

Module 4b: Reliability and Validity

Passport to Great Teaching – Creative Assessment  
Timothy S. Brophy

1

---

---

---

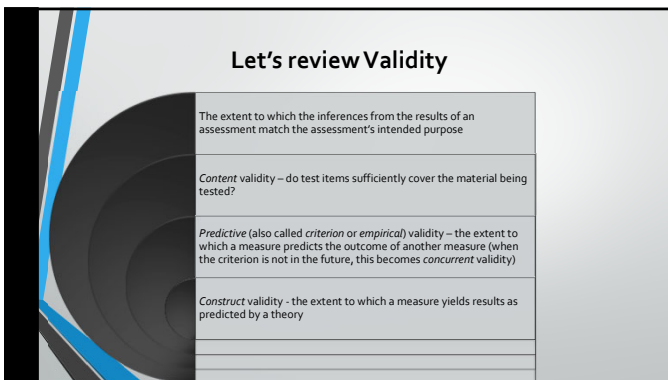
---

---

---

---

---



Let's review Validity

The extent to which the inferences from the results of an assessment match the assessment's intended purpose
Content validity – do test items sufficiently cover the material being tested?
Predictive (also called <i>criterion</i> or <i>empirical</i> ) validity – the extent to which a measure predicts the outcome of another measure (when the criterion is not in the future, this becomes <i>concurrent</i> validity)
Construct validity - the extent to which a measure yields results as predicted by a theory

2

---

---

---

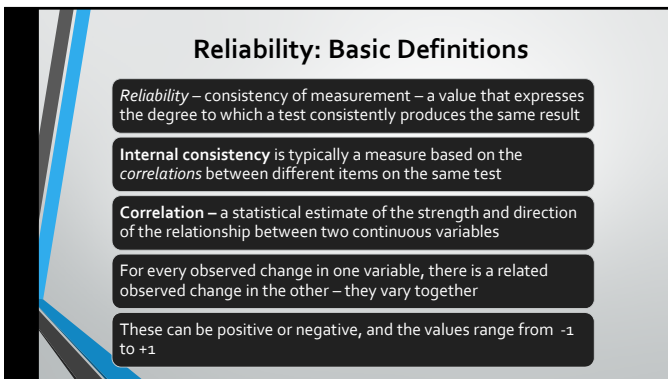
---

---

---

---

---



Reliability: Basic Definitions

- Reliability – consistency of measurement – a value that expresses the degree to which a test consistently produces the same result
- Internal consistency is typically a measure based on the correlations between different items on the same test
- Correlation – a statistical estimate of the strength and direction of the relationship between two continuous variables
- For every observed change in one variable, there is a related observed change in the other – they vary together
- These can be positive or negative, and the values range from -1 to +1

3

---

---

---

---

---

---

---

---

### Test Reliability

- This number is estimated in several ways
- Parallel forms – the Pearson *product-moment correlation coefficient*
- Test-retest reliability – correlate scores on two administrations of the same test
- Split-halves reliability - Flanagan's Formula
- Kuder-Richardson Formulae 20 and 21 ( $KR_{20}$  and  $KR_{21}$ ) – for tests with dichotomous items

4

---

---

---

---

---

---

---

---

### Calculating reliability in Teacher-made Tests

- One easy way is to use Flanagan's Formula to calculate a *split-half reliability*
- To do this, divide your quiz or test into 2 equal parts
- Calculate the variance for each part
- Apply this formula:  

$$R = 2(1 - \frac{s_a^2 + s_b^2}{s^2})$$
- You add the variances of part a and part b, divide the sum by the variance of the total test, subtract it from 1, and multiply that result by 2.
- The result is the split half reliability of your quiz.
- Good tests have reliability coefficients of .70 and higher.

5

---

---

---

---

---

---

---

---

### Let's work an example: Calculating Split-half Reliability

Let's calculate the split-half reliability for a history quiz.

Open the *Formulae for Reliability Studies* page in your resources folder. There are a number of reliability formulae on this page, which you are welcome to explore further on your own.

We will use *Flanagan's Formula* to calculate a reliability coefficient together. This is the formula for split-half reliability.

In this calculation, the teacher divides the test into two even parts, and uses the subscores on each part as well as the total score to calculate the reliability of the test.

This is an easy calculation for teachers and the most practical approach to calculating reliability for teacher made tests.

6

---

---

---

---

---

---

---

---

### Flanagan's Formula

- First, split the quiz questions evenly into two parts – part *a* and part *b*, and calculate the subscores for each student on each part. Then, calculate the variances of subscores on the two parts and then the variance of the total scores. Enter these numbers into this formula and calculate:

$$r = 2 \left( 1 - \frac{S_a^2 + S_b^2}{S^2} \right)$$

- Key:
  - $S_a^2$  is the variance of part a
  - $S_b^2$  is the variance of part b
  - $S^2$  is the variance of total scores

---

---

---

---

---

---

---

---

7

### Our example – History quiz results (max score = 10)

Name	Quiz part a	Quiz part b	Total Score
Joe	5	5	10
Mary	3	2	5
Cheryl	5	4	9
Charlene	4	4	8
Chris	2	3	5
Brian	3	4	7
LaTerrance	5	4	9
Kim	4	3	7
Robbie	3	5	8
Anwar	5	3	8

---

---

---

---

---

---

---

---

8

### Calculating the Variances

Using the [standard deviation online calculator](#), enter the scores for part *a*, part *b*, and the *total*, one at a time. You will get three variances. *Try pausing the video and doing this on your own before continuing.*

Here are the results:  
Part *a* = 1.21  
Part *b* = .9  
Total = 2.71

Next we place these figures into the formula.

---

---

---

---

---

---

---

---

9

### Calculating the Split-Half reliability

- Flanagan's Formula:  $r = 2 \left( 1 - \frac{s_a^2 + s_b^2}{s^2} \right)$
- Our data:
- $r = 2 \left( 1 - \frac{1.21 + .9}{2.71} \right) = 2 \left( 1 - \frac{2.11}{2.71} \right) =$   
 $2(1 - .78) = 2(.22) = .44$

10

---

---

---

---

---

---

---

---

### Interpreting our coefficient

The Split-Half reliability is .44

Reliability coefficients should be at .70 or higher for a test to be considered reliable

This quiz is not reliable – so this teacher should examine the performance of each item and revise

Next, we will examine how you can determine item difficulty and discrimination.

11

---

---

---

---

---

---

---

---

### Pause to Think

- A teacher has given a test in a Calculus 1 class. The teacher used the split-half reliability formula to calculate a reliability coefficient of .85. What does this tell you?

12

---

---

---

---

---

---

---

---