Chapter 5
Statistical Techniques for Making Cross-Cultural Comparisons on Organizational Instruments

Tanesia R. Beverly
Law School Admission Council, USA

ABSTRACT

Researchers tend to evaluate psychological instruments in terms of reliability (internal consistency) and construct validity (exploratory factor analysis and confirmatory factor analysis). In many instances, these instruments are used for cross-cultural comparisons such as gender and race—however, many of these studies do not provide evidence of measurement invariance or measurement equivalence. Measurement equivalence is a statistical property of an instrument that indicates that participants interpret and respond to the items similarly or that the same latent construct is being measured across observed groups of people. Partial measurement equivalence is a necessary condition for comparing latent mean differences across cultures. This area of construct validity is often neglected in the literature; therefore, this chapter aims to introduce the concept of measurement invariance. Additionally, it highlights the necessity of testing for measurement invariance when making cross-cultural comparisons on organizational leadership instruments.

INTRODUCTION

Researchers often use self-reported survey instruments (e.g., surveys and questionnaires) to measure different latent constructs (e.g., attitudes, grit, self-efficacy). These researchers utilize factor analytic procedures to determine the factor structure of survey or Likert scale data. Researchers spend a great deal of energy ensuring that their newly developed survey instrument has strong psychometric properties such as internal consistency and validity. However, these instruments are typically used to compare mean differences across observed demographic groups (e.g., gender, race, and nationality), but many research-

DOI: 10.4018/978-1-7998-7665-6.ch005

ers fail to ensure that their instrument meets the basic level of scrutiny to make such comparisons. To make valid group comparisons, researchers must ensure that the instrument has the same structure across groups or that the survey questions function similarly for all respondents regardless of their demographic characteristics. Some researchers apply traditional methods, such as ANOVA, when making cross-cultural comparisons based on a composite or summed score from the instrument. Statistical analyses like ANOVA and t-tests assume measurement invariance or measurement equivalence (Borsboom, 2006) or that the instrument is measuring the same latent constructs across multiple populations (Marsh & Hocevar, 1985). For instance, measurement equivalence ensures that differences in latent scores across observed demographic groups result from true latent differences rather than an artifact of the instrument. Measurement equivalence testing provides construct validity evidence and allows researchers to draw inferences about group similarities and differences on these measured constructs. Neglecting to test for measurement invariance can affect the validity of the inferences being made, including inflation of Type I and II error rates.

This chapter introduces the concept of measurement invariance (i.e., equivalence). It also highlights the necessity of testing for measurement invariance when comparing groups (i.e., gender, race, nationality) on new and existing survey instruments. In this chapter, cross-cultural comparisons refer to observable characteristics such as race, gender, ethnicity. The same procedures apply regardless of the groups being compared.

BACKGROUND

The next section details the statistical underpinnings of the measurement equivalence procedure. It starts by introducing the foundations of the model and then details the procedure for testing for measurement equivalence.

Measurement Model

The multiple-group confirmatory factor analysis (MGCFA) was introduced to test measurement equivalence across observable groups (Jöreskog, 1971). Jöreskog's (1971) hierarchical tests for measurement equivalence excluded the mean structure or tests for equality of intercepts and latent means. Sorböm (1974) extended this technique to incorporate means and covariance structures (MACS) or compare intercepts and latent means across groups. In the MACS, equality constraints for factor loadings and intercepts across groups are imposed to detect latent mean differences (Sorböm, 1974).

MGCFA allows researchers to test a priori hypotheses about model parameters across groups. The confirmatory factor analysis model (CFA) provides a framework for evaluating equality constraints across groups. The CFA is defined as:

$$X_g = \tau_g + \Lambda_g \xi_g + \delta_g \tag{1}$$

where X_g is the vector of observed scores, τg is the vector of intercepts or observed means on items, Λg is a matrix of factor loadings or the regression of the observed variables on the latent factors ξ, and