

Measurement invariance and construct validity

Denny Borsboom

University of Amsterdam

Abstract

It is difficult to escape the intuition that measurement invariance has something to do with construct validity, but is it also difficult to explicate exactly what this connection could be. A possible way of thinking about this relation is in terms of construct equivalence. Construct equivalence holds when the same construct is measured across subpopulations. Based on the concept of construct equivalence, one can discern two views on the relation between measurement invariance and construct validity. These positions differ in their conceptualization of psychological constructs. The first holds that measurement models implicitly define a psychological construct. This view implies that it is impossible to measure the same construct across subpopulations where a violation of measurement invariance occurs. Thus, measurement invariance is a necessary condition for construct equivalence. Second, one might see measurement models as explanatory, which means that they do not primarily define constructs, but refer to them. Constructs are then interpreted as traits that exist independent of any particular test used to measure them, and hence independent of any particular measurement model used. In this conceptualization, a general demand for measurement invariance is not so easily justified: for it is logically possible that we should measure the same construct in different subpopulations, but in a different way - and this might easily lead to violations of measurement invariance. Thus, in this view, measurement invariance is not a necessary condition for construct equivalence. Although the latter view seems to align more closely with commonly held positions in both the substantive and the psychometric literature, it is not unproblematic. For instance, it means that construct equivalence does not entail measurement invariance as a consequence. Therefore, the demand of measurement invariance cannot be posited as a universal methodological demand, but must be based on substantive, context-specific considerations. Implications of these findings are discussed.