Journal of Research in Personality 45 (2011) 535–548

Contents lists available at ScienceDirect

Journal of Research in Personality

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

Assessing within-person personality variability via frequency estimation: More evidence for a new measurement approach

Matthew S. Fleisher [a,*], David J. Woehr [b], Bryan D. Edwards [c], Kristin L. Cullen [d]

[a] Human Resources Research Organization, 66 Canal Center Plaza, Suite 700, Alexandria, VA 22314-1591, United States
[b] Department of Management, The University of North Carolina, Charlotte, 9201 University City Blvd., Charlotte, NC 28223-0001, United States
[c] Department of Management, Oklahoma State University, 310 Business Building, Stillwater, OK 74078-4011, United States
[d] Department of Psychology, Auburn University, 226 Thach Hall, Auburn, AL, 36849-5214, United States

ARTICLE INFO

Article history:
Available online 18 July 2011

Keywords:
Measurement
Reliability
Validity
Intraindividual variability
Consistency

ABSTRACT

Variability in personality has been recognized in recent years as an important aspect of personality both conceptually and empirically. A relatively new and efficient method of obtaining variability information is frequency-based personality assessment (Edwards & Woehr, 2007). The purpose of the present research is to further examine the viability of frequency-based personality measurement as an alternative to traditional (Likert-type) measurement and to assess the usefulness of the variance-based parameters. Toward this end, three studies are presented. Specifically, Study 1 examined relationships between a frequency-based measure of the Big Five personality traits and several motivational variables. Study 2 examined the moderating role of temporal consistency information (provided by frequency-based measurement) on relationships between personality and peer ratings of task performance. Study 3 compared the frequency-based measure to a Likert-type measure with respect to each measure's susceptibility to deliberate response distortion. Results indicated that consistency information increases the predictive validity of agreeableness and conscientiousness and that a frequency-based format is less susceptible to faking than a Likert-type format for conscientiousness, emotional stability, and openness to experience.

1. Introduction

Recently, researchers have become increasingly interested in methods of capturing the (in)consistency of behavior (e.g., Fleeson, 2001; Gibbons & Rupp, 2009; Heggestad, Gordon, & Reeve, 2009; Huang & Ryan, 2009; Scott, Colquitt, Paddock, & Judge, 2010; Weinblatt & Heller, 2009). Specifically, the examination of within-person variability (or consistency) is an important focus of research in a variety of areas (e.g., Baird, Le, & Lucas, 2006; Fleeson, 2001, 2007; Gibbons & Rupp, 2009; Uy, Foo, & Aguinis, 2010). Frequency estimation-based assessment offers a viable approach to the measurement of personality and within-person variability (Edwards & Woehr, 2007). The purpose of the present research was to evaluate the efficacy of frequency estimation for personality measurement with respect to the assessment of both mean level and consistency/variability. We compared a frequency estimation-based response format to a traditional Likert-based response format. Arguably, the viability and efficacy of this approach rests on two considerations. First, frequency estimation-based assessments of the *level* of specific personality traits should demonstrate psychometric characteristics (i.e., reliability and validity) similar to those of more commonly used measurement approaches (i.e., Likert-type response scales). Second, the frequency estimation-based approach should add relevant information pertaining to the assessment of personality variability that is not available from single administrations of Likert scales. We present the results of three separate studies designed to directly address these issues.

1.1. Personality variability

Recently, several researchers have called for a broader view regarding the measurement and conceptualization of personality. More specifically, within-person consistency of trait-relevant behavior has been put forth as an important variable to consider, in addition to individuals' mean level or relative standing on personality traits (Baird et al., 2006; Baumeister, 1991; Bem & Allen, 1974; Biesanz & West, 2000; Biesanz, West, & Graziano, 1998; Brown & Moskowitz, 1998; Cervone, 2004; Fleeson, 2001, 2007; Fleeson & Leicht, 2006; Fleeson, Malanos, & Achille, 2002; Mischel & Shoda, 1998). The general idea is that any given individual behaves differently on different occasions. Thus, an individual's

* Corresponding author. Address: Human Resources Research Organization, 66 Canal Center Plaza, Suite 700 Alexandria, VA 22314-1591, United States.
E-mail address: mfleisher@humrro.org (M.S. Fleisher).

0092-6566/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jrp.2011.06.009

behavior over time forms a distribution with respect to a given trait, and the various characteristics of this distribution (e.g., mean, variance, etc.) are relevant to the measurement of personality. Important information related to the variance of a given trait cannot be captured with a mean score alone. Accordingly, Fleeson (2001) called for movement away from an exclusive focus on assessments of central tendency in personality measurement to an examination of the density distributions of trait-relevant behavior.

Other reasons for investigating new measurement approaches for assessing personality include two primary criticisms leveled at traditional direct report, self-perceptions of personality: (a) weak correlations with behavior, and (b) susceptibility to socially desirable responding. Personality research in general has been criticized because the relationship between personality variables and behavior is relatively weak (e.g., Arthur, Woehr, & Graziano, 2001; Mischel, 1968; Peeters, Van Tuijl, Rutte, & Reymen, 2006) and this may be in part due to the limitations of traditional measurement systems. For example, the uncorrected mean validity coefficients between self-reports of conscientiousness and job performance reported in the meta-analytic literature are approximately .13 (Barrick & Mount, 1991; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Hurtz & Donovan, 2000). However, we should point out that some recent research has demonstrated stronger relationships between personality and behavior (e.g., Back, Schmukle, & Egloff, 2009; Fleeson & Gallagher, 2009). Self-report-based personality assessment has also been directly criticized because many items are rather transparent, making them susceptible to socially desirable responding (e.g., James, 1998; James & Mazerolle, 2002). To date, there is a sizable amount of empirical evidence indicating that response distortion reduces the accuracy of self-reports and also attenuates the criterion-related validity of personality tests (e.g., Bing, Whanger, Davison, & VanHook, 2004; Holden, 2007, 2008; White, Young, Hunter, & Rumsey, 2008).

A small but growing body of research has examined the role of within-person variability, or temporal consistency, as a potential moderator of the relationship between personality and relevant external criteria (e.g., Baumeister & Tice, 1988; Biesanz & West, 2000; Biesanz et al., 1998). This approach is based on the assumption that low correlations between measured personality and relevant outcomes might occur if personality influences behavior to a larger extent for some individuals than for others (Bem & Allen, 1974). For example, research indicates that behavioral consistency moderates the magnitude of self–other agreement on ratings of personality (Bem & Allen, 1974; Biesanz et al., 1998; Edwards & Woehr, 2007). So, accurate measurement of personality, as well as the prediction of behavioral outcomes, requires not only information about an individual's trait level, but also the extent to which that trait will influence behavior for the individual (i.e., degree of consistency). That is, two individuals may have equivalent levels of a given personality trait, but differ in *traitedness*, or on the degree of consistency with which this trait influences behavior (Baumeister & Tice, 1988). Baumeister and Tice postulated that more traited individuals will demonstrate more consistent behavior with respect to that trait than less traited individuals. Thus, individuals with higher levels of traitedness should yield stronger personality–behavior relationships.

Conceptually, within-person variability reflects the extent to which the same individual acts differently on different occasions. Historically, researchers have operationalized consistency in a number of different ways. Some researchers have viewed consistency as nomothetic cross-situational consistency and simply asked respondents to provide an estimate of their overall level of behavioral consistency with respect to a particular personality dimension. Amelang and Borkenau (1986), however, noted that this approach has very low test–retest reliability. Other researchers have operationalized within-person variability in terms of response patterns. Specifically, consistency across items assessing the same dimension has been used as an index of within-person variability (i.e., dimension-level standard deviation). There are, however, a number of problems associated with this approach, not the least of which is that it does not provide information on consistency within situations over time (Biesanz et al., 1998).

Another approach to operationalizing consistency stems from the work of Allport (1937), who emphasized consistency within situations over time (i.e., temporal response pattern stability). For example, Biesanz et al. (1998) found that the level of self–other agreement on the dimensions of extraversion and conscientiousness was moderated by the participants' response consistency. Their results suggested that high levels of temporal consistency of behavior within situations (i.e., responses to the same behavioral items across measurement sessions) increased the predictability of the personality dimensions by external observers. Other studies have demonstrated the efficacy of measuring behavioral consistency using a multiple measurement approach called experience sampling methodology (ESM) (e.g., Baird et al., 2006; Fleeson, 2001; Heggestad et al., 2009). However, there are disadvantages for each of the previous approaches to measuring behavioral consistency. Thus, Edwards and Woehr (2007) evaluated a novel approach to obtaining personality variability information—frequency-estimation based personality assessment. The following sections describe the basis for this approach and initial research examining its validity.

1.2. Frequency estimation as a basis for assessment

Research over the last four decades has established that humans may be capable of recalling event frequencies with appreciable degrees of accuracy (Cosmides & Tooby, 1996; Kane & Woehr, 2006; Steiner, Rain, & Smalley, 1993). For example, in a meta-analysis of the accuracy of frequency estimation, Kane and Woehr (2006) found a mean sample-size-weighted correlation of .73 between frequency estimates and actual frequency counts. Further, a stream of research within this domain has provided evidence that many cognitive heuristic-based biases (e.g., Tversky & Kahneman, 1973, 1974) largely disappear when judgments are made in terms of frequencies as opposed to probabilities (Gigerenzer, 1991; Sedlmeier, Hertwig, & Gigerenzer, 1998). It has been suggested that the high levels of accuracy and sensitivity observed in frequency estimation may occur because this process parallels the way in which people naturally encode, store, recall, and process behavioral and event frequencies (Cosmides & Tooby, 1996). Thus, frequency estimation may reduce cognitive load during the assessment process because it may be easier for respondents to recall event frequencies than to mentally calculate an average level across time (as required by typical Likert-type formats).

Measurement systems based on frequency estimation require individuals to report the absolute or relative frequency of occurrence for specific outcomes or behaviors over a specified time period (Kane, 1986, 2000). The result of such assessments is a frequency distribution that portrays the range of the individuals' behavior. Typical descriptive summary measures including central tendency and variability may be calculated from this distribution. To date, frequency-based measurement has been predominantly examined in the context of organizational performance appraisal ratings (Deadrick & Gardner, 1997; Kane, 1986, 1996, 2000; Kane & Lawler, 1979; Steiner et al., 1993; Woehr & Miller, 1997). Through empirical testing, several positive characteristics of distributional assessment have emerged: (a) individuals are able to detect differences in the distributional characteristics of performance even when mean levels are constant (Steiner et al., 1993); (b) frequency-based estimates generally contain less measurement error than traditional estimates of performance (Woehr & Miller,