

## Chapter 4: Personality

Robert Hogan  
Hogan Assessment Systems  
Robert B. Kaiser  
Kaplan DeVries Inc.

This chapter provides an overview of personality psychology as it applies to business practices. It is organized in five sections. The first concerns the definition of personality, a topic that is often discussed but seldom defined; the lack of agreed upon definitions is responsible for considerable unnecessary confusion. The second section concerns the kinds of personality assessments available to be used by business. The third section evaluates the pragmatics of personality assessment. The fourth section reviews the inevitable criticisms of personality assessment, and the final section speculates on future directions for personality assessment.

### **1. Defining Personality.**

Personality psychology consists of three related activities. The first involves efforts to conceptualize human nature—how people are alike. These discussions primarily concern motivation—a simple example would be the debate about Theory X (employees are lazy and need direction and discipline) versus Theory Y (employees are motivated by needs for fulfillment which are often frustrated by bad management). To the degree that management practices take motivation into account, these discussions are enormously important. Maslow's (1954) theory of motivation is the core of Theory Y and has been vastly popular over the years. Recent work in evolutionary psychology suggests that biologically grounded motives such as needs for social acceptance, status,

and a sense of meaning and order are important in the workplace; these survival-relevant motives are easy to measure and demonstrably valid (cf. Hogan, 2007).

The second major activity in personality psychology involves identifying the most consequential ways in which people differ from one another and then developing measures of those differences. Scores on these measures can then be used to make decisions about, for example, employee selection and promotion. There are thousands of published personality measures available, each of which concern dimensions that interest a test author (e.g., androgyny) but may not be relevant to the world of work. This bewildering variety of personality measures also creates much of the confusion surrounding the use of personality to predict occupational performance.

The third major activity concerns determining how individual differences in personality develop (cf. McAdams & Olson, in press). Results from this line of research have important implications for career coaching and guidance; they provide suggestions regarding how to assist and encourage others in ways that are informed by data as opposed to helping them in an ad hoc manner.

As for the definition of the term “personality” itself, we prefer to draw on MacKinnon’s important (1944) observation that personality should be defined in two ways. On the one hand, personality refers to “factors” inside people that explain their behavior. Different writers prefer different internal factors—egos, traits, temperaments, schemas, etc.—and we call this personality from the perspective of the actor. On the other hand, personality refers to the distinctive impressions that people make on others, which impressions are captured in trait words (friendly, energetic, etc.)—we call this personality from the perspective of the observer. We find it useful to summarize the first

definition of personality in terms of “identity” (i.e., how people think about themselves) and the second definition in terms of “reputation” (i.e., how others think about them). Reputation tells us how people typically behave; identity tells us why they behave that way. Much confusion in the literature results from not keeping these two definitions distinct.

To the degree that success in business involves getting the people issues right, personality psychology is an indispensable tool. The early giants of I/O psychology—Mark May, Ross Stagner, and Edwin Ghiselli—made exactly this argument. Somewhere along the way, doubts arose.

## **2. Business Applications of Personality Assessment.**

A. What kinds of measures are available? In business, personality is assessed in one of four standard ways. By far the most common and most problematical method of assessment is the interview. In the typical case, the interviewer asks the interviewee questions, and then “scores” the answers using a private and intuitive coding method. From time to time, organizations try to standardize interviews by requiring interviewers to ask each interviewee the same questions and “score” their responses in a standard manner. The problem is that interviewers quickly become bored with standardized lists of questions, and what happens next is called “interview creep”—the standardization steadily melts away. Despite the subjective and unsystematic nature of interviews, it is a fact of nature that organizations refuse to give them up; virtually every hiring process will include some sort of interview, and the interview results often override the results of more standardized assessment processes.

A second and less frequently used assessment method involves questionnaires containing statements to which people respond; the responses to the statements are aggregated and used to form scores on the dimensions covered by the questionnaire (e.g., extraversion). Some of these questionnaires, like the well-known California Psychological Inventory and the Myers-Briggs Type Indicator, are commercially published, professionally reviewed, internet enabled, and available in multiple languages. Other personality measures are home grown; there are thousands of ad hoc personality questionnaires available, each is scored in terms of idiosyncratic dimensions, most are of questionable quality, and it is impossible to review them here. Potential users of any personality measure should demand to see the technical manual supporting the use of the test. Tests should be compared and evaluated in terms of the validity data contained in the technical manuals.

Beginning in the early 1990s, substantial consensus emerged among researchers regarding the appropriate structure (dimensionality) for personality questionnaires. This agreed-upon structure is called the Five-Factor Model (FFM; Wiggins, 1996); it suggests that all existing inventories of normal personality measure the same five broad dimensions with more or less efficiency. Table 1 presents the Five-Factor Model.

The third way in which business organizations measure personality is with observer rating forms; observers are asked to rate or describe target persons using a defined set of rating categories. The best known of these procedures are the so-called 360 degree evaluations, where target persons are rated by superiors, peers, and subordinates using a standardized rating form. Virtually every consulting firm in the developed world has a proprietary 360 rating scheme. Despite their popularity, there is

no agreed upon structure or content to these rating forms, and it is impossible to review them here—however, see Bartram (2005) for suggestions about how to standardize them.

The fourth common way in which personality is measured is by means of performance on unstructured tasks. The key assumption for this kind of assessment is that, because the tasks are unstructured, the manner in which people respond to them will indicate how they typically respond in ambiguous situations, and in this way, they will reveal their “personalities” and performance capabilities. Projective tests such as the Rorschach and the Thematic Apperception Test (TAT) are two well-known and widely used measures of this type, of which there are many.

B. For What Jobs is Personality Assessment Most Valuable? Every job in every organization has a distinctive set of psychological demands. For example, long distance truck drivers need to work alone, remain vigilant over long periods of time, follow rules closely, and be impervious to boredom; in contrast, bomb disposal technicians must enjoy taking risks but be willing to follow standardized procedures exactly. In principle, personality assessment is valuable for selecting people into every job in every organization (cf. J. Hogan & Holland, 2003).

In reality, it is not technically feasible to use personality to select people into every job in every organization. This is because every organization considers itself to be unique, and every organization has many specialized jobs with unique titles. To develop personality-based selection procedures for all these jobs, and to do so correctly, would require more time and money than most organizations are willing to expend. It is necessary, therefore, to simplify the problem.

Our approach to simplifying this problem begins with the perceptive and heuristic Holland model (Holland, 1973). Holland proposes that all occupations can be sorted into one of six ideal types (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional) and that each occupational type is also a distinctive personality type—there is a characteristic personality “profile” associated with each type. Gottfredson and Holland (1996) showed that the Holland model can classify every job in the Dictionary of Occupational Titles. This then provides links between personality and most major jobs in the real world.

Hogan and Hogan (2007) build on this important finding. Their test, the Hogan Personality Inventory (HPI; Hogan & Hogan, 1995), is a measure of normal personality based on the FFM and designed to predict occupational performance. They have data comparing scores on the HPI with performance for every major job category in the U.S. economy. Using the principles of validity generalization (meta-analysis, synthetic validity, and job component validity; SIOP, 2003), and the Department of Labor Standard Occupational Classification System (US DoL, 1991), they provide estimates for the validity of the FFM dimensions of personality for predicting performance across all major job categories.

In principle, then, personality measures can be used to forecast performance in all jobs in every organization. In practice, however, entry level employees are typically screened using short and inexpensive measures of integrity/reliability, resilience/stress tolerance, and service orientation. Applicants for more senior positions in sales and management are screened with more extensive batteries.

C. How Are these Measures Developed? The best-known personality inventories (CPI, 16 PF, HPI) were developed following accepted professional standards and guidelines (SIOP, 2003). But the scoring systems for interview questions, observer rating forms (including 360 degree appraisals), and the various projective tests are typically ad hoc, and the measures are often used with little concern for reliability and validity.

D. For What Purposes are these Measures Used? In the workplace, personality assessment is used for three purposes. First, assessment is frequently used to provide individuals with feedback that is intended to increase their self-awareness and improve team performance. The Myers-Briggs Type Indicator in particular is widely used for these purposes; it tells people how they tend to perceive the world, as compared with how others see the world, and how these perceptual differences can lead to unintended misunderstandings and conflict. The so-called 360 feedback process is also used to enhance self-awareness. A target person is described by subordinates, peers, and superiors using a standardized rating form. The results are summarized and reported back to the target person, a process that is as often disheartening as it is enlightening—because most people tend to think more highly of themselves than others actually do (Heidemeier & Moser, 2009).

Second, and related to the preceding point, assessment results are often used for career guidance. Suppose a person wants a career in the entertainment industry, perhaps as a script writer. Script writers need to be verbally fluent, hard working, persistent, imaginative, and able to meet deadlines. The person will be assessed for these characteristics and his/her fit with the profile can be evaluated in terms of potential strengths and developmental needs. Management and leadership development programs

rely heavily on assessment for career guidance. The relevant profile for effective performance in managerial or leadership roles is reasonably well understood; in addition, a vast armada of leadership assessment material is commercially available. Leadership development is a large consulting industry with world wide demand.

Finally, personality assessment is used for personnel selection. As noted above, we can profile the psychological requirements of every job in the U.S. economy, we can match the assessment profile of an applicant to the required profile of any job and determine the likelihood of the person succeeding in the job. Employment interviews are an informal way of doing personality-based personnel selection; assessment centers and standardized psychometric batteries are a more formal way of doing it.

### **3. How Well Does Personality Assessment Work?**

A. The Measurement Problem. The question of how well personality assessment works is, of course, the bottom line issue for this chapter. A naïve reader might think this is a straightforward empirical question, but in academic life few important questions seem capable of being answered in a straightforward manner. Especially on this question, opinion is bitterly divided. Some writers (like the authors) are enthusiastic advocates of personality assessment. Others doubt that personality assessment works and vigorously criticize its use for selection purposes (cf. Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). As always there are conceptual confusions clouding these discussions. Persons new to this field will need three pieces of background information in order to understand the debate.

First, the accepted method for determining the validity of psychometric tests is a statistical procedure called meta-analysis, and the preferred version of meta-analysis is

called validity generalization (Hunter & Schmidt, 2004). The correlation between test scores and performance criteria in any single research study is regarded as one data point, and that data point is assumed to be “contaminated” by a variety of statistical artifacts. The process of meta-analysis collects as many studies as are available on the relationship between a type of test (or measurement dimension) and performance criteria. The results of all the studies are combined and then corrected for statistical artifacts, such as the degree of measurement error in the test and in the outcome measure. These “corrected” results are assumed to provide the best possible estimate of the “true” relationship between a type of test (or measurement dimension) and a class of outcomes (or criteria). To summarize, meta-analysis is used to answer the question of how well a test works in predicting occupational performance.

Second, meta-analysis is well suited for estimating the validity of measures of cognitive ability, usually referred to as General Mental Ability or GMA (cf. Schmidt, Shaffer, & Oh, 2008). The reason meta-analysis works well for measures of GMA is that the various measures are highly inter-correlated, which means that collectively they all represent the same large general factor. Because measures of GMA are so statistically similar, one can compare the results from different studies with little regard for the particular measure used by the researchers.

In the case of personality, however, it is essentially impossible to combine measures across studies. The number of scales or dimensions on the best known personality inventories varies widely, from three on the Eysenck Personality Inventory to 21 on the California Psychological Inventory. Furthermore, scales with the same names on different inventories (e.g., Agreeableness) are not highly correlated—unlike the

different measures of GMA—and they predict different outcomes differently. This fact makes it virtually impossible to combine studies across different personality inventories; consequently, standard meta-analytic studies of the validity of personality for predicting occupational performance underestimate that validity (e.g., Barrick & Mount, 1991). The solution to this problem is to do the meta-analyses using one inventory at a time (cf. J. Hogan & Holland, 2003).

The third problem concerns how to define job performance—more specifically, how to assign numbers to individual differences in performance—in a way that is comparable across organizations. Job performance is typically defined in one of three ways. In the first case it is defined in terms of performance in training (Schmidt, et al. 2008). When performance is defined this way, GMA always outperforms personality in terms of validity—because GMA is the best single predictor of training performance. In the second case, performance is defined in terms of supervisors' ratings of overall job performance. In our view, this is problematic because overall ratings of performance tend to be political judgments—they reflect how much a supervisor likes an employee rather than how well he/she is performing. In the third and ideal case, performance is defined in terms of the relevant components of the job (e.g., showing effort, maintaining personal discipline, facilitating team performance, etc.) which can be identified statistically (cf. Campbell, McCloy, Opler, & Sager, 1993). Once the relevant dimensions of performance have been defined and measured, the predictors should be aligned with these dimensions in ways that make conceptual sense. For example, one would not use a measure of GMA to predict ethical behavior because many smart people (e.g., financial managers) have problems with integrity. However, a measure of

Conscientiousness would be expected to predict ethical behavior because following rules is a key element of Conscientiousness.

When the performance dimensions have been defined explicitly and the predictors have been aligned with the performance dimensions, it becomes possible to evaluate how well personality assessment works. As always, these judgments are relative. Consider Table 2, which presents meta-analytic results for the validity of the seven most commonly used measures of occupational performance. These validity coefficients vary between .11 and .28, and these values should be considered when evaluating the validity of personality measures.

Consider next the finding by Judge, Colbert, and Ilies (2004) that the fully corrected, meta-analytically derived validity for predicting leadership with measures of GMA is .27. These data suggest that validity coefficients above .30 are unusual for any single predictor of job performance, leadership, or other occupational criteria. Our point is that, when evaluating the validity of personality measures for predicting occupational performance, one should bear in mind the size of the validity coefficients that can normally be expected for any class of measurement variables, and that number is somewhere between .11 and .30.

As a second example of how well personality measures work, consider Table 3, which presents the results of a careful meta-analytic study of the links between personality (defined in terms of the Five-Factor Model) and rated leadership (Judge, Bono, Ilies, & Gerhardt, 2002). The correlations are consistent with the best results in Table 2; they show that good leaders seem self confident (Emotional Stability), socially poised (Extraversion), visionary (Openness), trustworthy (Conscientiousness) and not

necessarily charming (Agreeableness). The multiple correlation, using all five dimensions in Table 3, is .48.

In the best meta-analytic study of personality and job performance in the published literature, J. Hogan and Holland (2003) focused on one inventory, and aligned the personality dimensions with the relevant criteria—they did not try to predict training with Adjustment (a measure of FFM Emotional Stability) or sales performance with Prudence (a measure of FFM Conscientiousness). As shown in Table 4, the observed validity coefficients ( $r$ ) for personality compare favorably with those in Table 2 for other predictors of job performance.

These tables show that personality reliably predicts job performance. For those who care about social justice, it is also important to note that, unlike GMA, personality measures are gender and race neutral so that minorities and women get the same scores as white males. Furthermore, personality predicts a wide range of significant life outcomes in addition to job performance. Table 5 provides a summary of these personality-related outcomes, which range from job performance and leadership to counterproductive job behavior, absenteeism, teamwork, job and career satisfaction, health behavior, and even life expectancy. For an extended discussion of the validity issue, see Hogan (2005).

#### **4. Standard Criticisms of Personality Assessment.**

The standard criticisms of personality assessment can be reduced to three fundamental claims: (1) the tests have minimal or trivial validity for predicting real world outcomes; (2) scores on the tests are contaminated by social desirability response bias; and (3) scores on the tests can be altered by deliberate faking (for an energetic presentation of these criticisms, see Morgeson, et al. 2007).

Concerning the claim that personality measures lack validity for predicting occupational performance, readers should consult Section 3 above and form their own opinions.

The second criticism—that scores on personality measures are contaminated by social desirability response bias—has been around for over 50 years; it has been extensively studied and consistently rejected. The criticism is based on the assumption that, when people read items on a personality inventory, they are primarily motivated to respond in a way that presents themselves in the best possible light—that is, respond to the items in ways that are socially desirable rather than in ways that reflect their “true” selves. The claim is that peoples’ natural bias toward socially desirable responding invalidates personality measures.

A fundamental insight of personality psychology is that there are individual differences associated with every generalization we make about people. This is true for the statement that people are motivated to make good impressions on others. Some people are and some are not and these differences can be assessed using measures of social desirability, of which many are available. It then becomes a straightforward task to determine the degree to which socially desirable responding affects the validity of personality assessment. There is a substantial and convergent literature on this point; papers by Ones, Viswesvaran, and Reiss (1996) and Schmitt and Oswald (2006) are among the best. These papers show that socially desirable responding is a reliable dimension of personality related to the FFM dimensions of Emotional Stability and Conscientiousness and correcting personality test scores for socially desirable responding has no affect on their validity.

The third claim, that faking invalidates the use of personality measures for employee selection, is a more general statement of the second claim. In a nutshell, the argument is that, when employees complete a personality measure as part of the pre-employment screening process, they “fake”—they distort their scores so as to improve their chances of being hired—and this tendency invalidates the use of personality measures for hiring purposes. The empirical literature on the faking issue is enormous, complex, tedious, and largely beside the point. The research shows that after completing a personality measure, when instructed, people can alter their scores on the second trial. The research also shows that, on average, job applicants tend to get “better” scores than job incumbents (Ellingson, Sackett, & Connelly, 2007). However, only one study in this literature uses real job applicants in an effort to determine the degree to which faking actually occurs in the selection process.

Hogan, Barrett, and Hogan (2007) tested a large group of applicants for a government job with the HPI. Over 5,000 of these people were denied employment and, six months later, re-applied for the same job and completed the HPI a second time. It is reasonable to assume that these people were motivated to improve their scores on the second occasion. The data indicated that 95% of the scores across 5,000+ people and seven scales remained the same. Of the 5% of scores that changed, 2.5% became worse, and 2.5% became better. The authors conclude that, in real employment settings, faking is not a problem.

## **5. Future Directions**

The development of the Five Factor Model in the 1980s triggered a wave of meta-analytic studies of the links between personality and job performance. These efforts

rejuvenated personality research in applied psychology. However, future development requires moving beyond the FFM, exploring alternative measurement methods, becoming clearer about the agenda for personality measurement, and developing better theory to guide empirical research.

### *Beyond the FFM*

The evidence is clear that personality, framed in terms of the FFM, predicts performance in virtually every job from entry level to leadership positions, as well as job satisfaction and career success. But most people understand that there is more to personality than the FFM. Two recent developments suggest ways to expand the personality domain.

Judge and colleagues propose a higher-order personality dimension of core self-evaluations (Judge, Locke, & Durham, 1997), defined as a broad judgment that people make about their basic worth and ability to influence events. Judge argues that this higher-order dimension is composed of four narrower themes—self-esteem, generalized self-efficacy, locus of control, and emotional stability. Research shows that these four themes do indeed form a single higher-order factor. Moreover, this higher-order factor, core self-evaluations, is a better predictor of job performance and job satisfaction than any of the four component variables. Most importantly, core self-evaluations add incremental validity over the FFM, indicating that they reflect unique work-related variance (Judge, Erez, Bono, & Thoresen, 2003). This research is important because it demonstrates that, by “thinking big”, we can identify broad, integrative dispositions that enhance our ability to predict important outcomes at work.

Recent research has focused on predicting ineffectiveness, counter-productivity (cf. Rotundo & Sackett, 2002) and managerial incompetence (Hogan, Hogan, & Kaiser, in press). This has led to an interest in dysfunctional aspects of personality, based on the notion that undesirable tendencies such as arrogance, passive-aggression, eccentricity, paranoia, and perfectionism are functionally distinct from the FFM. There are three published inventories of dysfunctional personality (Hogan & Hogan, 2008; Moscosco & Salgado, 2004; Schmit, Kilm, & Robie, 2000) designed to assess extreme variations in normal personality that are appropriate for use in the workplace. This is in contrast with measures of psychopathology such as the Minnesota Multiphasic Personality Inventory which are inappropriate for selection purposes.

This new research on dysfunctional dispositions is promising for two reasons. First, it provides insight into such phenomena as employee deviance, career failure, and bad leadership. Second, although the dimensions of dysfunctional personality are related to the FFM, they predict different outcomes and add incremental validity over the FFM in predicting workplace criteria (Hogan & Hogan, 2008; Schmit, Kilm, & Robie, 2000). Thus, the links between dysfunctional personality tendencies and occupational performance is a promising topic for future research.

#### *Alternative Measurement Methods*

The questionnaire method of assessing personality is well accepted in organizational research. However, certain trends suggest it is time to consider some alternatives. On the one hand, persistent concerns about faking have led to a search for alternatives to standard personality questionnaires. On the other hand, advances in understanding the cognitive processes underlying personality make alternative

assessment strategies more feasible. Thus, we believe future methods of personality assessment may be influenced by advanced projective techniques and theory-driven interview techniques.

*Projective techniques.* Projective tests are designed to assess covert belief systems by presenting test-takers with ambiguous stimuli into which they project their assumptions and expectations. These methods (e.g., the Rorschach, the Thematic Apperception Test) are characterized by subjective scoring systems with notoriously low reliabilities (Schneider, 2007). However, two recent methods combine the goal of assessing covert motives and beliefs with the standardized scoring procedures of traditional inventory methods.

James' (1998) *conditional reasoning* approach assumes that peoples' personalities reflect the way people they think. Most people believe that their own behavior is logical and they develop beliefs and biases that justify their actions (e.g., aggressive people assume other people want to harm them). Individuals who think differently also behave differently, and these differences can be captured with conditional reasoning problems.

James and colleagues have developed conditional reasoning measures of achievement motivation and aggression. These tests are standardized, can be administered in groups, and produce reliable scores. Moreover, these scores are essentially uncorrelated with traditional inventory measures of achievement motivation and aggression (Bing, Stewart, Davison, Green, McIntyre, & James, 2007; James, 1998). However, these scores correlate quite well with organizational criteria. These findings have prompted researchers to distinguish between the *explicit* components of personality (consciously accessible self-perceptions that can be assessed with traditional inventories)

and the *implicit* components of personality (covert factors that require indirect assessment). Research indicates that the explicit and implicit components have both unique and interactive effects in predicting such outcomes as persistence, task performance, and counterproductive behavior, with multiple correlations in the .40 to .60 range (e.g., Bing, LeBreton, Davison, Mgetz, & James, 2007; Bing et al, 2007).

Greenwald and Banaji's (1995) *implicit attitude test* (IAT) is another new method for measuring subconscious thoughts. Although originally designed to assess racist and sexist attitudes, the method has recently been used to measure personality (e.g., Steffens & Konig, 2006). IATs are administered via computer, require respondents to indicate whether two paired words or phrases are similar or different, and then use reaction time to measure the strength of association. For instance, respondents who quickly associate "me" with "industrious" are thought to be more conscientious than those who take longer to make the association.

Preliminary evidence suggests that IAT-based measures of the FFM are unrelated to traditional inventory-based measures of the FFM, and may add incremental validity in predicting behavioral outcomes (Steffens & Konig, 2006). However, this new area of research has not been extended to the workplace. Furthermore, IATs require a large number of trials to produce reliable scores, which ultimately may make this approach impractical for assessing the full range of personality dimensions (i.e., the FFM) in organizational applications.

*Interviews.* The employment interview is most popular method of personality assessment and personnel selection (Huffcutt, Conway, Roth, & Stone, 2001). However, most interviews are unstructured and assess the characteristics needed for job

performance in a haphazard manner. Nonetheless, there is evidence that observer ratings add incremental validity beyond traditional personality inventories in predicting performance (Mount, Barrick, & Strauss, 1994), which suggests that interview-based assessments are potentially valuable. Recent research indicates how interviews might be better designed to assess personality.

Binning, LeBreton, and Adorno (1999) note that employment interviews are high-fidelity settings for personality assessment because, like much of job performance, they occur in contexts involving spontaneous social interaction. But the potential of this method depends on interviews being carefully designed to measure job-related personality characteristics. This requires interviews to be designed around a personality-based job analysis and focused on certain personality dimensions but also to allow interviewer discretion to probe responses. For example, interviewers may need to probe for counterproductive personality tendencies (Blackman & Funder, 2002). Huffcutt, et al. (2001) provide evidence that systematic but flexible interviews designed to evaluate job-related personality attributes can yield criterion-related validities that compare favorably with inventory-based measures.

Employment interviews can also be used to verify inventory-based personality data from earlier phases of the selection process. Binning et al. (1999) suggested integrating these two types of assessment to increase predictive validity. In a subsequent selection study in a call center, they showed that combining personality scores with ratings from an interview designed to probe the inventory results added incremental validity in predicting turnover, and reduced turnover rates by more than 50% (Binning, Adorno, & LeBreton 2008).

*Remembering the goals of assessment*

When psychological assessment was invented in the late 19<sup>th</sup> century, the pioneers (e.g., Alfred Binet) were quite clear that the tests were intended to predict useful outcomes—in Binet’s case, academic performance. But Charles Spearman’s early research changed the focus of measurement in a way that many people found appealing but which, in our judgment, has proved to be a dead end. Spearman shifted the emphasis from predicting outcomes to measuring entities. And with the emphasis on measuring entities (e.g., “g”, the general factor in GMA, traits), the concern with predicting outcomes gradually faded away. By the 1950s, Binet’s pragmatic approach to assessment was seen as an example of “dustbowl empiricism”—a term of derision applied to empirically keyed instruments (the MMPI, CPI, etc.) whose focus was prediction.

We believe that future progress in personality assessment depends on improving the ability to predict organizational outcomes—productivity, turnover, theft, and so on. Currently researchers seem concerned with measuring personality constructs for their own sake. For example, the research on using interviews to assess personality primarily concerns how well interview ratings converge with other measures of personality (cf. Barrick, Patton, & Haugland, 2000), and not how this research enhances validity. Similarly, most research on the FFM concerns interpreting, replicating, and generalizing the five factor structure. This work has been useful in providing a robust taxonomy for classifying the dimensions of personality, but it has almost nothing to do with predicting outcomes that matter in the workplace.

A related problem has been the emphasis on increasingly sophisticated statistical methodologies that yield no real gain in validity. For example, Cherneyshenko, Stark,

Drasgow, and Roberts (2007) examined the use of item-response theory and ideal-point latent variable theory for personality scale construction. They report that personality items fit ideal-point models of non-linear item-construct relationships better than more traditional and much simpler linear models. However, there were no differences in the two procedures in terms of the ability to predict observed outcomes. Despite the difficulty of using the complex ideal-point approach, the authors recommended it as a more flexible and advantageous solution.

Current research uses increasingly complex methodological designs, measurement methods, and data-analytic techniques (Aguinis, Pierce, Bosco, & Muslin; 2009). However, these methods—structural equation modeling, item-response theory, and multilevel and longitudinal latent variable models—have not improved our ability to predict important criteria. The goal of assessment is not to measure entities but to predict outcomes; the former only matters if it enhances the latter.

### *Better Theory*

Clearer thinking, not more sophisticated statistical methods, will improve the prediction of workplace outcomes. Two examples support this view.

*Predictor-criterion alignment.* Critics of personality assessment for employee selection base their criticism on the modest validity coefficients reported in meta-analyses (e.g., Morgeson et al., 2007). For example, Schneider (2007) estimates the overall validity for measures of conscientiousness to be in the .20s (Schneider, 2007). However, meta-analytic research rarely distinguishes between different kinds of job performance criteria. So, for example, the positive correlations between conscientiousness and rule-following are simply added to the negative correlations

between conscientiousness and creativity, which cancel each other and under-estimate the “true” validity of conscientiousness. In addition, the conscientiousness dimension has two components: responsibility and achievement-orientation. These components correlate in opposite directions with criteria such as persistence on a futile task, and when they are combined in a global measure of conscientiousness, they yield a zero correlation (e.g., Moon, 2001). However, when they are correlated separately with the criterion, they yield significant, but opposite, validity coefficients—responsible people persist at problems that can not be solved; achievement oriented people move on to those that can be solved.

Poor criterion measurement is widespread in applied research (cf. Murphy, 2008). When supervisors’ ratings of employee performance reflect how much they like the employees, the correlations between these ratings and the relevant personality dimensions will underestimate the personality-performance relationship. Future research should pay closer attention to the adequacy of performance criteria and use theory to align criterion measures with relevant personality dimensions.

*All personality measures are not created equal.* Meta-analyses assume that personality scales from different inventories with the same name are equivalent and measure the same dimensions. However, this assumption is not justified conceptually or empirically. For example, Gruzca and Goldberg (2007) compared the validity of 11 different personality inventories for predicting three types of behavioral criteria. Overall, the average validities were similar across instruments; however, scales from some inventories predicted certain criteria better than scales with similar names from other inventories. The researchers called for more comparative studies of the ability of different

inventories to predict specific classes of behavior. Such research would be invaluable for organizational practitioners who must choose among overtly similar inventories. An alternative is to conduct meta-analyses using measures from only one inventory (cf. J. Hogan & Holland, 2003).

*Other opportunities.* Few organizational researchers have studied non-linear and interactive effects between two or more personality dimensions. However, a small number of studies of these effects support the contention that these more complex relationships yield incremental predictive validity (e.g., Benson & Campbell, 2007; Witt, Burke, Barrick, & Mount, 2002).

Prediction methods that consider entire personality profiles are also an interesting research topic. Such multivariate approaches are potentially useful because single personality dimensions may yield relatively small relationships with a criterion, but when considered together, they often yield sizable multiple correlations. For example, the combined FFM predicts leadership with a multiple  $R$  of .48 (Judge, et al., 2002). This compares quite favorably to the meta-analytic estimate of .27 for the relationship between cognitive ability and leadership (Judge, et al., 2004).

Holistic methods can also put the person back in *personality* by using individuals rather than variables as the unit of analysis. It would be useful to study how *configurations* of personality characteristics predict outcomes because, historically, we have studied how single personality variables are related to criteria, but in practice we make decisions about people, who are constellations of personality characteristics. Unfortunately, we currently know much more about variables than we do persons.

**Last Thoughts.** Persons new to psychology may be surprised to learn that academic psychology is rather hostile to personality psychology. It all started in the late 19<sup>th</sup> century. The original pioneers (e.g., Francis Galton) maintained that personality is largely innate, under genetic control, and very hard to change. This view supported a racially biased, politically conservative philosophy, and generated a furious liberal backlash led by Franz Boas, the father of cultural anthropology. Boas devoted his career to promoting the idea that peoples' behavior should be explained in terms of their culture. American social scientists, responding to Boas' call, argued that social, historical, economic and situational factors explain what people do, which makes their behavior changeable in principle. Behaviorism, the guiding metaphysic of American psychology for perhaps 80 years, is the logical extension of this view—i.e., what people do depends on where they are not who they are (cf. Mischel, 1968). This view is consistent with a liberal political philosophy and the spirit of democracy—and by definition is rooted in ideology.

By attending to two important distinctions, much of the historical antipathy to personality psychology should go away. First, consider the distinction between identity—how people think and talk about themselves—and reputation—how other people think and talk about that person. The study of identity has not been very productive, but the study of reputation has been immensely productive. This is so because the best predictor of future behavior is past behavior, reputation is a summary of a person's past behavior, and by definition it is the best data source we have about that person's future behavioral tendencies. So, if we direct personality research to the study

of reputation, we are on very strong empirical grounds, and can avoid issues of political ideology.

Second, consider the distinction between defining the goals of assessment as: (1) measuring psychological entities (traits, intelligence, etc.); or (2) predicting significant outcomes. We believe that assessment has a job to do, and that job is to predict non-test behavior as observed by colleagues and customer. If so, then it is a relatively simple matter to evaluate the merits of personality measurement—it is far from perfect, but better than any other single alternative. Test users (organizations) are usually happy to be able to forecast stable patterns and trends in behavior (reputation) and not worry about how individuals see themselves (identity). Obviously prudent researchers would combine assessment methods to maximize prediction, but that is not our point here. Our point is that the antipathy toward personality psychology is unfounded based on the available data.

## References

- Aguinis, H., Pierce, C., Bosco, F. A., & Muslin, I. S. (2009) First decade of organizational research methods: Trends in design, measurement, and data-analysis topics. *Organizational Research Methods, 12*, 69-112.
- Arthur, W., Jr., Day, E.A., McNelly, T.L., & Edens, P.S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology, 56*, 125-154.
- Barrick, M. R., & Mount M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology, 44*, 1-26.
- Barrick, M. R., Patton, G. K., & Haugland, S. N. (2000). Accuracy of interview judgments of job applicant personality traits. *Personnel Psychology, 53*, 925–951.
- Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90*, 1185-1203.
- Benson, M. J. & Campbell, J. P. (2007). To be, or not to be, linear: An expanded representation of personality and its relationship to leadership performance. *International Journal of Selection and Assessment, 15*, 232-249.
- Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their correlates. *Journal of Applied Psychology, 92*, 410-424.
- Bing, M. N., Stewart, S. M., Davison, H. K., Green, P. D., McIntyre, M. D., L. R. James, (2007). An integrative typology of personality assessment for aggression: Implications for predicting counterproductive workplace behavior. *Journal of Applied Psychology, 92*, 722-744.
- Bing, M. N., LeBreton, J. L. , Davison, H. K. , Migetz, D. Z. , & James, L. R. (2007). Integrating implicit and explicit social cognitions for enhanced personality assessment: A general framework for choosing measurement and statistical methods. *Organizational Research Methods, 10*, 346-389.
- Binning, J. F., Adorno, A. J., & LeBreton, J. M. (2008, April). *Validity of inventory and interview assessments of person-job affective fit*. Presented at the annual Conference of the Society for Industrial and Organizational Psychology, San Francisco.
- Binning, J. F., LeBreton, J. M., & Adorno, A. J. (1999). Assessing personality. In R. W. Eder & M. M. Harris (Eds.), *The employment interview handbook* (105-123). Thousand Oaks, CA: Sage Publications, Inc.

- Blackman, M.C., & Funder, D.C. (2002). Effective interview practices for accurately assessing counterproductive traits. *International Journal of Selection and Assessment*, *10*, 109–116.
- Bliesener, T. (1996). Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational and Organizational Psychology*, *69*, 107-120.
- Campbell, D. P., McCloy, R. A., Oppler, S. H., & Sagar, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.
- Cherneyshenko, O. S., Stark, S., Drasgow, F., & Roberts, B. W. (2007) Constructing personality scales under the assumption of an Ideal Point response process: Toward increasing the flexibility of personality measures. *Psychological Assessment*, *19*, 1, 88-106.

- Cooper-Hakim, A., & Viswesvaran, C. (2002). A meta-analytic review of the MacAndrew Alcoholism scale. *Education and Psychological Measurement*, 62, 818-829.
- Degler, C. N. (1991). *In search of human nature: The decline and revival of Darwinism in American social thought*. New York: Oxford.
- Ellingson, J.E., Sackett, P.R., & Connelly, B.S. (2007). Personality assessment across selection and development contexts. *Journal of Applied Psychology*, 92, 386-395.
- Ghiselli, E. (1963). Managerial talent. *American Psychologist*, 18, 631-641.
- Gottfredson, G., & Holland, J.L. (1996). *Dictionary of Holland occupational codes* (3<sup>rd</sup> ed.). Odessa, FL: Psychological Assessment Resources.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Grucza, R. A., & Goldberg, L. R. (2007). The comparative validity of 11 modern personality inventories: Predictions of behavioral acts, informant reports, and clinical indicators. *Journal of Personality Assessment*, 89, 167-187.
- Heidemeier, H., & Moser, K. (2009). Self-other agreement in job performance ratings: A meta-analytic test of a process model. *Journal of Applied Psychology*, 94, 353-370.
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92, 1270-1285.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations. *Journal of Applied Psychology*, 88, 100-112.
- Hogan, R. (2007). *Personality and the fate of organizations*. Mahwah, NJ: Erlbaum.
- Hogan, R., & Hogan, J. (1995). *Hogan Personality Inventory manual* (2<sup>nd</sup> ed.). Tulsa, OK: Hogan Assessment Systems.
- Hogan, R., & Hogan, J. (2007). *Validity of the Hogan Personality Inventory for job family selection*. Tulsa, OK: Hogan Assessment Systems.
- Hogan, R. (2005). In defense of personality measurement: New wine for old whiners. *Human Performance*, 18, 331-341.
- Hogan, R., & Hogan, J. (2008). *Hogan Development Survey manual* (2<sup>nd</sup> ed.). Tulsa, OK: Hogan Assessment Systems.

- Hogan, R., & Hogan, J. (2001). Assessing leadership: A view of the dark side. *International Journal of Selection and Assessment*, 9, 40-51.
- Hogan, J., Hogan, R., & Kaiser, R. B. (in press). Management derailment: Personality assessment and mitigation. To appear in S. Zedeck (Ed.) *American Psychological Association Handbook of Industrial and Organizational Psychology*. Washington, DC: American Psychological Association.
- Holland, J. L. (1973). *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice-Hall.
- Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology*, 86, 897–913.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods*, 1, 131-163.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87, 765–780.
- Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership. *Journal of Applied Psychology*, 89, 542-555.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale (CSES): Development of a measure. *Personnel Psychology*, 56, 303-331.
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-Factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, 87, 530-541.
- Judge, T. A., Ilies, R., & Colbert, A. E. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, 89, 542–552.
- Judge, T. A., Locke, E. A., & Durham, C. C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior*, 19, 151-188.
- MacKinnon, D. W. (1944). The structure of personality. In J. McVicker Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 3-48). New York: Ronald Press.

- Maslow, A. (1954). *Motivation and personality*. New York: Harper and Row.
- McAdams, D. P., & Olson, B. D. (in press). Personality development: Continuity and change. *Annual Review of Psychology*.
- Moon, H. (2001). The two faces of Conscientiousness: Duty and achievement striving in escalation of commitment dilemmas. *Journal of Applied Psychology*, 86, 533–540.
- McDaniel, M.A., Whetzel, D.L., Schmidt, F.L., & Maurer, S.D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616.
- McDaniel, M.A., Hartman, N.S., Whetzel, D.L., & Grubb, W.L. (2007). Situational judgment test, response instructions, and validity: A meta-analysis. *Personnel Psychology*, 60, 63-91.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley
- Moon, H. (2001).
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60, 683-729.
- Moscoso, S., & Salgado, J. F. (2004). “Dark side” personality styles as predictors of task, contextual, and job performance. *International Journal of Selection and Assessment*, 12, 356-362.
- Mount, M. K., & Barrick, M.R. (2001). *Personal Characteristics Inventory manual*. Libertyville, IL: Wonderlic.
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the big five personality factors. *Journal of Applied Psychology*, 79, 272-280.
- Murphy, K. R. (2008). Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology*, 1, 148–160
- Ng, T.W., Eby, L.T., Sorensen, K.L., & Feldman, D.C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58, 367-409.
- Ones, D .S., Viswesvaran, C., & Reiss, A.D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660-679.

- Ones, D. S., Viswesvaran, C., & Schmidt, F.L. (1993). Comprehensive meta-analysis of integrity test validation. *Journal of Applied Psychology, 78*, 679-703.
- Ones, D. S., Viswesvaran, C., & Schmidt, F.L. (2003). Personality and absenteeism. *European Journal of Personality, 17* (Suppl.), 519-538.
- Pearlman, K., Schmidt, F. L., & Hunter, J.E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology, 65*, 373-406.
- Peeters, H., Van Tuijl, H.F.J., Rutte, C.G., & Raymen, J.M. (2006). Personality and team performance. *European Journal of Personality, 20*, 377-396.
- Personality tests poor predictors of job performance (2007, Dec.). Article retrieved June 3, 2009 from <http://www.management-issues.com/2007/12/13/research/personality-tests-poor-predictors-of-job-performance.asp>
- Roberts, B. R., Kuncel, N.R., Shiner, R., Caspi, A., & Goldberg, L.R. (2007). The power of personality: The comparative validity of personality traits, socio-economic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313-345.
- Rotundo, M. & Sackett, P.R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy capturing approach. *Journal of Applied Psychology, 87*, 66-80.
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology, 82*, 30-43.
- Schmidt, F. L., Shaffer, J. A., & Oh, I-S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology, 61*, 827-868.
- Schmit, M. J., Kilm, J. A., & Robie, C. A. (2000). Development of a global measure of personality. *Personnel Psychology, 53*, 153-193.
- Schmitt, N., & Oswald, F.L. (2006). The impact of corrections for faking on the validity of non-cognitive measures in selection settings. *Journal of Applied Psychology, 91*, 613-621.
- Schneider, B. (2007). Evolution of the study and practice of personality at work. *Human Resource Management, 46*, 583-610.
- Society for Industrial and Organizational Psychology (2003). *Principles for the validation and use of personnel selection procedures* (4<sup>th</sup> ed.). Bowling Green, OH: Author.

Steffens, M. C., & König, S. S. (2006). Predicting spontaneous big five behavior with implicit association tests. *European Journal of Psychological Assessment, 22*, 13-20.

Tett, R.P., Jackson, D.N., & Rothstein, M.G. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*, 703-742.

U.S. Department of Labor (2001). *Standard Occupational Classification*. Retrieved September 22, 2005 from [http://stats.bis.gov/soc\\_majo.htm](http://stats.bis.gov/soc_majo.htm).

Wiggins, J. S. (1996). *The five-factor model of personality*. New York: Guilford.

Witt, L. A., Burke, L., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of Conscientiousness and Agreeableness on job performance. *Journal of Applied Psychology, 87*, 164-169.

Table 1. The Five-Factor Model of Personality

<i>Factor</i>	<i>Definition</i>	<i>Low standing</i>	<i>High standing</i>
Extraversion	Combines ambition, which concerns taking initiative, being competitive, and seeking leadership roles, and sociability, which concerns seeming outgoing, energetic, and entertaining	Complacent and unassertive; withdrawn, reserved, and quiet	Proactive and assertive; talkative and socially engaged
Agreeableness	Concerns being sensitive, considerate, and skilled at maintaining relationships	Tough-minded, frank, and direct	Friendly, warm, and pleasant
Conscientiousness	Concerns being reliable, dependable, and rule-abiding	Non-conforming, impulsive, and flexible	Organized, dependable, and hard working
Emotional stability	Concerns self-confidence, composure, optimism, and stable moods	Tense, irritable, and negative	Confident, resilient, and optimistic
Openness	Concerns being curious, imaginative, visionary, and intellectually engaged	Focused and pragmatic	Creative and curious

Table 2. Validity of Other Assessments for Predicting Overall Job Performance

<b>Study</b>	<b>Predictor</b>	<b>r<sub>obs</sub></b>
A.	Conscientiousness Tests	.18
B.	Integrity Tests	.21
C.	Structured Interviews	.18
D.	Unstructured Interviews	.11
E.	Situational Judgment Tests	.20
F.	Biodata	.22
G.	General Mental Ability	.21
H.	Assessment Centers	.28

*Note:* r<sub>obs</sub> = mean observed validity; A = Mount & Barrick (2001); B = Ones, Viswesveran, & Schmidt (1993); C & D = McDaniel, Whetzel, Schmidt, & Maurer (1994); E = McDaniel, Hartman, Whetzel, & Grubb (2007); F = Bliesener (1996); G = Pearlman, Schmidt, & Hunter (1980); H = Arthur, McNelly, & Edens (2003).

Table 3. Relation between Five Factor Model of Personality and Leadership

<i>Dimension</i>	<i>k</i>	<i>N</i>	<i>r</i>	$\rho$
Extraversion	60	11,705	.22	.31
Agreeableness	42	9,801	.06	.08
Conscientiousness	35	7,510	.20	.28
Emotional Stability	48	8,025	.17	.24
Openness	37	7,221	.16	.24

*Note.* *k* = number of correlations;  $\rho$  = corrected correlation. Source: Judge, Bono, Ilies, & Gerhardt, 2002

Table 4. Summary of Hogan and Holland (2003) Results

<i>HPI Scale</i>	<i>k</i>	<i>N</i>	<i>r</i>	<i>ρ</i>
Adjustment	24	2,573	.25	.43
Ambition	28	3,698	.20	.34
Sociability	NA	NA	NA	NA
Likeability	17	2,500	.18	.36
Prudence	26	3,379	.22	.36
Inquisitive	7	1,190	.20	.34
Learning Approach	9	1,366	.15	NA

*Note.* HPI = Hogan Personality Inventory; *k* = number of correlations; *ρ* = corrected correlation. HPI Adjustment corresponds to FFM Emotional Stability; HPI Ambition and Sociability represent the two major components of FFM Extraversion; HPI Likeability corresponds to FFM Agreeableness; HPI Prudence corresponds FFM Conscientiousness; and HPI Inquisitive and Learning Approach correspond to two components of FFM Openness.

Table 5. Organizationally Significant Outcomes Predicted By Personality Assessment

Outcome	Research Support
Leadership	Judge, Bono, Illies, & Gerhardt, 2002
Job Performance	Barrick & Mount, 1991; Hogan & Holland, 2003; Tett, Rothstein, & Jackson, 1991; Salgado, 1997
Counter-productive work behavior	Berry, Ones, & Sackett, 2007
Absenteeism	Ones, Viswesvaran, & Schmidt, 2003
Teamwork and team performance	J. Hogan & Holland, 2003; Peeters, Van Tuijl, Rutte, & Raymen, 2006),
Job and career satisfaction	Judge, Heller, & Mount, 2002; Ng, Eby, Sorensen, & Feldman, 2005
Health behaviors and life expectancy	Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007