

Construct validation of measurements with typical response tasks

Sanneke J. Schouwstra

University of Amsterdam, Faculty of Social and Behavioural Sciences,
Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands
Email: schouwstra@pscw.uva.nl

Keywords: test development, construct validation, typical response tasks, construct representation, irrelevant variance, framework

Abstract

The current unified concept of validity (e.g. test standards of the APA, 1999; Anastasi, 1986; Messick, 1989) is difficult to handle when dealing with construct validation in practice. The Deductive Design (Schouwstra, 2000) provides in the need for a systematic theory-driven approach to the construct validation of measurements using typical response tasks, intended to measure theoretical constructs such as attitudes and personality. Construct validation is perceived as presenting and evaluating scientific arguments pertaining to the trustworthiness of a test score interpretation. These scientific arguments should focus at the two basic threats to construct validity (test standards of the APA, 1999; Messick, 1989): construct underrepresentation and irrelevant variance. The Deductive Design distinguishes itself from other validation procedures through the fact that these two basic issues are addressed from the outset of test development. Within the Deductive Design the first scientific argument to present consists of specifying how the interpretation is derived from the underlying theoretical notions: the *rationales*. This requires, first, formulating what the construct of interest is and what the plausible sources of irrelevant variance are, given the chosen measurement method. Then we should specify and model how we sought to insure during the test development that the test score is an adequate reflection of the construct of interest and how we sought to insure that the test score does not reflect those plausible sources of irrelevant variance. The second scientific argument is *empirical evidence* that the test-score reflects the whole of the construct and that no plausible irrelevant variance is present. Furthermore, a framework for construct validation should move beyond summing up the initial arguments supporting an interpretation or refuting an alternative interpretation. Therefore, the Deductive Design, in contrast to other validation procedures, incorporates an *evaluation* of the resulting scientific argumentation, that yields one *integrated judgement* of the trustworthiness and that indicates future lines of construct validation research. The Deductive Design has been used to evaluate the construct validation of self-efficacy scales and to develop a new self-efficacy measurement. The evaluation of construct validation studies demonstrates a general lack of attention for the basic issue of irrelevant variance and for the rationales underlying a test score interpretation. The application of the design for the validation of a new measurement instrument, on the other hand, demonstrates the usefulness of the design for obtaining an actual judgement of (un)trustworthiness. More importantly, the application demonstrates the added value of the framework. If a standard sequential validation procedure had been employed or the issue of plausible irrelevant variance had not been addressed from the outset, another, wrong, conclusion had been drawn about the trustworthiness of the test score interpretation. Our resistance to concluding that our measures are invalid and our inclination to rather search for theoretical explanations makes the a priori approach of the Deductive Design valuable and important.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in education (1999). Standards for educational and psychological testing. Washington, DC : American Educational Research Association.
- Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, 37, 1-15.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 13-103). New York: Macmillan Publishing Co.
- Schouwstra, S. J. (2000). *On testing plausible threats to construct validity*. Unpublished doctoral dissertation, University of Amsterdam, The Netherlands.