

Personality assessment: Does the medium matter? No ☆

Siang Chee Chuah *, Fritz Drasgow, Brent W. Roberts

*Department of Psychology, University of Illinois in Urbana-Champaign, 603 East Daniel Street, Champaign,
IL 61820, USA*

Available online 3 March 2006

Abstract

The equivalence of an Internet administration of personality tests with two other administration formats was assessed using Item Response Theory (IRT) and various other statistical methods. The analyses were conducted on measures of Neuroticism, Extroversion, Agreeableness, and Conscientiousness. A total of 728 participants took part in the study. Participants were randomly assigned to one of three administrative conditions: paper-and-pencil, proctored computer lab, and unproctored Internet. Analyses with IRT, factor analysis, criterion-related validity, and mean differences supported the equivalence of Internet and traditional paper-and-pencil administrations of personality tests.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Measurement equivalence; Personality assessment; Internet testing; Computer-based testing; Proctored testing; Big 5; IRT; DIF

1. Introduction

The proliferation of the Internet presents an opportunity to deliver questionnaires through a medium that offers tremendous benefits. It offers potential access to huge populations from diverse backgrounds and geo-political locations (Schmidt, 1997; Smith & Leigh, 1997). Tests can be administered at virtually any location where there is a computer and

☆ This research was partly supported by a grant from the National Institute of Aging R01 AG21178.

* Corresponding author.

E-mail address: chuah@uiuc.edu (S.C. Chuah).

Internet access. In that sense, the comfort of one's home may become a test center. Data are transmitted, literally at the speed of light, from the respondent to the test administrator through a network of computers connected by fiber-optic cables. Because the data are in a digital format, they may be quickly converted into a format suitable for data analysis (Davis, 1999). In addition, the resources needed to create a website to administer questionnaires can be relatively inexpensive (Schmidt, 1997).

However, with any new tool there lies the potential for abuse. As Butcher (1987) points out, people tend to be less critical when evaluating the legitimacy of computerized results. Computers are often seen as objective or impartial evaluators of test results. He argues that there is no assurance as to the quality of the computerized system. Without the appropriate input by trained professionals to develop or evaluate the adequacy of the system, the legitimacy of a computerized test is suspect.

A key concern for test developers and administrators who want to deploy their tests over the Internet is the equivalence of scores between paper-and-pencil administrations and Internet administrations. Before tests scores can be directly compared, equivalence between test formats needs to be established (Standards for Educational & Psychological Testing, 1999). Equivalence is important for several reasons. For example, the validity and reliability of the scale was most likely established using paper-and-pencil administration samples; however, we cannot simply assume that the Internet form has similar reliability and validity. Lacking empirical verification of cross-medium equivalence, the instrument needs to be reevaluated to assess its validity and reliability. Additionally, archival data is often used to establish population norms. These data are commonly collected using the paper-and-pencil version of the instrument. Norms are important because they allow one to interpret one's results in comparison with other studies across a broad range of test situations. If scores from an Internet are to be interpreted relative to norms obtained from a paper-and-pencil administration, equivalence is required.

This paper tests the equivalence of paper-and-pencil and Internet administrations by looking at the psychometric qualities and criterion validity of a personality measure across administration media. Item response theory (IRT) was used to assess the psychometric qualities of three different administrative methods of a personality questionnaire. Specifically, we examined the measurement equivalence of a set of adjectives from the Big 5 domain (Goldberg, 1992). We also assessed equivalence by means of multi-group factor analysis: A single factor loading matrix was simultaneously fit to data from all three administrations. We assessed criterion-related validity by comparing correlations between the Big 5 and the Health Behavior Checklist (HBC; Vickers, Conway, & Hervig, 1990) across administrative media. Previous research has already established the relationship between personality and health behaviors (Bogg & Roberts, 2004; Booth-Kewley & Vickers, 1994). The results and implications of the findings are discussed.

2. Previous research

A number of factors can influence responses to questionnaires administered in diverse test modalities. Previous studies comparing the equivalence of computer administered tests to paper-and-pencil tests have suggested that specific features of computerized tests may affect the outcome. A meta-analysis by Richman, Weisband, Kiesler, and Drasgow (1999) found that non-cognitive tests suffered from social desirability distortions under certain conditions. Social desirability distortion refers to the tendency of respondents "to answer questions in a

Download English Version:

<https://daneshyari.com/en/article/952098>

Download Persian Version:

<https://daneshyari.com/article/952098>

[Daneshyari.com](https://daneshyari.com)