

Activities: Reliability for Teachers

Reliability for Teachers Activity: What is test reliability?

This activity will help you answer the essential question:

- What is test reliability?
-

Activity 1: What is test reliability?

You may complete this activity individually or in groups

Part 1

Read and consider the following definitions of test reliability and then respond to the following:

Why is it critical that the large-scale assessments used to measure student learning be reliable?

1. "Test reliability refers to the consistency of scores students would receive on alternate forms of the same test."
From:
An Instructor's Guide to Understanding Test Reliability, Craig S. Wells and James A. Wollack, Testing & Evaluation Services University of Wisconsin, November 2003. Available at: <http://testing.wisc.edu/Reliability.pdf> . Retrieved March 7, 2012.
 2. "In research, the term reliability means "repeatability" or "consistency". A measure is considered reliable if it would give us the same result over and over again (assuming that what we are measuring isn't changing!)."
From:
The Research Methods Knowledge Base, 2nd Edition, William M. Trochim. Available at: <http://www.socialresearchmethods.net/kb>. Retrieved March 7, 2012.
 3. "Reliability is the degree to which students' results remain consistent over time or over replications of an assessment procedure. An important point to remember is that reliability is a necessary, but insufficient, condition for valid score-based inferences. That is, you cannot make valid inferences from a student's test score unless the test is reliable."
From:
Educational Assessment of Students, 6th Edition, by A. J. Atko and S. M. Brookhart. Prentice Hall, 2010.
 4. The three varieties of reliability evidence are stability, alternate form, and internal consistency:
 - Stability is the consistency of results between two time-separated testing occasions.
 - Alternate form is the consistency of results between two different forms of a test.
 - Internal consistency is the consistency in the way a test's items function.
-

Part 2

Read the following and then review the state and/or national assessments manuals from large-scale assessments that are administered at your school. What are the correlation coefficients? Are you confident that they are reliable enough?

There are a variety of ways to compute and interpret correlation-based reliability but what is important to know is that correlational procedures take two sets of scores from the same group of test-takers. Then those scores are analyzed to see how closely they are related. The result of this analysis is called the **correlation coefficient**. The correlation coefficient is also sometimes referred to as the Pearson correlation coefficient after its originator Karl Pearson. The correlation coefficient is usually represented by the small letter *r*. This can represent a positive relationship, no relationship, or a negative relationship.

You will encounter correlation coefficients whenever you are provided information about the quality of a commercial or state assessment. It isn't necessary for you to know how to calculate a correlation coefficient – but you do need to know what a correlation coefficient means. The closer to 1.0 the correlation coefficient is, the closer the relationship between the results of two sets of scores or the greater the **score consistency**. A correlation coefficient of zero signifies no relationship at all. If a correlation coefficient is below zero, the two scores show a negative relationship.

Test users have to decide if a test is reliable enough based up the score-consistency evidence available. Score consistency reliability is an estimate of a test's consistency derived from the test taker's scores. Teachers can decide how much they trust a national, state, or district test based upon the score consistency correlations that come with those tests.

Here you see a Table from *Mastering Assessment Booklet, Reliability: What Is It and Is It Necessary?* illustrates simple interpretations of correlation coefficients.

Correlation-Coefficients

Correlation Coefficient or <i>r</i>	Common Sense Interpretation
$r = 1.00$	A perfect positive relationship indicating the relative ranks of scores in two sets of data are identical.
$r = 0$	An indication of no relationship whatsoever between two sets of scores.
$r = -1.00$	A perfect negative relationship indicating the relative ranks of scores in two sets of data are completely reversed.

Adapted From:

Mastering Assessment: A Self-Service System for Educators; Reliability: What Is It and Is It Necessary? by W. James Popham, P. 9