

Chapter 4

An Illustration of the Actual Steps in Development and Validation of a Multi-Item Scale for Quantitative Research: From Theory to Practice

Dail Fields

Regent University, USA

ABSTRACT

This chapter describes in detail the process used to develop and validate a scale that measures servant leadership. The steps covered include construct identification from previous studies, review of previously proposed and developed measures, item selection, survey development, collection of data, scale identification, and evaluation of convergent, discriminant, predictive validity. The chapter provides a hands-on example of the steps required for scale measure development and assessment and includes description of the mechanics involved in completing each step of this process.

INTRODUCTION

Over the years, research resources have prescribed the steps involved for development and validation of multi-item scale measures for use in quantitative research (Arnold, Arad, Rhoades & Drasgow, 2000; Fields, 2002). Multi-item scale measures operationalize a construct, which in quantitative research means the scale provides a measure or single numerical of the construct and/or its constituent parts for each study subject. Multi-item summated scale measures include examples such as the Job Diagnostic Survey (Hackman & Oldham, 1980), The Burnout Measure (Pines & Aronson, 1988), and the Minnesota Satisfaction Questionnaire (Weiss, Dawis, England & Lofquist, 1967). These measures all use multiple items (statements or questions) to describe a respondent's perception of an attitude, condition or behavior to improve the internal consistency reliability of the scale and increase the chances that the

DOI: 10.4018/978-1-7998-7665-6.ch004

scale measure covers the meaning of the construct as interpreted by a variety of individuals (Cortina, 1993; Spector, 1992). These measures are very critical to successful quantitative research designs as they enable analysis of possible relationships between each construct and other constructs of interest and theoretical relevance (DeVellis, 2003; Fields, 2002).

There are seven steps specified in the measurement literature that provide critical reassurances to possible users that a scale measure of a construct in the social and behavioral sciences has been rigorously developed, evaluated, and is indeed ready to use (DeVellis, 2003; Kraimer, Seibert, & Liden, 1999; Spector, 1992; Walumbwa, Avolio, Gardner, Wernsing, & Peterson, 2008). To one extent or other, these steps are presented to students in the process of quantitative research training. But how does one go about performing each of these steps correctly? That is the focus of this chapter. Specifically, this chapter reviews and illustrates the ‘nuts and bolts’ involved in performing each of the steps required in the course of developing a measure of the construct describing a particular set of leadership behaviors.

BACKGROUND

Meyer and Allen (1997) presented requirements for the psychometric evaluation of measures used in quantitative research in the social and behavioral sciences. The process generally consists of two stages. The first is identification of a pool of possible descriptors of the construct of interest and analysis of the applicability of the descriptors based on the views of a sample of subjects familiar with the construct. The second stage consists of evaluating statistically the internal consistency of the new measure as well as its construct validity. The construct validity of a measure, which provides evidence that the derived scale in fact measures what it purports to measure, can be assessed by examining its correlations with other constructs and comparing these correlations with what is expected theoretically (Kerlinger & Lee, 2000). Specifically, construct validity assesses the extent to which a focal measure is significantly related to another validated of a very similar construct (convergent validity), is not related to other distinct different constructs (discriminant validity), and the measure is positively related to an outcome with which the construct is known to be associated (predictive validity). For example, a newly derived measure of job satisfaction should be significantly positively associated with other previously validated measure of job satisfaction, should be negatively associated with a measure of resentment towards and employer, and should be predictive of employee organizational commitment (Scarpello & Vandenberg, 1992).

In addition, confirmatory factor analysis is appropriate for investigating construct validity of multi-item scales because it allows for direct examination of the degree to which specific items jointly are associated with hypothesized factors (i.e. convergent validity) and display minimal cross-loadings on other factors (i.e. discriminant validity). For example, in a four-dimensional measure, if the dimensions do not have discriminant validity, the fit of a single-factor model will be no worse than will the fit of a four-factor model (Kraimer, Seibert, & Liden, 1999).

Starting with a theoretically based description of the construct, the steps generally advocated for development and validation of a quantitative measure of a construct include the following (Arnold, Arad, Rhoades, & Drasgow, 2000; Devillis, 2003; Strauss & Smith, 2009):

1. Use theory and previous empirical research to develop a ‘pool’ of possible statements or questions that could possibly describe the nature of the construct.

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/an-illustration-of-the-actual-steps-in-development-and-validation-of-a-multi-item-scale-for-quantitative-research/285188

Related Content

Simulating Light-Weight-Cryptography Implementation for IoT Healthcare Data Security Applications

Norah Alassaf and Adnan Gutub (2021). *Research Anthology on Blockchain Technology in Business, Healthcare, Education, and Government* (pp. 1468-1483).

www.irma-international.org/chapter/simulating-light-weight-cryptography-implementation-for-iot-healthcare-data-security-applications/268671

Machine Learning for Big Data

Supriya M. S. and Keerthana Sasidaran (2021). *Applications of Big Data in Large- and Small-Scale Systems* (pp. 56-76).

www.irma-international.org/chapter/machine-learning-for-big-data/273920

Gender-Based Tweet Analysis (GTA)

Dipti P. Rana and Navodita Saini (2021). *Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance* (pp. 255-267).

www.irma-international.org/chapter/gender-based-tweet-analysis-gta/280922

Edge-of-Things Computing-Based Smart Healthcare System

Diana Yacchirema, Carlos Palau and Manuel Esteve (2022). *Research Anthology on Edge Computing Protocols, Applications, and Integration* (pp. 299-320).

www.irma-international.org/chapter/edge-of-things-computing-based-smart-healthcare-system/304308

Measuring Clergy Effectiveness: The Development of the Clergy Effectiveness Scale

Samuel P. Dobrotka (2021). *Handbook of Research on Advancements in Organizational Data Collection and Measurements: Strategies for Addressing Attitudes, Beliefs, and Behaviors* (pp. 293-313).

www.irma-international.org/chapter/measuring-clergy-effectiveness/285203