Module 4c:Analyzing Item Difficulty and Discrimination

Passport to Great Teaching – Creative Assessment

Timothy S. Brophy

1

Item Difficulty and Discrimination

- Teachers who create tests for classroom use often seek to know how effective their tests are
- Item analysis provides important information about how well items function
- Item difficulty helps us to know the degree to which students get the answer correct
- Item discimination examines how the top scoring group of test takers compare to the low scoring group

2

Calculating Item Difficulty

Count the total number of students answering each item correctly.

For each item, divide the number answering correctly by the total number of students.

This gives you the proportion of students who answered each item correctly. This figure is called the item's *difficulty level.*

Caution: The *higher* the difficulty level the *easier* the item, and vice versa.

3

Item Discrimination Index (DI)

When we subtract the proportion of low scoring students who got the item right from the proportion of high scoring students who got it right, the remainder is the *discrimination index*

This is a measure of how well the item discriminates between the top scorers and the bottom scorers on the item

4

Calculating the DI

Step 1. Compute each student's score. (i.e., count the number of items each student got correct.)

Step 2. Divide the class into groups on the basis of total score (e.g., 50% high and 50% low, or 33% high, 33% middle, and 33% low, or 25% high, 50% middle, 25% low, etc.).

Step 3. The high and low groups must be the same size. The middle group is not needed in computing discrimination index. (e.g., 25% high and 25% low-50% middle not used.)

Step 4. For each item, count the number of students in the high scoring group that got the item right.

5

Calculating the DI

Step 5. Divide this number by the total number in the high scoring group, for each item.

Step 6. Repeat steps 4 and 5 for each item, this time using only the low scoring group.

Step 7. For each item subtract the proportion of low scoring students who got the item right from the proportion of high scoring students who got it right. The remainder is the *discrimination index.*

6

Using Difficulty and Discrimination Indices for Item Analysis

- Difficulty should range from about .25 to .75 for most items
- Why? Items that are below .25 are the most difficult, and items above .75 cross the probability threshold for guessing in items with 4 answer choices
- The lowest recommended DI is .20
- Why? If there is a <20% difference between the students in the top-scoring and low-scoring groups responding to the item correctly, the item is not discriminating between the high and low scorers as effectively

7

Let's work an example

8

Biology Test (*N* = *100*)

Score group →	Upper (n = 25)		Middle (n = 50)		Lower (n = 25)		Difficulty level	Discrimination index
Item number ↓	*n*	*p*	*n*	*p*	*n*	*p*		
1	20	.80	15		15	.60		
2	15	.60	30		5	.20		
3	24	.96	38		12	.48		

- A biology instructor created and administered a 13 item test to 100 students.
- The teacher wants to know how well the items functioned by examining their difficulty levels and discrimination indices.
- This is a table of the numbers of students who completed items 1, 2, and 3
- correct in the top and bottom 25% of the group, and the middle 50%.
- Let's calculate the difficulty and discrimination levels for each together.

9

Item 1. Calculating Difficulty and Discrimination

Score group	Upper (n = 25)		Middle (n = 50)		Lower (n = 25)	
Item number	*n*	*p*	*n*	*p*	*n*	*p*
1	20	.80	20		15	.60

- Top scoring group (n = 25): 20 got it correct, or 80%
- Middle group (n = 50): 20 got it correct
- Lowest scoring group (n = 25): 15 got it correct, or 60%
- Difficulty level: Add the total who got the item correct – 20+20+15 = 55, and divide by the number of students, 100
- 55/100 = **.55**
- Discrimination Index: Subtract 80% - 60% = 20%, for a DI of **.20**

10

Item 2. Calculating Difficulty and Discrimination

Score group	Upper (n = 25)		Middle (n = 50)		Lower (n = 25)	
Item number	*n*	*p*	*n*	*p*	*n*	*p*
2	15	.60	30		5	.20

- Top scoring group (n = 25): 15 got it correct, or 60%
- Middle group (n = 50): 30 got it correct
- Lowest scoring group (n = 25): 5 got it correct, or 20%
- Difficulty level: Add the total who got the item correct – 15+30+5 = 50, and divide by the number of students, 100
- 50/100 = **.50**
- Discrimination Index: Subtract 60% - 20% = 40%, for a DI of **.40**

11

Item 3. Calculating Difficulty and Discrimination

Score group	Upper (n = 25)		Middle (n = 50)		Lower (n = 25)	
Item number	*n*	*p*	*n*	*p*	*n*	*p*
3	24	.96	38		12	.48

- Top scoring group (n = 25): 24 got it correct, or 96%
- Middle group (n = 50): 38 got it correct
- Lowest scoring group (n = 25): 12 got it correct, or 48%
- Difficulty level: Add the total who got the item correct – 24+38+12 = 74, and divide by the number of students, 100
- 74/100 = **.74**
- Discrimination Index: Subtract 96% - 48% = 48%, for a DI of **.48**

12

Keeping or Revising items

13

Keeping and revising items

When an item meets the general acceptability guidelines for difficulty and discrimination, you should keep the item

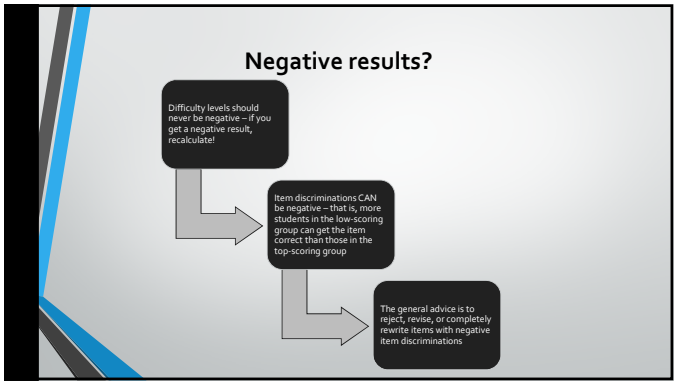
If it falls outside the guidelines, you make a judgment whether to keep, revise, or reject it

14

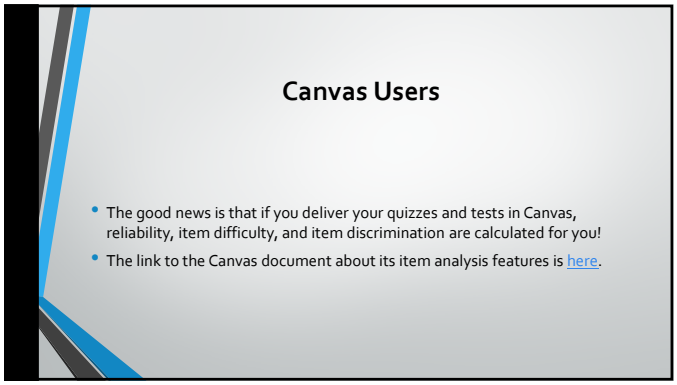
Guidelines for Item Difficulty and Discrimination Index statistics

- Difficulty level: the standard range is .25-.75
- Discrimination index: the standard is .20 or higher
- Item 1 – Difficulty, .55; DI, .20
- Item 2 - Difficulty, .50; DI, .40
- Item 3 - Difficulty, .74; DI, .48
- Keep or reject?

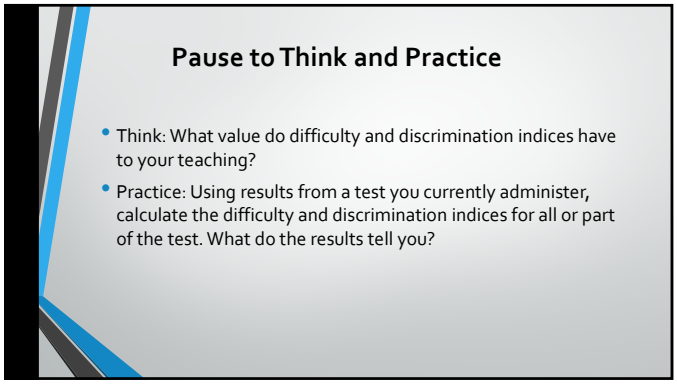
15


Negative results?
Difficulty levels should never be negative – if you get a negative result, recalculate!
Item discriminations CAN be negative – that is, more students in the low-scoring group can get the item correct than those in the top-scoring group
The general advice is to reject, revise, or completely rewrite items with negative item discriminations

16


Canvas Users
The good news is that if you deliver your quizzes and tests in Canvas, reliability, item difficulty, and item discrimination are calculated for you!
The link to the Canvas document about its item analysis features is here.

17


Pause to Think and Practice
Think: What value do difficulty and discrimination indices have to your teaching?
Practice: Using results from a test you currently administer, calculate the difficulty and discrimination indices for all or part of the test. What do the results tell you?

18