

What Role Does the Consequences of Testing Play in Validity?

Currently, the field of educational measurement appears to have reached broad consensus that validity is a judgment of the degree to which arguments support the interpretations and uses of test scores (Kane, 2006). However, the field is less clear on the role that the consequences of test score use play in validity.

Though validity is maintained as a unitary construct, writers argue that validity is called into question when the consequences of test score use can be linked to a flaw in the conceptualization of test interpretation. This flaw may be due to construct under-representation or inclusion of sources of construct irrelevant variance.

An example of the negative consequences of test score use linked to construct under-representation was offered by Shepard (1997). In preparing for the MCAT, pre-med students typically concentrated on taking science courses. Deans and admission officers were concerned that a consequence of taking the MCAT was that students admitted to medical school were too narrowly educated. Therefore, the interpretation of MCAT scores as a level of preparation for medical practice and the use of the

scores to admit students to medical school was seen as a flaw in the conceptualization.

An example of the negative consequences of test score use linked to construct-irrelevant variance is the possible tendency for raters to score typed essays more stringently than hand-written essays. A number of possible scoring biases may contribute to rater stringency including the tendency for typed essays to appear shorter than identical hand-written responses and more obvious writing errors because of the greater ease of reading typed responses compared to hand-written responses.

“The wall separating the consequences of test score use from other evidence relevant to the validity of test score interpretation is artificial.”

But the inclusion of consequences in a discussion of validity evidence remains controversial. Green (1998) and Reckase (1998) argue that evidence of a cause-and-effect relationship is impossible to collect under the conditions of an operational testing program.

In contrast to arguments that collecting evidence of test score use consequences is too burdensome, Mehrens (1997) argues that the analyses of the effects of test score use should not be included as validity

evidence, and that the psychometric community should narrow the use of the term validity to evidence of the accuracy of inferences regarding test scores.

The wall separating the consequences of test score use from other evidence relevant to the validity of test score interpretation is artificial. The consequences of test score use are evidence relevant to test score interpretation when the consequences can be linked to construct under-representation or construct irrelevant variance. These consequences can be excluded by fiat from evidence relevant to the inference that a score is a reasonable indicator of the amount of a construct possessed. But such a wall is fragile and easily breached.

- Paul D. Nichols
- Natasha Williams

The Pearson Research Report, "*Evidence of Test Score Use in Validity: Roles and Responsibilities*" is provided online at:
www.PearsonEdMeasurement.com/Research/Research/.

REFERENCES

- Green, D. R. (1998). Consequential aspects of the validity of achievement tests: A publisher's point of view. *Educational Measurement: Issues and Practice*, 17(2), 16-19, 34).
- Mehrens, W. A. (1997). The consequences of consequential validity. *Educational Measurement: Issues and Practice*, 16(2), 16-18.
- Reckase, M. (1998). Consequential validity from the test developer's perspective. *Educational Measurement: Issues and Practice*, 17(2), 13-16.
- Shepard, L. A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16(2), 5-8, 13, 24.

The Educational Measurement (EM) group of Pearson is the most comprehensive provider of educational assessment products, services and solutions. A trusted partner in district, state and national assessments for more than 50 years, Pearson helps administrators, teachers, parents and students use assessment and research to promote learning and academic achievement.

More information is available at www.PearsonEdMeasurement.com.