

National Council on Measurement in Education (NCME)  
Position Statement on Theories of Action for Testing Programs

The information produced by testing programs is intended by sponsors and designers to be used for specific purposes. Among these purposes may be determining students' performance on standards-based educational assessments, placing students into appropriate levels of instruction or training, comparing countries with respect to their students' levels of knowledge or skill, or outside of school testing, identifying candidates having the minimum competency required to enter a profession.

Sometimes, however, testing program sponsors, designers, or users have goals that extend beyond providing scores and related interpretive text. In such cases, an explicit objective is to contribute directly to change in individuals or institutions. For example, many state K-12 testing program sponsors intend, through test design and score use, to cause change in the behavior of school leaders and teachers. The intended change may occur through the reallocation of resources or the shifting of instructional practices toward improving particular student competencies. In these instances, the testing program takes on the role of an intervention that may, in some ways, be as important as its measurement function. As such, the claims of testing program sponsors are different from the claims associated with more conventional testing programs.

If a testing program is to serve as a "change agent," it is helpful to delineate how the program needs to operate so that the desired change is most likely to occur. Known in the program evaluation literature as a "theory of action," such documentation includes an explicit identification of:

- the testing program's constituent parts (in the case of a K-12 testing program, e.g., domain definition, test design, reports, and formative, interim and summative assessments);
- its intended (short- and long-term) outcomes;
- the causal mechanisms posited as responsible for the intended change; and
- the ideal implementation of the program that is believed to be most likely to lead to those outcomes.

A good theory of action anticipates both what needs to be in place to have the desired effects, but also anticipates ways that an improper implementation of the program may lead to unintended negative consequences. In some cases evidence already exists to support a theory of action, and if so, this support should be made explicit. In other cases, evidence may be lacking, and this lack should be acknowledged. In either case, a theory of action must be regularly evaluated with empirical evidence that examines whether it is operating as intended. This evaluation may lead to a revision to the theory of action.

In sum, when program sponsors, designers, or users intend a testing program to effect change, NCME recommends that they document and disseminate arguments and evidence in support of the expected (causal) relationships among the program's constituent parts, the implementation actions, and the intended outcomes, and also plan for ongoing evaluation to detect and mitigate unintended, negative consequences. In addition, NCME recommends that sponsors, designers, and users of other testing

programs make clear how their programs will be designed, implemented, and monitored to facilitate achieving their intended goals and avoiding undesirable outcomes.<sup>1</sup>

## References

- Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning (CBAL): A preliminary theory of action for summative and formative assessment. *Measurement: Interdisciplinary Research and Perspectives*, 8(2-3), 70-91. NCSC. (2016).
- Haertel, E. H. (2013). How is testing supposed to improve schooling? *Measurement: Interdisciplinary Research and Perspectives*, 11, 1-18.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1-73.
- National Center and State Collaborative (NCSC). (2016). NCSC's Theory of Action and Validity Evaluation Approach. NCSC Brief (No. 9). Retrieved from <http://www.ncscpartners.org/Media/Default/PDFs/Resources/NCSCBrief9.pdf>
- Sireci, S. G. (2015). A theory of action for validation. In H. Jiao & R. Lissitz (Eds.). *The next generation of testing: Common core standards, Smarter-Balanced, PARCC, and the nationwide testing movement* (pp. 251-269). Charlotte: Information Age Publishing Inc.

## About NCME

The National Council on Measurement in Education (NCME) is a professional organization for individuals involved in assessment, evaluation, testing, and other aspects of educational measurement. Members are involved in the construction and use of standardized tests; new forms of assessment, including performance-based assessment; program design; and program evaluation. NCME members include university faculty; test developers; state and federal testing and research directors; professional evaluators; testing specialists in business, industry, education, community programs, and other professions; licensure, certification, and credentialing professionals; graduate students from educational, psychological, and other measurement programs; and others involved in testing issues and practices.

Approved by the NCME Board of Directors, July 26, 2018

---

<sup>1</sup> Arguments supporting this recommendation can be found in Haertel (2013) and Kane (2013). Examples of this type of documentation can be found in Bennett (2010), NCSC (2016), and Sireci (2015).