < D < 30\%$	<i>ok</i>	<i>review</i>	<i>ok</i>
$D \geq 30\%$	<i>ok</i>	<i>ok</i>	<i>ok</i>

Distracter Analysis: Definition

- Compare the performance of the highest- and lowest-scoring 25% of the students on the distracter options (i.e. the incorrect answers presented on the exam.)
- Fewer of the top performers should choose each of the distracters as their answer compared to the bottom performers.



Distracter Analysis: Examples

Item 1	A*	B	C	D	E	Omit
% of students in upper ¼	20	5	0	0	0	0
% of students in the middle	15	10	10	10	5	0
% of students in lower ¼	5	5	5	10	0	0

Item 2	A	B	C	D*	E	Omit
% of students in upper ¼	0	5	5	15	0	0
% of students in the middle	0	10	15	5	20	0
% of students in lower ¼	0	5	10	0	10	0

(*) marks the correct answer.

Distracter Analysis: Discussion

- What is the purpose of a good distracter?
- Which distracters should you consider throwing out?



Item Analysis Report

Item Analysis for Chi Cui

277 1 Order ID and group number

ITEM	percentages						counts						ITEM
	A	B	C	(D)	E	Omit	A	B	C	(D)	E	Omit	
Q 1 Upper 1/4	1	0	0	24	0	0	1	0	0	18	0	0	Upper 1/4 Q 1
Q 1 Mid 1/2	16	0	2	30	0	0	12	0	2	23	0	0	Mid 1/2 Q 1
Q 1 Lower 1/4	6	0	0	18	0	0	5	0	0	14	0	0	Lower 1/4 Q 1
Q 1 Total	24	0	2	73	0	0	18	0	2	55	0	0	Total Q 1
Point Biserial:	0.182		Discrimination Index: 0.213										
ITEM	A	B	(C)	D	E	Omit	A	B	(C)	D	E	Omit	ITEM
Q 2 Upper 1/4	0	0	25	0	0	0	0	0	19	0	0	0	Upper 1/4 Q 2

- The left half shows percentages, the right half counts.
- The **correct option** is indicated in parentheses.
- **Point Biserial** is similar to the discrimination index, but is not based on fixed upper and lower groups. For each item, it compares the mean score of students who chose the correct answer to the mean score of students who chose the wrong answer.

Exercise: Interpret Item Analysis

- Review the sample report.
- Identify any exam items that may require revision.
- For each identified item, list your observation and a hypothesis of the nature of the problem.



Reference

Oosterhof, A. (1990). *Classroom Applications of Educational Measurements*. Merrill, Columbus, OH.

