

An item and construct bias analysis of two language versions of a verbal analogies scale

Rizwana Roomaney and Elize Koch

Abstract

The Woodcock Muñoz Language Survey is a test of cognitive academic language proficiency that has been adapted from English into Xhosa by a South African team of researchers. This study was primarily concerned with the Verbal Analogies Scale of the Woodcock Muñoz Language Survey and aimed to extend previous research on the equivalence of the two language versions of the scale. The study employed a monolingual two-group design consisting of 150 mainly English-speaking and 149 mainly Xhosa learners in Grades 6 and 7. The first research objective was to investigate item bias (or differential item functioning items) in the Visual Analogies Scale across the Xhosa and English versions using logistic regression and Mantel–Haenszel statistical techniques. Five items were identified as differential item functioning. The second objective was to evaluate the construct equivalence of the two versions by conducting a factor analysis after removing the differential item functioning items from the scale. Two factors were identified. The first factor displayed significant loadings across both language versions. The second factor was stable for the English version but not for the Xhosa version. Results were supported by calculating a Tucker’s phi coefficient for both factors. It was therefore concluded that Factor 1 is structurally equivalent across the two language versions but that Factor 2 was not structurally equivalent. Thus, the detection and removal of differential item functioning items did not result in structural equivalence.

The main focus of this study was an investigation into the equivalence of two language versions (English and Xhosa) of a Verbal Analogies (VA) Scale that is used for language testing in an additive bilingual education programme. This focus developed out of a broader need for tests that can be used in a multilingual South African (SA) society that are valid for use across language groups (Foxcroft, 1997). A requirement for tests that are valid across groups is built into current SA legislation (Employment Equity Act, 1998).

A number of studies on psychological and educational tests in the SA context demonstrated the level and extent of bias that is present in current available tests and supports this requirement and the focus of the study (Abrahams, 1996, 2002; Claassen, 1993; Koch, 2007; Meiring, Van de Vijver, Rothmann, & Barrick, 2005). One way of going about producing tests that are valid for use across groups is to have a test available in more than one language and to produce evidence that the scores of the two (or more) versions have the same meaning. These concepts and the methodology will be explored

- Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée*, 54, 119–135.
- Watkins, M. W. (2002). *Coefficient of congruence (Rc)* [Computer software]. State College, PA: Author.
- Woodcock, R. W., & Muñoz-Sandoval, A. F. (2001). *Woodcock-Muñoz Language Survey: Normative update*. Itasca: Riverside Publishing Company.