Personality at Work: Raising Awareness and Correcting Misconceptions

Deniz S. Ones
Department of Psychology
University of Minnesota

Chockalingam Viswesvaran
Department of Psychology
Florida International University

Stephan Dilchert
Department of Psychology
University of Minnesota

Personality variables have always predicted important behaviors and outcomes in industrial, work, and organizational psychology. In this commentary, we first review empirically supported structural models of personality that show the following: (a) Personality traits are hierarchically organized, (b) the Big Five are not orthogonal, (c) abnormal personality measures assess the same continuum of traits as normal adult personality measures, and (d) there are compound personality traits that are especially useful in the prediction of organizational behaviors. Second, we provide a brief overview of meta-analyses of compound personality variables. The highest operational validities of single scales (.40s) are associated with personality measures assessing broad, compound personality characteristics, such as integrity, violence potential, customer service orientation, and managerial potential, that incorporate aspects from multiple dimensions of the Big Five. Third, we also review meta-analytic evidence that has linked personality attributes to other important organizational attitudes and behaviors, such as job satisfaction, motivation, and leadership, with multiple correlations for the Big Five in the .40 to .50 range. Fourth, we discuss the important role that meta-analysis has had in establishing the predictive and explanatory value of personality variables. We conclude with some caveats and directions for future research.

Correspondence should be sent to Deniz S. Ones, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455. E-mail: deniz.s.ones-1@tc.umn.edu
Personality variables have always predicted important behaviors and outcomes. In the past 2 decades, large-scale meta-analyses have documented the pervasive influence of personality constructs in virtually all aspects of organizational behavior. In this commentary, we first summarize the current state of the art in research on personality structure and measurement. Second, we provide a brief overview of the criterion-related validities of personality traits for overall job performance and other organizational behaviors (e.g., motivation, leadership, job satisfaction) based on extensive meta-analytic evidence. Third, we discuss and clarify the role that meta-analysis has had and ought to have in summarizing the mountains of data supporting the use of certain personality variables in personnel selection. We conclude our commentary by identifying some future directions for personality research in industrial, work, and organizational (IWO) psychology.

PERSONALITY STRUCTURE AND MEASUREMENT

Personality traits are enduring dispositions and tendencies of individuals to behave in certain ways. Personality is not one single thing. Instead, personality refers to a spectrum of individual attributes that consistently distinguish people from one another in terms of their basic tendencies to think, feel, and act in certain ways. The enduring nature and consistency of personality characteristics are manifested in predictable tendencies of individuals to behave in similar ways across situations and settings. There are thousands of personality characteristics that can be used to distinguish individuals from one another.

When individuals respond to single items on personality inventories, they are providing self-descriptions about their own behaviors and reputations. Specific items refer to single instances of attitudes, thoughts, feelings, and behaviors. Psychologically similar items cluster together to make up specific traits. Similar traits can be grouped together to make up higher order personality dimensions. Factor analysis of self- and peer ratings on single trait descriptors or items is often used to derive such higher order dimensions or factors. An individual’s “personality” is best understood as a profile of high and low levels on such dimensions and facets, rather than an average score across different traits. Next, we briefly describe contemporary research on the structural organization of personality constructs to raise awareness about the meaning of personality scale scores.

Personality Constructs Are Hierarchically Organized

In the past several decades, research has accumulated that clusters thousands of personality characteristics into meaningful, interrelated groups. It has been shown that these interrelated clusters are hierarchically organized. The past 20 years or so
have witnessed the emergence and wide acceptance of the Five-factor model of personality (FFM). The Big Five factors are Emotional Stability, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. These factors have emerged from lexical studies of phenotypic personality traits (e.g., Goldberg, 1993), and from joint factor analyses of personality instruments built to assess the FFM and personality instruments created based on other theoretical perspectives (e.g., Gough’s folk concepts, Murray’s needs, psychopathology; see Dilchert, Ones, Van Rooy, & Viswesvaran, in press, for an account of the history, description, scientific and applied utility of the FFM).

Data from multiple personality inventories and thousands of test takers have provided consistent evidence for the hierarchical organization of personality. At the lowest level are individual responses to test items. Items that cluster together are indicators of specific attributes that may be referred to as personality subdimensions or facets. Facets that share psychological meaning, and most likely similar etiology, combine to define personality factors. For example, Extraversion is a broad factor that encompasses subdimensions or facets such as sociability (gregariousness), energy (activity), and assertiveness (dominance). There are individual scale items that uniquely assess each of these Extraversion facets. For example, “I enjoy parties” is an item that can be found on sociability scales and “I am good at persuading others” is an item that can be found on assertiveness and dominance scales. Note that the overall factor of Extraversion is defined by the common variance that is shared across its facets. There are some additional features of the personality factor structure that are important to highlight as they will have implications for the explanation and prediction of behavior in the world of work.

The Big Five Are Not Orthogonal

Often, the Big Five factors of personality are characterized and described as if they are virtually orthogonal. They are not; the Big Five factors correlate with one another. Based on the meta-analysis conducted by Ones (1993), Ones, Viswesvaran, and Reiss (1996) reported a meta-analytically derived matrix of intercorrelations among the Big Five factors. Unreliability-corrected correlations among three of the factors (Conscientiousness, Agreeableness, and Emotional Stability) were in the .20s. These correlations have implications for the presence of psychologically meaningful higher order factors. In Digman’s (1997) factor analyses of 14 matrices reporting intercorrelations among the Big Five factors, two higher levels were supported. The first was a higher order construct of socialization and getting along in a society with rules, norms, and conventions. Conscientiousness, Agreeableness, and Emotional Stability defined this higher order factor. Digman referred to this factor as “factor alpha.” Extraversion and Openness were related to a higher order factor of “getting ahead,” termed factor beta. Recent meta-analytic investi-
gations have reconfirmed the presence of these two higher order factors in the personality hierarchy (Markon, Krueger, & Watson, 2005).

Figure 1 presents the structure of personality as it is currently understood and described based on multiple empirical studies and meta-analyses. The figure is based on a recent article by Markon et al. (2005), but has been adapted to include all levels of the hierarchy. Note that the precise numbers of facets that make up each of the Big Five factors are not fully established (but see Roberts, Chernyshenko, Stark, & Goldberg, 2005, for an empirical approach addressing facets of Conscientiousness, for example). Figure 1 implies that responses to individual personality items capture variance from multiple levels of the hierarchy. For example, an affirmative response to the item “I quickly get loud in a fight” captures variance from factor alpha, from disinhibition, from disagreeable disinhibition, from agreeableness, and from the compliance facet of agreeableness, and unique variance that is specific to this item (unshared with latent traits at a higher level of the hierarchy).

Measuring personality traits is not simple. Personality items can capture trait variance from multiple factors and facets of the Big Five. Scales composed of such items assess compound personality traits (Hough & Ones, 2001). For example, ambition is a compound trait composed of aspects of Conscientiousness and Extraversion (see the working taxonomy in Hough & Ones, 2001, for other examples of compound traits). Whenever compound traits have been examined in meta-analyses or large-scale studies, the operational validities associated with scales assessing them have outperformed those obtained from the Big Five factor scales.

Profile scores represent a composite of several traits that are differentially weighted and combined. In a way, multiple regression can be viewed as a linear profile of independent traits. As Hogan (this issue) notes, several studies have found substantial validities for compound traits in predicting organizational behavior. In fact, the lack of awareness of these multiple correlations seems to be the reason for several recent criticisms of personality’s so-called low predictive value. The proponents of such criticisms invariably point out that reliability corrected correlations are only in the .20s, for even the best of the Big Five factors—overlooking the substantial criterion-related validities of compound traits reported in the literature (discussed later).

Utility of a Construct Domain Cannot Be Judged by its Poorest Measures

Critics of personality measures in IWO psychology (e.g., Murphy & Dzieweczynski, this issue; Murphy Paul, 2004) single out a number of personality measures when discussing the utility of personality. Often, these measures are not the best representatives of psychometrically sound personality measures and are not com-
patible with contemporary understanding of the construct domain of personality (e.g., Myers–Briggs Type Inventory). We should not judge the merits of a construct domain by its poorest indicators, especially when these indicators have nothing to do with modern-day conceptualizations of the construct domain, although they might be commonly used by less sophisticated researchers.

**Older Criterion-Related Validity Evidence for Some Inventories Should Not Be Overlooked Because They Were not Initially Constructed to Assess the Big Five**

Murphy and Dzieweczynski (this issue) claim that the Sixteen Personality Factor Questionnaire and the California Personality Inventory do not assess personality dimensions that are “known to be consistently related to job performance or other organizational criteria” and that the validity evidence for the Minnesota Multiphasic Personality Inventory (MMPI) is weak. There are hundreds of studies that provide criterion-related validity evidence for these instruments (a list of references is available from the authors). Scales from these inventories have demonstrated utility in selection settings.

Finally, an additional frequent misconception encountered is that clinical personality measures are not personality inventories “per se” (see Murphy & Dzieweczynski, this issue). Clinical and abnormal personality measures assess the same traits as normal personality measures. Markon et al. (2005) pointed out the following:

In personality and clinical psychology, there is now broad consensus that normal and abnormal personality variation can be treated within a single, unified structural framework […] that normal and abnormal personality are strongly related at the etiologic level …, and that abnormal personality can be modeled as extremes of normal personality variation. (p. 139)

In IWO psychology applications, clinical scales such as those found on the MMPI have demonstrated validities on par with those obtained for the Big Five or their facets (e.g., Cullen & Ones, 2001; Ones, Viswesvaran, Cullen, Dilchert, & Drees, 2003).

**POWER OF PERSONALITY IN PREDICTION AND EXPLANATION**

Personality variables have been criticized for having low levels of criterion-related validity and therefore limited utility in occupational settings. Such arguments either (a) inappropriately average validities across theoretically and etiologically
different personality traits, in effect treating personality as a unitary construct; or (b) presume that all the predictive variance in the personality domain is limited to the variance associated with the Big Five dimensions, or both. In the following section we aim to raise the awareness among IWO psychologists about the magnitude of effects associated with personality variables, especially those that have greater breadth than any of the Big Five.

Compound Personality Variables Have Substantial Validities for Predicting Job Performance

Not all personality traits are created equal in terms of their predictive and explanatory value. Highest validities for predicting overall job performance using predictors from the personality domain are found for compound personality variables. Table 1 lists the operational validities for compound personality variables based on meta-analytic investigations. The table is organized to report the compound trait name, the associated multiple Big Five personality constructs assessed by its scales, the total sample size reported in the meta-analysis, and the operational validities for predicting supervisory ratings of job performance. The data reported in Table 1 are drawn from Ones, Viswesvaran, and Schmidt (1993) for integrity tests; from Ones and Viswesvaran (2001a, 2001b) for customer service, violence, drug and alcohol, and stress tolerance scales; and Ones, Hough, and Viswesvaran (1998) for managerial potential scales. Note that all these compound trait scales are squarely personality measures, as 70% to 100% of the variance in them are

<table>
<thead>
<tr>
<th>Type of Scale</th>
<th>Personality Traits Assessed</th>
<th>Source</th>
<th>N</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer service orientation</td>
<td>A, ES, C</td>
<td>Ones and Viswesvaran (2001b)</td>
<td>6,944</td>
<td>.39</td>
</tr>
<tr>
<td>Violence and aggression</td>
<td>C, A, ES</td>
<td>Ones and Viswesvaran (2001a)</td>
<td>4,003</td>
<td>.41</td>
</tr>
<tr>
<td>Stress tolerance</td>
<td>ES, A, C</td>
<td>Ones and Viswesvaran (2001b)</td>
<td>1,010</td>
<td>.41</td>
</tr>
<tr>
<td>Drug and alcohol</td>
<td>C, ES, A</td>
<td>Ones and Viswesvaran (2001b)</td>
<td>1,436</td>
<td>.19</td>
</tr>
<tr>
<td>Managerial potential</td>
<td>ES, EX, C</td>
<td>Ones, Hough, and Viswesvaran (1998)</td>
<td>11,009</td>
<td>.42</td>
</tr>
</tbody>
</table>
accounted for by three of the Big Five dimensions (e.g., see Ones, 1993, for integrity tests).

The operational validities in Table 1 are sizable. Note that none of these operational validities have been corrected for unreliability in the predictor, as the interest is in how these compound personality scales perform in applied personnel selection, and not how the constructs they assess predict overall job performance.

There are several important conclusions that emerge from Table 1. First, the operational validities associated with broad, compound personality variables are substantial, particularly higher than those reported for any of the Big Five for the same criterion (overall job performance). Validities for most compound personality traits are in the .35 to .40 range when predicting overall job performance. Second, the criterion-related validities of the compound personality variables presented in Table 1 are among the highest for individual differences traits predicting overall job performance. In fact, only general mental ability has superior criterion-related validities for the same criterion. Multiple meta-analyses, incorporating data from thousands of individuals, have established that compound personality variables have substantial validity in the prediction of overall job performance.

The Big Five Personality Variables Have Sizable Operational Validities for Predicting Job Performance and Other Criteria at Work

Murphy and Dziewczynski (this issue) argue that “the validity of measures of broad personality traits is still low.” First, there is the issue of what is meant by the validity of a measure. The natural question to ask is “validity for what?” Validities of personality traits, broad and narrow, depend on the particular criteria against which they have been validated (e.g., overall job performance, counterproductive work behaviors, training performance). In this case, we presume that Murphy and Dziewczynski are referring to overall job performance. Then, a second question should be asked: what is meant by broad personality traits? As we pointed out earlier, compound traits are broad in their coverage of the Big Five. But Murphy and Dziewczynski are referring to the validity of the Big Five when they express their concerns. Perhaps then, it would be informative to also describe how the Big Five dimensions of personality as a distinct set of multiple dimensions predict various criteria.

We recently computed multiple correlations for the Big Five as a set with the following criteria: overall job performance, objectively assessed performance, citizenship performance, teamwork, training performance, leadership, performance motivation, job satisfaction, “getting along,” and “getting ahead.” We obtained all validities and predictor intercorrelations from meta-analytic investigations. The operational multiple Rs (not corrected for predictor unreliability) were remarkable. Operational validities for overall job performance and its various facets were
in the .30s, but are as high as .47 for teamwork, .44 for training performance, and .43 for citizenship performance. Even for objectively measured performance, the validity of the Big Five as a set is a respectable .28. Validities for organizational behavior criteria (e.g., job satisfaction, leadership) are in the .40 to .50 range. A detailed table reporting these multiple correlations may be obtained by writing to the authors of this article. In sum, then, validities for the broad Big Five personality dimensions as a set are substantial.

There Are Solid Theoretical Explanations of How and Why Personality Variables Come to Predict Job Performance and Other Behaviors at Work

Several theoretical accounts of how specific personality variables come to predict behaviors and outcomes of interest have been put forth and empirically tested. For example, Barrick, Mount, and Strauss (1993) hypothesized and found empirical support for the idea that sales representatives high in Conscientiousness would set goals and be more committed to them than their less conscientious counterparts (resulting in greater sales volume and higher supervisory ratings of job performance). Broader theories of personality traits in IWO psychology have been offered by Judge and colleagues for core self-evaluations (Judge & Bono, 2001; Judge, Bono, Erez, Locke, & Thoresen, 2002; Judge, Van Vianen, & De Pater, 2004), by Ones and her colleagues (Hogan & Ones, 1997; Ones & Viswesvaran, 1996), along with Roberts and colleagues (Roberts & Bogg, 2004; Roberts, Bogg, Walton, Chernyshenko, & Stark, 2004; Roberts et al., 2005) for Conscientiousness, and Ones (1993), Hogan and Hogan (1989), and Gough (1948), for integrity. Others have laid out careful theories of work behaviors such as organizational citizenship behaviors (Konovsky & Organ, 1996; Organ, 1994; Organ & Lingl, 1995), contextual performance (Borman & Motowidlo, 1997; Hogan, Rybicki, Motowidlo, & Borman, 1998; Motowidlo, Borman, & Schmit, 1997), counterproductive work behaviors (Cullen & Sackett, 2003; Ones & Viswesvaran, 2003; Sackett & DeVore, 2001), motivation to perform (Judge & Ilies, 2002), and so forth, that give prominent roles to various personality variables.

THE ROLE OF META-ANALYSIS

The resurgence of interest in personality variables in IWO psychology has been fueled by applications of meta-analysis (Hough & Ones, 2001). Meta-analysis has enabled the detection of broad patterns of relations between personality variables and organizational behaviors. All data come encrypted with errors (both measurement and sampling) and need to be deciphered (Schmidt, 1992). The application of these meta-analytic methods (e.g., Barrick & Mount, 1991; Ones et al., 1993) have
clarified fundamental relations between personality variables and organizational behavior and this clarification in turn has spurred the development of theories of organizational behavior. However, in their comments, Hogan (this issue) along with Murphy and Dzieweczynski (this issue) raise some concerns about the role of meta-analysis in this area. We address two of them here.

Inventory-Specific Meta-Analyses Are not Sufficient to Build a Science Around Constructs

Some argue that meta-analyses in the personality domain should not include data from multiple inventories, but should focus only on specific inventories. There are two versions of this argument. One argument posits that when data from well-developed inventories are combined with poorly developed measures, the result will be downwardly biased estimates of operational validity. Note that this version of the argument does not argue for inventory specific analysis; it merely calls for excluding studies conducted with poor measures. There are two reasons for not doing so. First, the decision of which measures are poor is subjective. No single indicator provides perfect construct validity. Second, if a reason for this decision can be articulated, then it can be empirically tested. For example, if a hypothesis states that personality inventories developed based on the Big Five framework are more construct valid and are better measures than inventories based on other models, the meta-analyst can actually test this hypothesis by subgrouping analyses based on studies done with measures of Big Five and studies using other measures. Even if there are unspecified differences, the meta-analyst will be alerted to the presence of these differences by substantial residual variance, after corrections for statistical artifacts have been applied. Meta-analysis does not merely provide a mean estimate for the effect under investigation, but also an examination of its true variability across measures, jobs, settings, and the like. As Murphy and Dzieweczynski (this issue) correctly note, “variance in the validity of personality measures is small,” even when investigated in meta-analyses of multiple inventories.

Hogan (this issue), however, endorses a stronger version of the argument for inventory-specific meta-analysis. According to his example, personality scale content differs by personality inventory even if scales often appear to assess the same trait. He argues that “Agreeableness” in the revised NEO Personality Inventory (NEO PI-R) describes individuals who will try not to give offense, whereas in the Hogan Personality Inventory (HPI), “Likeability” refers to people who are actively charming. He argues that meta-analyses that combine results from these two scales are therefore meaningless.

Murphy and Dzieweczynski (this issue) raise a related point albeit in a different context. They note that due to the positive manifold in the cognitive ability domain, the exact items used to assess cognitive ability do not matter. And actually,
we see the solution to Hogan’s (this issue) problem here. Because multiple indicators of Agreeableness (such as the NEO PI-R Agreeableness and the HPI Likeability scales) are positively correlated and this shared variance predicts important outcomes, combining across inventories is essential and necessary for accurate estimates of relations between constructs and outcomes.

Science progresses by building theories around constructs, not measures. Knowledge is built by triangulation across measures, not by focusing on measure-specific variance. The variance in any measure can be partitioned into variance due to the underlying, abstract construct (in this case Agreeableness), variance due to the specific measure under investigation (not giving offense or being actively charming), and error variance (both random and systematic). The basis of a latent construct is the variance shared across its measures. Incidentally, this is the logic behind modern latent variable models (Maruyama, 1998). If a researcher wants to test the hypothesis that what is unique to the Likeability scale of the HPI (a personality inventory that we hold in very high esteem) is an important variable, he or she should collect data to show that specific variance associated with the measure, after controlling for variance it shares with the general factor of Agreeableness, predicts important organizational behaviors and outcomes. We should not limit our science to inventory specific investigations.

Bidirectionality Should not Be a Problem in Personality Meta-Analyses

Murphy and Dzieweczynski (this issue) write that “bidirectionality is problematic for methodologies such as traditional meta-analysis, where the personality–performance relation will be underestimated if the sign of the correlation between personality and performance varies from study to study.” Hogan (this issue) correctly points out that validities estimated in some meta-analyses are underestimates because they do not take into account the possibility that a personality variable (e.g., Conscientiousness) could be positively correlated with one outcome but negatively correlated with another outcome. By averaging them, one would arrive at a small number. The conclusion should be obvious: If there is a hypothesis that a particular personality variable will be positively correlated with one outcome but negatively correlated with another outcome, the two outcomes should not be combined in the same analysis. The hypothesis should be tested empirically by subgrouping analyses by criteria. It goes without saying that this simple fact applies to meta-analyses of all predictor domains, not just personality.

Meta-analytically derived validity estimates can be distorted if researchers combine measures (both on the predictor and on the criterion side) that do not assess the same construct. In those instances, the poor quality of a meta-analysis is the result of bad judgment calls on the part of the researcher. In other instances,
poor meta-analyses are the result of a lack of understanding of meta-analytic techniques, or psychometric corrections in general.

Bad judgment calls can be avoided by providing researchers with a proper taxonomy of personality variables that will provide insight into the differences between Big Five factors, facets, compound traits, and even profiles. Such sound taxonomies are essential in making informed decisions on when it makes sense to combine scales from different personality inventories to contribute to the same analysis. One attempt to provide guidance in this way was undertaken by Hough and Ones (2001). Similar taxonomies are provided on the criterion side, providing rational and empirically derived clusters of criteria that can be meaningfully combined as measures of the same underlying construct (e.g., see the work by J. P. Campbell and colleagues along with Viswesvaran and colleagues on taxonomies of job performance across different jobs).

Interestingly, the proponents of bidirectionality have argued that even when specific hypotheses cannot be postulated, negative validities can indicate bidirectionality and statistical models should be used to make corrections. The merits of these models have been debated elsewhere (e.g., Ones, Mount, Barrick, & Hunter, 1994), but suffice to say that across several different meta-analyses there is sufficient convergence in estimates of the validities of personality traits for predicting organizational outcomes.

CAVEATS AND UNANSWERED QUESTIONS

Personality variables have substantial validity and utility for prediction and explanation of behavior in organizational settings. There are theories that explain how and why personality variables come to influence the behaviors and outcomes they do. Clearly, individual differences in personality are relevant and important. However, there are a number of caveats we should highlight. For example, the reality is that cognitive ability is the stronger predictor of overall job performance but that personality also plays an important role in explaining behavior. Some (e.g., Borman & Motowidlo, 1997) have argued that personality predicts contextual performance better than cognitive ability, whereas cognitive ability predicts task performance better than personality variables. Even if this is true, no organization is likely to forego hiring on task performance (Gottfredson, 2002); also, equal validity does not mean that the same kinds of individuals would be hired using different predictors (Kehoe, 2002). Furthermore, in a large-scale meta-analysis, Alonso, Viswesvaran, and Sanchez (2001) found that (a) cognitive ability predicts task performance better than personality, and (b) cognitive ability in fact also predicts contextual performance better than personality variables.

Hogan (this issue) comments that there are no group differences in personality test scores and as such their use is advantageous to organizations concerned with
workforce diversity. Cognitive ability tests display large ethnic group mean-score differences that can result in adverse impact for the low scoring group when organizations choose to be selective. However, group differences and adverse impact do not tell the whole story: cognitive ability tests do not overestimate or underestimate performance based on group membership. That is, ample research exists to show that cognitive ability tests are not predictively biased. Empirical tests of predictive bias with personality tests are rare. The only study we are aware of was conducted in a military setting (Saad & Sackett, 2002) and, as such, its generalizability to civilian settings is an open question.

CONCLUSIONS

In personnel decisions, personality variables have substantial criterion-related validity and therefore, utility. Were we to abandon the use of existing personality measures, the question of alternatives remains (see also Hogan, this issue). Perhaps it is worthwhile to point out that many of IWO psychology’s mainstay tools such as assessment centers, structured and unstructured interviews, and biodata instruments, to name but a few, all partially assess personality (Collins et al., 2003; Huffcutt, Conway, Roth, & Stone, 2001; Mount, Witt, & Barrick, 2000; Salgado & Moscoso, 2002). The operational validities for compound personality variables are in the .40s. Such a level of validity places these personality constructs among our best predictors and best explanatory variables in IWO psychology. Similarly, the Big Five as a set produce operational validities in the .30s and .40s for a range of important criteria. These conclusions are supported by more than two dozen meta-analyses, incorporating thousands of individual studies, spanning decades of research. They explain and predict human behavior in general, and employee behavior in particular. Any assertion to the contrary demands extensive and comparable evidence that critics have so far been unable to provide!

REFERENCES


