

Test Bias as a factor of context:

As I read “Test Bias or Cultural Bias: Have We Really Learned Anything?” by Stephen J. Schellenberg (2004), I think about the difficulty of building a pre-test/post-test for my term project. Developing effective test questions that consider content in the course and are clear, concise, and applicable turns out to be extremely tedious. But in order for a test to have validity and suit the needs of the tester, and equity to suit the needs of the subjects is a daunting process.

The process of developing an appropriate test could be a little less challenging if one were building a test based on predetermined standards. That way you would test to the standards that are already in a form that is more easily converted to a question with specific answers. But to have to build a test for a course, and tease out appropriate and applicable tidbits of information to challenge a learner, and remain within reach of the comprehension and retention of an average person taking the course, was altogether another thing.

Schellenbert cites Eels et al (1951) to summarize the three general explanations that were used in the mid 20th century:

1. Subjects scoring well on tests are genuinely superior in inherited genetic equipment.
2. Large-scale group differences are probably the result of a faulty test.
3. High scores are generally the product of a superior environment and low scores the product of a poor one.

The three explanations are cringeworthy from a modern perspective, but at least the second two offer some useful guidance on considering the efficacy and environment of an evaluative instrument. On the one hand, is a test seeking understanding of a taker’s ability to learn (e.g., IQ test)? Or is it assessing the taker’s final knowledge and abilities (e.g., achievement tests). Each query has its place in the spectrum of learning and evaluation, but one is a very different purposes from the other. In any case, validity must be assessed based on the population being tested and the purpose of the test.

It testing inherently biased? I think it depends on the purpose of the test and how the test has been developed to evaluate a specific population for a specific purpose. So, tests can work, but standardized tests must by definition be inherently biases due the fact that the populations being tested and the purposes for testing are inherently diverse and not standard at all. This leads me to the conclusion that learning and equity depends on the teachers and instructional institutions. Gauging learning success is incumbent upon institutions of higher learning or end-users of the trained and/or educated subjects (i.e., agencies, businesses, institutions, etc., seeking to recruit and discern among applicants).

Ultimately, I come away from the experience of building a test, and the controversy of test bias with the conclusion that in order to assess a population that participates in a particular learning treatment, one must evaluate the subjects. The evaluation can include a test, but the test must be appropriate to the population and the treatment, as well as to the goals of the outcome. Validity across the board is a key determining factor.

References:

Eels, K., Davis, A. Havighurst, R. J. Herrick, V. E., and Tyler, R. W. (1951) Intellectual and cultural differences: a study of cultural learning and problem-solving. Chicago: University of Chicago Press.

Russ-Eft, D., Preskill, H. 2009. *Evaluation in Organizations: A Systematic Approach to Enhancing Learning, Performance, and Change* (2nd Ed.) New York: Basic books. ISBN-10: 0465018661, ISBN-13: 978-0465018666

Schellenberg, S., (2004, April 14). "The Achievement Gap: Test Bias or School Structures?." A paper presented as part of the symposium sponsored by the National Association of Test Directors as part of the Annual Meeting of the National Council for Measurement in Education San Diego, California.