Personality and Individual Differences 52 (2012) 37–40

Contents lists available at SciVerse ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

A putatively general factor of personality (GFP) is not so general: A demonstration with the NEO PI-R

Ronald R. Holden *, Zdravko Marjanovic

Queen's University, Kingston, Ontario, Canada

ARTICLE INFO

Article history:
Received 25 May 2011
Received in revised form 23 August 2011
Accepted 24 August 2011
Available online 28 September 2011

Keywords:
General factor of personality
Personality structure
NEO PI-R
Big Five

ABSTRACT

Recent research suggests that a general factor of personality (GFP) represents the zenith of a hierarchy of personality structure. For a roommate sample of 602 students, we evaluate the presence and validity of a general factor of personality in a Big Five measure. Findings indicate that a first factor, similar to what has been putatively labeled a GFP, can be extracted from self-report and observer-report, that this self-report first factor has validity for predicting an alleged observer-report GFP, and that this validity is not attributable to socially desirable responding. However, despite the existence of a valid first factor, it is not a general factor of personality because it fails to summarize adequately the complete systematic variance in the structure of personality.

1. Introduction

Recently, there is a growing amount of research proffering the perspective that there is a general factor of personality (GFP) that is analogous to *g*, the general factor in intellectual abilities (Just, 2011). Across a variety of Big Five personality measures and sampling from various populations, Musek (2007) and Rushton and Irwing (2008) indicate a single higher-order factor that sits at the summit of a hierarchical personality structure. Similar results have been found for other inventories including the Comrey Personality Scales, the Minnesota Multiphasic Personality Inventory-2, and the Multicultural Personality Inventory (Rushton & Irwing, 2009a), the Multidimensional Personality Questionnaire (Rushton & Irwing, 2009b), the Guilford-Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory (Rushton & Irwing, 2009c), and the Millon Clinical Multiaxial Inventory-III, the Dimensional Assessment of Personality Pathology, and the Personality Assessment Inventory (Rushton & Irwing, 2009d).

Dissenting perspectives exist, however, with some researchers reporting no evidence for the GFP (de Vries, 2011), some suggesting alternative explanations (e.g., non-hierarchical models producing equivalent personality scale correlation matrices; Ashton, Lee, Goldberg, & de Vries, 2009; Revelle & Wilt, 2010), and others indicating that a major general factor is explainable as a social desirability response style (Backstrom, Bjorklund, & Larsson, 2009). Rebuttals to these contrary views have shown that a GFP remains intact even when controlling for a social desirability response bias (Rushton & Erdle, 2010) and that a GFP has criterion validity for predicting job performance (van der Linden, te Nijenhuis, & Bakker, 2010). The current research contributes to the debate on a general factor of personality by employing criteria that are external to self-report, by using multiple measures of socially desirable responding, and by analyzing data with a technique allowing for an alternative examination of whether the GFP exists at a personality structure apex that completely summarizes self-report in the prediction of relevant criteria. In this contribution, we do not dispute Rushton and Irwing (2009a–d) regarding the existence of a valid linear combination to which all Big Five dimensions contribute. Rather, we take issue with their view that this combination is a general summary that is at the apex of personality structure.

2. Data source

Data are from Holden and Passey (2010) and comprise the self-report NEO Personality Inventory (NEO PI-R) responses of 602 university students and corresponding roommate criterion ratings from the observer form of the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992). The 240-item NEO PI-R and 60-item NEO-FFI each include Big Five scales of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. In addition, participants self-reported on five measures of socially desirable responding: the Self-Deceptive Enhancement scale and the Impression Management scale of the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1998), the Marlowe–Crowne Social Desirability Scale (Crowne & Marlowe, 1960), Jackson's (1984)

* Corresponding author. Address: Department of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6. Tel.: +1 613 533 2879; fax: +1 613 533 2499.
E-mail address: holdenr@queensu.ca (R.R. Holden).

0191-8869/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.paid.2011.08.024

Personality Research Form Desirability scale, and the validity index of the Holden Psychological Screening Inventory (HPSI; Holden, 1996). Holden and Passey (2010) indicate that all scale reliabilities were consistent with those reported in respective test manuals. Respondents were paid for participating, had a mean age of 18.80 years (range 17–27), and had known their respective roommates for at least 3 months.

3. Analyses and results

3.1. Support for the GFP

Various techniques have been used to compute scores for the GFP including summing Big Five scales with the Neuroticism scale reverse-keyed (Erdle & Rushton, 2011), calculating scores on the first unrotated principal component (de Vries, 2011; Musek, 2007; van der Linden et al., 2010; Weiss, Adams, & Johnson, 2011), and using hierarchical factor analysis (de Vries, 2011). Following Musek's (2007) original work, we employed principal components analysis because it is the most widely used method of factoring and, like Musek, yielded practically identical results with other methods of generating GFP scores, such as principal axis factoring (GFP scores based on principal components and principal axis factoring correlated .985), maximum likelihood factoring (scores based on principal components and maximum likelihood methods correlated .955), and scores based on simply summing domain scales with the Neuroticism scale reflected (scores based on principal components and raw scale sums correlated .990).

For the GFP on the self-report NEO PI-R, scores were computed for the first principal component of the five domain scales (Table 1). That component, accounting for 32.54% of the five scales' variance, is highly similar to the component reported by Musek (2007) and van der Linden et al. (2010). Separately, a GFP for the observer-report was similarly computed in the NEO-FFI data. That component, accounting for 35.77% of the variance in its five scales, is interpretively similar to the one extracted for self-report in the NEO PI-R. In another distinct analysis, scores on a general factor of socially desirable responding were computed for the first principal component underlying the five response style scales. This component accounted for 49.19% of its five scales' total variance.

To evaluate the validity of the self-report GFP, we correlated first principal component scores from the self-report NEO PI-R with corresponding first component scores from the observer-report NEO-FFI. A correlation of .36 ($p < .00001$) indicated validity of more than a medium effect size (i.e., .30; Cohen, 1992) for the self-report GFP. This correlation compares favorably with those validities found for many personality scales. When scores on the first component for the socially desirable responding scales were partialled out, the association between the self-report GFP and observer-report GFP remained significant, $r(592) = .26$, $p < .00001$, indicating that the validity of the GFP was not attributable to a socially desirable response style. We also computed the Gower similarity statistic (Barrett, 2010; Gower, 1971), estimating agreement between the self-report and observer-report GFPs relative to the maximum possible absolute discrepancy, which indicated that self-report and observer-report agreed to within 88% of each other's values ($p < .01$). Thus, in support of the GFP, we replicate what others have found for its interpretive nature, we show its moderate validity using a criterion external to self-report, and we demonstrate that this validity is not produced by a common response style.

3.2. Opposition to the GFP

Although the analyses above are generally supportive of a GFP, there is an important discrepancy between the above results and the individual domain scale validities reported by Holden and Passey (2010). For the domain scales, Holden and Passey (2010) reported correlations between self-report and observer-report that ranged from .42 for Neuroticism to .52 for Conscientiousness. With a mean validity of .46, this indicates that an average of 21% of the criterion variance was accounted for at the individual domain scale level. When compared to the 13% of the variance in the observer-report accounted for by the self-report GFP, it is evident that by simplifying our rendering of personality structure, explanatory power is being lost in the tradeoff.

Given that the validity for the GFP and the individual scale validities reported by Holden and Passey (2010) are not directly comparable, we undertook a canonical correlation analysis relating self-report and observer-report scales to articulate more finely the associations between self-report and observer-report. In canonical correlation, optimally weighted linear combinations of predictor (e.g., self-report) and criterion (e.g., observer-report) variables are constructed so as to maximize associations between predictors and criteria. The number of possible canonical correlations is the minimum of either the number of predictors or criteria. Linear combinations are sequentially generated that maximize the correlations between predictors and criteria subject to these solutions being mathematically orthogonal to earlier generated solutions. For statistical significance, the pool of canonical correlations is tested and, if significant, the largest linear combination is removed and the remaining pool retested. This continues sequentially until the remaining pool of canonical associations is no longer significant or the pool is exhausted. If there were a GFP that adequately summarized personality structure, only the pool comprising all canonical correlations should be significant and, after the removal of the largest function, the remaining pool should no longer manifest statistically significant associations between predictors and criteria (see Sherry & Henson, 2005).

For the five self-report scales and five observer-report scales, five canonical correlations were extracted and their associated linear combinations or variates are displayed in Table 2. With

Table 1
Loadings on first unrotated principal component.

Scale	NEO PI-R self-reported personality	Scale	Self-reported socially desirable responding	Scale	NEO-FFI observer-reported personality
Neuroticism	−.69	Self-Deceptive Enhancement	.35	Neuroticism	−.69
Extraversion	.68	Impression Management	.73	Extraversion	.65
Openness	.33	Marlowe–Crowne	.79	Openness	.15
Agreeableness	.52	Jackson Desirability	.76	Agreeableness	.78
Conscientiousness	.56	HPSI total	−.78	Conscientiousness	.52
Variance Accounted For	32.54%		49.19%		35.77%