# Predicting Supervisor Ratings of Job Performance Using the NEO Personality Inventory

RALPH L. PIEDMONT HAROLD P. WEINSTEIN Caliper Corporation Princeton, NJ

ABSTRACT. The resurging interest in identifying personality predictors of job performance has led researchers to use the five-factor model of personality as an important conceptual tool. Two recent meta-analyses illustrated the potential predictive power of the model within an employment context. In this study, we evaluated the relations between a well-validated self-report measure of the five-factor model and supervisor ratings of performance. Fifty-two women and 159 men completed the NEO Personality Inventory (Costa & McCrae, 1985) and were independently rated on various performance dimensions by their supervisors. Scores on the Conscientiousness scales correlated with all the performance ratings over this diverse number of occupational groups. Low Neuroticism and high Extraversion scores also predicted high performance. Low straightforwardness (a facet of agreeableness) was associated with ability to accomplish work-related goals and to adapt to changing work conditions. The results support the utility of the five-factor model as a predictor of job performance.

SOME EARLY RESEARCH has been critical of the role of personality variables in predicting job performance (Guion & Gottier, 1965; Schmitt, Gooding, Noe, & Kirsch, 1984). More recent evidence suggests that such individual difference variables can make a substantive contribution to the selection equation (Day & Silverman, 1989; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Lord, De-Vader, & Alliger, 1986). Meta-analysis by Tett, Jackson, and Rothstein (1991) demonstrated that empirical grounds exist for optimism concerning the use of

Address corrrespondence to Ralph L. Piedmont, 1202 Brixton Road, Baltimore, MD 21239–1218 or to Harold P. Weinstein, Caliper Corporation, 741 Mt. Lucas Road, Princeton, NJ 08540.

personality measures to predict job performance. Mean predictive validity coefficients as high as .38 were found over 94 different samples.

Perhaps the most theoretically meaningful research in this area has centered on the five-factor model of personality. This taxonomy represents an empirically derived model of personality, consisting of the dimensions of *neuroticism* (the tendency to experience negative affect, such as anxiety, depression, and hostility), *extraversion* (the quantity and intensity of interpersonal interaction), *openness* (the proactive seeking and appreciation of new experiences), *agreeableness* (the quality of one's interpersonal interactions along a continuum from compassion to antagonism), and *conscientiousness* (the amount of persistence, organization, and motivation in goal-directed behaviors). These five dimensions are stable over time and robust (Digman, 1990; Piedmont, McCrae, & Costa, 1991). Most important, research has established that these five dimensions are sufficient to describe the information contained in other personality systems (McCrae & Costa, 1989a).

Because the five-factor model represents constructs abstracted at a higher level than traits, this model provides a larger framework for understanding the many individual difference variables that describe personality (Goldberg, 1993; McCrae & John, 1992). Whereas these large dimensions are relatively homogeneous, their measurement can orient researchers to the kinds of specific psychological qualities that may be involved in any process or outcome. In their meta-analytic review, Barrick and Mount (1991) used the five-factor model as the organizing template for their analyses. They demonstrated that conscientiousness showed consistent associations with job performance over diverse occupational groups. Extraversion and openness showed associations with more specific job performance criteria (e.g., training proficiency). Neuroticism and agreeableness showed no appreciable predictive capacity.

Barrick and Mount's (1991) study was a 35-year view on personality-based prediction research, which was able to identify some salient relations between personality and work performance because it made appeal to the five-factor model. Their results highlight both the interpretive and empirical advantages to be gained by organizing personality-related information within the parameters of this model. However, the findings that emerged from this meta-analytic study summarize diverse effects and do not necessarily represent any specific variable.

For example, measures of each of the five major personality factors were formed by aggregating, across various studies, scales from different personality inventories that were rationally determined by Barrick and Mount to reflect each of these constructs. Thus, the personality constructs involved in Barrick and Mount's study can only be termed rough approximations of the five-factor model dimensions. Therefore, the lack of correlations between performance and Barrick and Mount's neuroticism and agreeableness composite scales may be the result of either a reliance on many different personality measures with diverse levels of validity or a consequence of the fact that none of the personality scales used in

the aggregations were explicitly designed and validated as measures of the five factors.

In a complementary meta-analysis by Tett et al. (1991), scales were clustered into the five major personality categories on the basis of previous factor analytic studies (e.g., Costa & McCrae, 1988a). The use of this empirically based sorting method contributed to finding validity coefficients, on average, twice as high as those found by Barrick and Mount (1991). In fact, scores on the derived neuroticism and agreeableness scales were found to be very substantial predictors of job performance.

Again, because no explicit measure of the five-factor model was included in the Tett et al. study, we could not determine whether the improved predictiveness was the result of better precision afforded by a more accurately derived measurement model or an idiosyncratic consequence of the scales sampled. One purpose of the present study was to attempt to put more form to the constructs identified by these meta-analyses by using a standardized, well-validated, self-report measure of the five factors—the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985). This instrument provides a clear psychometric articulation of these five dimensions, and we hoped that such precision might realize larger and more numerous correlations with performance. By regressing supervisor ratings of workers' performance on these five dimensions, our study provides a first step in determining the actual strength of association between these five factors and job performance.

Furthermore, as Barrick and Mount (1991) recommended, given the strong association between conscientiousness and job performance, we deemed it useful to delineate further the boundaries of conscientiousness to determine the kinds of specific qualities that may be underlying performance. Tett et al. (1991) argued that using such specific trait assessments would allow for greater interpretive clarity and predictive precision. Thus, we also included the new NEO-PI facet scales for the traits of agreeableness and conscientiousness (Costa, McCrae, & Dye, 1991), which more clearly specify the kinds of personality dynamics subsumed by these larger dimensions. For example, the conscientiousness dimension includes the facet scales of Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, and Deliberation; these scales survey an individual's sense of personal organization, self-control, and aspiration level. Are all necessary for success in the workplace? We hypothesized that one's sense of aspiration (measured by the Achievement Striving and Competence facet scales) would be most predictive of success in all performance contexts, whereas personal organization (Order and Dutifulness facet scales) and self-control (Deliberation and Self-Discipline facet scales) would be relevant only for certain types of performance.

In the current study, we also made an explicit attempt to include performance dimensions that may be related to personality qualities other than conscientiousness. For example, the ability to relate well to co-workers and customers as well as the ability to cope with job stress and pressure should be related to a person's levels of extraversion and neuroticism, respectively (see Costa & McCrae, 1989b; Piedmont, 1993). The incorporation of such performance dimensions may help to illustrate how all five factors can be useful in predicting performance by sampling work-related behaviors outside the usual domain of productivity. Finally, when evaluating how well a score predicts an outcome, it is important to show not only statistical but also practical significance. As already noted, in order to evaluate the functional utility of the NEO-PI as a predictor of work performance, we regressed scores on the supervisor ratings of performance on the NEO-PI. The resulting multiple correlations were then evaluated in the context of Taylor and Russell's (1939) prediction tables to determine the kinds of predictive gains one can expect when making selection decisions on the basis of these personality scores.

## Method

# Subjects

Subjects were 52 women and 159 men who were employed at the time of testing. They were engaged in a wide range of occupations, including customer service, sales, management (lower, middle, and upper), and finance. Sales and customer service positions were the most frequently represented in this sample (49% and 24%, respectively). For those whose year of hire was known, 51% began working for their current employer in or after 1990. Individuals had to be at their current job for a minimum of 6 months to be included in this study. Concerning race, 88% were White, 2.5% Black, 1.5% Asian, 4% Hispanic, and the remaining 4% did not report their racial identification. Although this sample is representative of the kinds of companies we service, it may not be a good reflection of national corporate demographics.

#### Measures

NEO-PI. Developed by Costa and McCrae (1985), this 181-item questionnaire was developed through rational and factor analytic methods to measure the five major dimensions of personality. It is the only commercially available instrument designed specifically to measure the five-factor model. For the Neuroticism, Extraversion, and Openness to Experience domains, there are six 8-item facet scales designed to capture the qualities that are subsumed by these domains. The Agreeableness and Conscientiousness domains are measured with global 18-item scales. Items are answered on a 5-point scale ranging from strongly agree (1) to strongly disagree (5), and scales are balanced to control for the effects of acquiescence. Internal consistency for the domain scales range from .76 to .93, and scores for adults are extremely stable, with 3- and 6-year retest coefficients ranging from .63 to .83 (Costa & McCrae, 1988b). The NEO-PI has been extensively validated

in studies with other self-report inventories (e.g., McCrae & Costa, 1989b; McCrae, Costa, & Piedmont, 1993; Piedmont et al., 1991) and with alternative measures of the five-factor model (e.g., Goldberg, 1989; Trapnell & Wiggins, 1990). Scales have shown evidence of convergent and discriminant validity across instruments, methods, and observers and have been related to a number of life outcomes, including frequency of somatic complaints, ability to cope with stress, and response to psychotherapy (Costa & McCrae, 1989a; Miller, 1991). Reviews of the instrument have been provided by Hogan (1989) and Leong and Dollinger (1990).

Agreeableness and Conscientiousness facets. In an attempt to fully specify the qualities measured by the five factors, Costa and McCrae (1992) recently developed facet scales for the Agreeableness and Conscientiousness domains. The facet scales for Agreeableness (with their respective alpha reliabilities) include Trust (.79), Altruism (.75), Straightforwardness (.71), Compliance (.59), Modesty (.67), and Tendermindedness (.56). For Conscientiousness, the facet scales are Competence (.67), Order (.66), Dutifulness (.62), Achievement (.67), Self-Discipline (.75), and Deliberation (.71). Evidence of convergent and discriminant validity using multiple information sources and instruments does exist (McCrae & Costa, 1992; Piedmont & Weinstein, 1993). These 12 eight-item scales are also measured on a 5-point scale ranging from strongly agree (1) to strongly disagree (5). Items are balanced to control for acquiescence. Overall Agreeableness and Conscientiousness domain scores are created by summing scores on the respective facet scales.

Supervisor performance ratings. The immediate supervisor for each participant completed a performance evaluation, developed by us, that was no part of subjects' normal performance appraisal process. Supervisors did not have access to any of the self-ratings prior to making their evaluations. The rating form asked supervisors to rate subjects on a 5-point scale ranging from unsatisfactory (1) to excellent (5) on 12 relevant performance items that constituted three performance scales. These scales were labeled interpersonal relations (items included "communicates ideas clearly," "relates well to supervisors," "team player," and "service minded"), task orientation (included the items "self-starter," "hard working," "detail skills," and "gets things done"), and adaptive capacity (included the items "learns and adapts readily," "copes effectively with setbacks," "functions well in unstructured situations," and "plans, coordinates, and follows up work of others"). If an item was not applicable, a rating of N/A was made. Individuals who were not rated on each item of a scale did not receive a total score and were excluded from any analyses involving that scale. In a separate sample of 100 individuals, alpha reliabilities for these three dimensions were .80, .82, and .85 for the interpersonal relations, task orientation, and adaptive capacity scales, respectively. Supervisors also made a global rating of performance on a 4-point scale

ranging from below average (1) to excellent (4). Each subject was rated by only one supervisor.

#### Procedure

Companies were randomly selected from our client list and invited to participate in this study. If a client agreed to participate, a list of individuals who had tested with us within the past year were identified and sent the NEO-PI and the Agreeableness and Conscientiousness facet scales. Their supervisors were also sent, among other materials, the performance measure to complete on the subject. All information was then mailed back to us for processing.

## Results

Table 1 contains the correlations between the NEO-PI domain and facet scales and supervisor performance ratings. As can be seen, there are a number of significant correlations. As expected, the conscientiousness facets correlated with the performance ratings in theoretically consistent ways. For example, competence, achievement striving, and self-discipline appear to be qualities that underlie successful performance in all the areas rated. No doubt a personal sense of efficacy, ambition, and persistence are perceived as essential qualities for success on the job across diverse occupational groups.

Extraversion scores were significantly correlated with successful ratings on the interpersonal relations scale, r(208) = .20, p < .01. Those same interpersonal skills also were correlated with how successful a subject was rated on the other two performance dimensions. However, the organization and focus (i.e., adaptive capacity and task orientation, respectively) necessary to obtain performance goals also requires the more antagonistic elements of low straightforwardness. Finally, neuroticism emerged as a significant predictor of ratings concerning interpersonal relations, r(208) = -.16, p < .05, and adaptive capacity, r(208) = -.17, p < .05.

Four separate regression analyses were performed, using each performance scale as the criterion and the total Neuroticism, Extraversion, and Openness scores from the NEO-PI and the overall Agreeableness and Conscientiousness scores based on the facet scales as the predictors. The five personality scores were entered as one block of variables in the regression analysis, and the results indicate the strength of association between the scores and the performance criteria. Be-

<sup>&</sup>lt;sup>1</sup> This finding may be a function of the sample included here. As noted earlier, the majority of subjects were from the sales and customer service areas—positions that have a strong interpersonal component built into the job. Being able to interact with others in a positive way is one of the keys to successful job performance.

TABLE 1 Correlations Between NEO-PI Domain and Facet Scores and Supervisor Ratings of Performance

Scale	Supervisor ratings			
	Overall rating	Interpersonal relations	Task orientation	Adaptive capacity
NEO-PI domain scores				
Neuroticism	12	16*	09	17*
Extraversion	.07	.20**	.16*	.19*
Openness	.04	.07	03	.07
Agreeableness	13	.07	10	14
Conscientiousness	.19**	.23**	.29**	.21**
NEO-PI facet scales Agreeableness				
Trust	.00	.12	.10	.02
Straightforwardness	19**	07	14*	21**
Altruism	05	.10	03	07
Compliance	05	.07	08	08
Modesty	.03	08	.00	09
Tendermindedness	.02	.00	.01	02
Overall Agreeableness	07	.03	04	12
Conscientiousness				
Competence	.15*	.16*	.16*	.21**
Order	.12	.21**	.21**	.11
Dutifulness	.11	.06	.16*	.04
Achievement Striving	.23**	.26**	.26**	.27**
Self-Discipline	.21**	.18**	.28**	.24**
Deliberation	.11	.14*	.11	.07
Overall Conscientiousness	.20*	.20**	.26**	.20**

Note. N = 207-211.

cause regression analyses are sample specific, they tend to overestimate the strength of the population relationship. Thus we report only the shrunken correlations, which correct for the number of predictors and sample size, because they provide a less biased estimate of the true magnitude of association. Shrunken multiple correlations of .18, .31, .19, and .26 between the NEO-PI scores and the overall performance, task orientation, interpersonal relations, and adaptive capacity performance ratings, respectively, were obtained, F(5, 200) = 2.39, 5.27, 2.48, and 3.93, p < .05, .0001, .05, and .002, respectively. Assuming a 50% base rate (i.e., predicting who will score above the median level of job performance vs.

<sup>\*</sup>p < .05 (two-tailed). \*\*p < .01 (two-tailed).

those who will perform below the median) and using a 30% selection ratio,<sup>2</sup> one could expect to accurately classify anywhere from 59% to 64% of any group of individuals (Taylor & Russell, 1939). This represents a 9% to 14% gain over chance, depending on the performance dimension.

# Discussion

The results of this study clearly replicate previous meta-analytic findings (Barrick & Mount, 1991). The conscientiousness dimension scores correlated consistently with job performance across the wide range of occupational categories. The new facet scales for conscientiousness show that it is competence, achievement striving, and self-discipline that are the specific personality qualities that underlie success. The other conscientiousness facets may be relevant in specific performance categories. Extraversion was also a significant predictor of success. Barrick and Mount found this dimension to be relevant particularly for managerial and sales positions (such individuals constitute half of this sample). High scores on the Extraversion scale indicate a warm, engaging, positive, and concerned orientation toward others that facilitates many types of interpersonal interactions.

The correlations between the Agreeableness facets and performance help to provide an additional insight into the interpersonal activities of these highly extraverted individuals as they strive to meet their performance objectives: They not only like being with others but also are willing to engage in whatever social manipulations are necessary to attain their goals (see Trapnell & Wiggins, 1990; Wiggins, 1979). In a group situation, individuals high on extraversion and low on straightforwardness may be perceived as leaders or as charismatic or both. They certainly exude a strong sense of self-confidence and can be cunning when it comes to reaching their goals.

The correlations with neuroticism highlight the relevance of a personality dimension not found in Barrick and Mount's (1991) study. Hardy, emotionally stable, even-tempered individuals were rated by their supervisors as being able to maintain consistent relationships with others and able to cope with the myriad of demands associated with any given job. The emergence of neuroticism as a predictor of performance replicates the Tett et al. (1991) findings and calls attention

<sup>&</sup>lt;sup>2</sup>A selection ratio of 30% represents approximately the top third of a distribution of candidates. This value was arbitrarily selected so that the kinds of hit rates that, on average, one could expect for predicting the criterion measures in this study could be illustrated. However, using a full range of selection ratios, from 5% to 50%, one could expect hit rates for these four scales to range from 56% (using a 50% selection ratio when the overall performance rating is the criterion) to 74% (using a 5% selection ratio when the task orientation scale is the criterion) when the five factors are used as predictors. See Taylor and Russell (1939) for the expected hit rates for other selection ratios. Thus, depending on the criterion, using validated measures of the five factors can provide an increase in predictive accuracy anywhere from 6% to 24% over that expected by chance.

to the need for researchers to use psychometrically unambiguous measures of the five-factor model. Reliance on scales rationally determined to reflect these particular personality dimensions may not provide the level of sensitivity required for the reliable detection of the construct. Instruments like the NEO-PI can be useful in identifying both general personality dimensions and specific individual-differences traits necessary in job performance.

On the basis of these results and those mentioned previously, it appears that all five factors are necessary to account for personality's contribution to supervisors' ratings of occupational success. Of course, such ratings have little directly to do with the magnitude of production, sales, or profits an individual may generate. The performance criteria used here may best be termed as reflecting an employee's orientation toward, and general capacity for, work. Thus, the value of measuring personality as a predictor of job success is dependent on the significance of these performance capacities in a specific job context. By carefully examining the work environment, future research may be able to more fully document how personality contributes to the corporate process in both direct (e.g., productivity and worker efficiency) and indirect (e.g., employee satisfaction and product quality) ways. The results presented here demonstrate that a measurement tool empirically based on the five-factor model will provide a sound vehicle for documenting such effects.

This study adds to the emerging consensus that personality has something substantive to contribute to the job-success prediction equation. The regression analyses showed that the five-factor model is significantly related to supervisor performance ratings. When the magnitude of these associations are evaluated in terms of the tables provided by Taylor and Russell (1939), it is clear that meaningful gains in predictive efficacy can be expected. Because personality variables provide information about people over and above that obtained from cognitive ability tests (Day & Silverman, 1989), combining information from the five-factor model with data obtained from cognitive tests and work histories may provide a powerful predictive combination for use in the job selection process.

#### REFERENCES

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1–26.

Costa, P. T., Jr., & McCrae, R. R. (1985). The NEO Personality Inventory manual. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1988a). From catalogue to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology*, 55, 258–265.

Costa, P. T., Jr., & McCrae, R. R. (1988b). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal* of Personality and Social Psychology, 54, 853–863.

Costa, P. T., Jr., & McCrae, R. R. (1989a). NEO-PI/NEO-FFI manual supplement. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1989b). Personality, stress, and coping: Some lessons

from a decade of research. In K. S. Markides & C. L. Cooper (Eds.), Aging, stress, and health (pp. 269-285). New York: Wiley.

Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory and NEO Five-Factor Inventory, professional manual. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Indi*vidual Differences, 12, 887–898.

Day, D. V., & Silverman, S. B. (1989). Personality and job performance: Evidence of incremental validity. Personnel Psychology, 42, 25–36.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. Annual Review of Psychology, 41, 417–440.

Goldberg, L. R. (1989, June). Standard markers of the Big Five factor structure. Paper presented at the Invited Workshop on Personality Language, University of Groningen, Groningen, The Netherlands.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. American Psychologist, 48, 26–34.

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. Personnel Psychology, 18, 135–164.

Hogan, R. (1989). Review of the NEO Personality Inventory. In J. C. Conoley & J. J. Kramer (Eds.), The 10th mental measurements yearbook (pp. 546-547). Lincoln, NE: Buros Institute of Mental Measurements.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581–595.

Leong, F. T. L., & Dollinger, S. J. (1990). NEO Personality Inventory. In D. J. Keyser & R. C. Sweetland (Eds.), Test critiques (Vol. 8, pp. 527–539). Austin, TX: PRO-ED.

Lord, R. G., DeVader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, 71, 402–410.

McCrae, R. R., & Costa, P. T., Jr. (1989a). More reasons to adopt the five-factor model. American Psychologist, 44, 451–452.

McCrae, R. R., & Costa, P. T., Jr. (1989b). Reinterpreting the Myers–Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality*, 57, 17–40.

McCrae, R. R., & Costa, P. T., Jr. (1992). Discriminant validity of the NEO-PIR facet scales. Educational and Psychological Measurement, 52, 229–237.

McCrae, R. R., Costa, P. T., Jr., & Piedmont, R. L. (1993). Folk concepts, natural language, and psychological constructs: The California Psychological Inventory and the five-factor model. *Journal of Personality*, 61, 1–26.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175–215.

Miller, T. (1991). The psychotherapeutic utility of the five-factor model of personality. Journal of Personality Assessment, 57, 415–433.

Piedmont, R. L. (1993). A longitudinal analysis of burnout in the health-care setting: The role of personal dispositions. *Journal of Personality Assessment*, 61, 457–473.

Piedmont, R. L., McCrae, R. R., & Costa, P. R., Jr. (1991). Adjective check list scales and the five-factor model. *Journal of Personality and Social Psychology*, 60, 630–637.

Piedmont, R. L., & Weinstein, H. P. (1993). A psychometric evaluation of the new NEO– PIR facet scales for agreeableness and conscientiousness. Journal of Personality Assessment, 60, 302–318.

- Schmitt, F. L., Gooding, R. A., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407–422.
- Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection. *Journal of Applied Psychology*, 23, 565–578.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703–742.
- Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality (IASR-B5). Journal of Personality and Social Psychology, 59, 781-790.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395–412.

Received September 20, 1993

Reprinted from The Journal of Psychology published by HELDREF PUBLICATIONS, 1319 Eighteenth Street, NW, Washington, DC 20036-1802. (202) 296-6267 (202) 296-5149 FAX