
Validity and Validation

Validity is the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses. Validation is the accumulation of evidence. These definitions come from the 1999 and 2014 editions of the *Standards for Educational and Psychological Testing*.

Although it was perhaps first suggested by Loevinger in the 1950s, it wasn't until Messick's work in the 1980s when the field began to treat validity as a unitary concept, such that all forms of validity evidence support score interpretation. Messick clarified this unitary conceptualization of validity in that there aren't different types of validity, but different forms of validity evidence. This also is consistent with the approach taken in the landmark article on construct validity by Cronbach and Meehl (1955) declaring the need for a clearly articulated proposed interpretation, followed by deep analysis of evidence and the consideration of alternative explanations – although their intent was not to declare validity as a unitary concept. During this time (late 1980s), Cronbach further presented validation as the defense of an argument, suggesting that a validity argument can be developed to support intended interpretations and uses of test scores, similar to Messick's idea that we need to evaluate the adequacy and appropriateness of test score inferences, also found in the Cronbach and Meehl article. All of this was clarified by the end of the 1980s.

Kane (1992) synthesized this earlier work into a more direct model of validity as argument, followed by several articles with examples and clarifications. He was the one to introduce the idea of the Interpretive Argument (2006), now called the Interpretation and Use Argument (2013a, 2013b), which is relatively new and still gaining presence in the validity literature. However, the conceptualization of validity as argument has had presence for over three decades.

With the 1999 edition of the *Testing Standards* and reaffirmation in the 2014 edition, the field moved away from references to specific types of validity (construct validity, content validity, criterion validity), to adopt the consistent approach describing validity evidence as it supports specific interpretations and uses of test scores. This promotes the gathering of relevant evidence to support specific claims, rather than collecting validity evidence for the sake of validity evidence.

Key articles in this history:

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington DC: American Educational Research Association.
- Cronbach, L.J. (1982). Five perspectives on validity argument. In H. Wainer & H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum.
- Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Kane, M.T. (1992). An argument-based approach to validation. *Psychological Bulletin*, 112, 527-535.
- Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: American Council on Education and Praeger.
- Kane, M.T. (2013a). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Kane, M.T. (2013b). The argument-based approach to validation. *School Psychology Review*, 42(4), 448-457.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694.
- Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Researcher*, 10, 9-20.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed.). New York, NY: American Council on Education and Macmillan.