The Five-Factor Model of Personality and Managerial Performance: Validity Gains Through the Use of 360 Degree Performance Ratings

In-Sue Oh University of Alberta Christopher M. Berry Texas A&M University

This study investigated the usefulness of the five-factor model (FFM) of personality in predicting two aspects of managerial performance (task vs. contextual) assessed by utilizing the 360 degree performance rating system. The authors speculated that one reason for the low validity of the FFM might be the failure of single-source (e.g., supervisor) ratings to comprehensively capture the construct of managerial performance. The operational validity of personality was found to increase substantially (50%–74%) across all of the FFM personality traits when both peer and subordinate ratings were added to supervisor ratings according to the multitrait—multimethod approach. Furthermore, the authors responded to the recent calls to validate tests via a multivariate (e.g., multitrait—multimethod) approach by decomposing overall managerial performance into task and contextual performance criteria and by using multiple rating perspectives (sources). Overall, this study contributes to the evidence that personality may be even more useful in predicting managerial performance if the performance criteria are less deficient.

Keywords: personality, five-factor model of personality, 360 degree performance ratings, managerial performance

Managerial performance is largely composed of task completion and relationship building (e.g., Conway, 1999). Both of these performance dimensions have intuitive links to various attributes of persons, such as personality traits (e.g., it is intuitive that traits such as Conscientiousness or Agreeableness should be related to engaging in task completion or relationship-building behaviors, respectively; Barrick & Mount, 1991). However, empirical research has not resulted in strong support for these intuitive links between self-report personality and managerial performance, as operational validities have generally been less than .20 (Barrick, Mount, & Judge, 2001; Ones, Hough, & Viswesvaran, 2005). We note, though, that previous empirical research examining personality-managerial performance relationships has generally relied on single-source supervisor ratings as operational measures of managerial performance. Such ratings neglect that managerial performance is multilevel, in that managers work with various individuals at different levels of the organization (e.g., supervisors, peers, subordinates) and that certain managerial performance behaviors (e.g., providing coaching and career advice) may be more easily observed by individuals at one organizational level (e.g.,

subordinates) than by individuals at other organizational levels (e.g., supervisors). Thus, the possibility exists that the relationships of personality traits with managerial performance may be underestimated due to criterion deficiency resulting from the use of single-source supervisor ratings. We make the case that the criterion deficiency issue may be partially addressed via the use of 360 degree performance rating systems that utilize multisource ratings to assess managers' performance (e.g., Mount, Judge, Scullen, Sytsma, & Hezlett, 1998). Therefore, the present study investigated the extent to which five-factor model (FFM) personality traits relate to managerial performance as assessed by 360 degree ratings. We begin by discussing the multidimensional and multilevel nature of managerial performance and then move to the main propositions of the present study: first, that it is important to include multiple perspectives in performance measurement and second, that use of multiple perspectives has implications for measuring relationships between FFM personality traits and managerial performance.

Managerial Performance

Managerial performance is often defined as "managerial behaviors believed to be optimal for identifying, assimilating and utilizing resources [including human resources] toward sustaining the organizational unit for which a manager has responsibility" (Nilsen, 1995, p. 2; see Campbell, Dunnette, Lawler, & Weick, 1970). Research and theory have often converged on a set of performance factors that map onto broad managerial task and contextual performance factors. Managerial task performance includes behaviors focused on structuring work and getting things done, and managerial contextual performance includes behaviors focused on facilitating the psychological and social contexts of work and getting along with others (Van Scotter & Motowidlo, 1996). For instance, Borman and Brush's (1993) taxonomy of managerial performance

In-Sue Oh, Department of Strategic Management and Organization, School of Business, University of Alberta, Edmonton, Alberta, Canada; Christopher M. Berry, Department of Psychology, Texas A&M University.

Both authors contributed equally to this article. We thank Ronald Page and Assessment Associates International for providing the data for this study and Frank Schmidt and Kibeom Lee for their helpful comments on an earlier draft of this article.

Correspondence regarding this article should be addressed to In-Sue Oh, Department of Strategic Management and Organization, School of Business, University of Alberta, 2-32C Business Building, Edmonton, Alberta T6G 2R6, Canada, or Christopher M. Berry, Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235. E-mail: insue@ualberta.ca or cmberry@tamu.edu

proposed 18 factors. Viswesvaran and Ones (2000, p. 219) noted that the 18 factors "can further be grouped into four broad managerial performance dimensions": (a) leadership and supervision (e.g., supervising employees to accomplish their goals); (b) technical behaviors and mechanics of management (e.g., administration such as planning, implementing, monitoring task activities); (c) interpersonal facilitation (e.g., giving individualized consideration and attention to subordinates); and (d) job dedication (e.g., doing whatever it takes to get a job done, exerting extra effort when facing crises). Johnson (2003, p. 92) further argued that Viswesvaran and Ones' four dimensions can be classified into task and contextual performance factors. That is, (a) leadership and supervision and (c) technical behaviors and mechanics of management fall into the category of task performance and (b) interpersonal relations and communications and (d) useful behaviors and skills (e.g., handling crises) fall nicely into Van Scotter and Motowidlo's (1996) two factors of managerial contextual performance: interpersonal facilitation and job dedication, respectively.

Other researchers' taxonomies of managerial performance may be classified into two factors, as discussed above (e.g., Campbell, McCloy, Oppler, & Sager, 1993; Scullen, Mount, & Goff, 2000; Tett, Guterman, Bleier, & Murphy, 2000). For example, Campbell et al. (1993) proposed an eight-factor model of job performance, with two of these factors related to managerial performance: supervision/leadership and management/administration. According to Campbell et al., the former represents "behaviors directed at influencing the performance of subordinates through face-to-face interpersonal interaction and influence" (1993, p. 48), and the latter represents behaviors directed at "major elements in management that are distinct from direct supervision" (p. 48; i.e., things such as monitoring progress or obtaining additional resources). Thus, management/administration duties represent more core job tasks with behaviors focused on structuring work and getting things done, and supervision/leadership duties have a more interpersonal focus. Behaviors focused on structuring work and getting things done resemble managerial task performance, and performance behaviors with a more interpersonal focus resemble managerial contextual performance.

What these models have in common is a set of managerial performance behaviors that map onto broader task performance factors (structuring work and getting things done) and contextual performance factors (facilitating the psychological and social contexts of work and getting along with others). This, of course, does not mean that such a two-factor model is the only appropriate model of managerial performance. However, the above review of the managerial performance literature demonstrates that the broad managerial task and contextual performance factors are repeatedly observed across the aforementioned managerial performance models and taxonomies currently available. Thus, it seems reasonable to classify managerial performance into two higher order task and contextual performance factors, so the present study will hereafter use such a model of managerial performance.

In addition to being multidimensional, managerial performance is multilevel in that managers work with various individuals at different levels of the organization (e.g., subordinates, peers, supervisors). Managerial performance behaviors may differ systematically depending on the organizational level of the person with whom the manager is interacting. For instance, the interactions of a manager with his or her supervisor may commonly involve

behaviors such as reporting on performance and progress, carrying out orders, and providing opinions about a given course of action; the interactions of that manager with his or her subordinates may more commonly involve behaviors such as providing direction and structure, influencing, and facilitating relationships between subordinates. Thus, managerial performance behaviors may differ systematically, depending on the organizational level of the person toward whom the behavior is directed. Therefore, organizational members at different levels in relation to a given manager will likely have unique perspectives on that manager's performance.

Why It Is Important to Include Multiple Perspectives in Managerial Performance Ratings

Single-source ratings from supervisors are the most common form of performance measurement (e.g., Murphy & Cleveland, 1995). Supervisor ratings are certainly an important source of performance information, but they do not likely adequately capture all of the relevant variance in true levels of manager ratees' job performance, especially given the inherently multidimensional and multilevel properties of managerial performance. Thus, both researchers and practitioners have recommended the use of 360 degree performance ratings (a.k.a., multisource or multirater ratings) in assessing managerial performance (Mount et al., 1998). The term 360 degree performance ratings refers to performance ratings of a ratee (usually a manager) made by multiple individuals with varying relationships with the (manager) ratee. The most common sources for 360 degree ratings are a manager's supervisors, peers, and subordinates, along with self-ratings (Conway & Huffcutt, 1997; Craig & Hannum, 2006). Although much research has addressed how well 360 degree ratings work for the purposes of personal development or administrative decisions (see Morgeson, Mumford, & Campion, 2005, for one review), we focus on their properties as validation criteria and their potential for increasing understanding of personality-managerial performance relationships. Indeed, a fundamental assumption underlying collection of ratings from multiple sources is that each rating source has a unique, yet potentially valid, perspective on the ratee's performance; thus, ratings from multiple sources will tap a greater proportion of the true performance domain than any single rating source (e.g., Borman, 1997; Craig & Hannum, 2006; LeBreton, Burgess, Kaiser, Atchley, & James, 2003). To the degree that this is true, 360 degree ratings should represent a less deficient criterion than supervisor ratings alone, and thus 360 degree ratings hold the potential for increased prediction by appropriate psychological tests.

The socioanalytic approach (Hogan, 1991; Hogan & Holland, 2003) to job performance and personality represents one theoretical perspective on why managerial performance ratings from different sources (i.e., members of different organizational levels, such as supervisors vs. peers vs. subordinates) may each provide unique performance information. The socioanalytic approach is designed to explain individual differences in success at work and is based on two broad assumptions: that people live and work in groups and that groups are structured in terms of status hierarchies. These assumptions suggest two broad motives that translate into behavior: the motivation to get along with other members of the group and the motivation to achieve status with other members of the group (i.e., get ahead). Hogan and Holland (2003) suggest that

the importance of these two motives can be expressed "in Darwinian terms: People who cannot get along with others and who lack status and power have reduced opportunities for reproductive success" (p. 100). Thus, it stands to reason that characteristic patterns of behavior in the workplace (and life in general) will be driven by these two motives. However, although most people will try to get along and get ahead in the workplace, there are individual differences in how successful persons are at translating these motives into behaviors that help them get along with others and/or get ahead. This principle holds for managers, and thus some managers will be more successful than others at getting along and/or getting ahead. The characteristic patterns of behavior of managers will gain them reputations in the eyes of performance raters (e.g., supervisors, peers, subordinates) regarding the degree to which the managers are successful at getting along and getting ahead.

Further, Hogan and Shelton (1998) made the case that a performance rater's judgments of a manager's performance are based, in part, on the degree to which the manager meets the rater's expectations and promotes the rater's agenda. Hence, if the expectations and agendas of raters differ by rating source, raters from different organizational levels will be more likely to attend to different aspects of the manager's behaviors; thus, ratings from different rating sources will likely reflect unique sets of performance behaviors to some degree. For example, a certain manager may have been especially effective at manifesting his or her drive to get along with others by engaging in behaviors that help build relationships with subordinates, and thus this manager is seen as someone who is especially effective at getting along with subordinates. This manager's subordinates may think that this considerate manager is an effective manager, given the benefits of such manager behavior to the subordinate rater (e.g., it is more enjoyable to work for a manager who motivates and influences through relationship building). On the other hand, a supervisor rater may value productive managers (getting ahead), as these manager attributes are the ones in line with the supervisor rater's personal agenda (i.e., the supervisor is more likely to be rewarded for supervising productive managers than managers whose employees enjoy working for them), and may be less concerned with whether the manager engages in relationship building with subordinates. Thus, for a given manager ratee, the supervisor and subordinate ratings may have different psychological meaning, as they are more heavily weighted toward different pieces of information. Other theoretical perspectives on multisource performance ratings support the same idea that ratings from different sources have the potential to reflect unique managerial performance variance (e.g., Lance, Baxter, & Mahan, 2006; Tett & Burnett, 2003). Therefore, when measuring managerial performance, it is important to attend to the unique information that may be gained from multiple performance ratings perspectives.

According to Borman (1997), this unique information could be reflective of one or more of the following mechanisms: that raters from different organizational levels rely on a different model of performance (e.g., rely on different performance dimensions, define dimensions differently, weight dimensions differently) or simply observe (through opportunity or motivation) a different subset of ratee behavior. Note that this does not mean that ratings from different sources will be unique. However, to the degree that the above ideas are true, ratings pooled across multiple sources have

the potential to represent a less deficient and more predictable criterion than do ratings from a single source (e.g., supervisors).

Two lines of empirical evidence are supportive of the socioanalytic approach's assertion that performance ratings from multiple sources may account for additional true variance in performance. One is the relative magnitude of correlations of ratings between and within sources (i.e., greater within-source than between-source agreement would signal the potential for incremental validity). A meta-analysis by Conway and Huffcutt (1997) found that mean interrater reliabilities were .50 for supervisors, .37 for peers, and .30 for subordinates; mean supervisor-peer (.34), supervisorsubordinate (.22), and peer-subordinate (.22) correlations were generally lower. Thus, correlations between performance ratings from raters of different organizational levels are generally lower than correlations between ratings within sources. The second supportive line of evidence is that ratings from different sources form "method" factors that have different patterns of correlations with a common set of correlates (e.g., Beehr, Ivanitskaya, Hansen, Erofeev, & Gudanowski, 2001; Conway, Lombardo, & Sanders, 2001; Lance et al., 2006; Sala & Dwight, 2002). These different correlates imply that ratings from each source have different psychological meaning. Given this empirical support for our above assertions based on socioanalytic theory, it is worth investigating whether 360 degree ratings may yield a less deficient and more predictable criterion than single-source supervisor ratings.

Implications of Measuring Managerial Performance via 360 Degree Ratings on FFM Personality–Performance Relationships

The present study makes the case that the use of 360 degree performance ratings may yield a less deficient and more predictable criterion than that yielded by single-source supervisor ratings, resulting in stronger relationships between the FFM of personality and managerial performance. The FFM of personality comprises the five broad personality traits: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability. Although a number of researchers have challenged the FFM on a number of grounds (e.g., not the correct number of factors, Hough, 1992; Lee & Ashton, 2004; inappropriate factor analytic methods, Block, 1995), it is widely considered a useful taxonomy of personality traits that is generally supported by research and theory (Barrick et al., 2001).

Hogan and Holland's (2003) socioanalytic perspective on the links between personality traits and job performance explains why there should be relationships between FFM personality traits and managerial performance behaviors. Hogan and Holland (2003) made the case that both performance behaviors and FFM personality traits can be mapped onto the two motivational dimensions of "getting ahead" and "getting along." Similarly, we make the case that managerial performance behaviors can be classified into two higher order task and contextual performance factors. Task performance, with its focus on structuring work and getting things done, corresponds to the "getting ahead" motive, and contextual performance, with its focus on facilitating the psychological and social contexts of work and getting along with others, corresponds to the "getting along" motive (Hogan & Holland, 2003). Similarly, Hogan and Holland (2003) listed as "getting along" traits the traits of Emotional Stability (e.g., emotionally stable persons will be

more positive), Conscientiousness (e.g., conscientious persons will have more predictable behavior), and Agreeableness (e.g., agreeable persons will be more sensitive to others). We note that the sociability facet of Extraversion implies a strong interpersonal component to Extraversion, making it likely that Extraversion may also be a "getting along" trait. Hogan and Holland listed as "getting ahead" traits the traits of Emotional Stability (e.g., emotionally stable persons will be more confident), Extraversion (e.g., more extraverted persons will tend to be ambitious), and Openness to Experience (more open persons will tend to be curious and eager to learn). We note that Conscientiousness could be conceptualized as a getting ahead trait as well, given that it is partially reflective of a need for achievement. Thus, the socioanalytic approach suggests that getting ahead traits should be predictive of engaging in getting ahead behaviors (i.e., task performance) and that getting along traits should be predictive of engaging in getting along behaviors (i.e., contextual performance). Figure 1 outlines these expected relationships. In particular, Emotional Stability, Extraversion, and Conscientiousness are expected to be related to both task and contextual performance; Openness is expected to be related to only task performance, and Agreeableness is expected to be related to only contextual performance. For example, the levels of confidence (Emotional Stability), ambition (Extraversion), eagerness to learn new things (Openness), and need for achievement (Conscientiousness) shown by managers should have bearing on the degree to which they engage in managerial task performance (getting ahead) behaviors.

Further, two implications of the socioanalytic perspective combine to support the idea that including additional managerial performance rating perspectives (e.g., peers and subordinates) beyond the typical single-source supervisor ratings may yield a less deficient and more predictable criterion than that yielded by single-source supervisor ratings, resulting in stronger personality–performance relationships. First, FFM traits may differ in the degree to which they are predictive of getting ahead versus getting along performance behaviors (Hogan & Holland, 2003). Second, raters from different organizational levels (e.g., supervisors vs. peers vs. subordinates) will likely have different expectations of and agendas associated with managers; this would make raters from differ-

ent organizational levels more likely to encounter and attend to different managerial behaviors when they make their ratings (Hogan & Shelton, 1998). For instance, the agendas of supervisors likely reflect a desire for production from the managers beneath them, so getting ahead FFM traits—such as Emotional Stability, Extraversion, Openness to Experience, and Conscientiousness should be predictive of supervisors' managerial performance ratings (more so for task, rather than contextual, performance). On the other hand, subordinates' and peers' agendas likely reflect more of a desire for considerate managers who are easy to get along with, so getting along FFM traits—such as Emotional Stability, Conscientiousness, Agreeableness, and Extraversionshould be predictive of peers' and subordinates' performance ratings (more so for contextual, rather than task, performance). Figure 1 also shows these expected relationships between the FFM traits and managerial performance ratings from different sources. Both getting ahead (task performance) and getting along (contextual performance) behaviors are important managerial performance behaviors, so it stands to reason that ratings from different sources (organizational levels) may capture these multiple managerial performance factors more completely than do single-source supervisor ratings. Thus, single-source ratings (e.g., supervisor ratings) likely represent a deficient and, depending on the FFM trait in question, less predictable criterion than ratings pooled across multiple sources.

Present Study

Despite theoretical reasons for expecting relationships between FFM personality and managerial performance (e.g., Hogan & Holland, 2003), marginal support, at best, has been found for these intuitive links by meta-analyses, with criterion-related validities generally being less than .20 (Barrick et al., 2001; Ones et al., 2005). However, the low estimates of criterion-related validity found in these meta-analyses may partially be a function of reliance on single-source supervisor ratings of overall managerial performance as criteria. Although supervisor ratings are certainly important sources of performance information, sole reliance on supervisor ratings may not fully account for the fact that manage-

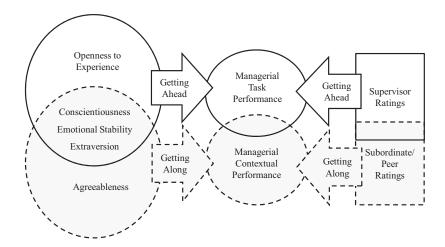


Figure 1. Strongest relationships between FFM traits and performance ratings from different sources expected from a socioanalytic perspective. FFM= five-factor model.

rial performance is multilevel, so that raters from levels (perspectives) other than that of the immediate supervisor may be better suited to observe certain aspects of managerial job performance. The 360 degree performance rating system was designed to address this criterion deficiency problem and to assess multidimensional managerial performance from multiple sources and/or perspectives (Mount et al., 1998). Yet few, if any, studies have investigated how personality relates to managerial performance assessed by the 360 degree performance rating system. Given this, the present study investigated the extent to which FFM personality traits relate to two-factor managerial performance assessed from four different perspectives using the 360 degree performance rating system versus supervisor ratings alone. It is particularly useful to examine the gain in validity when peer, subordinate, and self ratings are added to supervisor ratings. We think that this type of research will enrich our understanding of the personality and performance linkage, which often is deemed as possessing little practical value (e.g., Morgeson et al., 2007).

Before we present the details of the present study, one important point requires discussion. That is, in the present study we examined relationships between FFM personality and managerial performance using data drawn from the same sample as that in Berry, Page, and Sackett (2007). Thus, we wish to be clear about the ways in which the present study goes beyond Berry et al. First, Berry et al. focused on the effects of response distortion on personalityperformance relationships, and the present study more clearly focused on understanding the relationship between FFM personality and 360 degree managerial performance ratings. Second, Berry et al. reported relationships between the FFM and overall job performance, and the present study distinguished between task and contextual performance factors for a more nuanced analysis. Third, Berry et al.'s measure of job performance was a simple mean average of ratings from supervisors, peers, and subordinates, and the present study used confirmatory factor analysis to more clearly explore the structure of ratings from four different perspectives (the three perspectives from Berry et al. plus self-ratings). Fourth, Berry et al. used multiple regression when examining relationships between personality and performance, and the present study used structural equation modeling when examining personality-performance relationships while accounting for variance due both to different performance rating perspectives (method factors) and to performance rating criteria (trait factors). Fifth, the present study compared the operational validities of the FFM personality traits when using as criteria supervisor ratings alone versus 360 degree ratings. No such comparison was attempted by Berry et al.

Method

Sample

The sample of this study comprised 277 managers at a large energy company. They were "middle managers" who occupied a diverse group of positions at organizational levels above front-line supervisor but below the level of vice president. Managers generally had between 40 and 100 employees below them. Some were plant general managers, and others held managerial positions in a wide range of departments such as human resources, information technology, finance, and public affairs and regulatory services.

The managers participated as part of a leadership development program, in which participants completed a FFM-based inventory and their performance levels were assessed on a customized 360 degree performance rating system. Of the 277 managers, 264 provided their gender (73.6% male), and 150 provided their approximate age (M=50.7 years, SD=7.8). Full data were collected for between 239 and 261 managers on the FFM and job performance ratings.

Personality

Participants' FFM personality traits were measured on a 5-point Likert scale with the Work Behavior Inventory (WBI; Page, 2007). The WBI is a 240-item occupational-purpose personality instrument comprising 20 scales, 18 of which are 12-item facet-level personality scales that map onto the FFM traits. See the Appendix for a listing of the WBI personality scales and their definitions. The intercorrelations of FFM personality variables (and all study variables) and their means, standard deviations, and coefficients alpha for this sample are shown in Table 1.

The WBI User's Manual (Page, 2007) details criterion-related, convergent, and discriminant validity evidence for the WBI. For instance, Page (2007) described results from a study in which 509 participants completed the WBI. Principal-components analysis of participants' responses revealed five principal components corresponding to the FFM, with each of the 18 scales having loadings greater than .50 on their appropriate FFM scales. Page also described results of a study in which 229 undergraduates were administered both the WBI and the Hogan Personality Inventory (HPI). The general pattern was for WBI scales to be strongly correlated with HPI scales with which they should be correlated (convergent validity) and relatively uncorrelated with HPI scales with which they should not be correlated (divergent validity). For instance, the Sociability scale of the WBI had an average correlation of .48 with the HPI scales Entertaining, Likes People, Leadership, No Social Anxiety, Likes Parties, Likes Crowds, and Self Confidence but had an average correlation of .14 with all other HPI scales. Similar patterns were found for each of the WBI scales (see Page, 2007). In terms of criterion-related validity evidence, in addition to results for the sample used in the present study, Page (2007) lists significant validity coefficients for FFM scales in samples of state patrol officers (n = 160), long-haul truck drivers (n = 122), and Department of Corrections supervisors (n = 74).

Managerial Performance

An average of 13.44 raters (one supervisor, one self, 6.75 peers, and 4.69 subordinates) rated each ratee (manager) in terms of 19 managerial performance competencies. Using the conceptualization of task versus contextual managerial performance as outlined in the Introduction of this article, In-Sue Oh and Christopher M. Berry separately classified the 19 competencies into the two performance categories. We initially agreed on 15 out of 19 competencies and after further discussion agreed on the classification of 17 out of 19 competencies. Two competencies we did not agree on (Delegation, Leading/Modeling/Visioning) were deleted from subsequent analyses. Managerial task performance included competencies mainly related to getting things planned, organized, done, and controlled (e.g., Results Driven, Problem Analysis, Process Management), and contextual performance included competencies mainly related to supporting the social and psychological contexts

Table 1
Means, Standard Deviations, Reliabilities, and Correlations Between Study Variables

Variable	M	SD	N	1	2	3	4	5	6	7	8	9	10	11	12	13
Supervisor task performance	6.18	1.03	259	.92												
2. Peer task performance	6.18	0.66	275	.37	.93											
3. Sub task performance	6.84	0.79	239	.13	.27	.93										
4. Self task performance	6.46	0.91	268	.21	.19	.02	.88									
5. Supervisor contextual																
performance	6.20	1.02	260	.89	.36	.12	.18	.91								
6. Peer contextual performance	6.06	0.64	275	.28	.88	.21	.11	.33	.90							
7. Sub contextual performance	6.68	0.86	239	.05	.20	.89	.01	.07	.22	.93						
8. Self contextual performance	6.51	0.90	268	.10	.14	.03	.82	.11	.16	.08	.87					
9. Openness	2.93	0.13	261	.15	.26	.01	.33	.14	.16	04	.26	.94				
10. Conscientiousness	3.04	0.14	261	.13	.20	.05	.27	.16	.14	.05	.23	.44	.92			
11. Emotional Stability	3.13	0.16	261	.14	.24	.02	.28	.19	.21	.03	.29	.55	.50	.93		
12. Extraversion	2.93	0.14	261	.17	.21	.01	.32	.21	.17	.03	.33	.60	.56	.55	.95	
13. Agreeableness	3.08	0.13	261	.01	.06	.04	.27	.06	.11	.12	.34	.29	.48	.58	.46	.92

Note. Ns = 239–275. If a correlation is greater than .10, its 95% confidence interval does not include zero (-.10 < .00 < .10) and, thus, it is statistically significant at p = .05. Coefficients alpha for 360 degree performance ratings and the FFM personality model are along the diagonal. For each of the FFM personality traits, the equally weighted mean and standard deviation of its facets are reported. FFM = five-factor model; Sub = subordinate.

of work (e.g., Team Building, Coaching/Mentoring, Negotiating Resolution). Regarding our two-factor model, confirmatory factor analysis (CFA) results of ratings on these final 17 competencies generally fit the data relatively well except for root mean square error of approximation (RMSEA), RMSEA = .104–.162, standardized root mean square residual (SRMR) = .059–.071, comparative fit index (CFI) = .94–.96, normed fit index (NFI) = .93–.95, nonnormed fit index (NNFI) = .93–.95, within each rating perspective. These CFA results also supported the fit of our two-factor task vs. contextual performance model over a one-factor model ($\Delta\chi^2$ = 19.75, 56.81, 3.91, 127.26 with Δdf = 1, ps < .05 for self, peer, supervisor, and subordinate ratings, respectively). That is, we found that managerial performance for each perspective was better represented by two factors (task and contextual factors) than by a single/general factor.

The ICC(1) estimates were on average .63 and .62 for peer and subordinate ratings of managerial task and contextual performance, respectively; more than 60% of the variance in individual peer or subordinate ratings of managerial performance is attributable to group membership. The ICC(2) estimates were on average .92 and .88 for peer and subordinate ratings, respectively; peers' or subordinates' mean ratings reliably distinguished among the managers. Because, in the 360 degree feedback literature, researchers are typically interested in the average of multiple raters' ratings and individual-level measures assess individual differences rather than some common group referent phenomenon, such as group efficacy or work-unit climate, (high) ICC(2), or ICC(K), estimates rather than (high) ICC(1) estimates may be more appropriate for use when justifying the aggregation (Bliese, 2000; Ployhart, Weekly, & Ramsey, in press). Given that the use of aggregate ratings in this study was intended to provide important inferences about specific managers' developmental potentials, higher than typical levels of interrater consensus and reliability are desirable and needed for justifying the aggregation (Bliese, 2000; see LeBreton & Senter, 2008 for more details). Our data meet this condition.

Results

Structure of 360 Degree Managerial Performance Ratings

We generated two managerial performance scores (one for task performance and the other for contextual performance) for each of the three distinct perspectives (supervisors, peers, and subordinates). We note that although we included the fourth rating perspective (self-ratings) in some later analyses for comparison purposes, we did not include self-ratings in these analyses because of the potential for same-source bias. Given the multiple-criteria and multiple-perspective nature (two performance criteria times three perspectives) of these analyses, we needed to determine the most appropriate higher order factor structure for managerial performance ratings. One way to conceptualize these ratings is as a multitrait-multimethod (MTMM) matrix, wherein the two performance criteria are the "traits" and the three rating perspectives are the "methods." CFA is especially suited for investigating a MTMM matrix, given that it can specify each variable to contain trait (criterion type) variance, method (perspective type) variance, and unique (error) variance and that it enables us to compare alternative models that might explain the structure of the data.

¹ We further examined other model fit indices less affected by sample size (i.e., CFI). CFIs were greater (ranging from .003 to .005) for the two-factor model for all perspectives except for the supervisor ratings. In the context of testing measurement invariance in multigroup structural equation modeling, several researchers suggested using the difference in CFI between hierarchically nested models for model selection (Cheung & Rensvold, 2000; Meade, Johnson, & Braddy, 2008). Meade et al. (2008) recommended using the cutoff value of .002 to determine if the more constrained model should be selected. Although the hierarchically nested models examined in the current study do not directly pertain to the issue of measurement invariance, we refer to such a cutoff value to aid our model comparison process. This test is often used as a supplementary test for the chi-square difference test. More detailed CFA results are available from In-Sue Oh.

On the basis of the above discussion and following the analytical strategies of Bentler (2007), Lim and Ployhart (2006), and Widaman (1985), we specified and compared a number of alternative models of performance ratings. Model 1 is a single-factor model that represents a single criterion and ignores the threeratings perspectives, $\chi^2(9) = 382.42$, p < .01, RMSEA = .438, SRMR = .264; CFI, NFI, and NNFI are improper. Model 2 is a two-factor model that represents the two correlated performance criteria and ignores the three-ratings perspectives, $\chi^2(8) = 346.02$, p < .01, RMSEA = .442, SRMR = .237, CFI; NFI and NNFI are improper. Model 3 is a three-factor model that represents the three correlated rating perspectives and ignores performance criteria, $\chi^2(6) = 34.08, p < .01, RMSEA = .147, SRMR = .102, CFI =$.957, NFI = .949, NNFI = .892. Model 4 is a four-factor model that represents a single overall performance criterion as a substantive factor and also accounts for the three uncorrelated perspectives as method factors, $\chi^{2}(6) = 35.50, p > .01, \text{RMSEA} = .151,$ SRMR = .108, CFI = .955, NFI = .947, NNFI = .888. Model 5 is a five-factor model that represents the two correlated performance criteria (task and contextual factors) as substantive factors and accounts for the three uncorrelated perspectives as method factors, $\chi^2(5) = 6.73$, p = .24, RMSEA = .040, SRMR = .025, CFI = .998, NFI = .990, NNFI = .993. Models 6-1 and 6-2 are four-factor models that represent the three uncorrelated (Model 6-1) or correlated (Model 6-2) rating perspectives as substantive factors and account for a single performance criterion as a method factor (these models did not converge). Models 7-1 and 7-2 are five-factor models that represent the three uncorrelated (Model 7-1) or correlated (Model 7-2) rating perspectives as substantive factors and account for the two uncorrelated performance criteria as method factors (Models 7-1 and 7-2 did not converge).

For Models 3–7, we constrained each method factor to influence its associated indicators with equal factor loadings to promote simplicity and model identification (Bentler, 2007). Further, method factors were fixed to be uncorrelated for Models 4 and 5 (as the methods factors were the three rating perspectives) but were freed to correlate with each other for Models 6-1 and 7-1 (as the methods factors were the two performance factors, which were strongly correlated). However, substantive factors were freed to correlate with each other, and we assumed these substantive factors to be uncorrelated with method factors for purposes of model identification for both models, following Lim and Ployhart's (2006) and Widaman's (1985) recommendation. These alternative models were tested with the maximum likelihood methods of LISREL 8.54 and compared with multiple fit indices. All fit indices and a series of chi-square tests indicated that Model 5 fit the data significantly better than did the other models.² We further tested one variant model of Model 5, a five-factor model that represents the two correlated performance criteria (task and contextual factors) as substantive factors and accounts for the three correlated perspectives as method factors, but this model did not converge. Subsequently, Model 5 was the model used for 360 degree performance ratings in subsequent analyses examining the multivariate relationships between the FFM and 360 degree performance ratings. The factor structure and completely standardized factor loadings of Model 5 are given in Figure 2.

For Models 1–3, unique (error) variance was much larger because either trait (performance) variance or method (source/perspective) variance was specified and the remainder of variance

was treated as unique variance. That is, these models did not fully take advantage of the MTMM nature of the data. In contrast, for Models 4 and 5, both trait variance and method variance were specified and the remainder of variance was treated as unique variance, so that the portion of error variance in total variance should be smaller. That is, these two models allow determinations of the degree to which both methods and traits account for covariance among the measures and, thus, fully take advantage of the 360 degree performance ratings (see Mount et al., 1998, for more details). Models 6 and 7 did not converge. These models may be less realistic, given that the factor loadings of performance factors (in these models, treated as methods factors) onto the three sources were fixed to be equal and the two performance factors were fixed to be uncorrelated. However, these constraints were needed for model identification purposes.

Multivariate Relationships Between the FFM Personality Traits and 360 Degree Managerial Performance Ratings

Our purpose in the following analyses was to determine if including additional rating sources (i.e., peers, subordinates, and self) beyond supervisors increased estimates of FFM validity. Thus, for the purpose of establishing a baseline with which to compare validity when using multisource ratings, we first tested a simple structural equations model that includes only supervisor ratings as criteria (i.e., similar to the model in Figure 3 but with only supervisor ratings as indicators of the managerial task and contextual performance factors). For this model, we fixed error variance of task and contextual performance using reliability information shown in Table 1 [variance \times (1 - reliability)]. Then, the FFM traits were added to the performance ratings model described above. When estimating true (operational) validity using LISREL, we did not correct for measurement errors in the FFM personality traits by using a single indicator for each personality trait without adjustment for error variance. Given that this model is statistically saturated, no overall fit indices were provided. As shown in Table 2, Extraversion ($r_c = .18$), Openness ($r_c = .16$), Emotional Stability ($r_c = .15$), and Conscientiousness ($r_c = .14$) were significantly related to task performance, whereas Extraversion ($r_c = .22$), Emotional Stability ($r_c = .20$), Conscientiousness $(r_c = .17)$, and Openness $(r_c = .15)$ were significantly related to contextual performance.

In the next model, we added subordinate and peer ratings to the supervisor ratings (i.e., this is the model depicted in Figure 3). The FFM traits were then added to the performance ratings model described above. When estimating true (operational) validity using LISREL, we did not correct for measurement errors in the FFM personality traits by using a single indicator for each personality trait without adjustment for error variance. This model fit the data well, $\chi^2(25) = 24.16$, p = .51, RMSEA = .000, SRMR = .038, CFI = 1.000, NFI = .985, NNFI = 1.001. Table 2 shows the

² We further tested a less restricted version of Model 5 (Model 8) wherein in addition to correlated performance factors, method factors were allowed to freely intercorrelate (so-called correlated trait-correlated method model), given that the intercorrelations among method factors (perspectives) were not zero (Conway, Lievens, Scullen, & Lance, 2004). These models are hierarchically nested with each other. Model 8 did not converge.

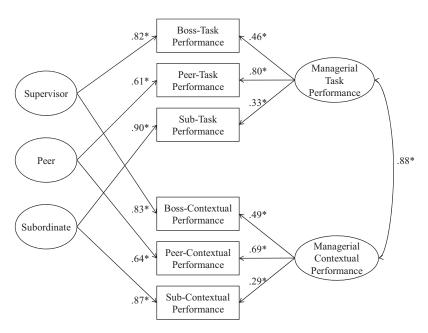


Figure 2. The structure of managerial performance measured via 360 degree ratings. Latent variables are scaled to unity. Disturbance terms are not shown for clarity. All factor loadings and path coefficients are completely standardized. Sub = subordinate. * p < .05.

operational (true) validities of FFM personality traits for task and contextual performance criteria estimated using supervisor, peer, and subordinate ratings. Openness ($r_c = .32$), Emotional Stability ($r_c = .31$), Extroversion ($r_c = .28$), and Conscientiousness ($r_c = .27$) were significantly related to task performance, whereas Emotional Stability ($r_c = .35$), Extraversion ($r_c = .32$), and Conscientiousness ($r_c = .27$) were significantly related with contextual performance. When peer and subordinate ratings were added,

operational validities increased between 36% and 88% compared with the results with supervisor ratings alone.

However, the results using supervisor plus peer plus subordinate ratings as criteria are somewhat different when we examine the standardized regression coefficients (see Table 3 and Figure 3). There are at least two important things to note about the regression results. First, the standardized regression weights were considerably smaller (absolute values in the .03–.27 range) than the oper-

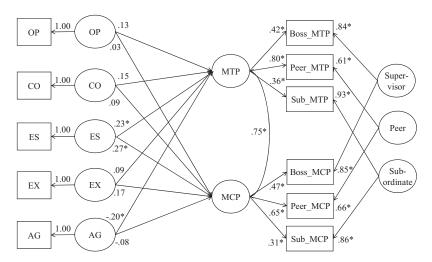


Figure 3. The relationships between the FFM personality traits and managerial performance. Disturbance terms and the intercorrelations among the FFM are not shown for clarity. All factor loadings and path coefficients are completely standardized. OP = Openness; CO = Conscientiousness; ES = Emotional Stability; EX = Extraversion; AG = Agreeableness; MTP = managerial task performance; Sub = subordinate; MCP = managerial contextual performance. * p < .05.

Table 2
Comparison of Operational (True) Validities of FFM Personality Traits for Managerial Performance

		Manageri	al task perform	iance		Managerial contextual performance					
	1	2	3			6	7	8			
Predictor	Supervisor ratings alone	Peer and sub ratings added	Self-ratings further added	2 vs. 1 (3 vs. 1)	5 3 vs. 2	Super-visor ratings alone	Peer and sub ratings added	Self-ratings further added	9 7 vs. 6 (8 vs. 6)	10 8 vs. 7	
Openness	.16*	.32*	.45*	50 (64)%	29%	.15*	.24*	.37*	38 (59)%	35%	
Conscientiousness	.14*	.27*	.38*	48 (63)%	29%	.17*	.27*	.39*	37 (56)%	31%	
Emotional Stability	.15*	.31*	.42*	52 (64)%	26%	.20*	.35*	.49*	43 (59)%	29%	
Extraversion	.18*	.28*	.43*	36 (58)%	35%	.22*	.32*	.50*	31 (56)%	36%	
Agreeableness	.01	.08	.23*	88 (96)%	65%	.06	.19*	.43*	68 (86)%	56%	
Adjusted R	.23*	.40*	.53*	43 (57)%	25%	.26*	.39*	.58*	33 (55)%	33%	

Note. N = 239. Operational (true) validity is the LISREL estimated correlation corrected for measurement error in the criterion measure. Columns 1 and 6 represent operational validity for performance rated by only the supervisor perspective; Columns 2 and 7 represent operational validity for performance rated by three perspectives (supervisor, peer, and subordinate); Columns 3 and 8 represent operational validity for performance rated by all four perspectives (supervisor, peer, subordinate, and self); Columns 4 and 9 represent the percentage of decrease in operational validity from Columns 2 to 1 (2 vs. 1) and from Columns 3 to 1 (3 vs. 1) and from Columns 7 to 6 (7 vs. 6) and from Columns 8 to 6 (8 vs. 6), respectively; Columns 5 and 10 represent the percentage of decrease in operational validity from Columns 3 to 2 (3 vs. 2) and from Columns 8 to 7 (8 vs. 7), respectively. Sub = subordinate.

* p < .05.

ational validities discussed above. This is a function of the especially high interrelationships between the FFM traits in the present data set, which limited each trait's unique variance when the other traits were held constant. For instance, although Openness had the highest operational validity for predicting task performance, Openness had a nonsignificant standardized regression weight. This is likely due to the high correlation between Openness and Extraversion (r = .60) in the current data; in fact, this is the highest correlation among all intercorrelations among the FFM traits. That is, the high correlation between Openness and Extraversion suppressed/lowered their unique relationships (expressed as standardized regression coefficients) with criterion ratings (Pedhazur, 1982). Thus, although operational validity of single FFM traits may be relatively high, the organization in the present study would likely be wasting resources if it selected applicants on the basis of all FFM traits, as their unique variance is limited and maximum predictability was achieved with one or two FFM traits.

Second, the regression results speak to the relative magnitude of the unique relationships that each FFM trait had with multisource performance ratings. Emotional Stability ($\gamma = .23$, SE = .11), lack of Agreeableness ($\gamma = -.20$, SE = .10), and Conscientiousness ($\gamma = .15$, SE = .09) were found to be the three most important predictors of task performance, whereas Emotional Stability ($\gamma = .27$, SE = .12), Extraversion ($\gamma = .17$, SE = .12), and Conscientiousness ($\gamma = .09$, SE = .10) were the three most important predictors of contextual performance.

In the third model, self-ratings were added to the supervisor, peer, and subordinate ratings. The inclusion of self-ratings for validation purposes is controversial, because common method variance may inflate validity estimates (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; see also Harris & Schaubroeck, 1988). However, it is likely not the case that self-ratings completely reflect an inflation bias, so we included self-ratings in the final model for informational purposes, as shown in Table 2. The

Table 3
Standardized Path Coefficients From FFM Personality Traits to Managerial Performance

			Manage	rial task p	erformanc	e		Managerial contextual performance						
Criterion	r_c	cr_c	γ	$SE(\gamma)$	LL	UL	$r_{c12.44}$	r_c	cr_c	γ	$SE(\gamma)$	LL	UL	$r_{c12.44}$
Predictor														
Openness	.32*	.28	.13	.10	03	.29	.20	.24*	.20	03	.10	20	.14	.19
Conscientiousness	.27*	.25	.15	.09	.00	.30	.17	.27*	.26	.09	.10	09	.26	.21
Emotional Stability	.31*	.27	.23*	.11	.05	.40	.19	.35*	.32	.27*	.12	.07	.47	.25
Extraversion	.28*	.25	.09	.10	07	.25	.23	.31*	.30	.17	.12	02	.36	.28
Agreeableness	.08	.07	20*	.10	36	05	.01	.19*	.21	08	.10	25	.09	.08
Adjusted R	.40*							.39*						

Note. N=239. $r_c=$ LISREL estimated correlation coefficients corrected for measurement error in the criterion measure (comparable to operational validities); $cr_c=$ unit-weighted composite correlation corrected for measurement error in the criterion measure using composite reliability; $\gamma=$ LISREL estimated standardized path (regression) coefficient; SE= standard error; LL (UL) = lower (upper) limit of the 95% confidence interval; $r_{c12.44}=$ operational validities that would be obtained if 12.44 supervisor ratings were used.

* p<.05.

addition of self-ratings to supervisor, peer, and subordinate ratings increased FFM operational validity between 26% and 65%, and the addition of self-ratings to supervisor ratings alone increased operational validity between 56% and 96%.

In order to ensure that the above results were not due to our methodological choice (the MTMM approach using CFA), we computed composite correlations between the three different ratings perspectives within each criterion and each of the FFM personality factors as well as composite reliabilities for each criterion following the procedures in Nunnally and Bernstein (1994). Then, we corrected the composite correlations for measurement error in the criterion using the estimated composite reliability to obtain composite operational validity. The composite operational validities are consistently a bit smaller (on average 11% and 5% for task and contextual performance each) than the corresponding operational validities estimated using the MTMM model; this is likely because such composite correlations assume unit weighting of each of the ratings perspectives, and the structural equation models do not. Regardless, the composite correlations showed the same pattern of results in relation to managerial performance (see Table 2) and were still substantially larger than typical correlations using single-source ratings. These results provided further support for the usefulness of personality in predicting managerial performance measured by 360 degree ratings.

Discussion

Summary of Findings

We investigated the usefulness of the FFM personality traits in predicting two related, yet distinct, aspects of managerial performance (task vs. contextual) assessed via a 360 degree performance rating system. We postulated that the use of 360 degree ratings would allow more comprehensive coverage of the job performance domain than would the use of supervisor ratings alone. As a result, we found that the operational (true) validity of personality increased substantially across all of the FFM personality traits when 360 degree ratings were used as criteria, as opposed to when supervisor ratings were used as criteria. This increase in validity remained even when self-ratings were excluded, suggesting the increase was not simply a function of same source bias. The benefit of the use of the 360 degree performance rating system was also apparent when the overall Rs across two performance rating models, supervisor ratings only and supervisor ratings plus peer and subordinate ratings, were compared. The overall R of the FFM increased from .23 to .40 (a 74% gain) for task performance and from .26 to .39 (a 50% gain) for contextual performance. The pattern of increases in R is overall analogous to that for operational validity. These greater validity estimates with peer and subordinate ratings included should not be too surprising, given that ratings from each of the perspectives correlated with at least some of the FFM traits and that ratings from the three perspectives had an average intercorrelation of only .22, meaning potential for incremental validity is high. These results suggest that low validities of FFM personality traits may be overcome when more diverse and comprehensive rating perspectives are used in measuring performance. This matter should be addressed in future research.

Regarding specific FFM traits, the true (operational) validities of Emotional Stability, Extraversion, Openness, and Conscien-

tiousness were found to be especially high across performance models. Each of these FFM traits had operational validities (with supervisor, peer, and subordinate ratings used as criteria) in the .24 – .35 range against both task and contextual performance. These findings strongly support previous meta-analytic findings linking each of these FFM traits to managerial performance and/or leadership effectiveness (Barrick & Mount, 1991; Barrick et al., 2001; Judge, Bono, Ilies, & Gerhardt, 2002), although the magnitudes of operational validities in the present study when multisource ratings were used were generally higher than the meta-analytic operational validities resulting from single-source supervisor ratings. The operational validity of Agreeableness also increased substantially when peer and subordinate ratings were added to supervisor ratings, although the operational validity of Agreeableness was still low to moderate. The univariate operational validity of Agreeableness was positive, but the relationship between Agreeableness and managerial performance reversed and became negative when managerial performance was simultaneously regressed on all of the FFM traits (see Table 3). This suggests that the small positive effect of Agreeableness on managerial performance was actually due only to shared variance between Agreeableness and the other FFM traits that positively predicted managerial performance. Once this shared variance was controlled, Agreeableness was significantly negatively related to task performance and was unrelated to contextual performance. This result suggests that Agreeableness may not always be a desirable trait for managers and substantiates meta-analytic evidence (Judge et al., 2002). Judge et al. (2002) even suggested that Agreeableness may not be a valid and positive predictor of leader effectiveness except in more getting alongfocused environments, such as community or charity institutes.

It is noteworthy that each of the FFM traits had sizable relationships with at least one aspect of managerial performance once peer and subordinate ratings had been added to supervisor ratings. As outlined in the Introduction, theoretical perspectives such as the socioanalytic approach (Hogan & Holland, 2003) have postulated links between each of the FFM traits and job performance, although empirical results have not always been supportive of these links. The results of the present study suggest that this failure of previous research to support the theoretical supposition that each of the FFM traits should be related to job performance may have been a function of reliance on single-source supervisor ratings. When multiple ratings sources were used to measure managerial performance, each of the FFM traits had sizable relationships with managerial performance, as suggested by the socioanalytic approach.

These relationships appear to be at least partially a function of performance ratings from different sources capturing unique variance in managerial performance. For instance, ratings from supervisors, peers, and subordinates were each meaningfully related to at least some of the FFM traits, but each had differing patterns of relationships with the FFM traits (see Tables 1 and 4). Further, these patterns of differential relationships across ratings sources do not appear to be completely random. The socioanalytic approach suggests that supervisors' managerial performance ratings would be influenced by getting ahead traits of managers and that subordinates' and peers' ratings would be influenced by getting along traits. Empirically, as shown in Table 4, supervisor ratings were significantly related to two getting ahead traits (Emotional Stability and Extraversion) and negatively related to a getting along trait

Table 4
Differential Relationships of FFM Personality Traits With Different Perspectives of Managerial Performance

Criterion		Ma	nagerial tas	sk performa	ance		Managerial contextual performance					
	Supervisor		Peer		Subordinate		Supervisor		Peer		Subor	dinate
	γ	$SE(\gamma)$	γ	$SE(\gamma)$	γ	$SE(\gamma)$	γ	$SE(\gamma)$	γ	$SE(\gamma)$	γ	$SE(\gamma)$
Predictor												
Openness	.03	.08	.14*	.08	.00	.08	04	.08	.03	.08	09	.08
Conscientiousness	.07	.08	.11	.08	.06	.08	.07	.08	.03	.08	.02	.08
Emotional Stability	.12	.09	.19*	.09	01	.09	.17*	.09	.17*	.09	02	.09
Extraversion	.13	.09	.05	.08	03	.09	.17*	.09	.06	.09	.02	.09
Agreeableness	16*	.08	16*	.08	.03	.08	14*	.08	03	.08	.14*	.08
Adjusted R	.23*		.33*		.06*		.27*		.23*		.14*	

Note. N = 239. FFM = five-factor model; $\gamma = \text{LISREL}$ estimated standardized path (regression) coefficient; SE = standard error. * p < .05.

(Agreeableness). Subordinate ratings, on the other hand, were meaningfully related only to Agreeableness, a getting along trait, and two of the three significant correlates of peer ratings were getting along traits (Emotional Stability and Agreeableness). The fact that the differential relationships appear to be, at least partially, a function of the propositions of the socioanalytic perspective further supports the idea that ratings from each rating source are capturing important, unique managerial performance variance. Thus, a clear implication of the present study is that researchers investigating the validity of personality traits should pay close attention to the potential that criterion problems, such as criterion deficiency, have to confound validity estimates. Supervisor ratings were not the only valid source of performance information in the present study, and this is not surprising given the multilevel nature of managerial performance. If the present study had relied on single-source supervisor ratings of performance, a bleaker picture of the validity of FFM personality traits would have emerged.

We note that an examination of the correlational results in Table 1 and the multivariate results in Table 4 suggests that the gains in operational validity were actually due more to the addition of peer ratings than of subordinate ratings. That is, the average bivariate correlations between the FFM traits and peer ratings (.16-.19, depending on whether task vs. contextual performance was the criterion) were slightly higher than those of supervisor ratings (.12–.15), and the average correlations for subordinate ratings were quite low (.03-.04). The one exception was with the FFM of Agreeableness, where the gain in validity against contextual performance was likely driven by the relationship between Agreeableness and subordinate performance ratings. Whether the unique contribution of peer versus subordinate ratings is idiosyncratic to the present study is a question for future research. Regardless, the finding that peer ratings were able to capture such a sizable amount of unique performance variance is important and supports the idea that adding ratings sources beyond supervisors may enhance both performance measurement and personality validity estimation.

We found a similar pattern of results using standardized path (regression) coefficients relative to those found using operational validities. It is noted that, if one's interest is in determining the operational validity of any single FFM trait, sole reliance on standardized path coefficients is less informative, due to high intercorrelations of the FFM traits in the current data set (Dilchert

& Ones, 2008; Saucier, 2002). That is, the standardized path coefficients speak more to the relative importance of each of the FFM traits when the others are held constant (i.e., unique relationships) and less to the operational validity of any single trait. Thus, the most important implications of the regression results pertain to relative magnitudes of FFM validity. According to the regression results (see Table 2), in terms of predicting task performance, Emotional Stability was the best predictor, followed by (lack of) Agreeableness and Conscientiousness, in that order. In terms of predicting contextual performance, Emotional Stability was the best predictor, followed by Extraversion. These results suggest that the organization in the present study would likely be wasting resources if it used all five of the FFM traits to predict managerial performance, as only some of the FFM traits had unique relationships once the other FFM traits had been accounted for.

Theoretical Implications

There have long been calls in the organizational sciences for more attention to be paid to the effects of criterion attributes on study conclusions (e.g., Austin, 1964; Campbell et al., 1993). Our results suggest that part of the reason for self-report personality's low-to-moderate criterion-related validity may be an overreliance on single-source supervisor ratings of performance. This finding highlights the idea that before our field accepts the pessimistic view of self-report personality validity put forth by some (e.g., Morgeson et al., 2007), more attention must be paid to alternative explanations for such low validity. That is, it may indeed be the case that the low-to-moderate validity is due to problems with the way that we measure personality. However, before we can accept such a conclusion, other study factors that could influence criterion-related validity must be ruled out. A half-century ago, Thorndike noted the importance of this issue by arguing that "the most fundamental and most difficult problem in any selection research program is to obtain satisfactory criterion measures of performance on the job against which to validate selection procedures" (1949, p. 199). Our study demonstrates that one such influence on validity that clearly requires further exploration is criterion deficiency.

Our results also converge with and support some of the propositions put forth in Hogan and Holland's (2003) socioanalytic

theory. That is, the propositions of the socioanalytic approach to the relationships between personality and performance suggest that ratings from different perspectives (e.g., supervisors vs. peers) have the potential to each capture unique performance variance. To the degree that this is true, adding rating sources beyond supervisors will result in a less deficient and more predictable criterion. As discussed above, our results suggested that this was indeed the case, as (a) each of the rating perspectives had differential patterns of relationships with the FFM traits; (b) these differential patterns were similar to the predictions of the socioanalytic approach regarding relationships between getting ahead versus getting along traits and performance ratings from supervisors versus peers versus subordinates (i.e., getting ahead traits related to supervisor ratings, whereas getting along traits related to peer and subordinate ratings); and (c) adding peer and subordinate ratings to supervisor ratings significantly increased estimates of FFM validity. Our results support the idea that multisource performance ratings have the potential to represent a less deficient criterion. There is a need for further research using multisource ratings in different settings and with different predictors to substantiate and extend our findings.

Our results also have implications for whether differences between sources/perspectives in ratings should be conceptualized as error. This is a key issue when using interrater reliability coefficients as estimates of the reliability of performance ratings, as a crucial assumption in the use of interrater reliabilities is that any differences in ratings between raters are due to random error. That is, the interrater reliability coefficient is essentially the average intercorrelation between ratings from two or more raters and any unshared variance between raters is treated as measurement error; so, if there are systematic differences between raters' ratings and these differences reflect true performance variance (e.g., as a function of supervisors being exposed and attending to a different set of ratee performance behaviors than are subordinate raters), interrater reliabilities will underestimate the reliability of ratings. This is especially problematic when validity coefficients are corrected using interrater reliabilities, as the lower the reliability, the larger the correction for attenuation. Our data were structured such that we had only mean peer and subordinate ratings and had ratings from only one supervisor for each manager, and so we have no estimate of within-rating-source variability in our data set. Thus, our results cannot speak directly to differences between raters within sources (e.g., differences between two supervisors' ratings). However, our results suggest that differences between raters across sources (e.g., differences between supervisor and peer ratings) cannot necessarily be considered purely a function of measurement error. That is, ratings from different rating sources each appeared to capture some unique performance variance, likely because raters from different sources were either more exposed to or tended to attend to different managerial performance behaviors. Thus, on the basis of the results of our study, when interrater reliability is estimated using correlations between ratings from raters across multiple sources (e.g., supervisors, peers, and subordinates), correcting validity estimates for measurement error using interrater reliability should not be done blindly. Those wishing to correct using such interrater reliabilities should exercise caution by first examining the degree to which rating scores from different rating sources capture unique performance variance, using the techniques employed in our study. One should assume that differences between ratings across sources are a function of measurement error and correct validity coefficients using interrater reliabilities if ratings from the different sources do not reflect unique performance variance.

Practical Implications

Two practical implications of the present study stand out in particular. First, in light of our results, the low-to-moderate levels of criterion-related validity typically found for self-report personality must be rethought. At the narrowest level, organizational researchers and practitioners validating FFM personality tests against supervisor ratings criteria must consider whether organizations would be provided with a more accurate estimate of self-report personality validity if additional rating sources were added. At a broader level, organizational researchers and practitioners should consider the degree to which other criterion attributes (e.g., other forms of deficiency, criterion contamination) may be affecting self-report personality validity.

Second, if organizational researchers or practitioners are considering using 360 degree ratings as criteria in validation research (for personality or other predictors), our results demonstrate that special care should be taken when choosing which rating sources to include. For one, although subordinate ratings were related to Agreeableness, they generally appeared to be much less related to the FFM traits than did peer ratings (see the adjusted R for each source at the bottom of Table 4). Thus, if future research continues to find that peer ratings are more predictable than subordinate ratings, organizations may wish to consider not including subordinate ratings for use in validation criteria in an effort to maximize the cost effectiveness of their validation tools. The inclusion of self-ratings of performance for validation purposes is probably an even larger concern. Self-ratings of performance are typically included in 360 degree ratings. Our results (see Table 2) suggest that including self-ratings as criteria greatly increases estimates of FFM validity, but it was not possible for us to determine the degree to which this is a function of self-ratings capturing additional true performance variance, common method variance, or both. This is not to suggest that self-ratings are worthless. They likely have developmental benefits for the manager, and there is likely some true performance variance captured by self-ratings. However, we view the common method variance issue as particularly troubling for validation research and do not encourage researchers and practitioners to include self-ratings in multisource performance rating criteria for validation and evaluative (vs. developmental) purposes.

Additional Issues, Limitations, and Directions for Future Research

One noteworthy limitation is that we used aggregated ratings (means) for both peer and subordinate perspectives. Mount et al. (1998, p. 573) argued and found that method variance is more strongly associated with individual raters than with their level (perspectives). We could not fully account for the source of method variance associated with individual raters' idiosyncrasies because we had available to us only mean peer and subordinate ratings instead of individual peer and subordinate ratings. Future studies should be more carefully designed to address this topic.

Thus, our estimates of the personality-performance relationship may be on the conservative side due to undercorrection for this individual rater-related error variance.

An additional possible concern is that the increased usefulness of 360 degree performance ratings may simply be due to the increase in reliability from the greater number of raters rather than the incremental validities of unique perspectives from rating sources. To assess the degree to which validity gains from 360 degree ratings were simply a function of reliability, we first used the Spearman–Brown formula to estimate what the reliability of performance ratings would be if 12.44 supervisors had rated each manager, instead of only one supervisor (as was the case in the present study). We used 12.44 as the number of raters, as this was the average number of raters (peer, subordinate, and supervisor but excluding self-ratings) who rated each manager in our study. Because only one supervisor rated each manager, we used Viswesvaran, Ones, and Schmidt's (1996) meta-analytic value of .52 as our estimate of single-supervisor-rating reliability in the present study which would increase to a reliability of .93 if 12.44 supervisors rated instead of just one. We note that some (e.g., Murphy & DeShon, 2000) have made the case that interrater reliability underestimates reliability of performance ratings; to the degree this is true, the analyses we present using .52 will actually overestimate the reliability effect in our present study. Correcting singlesupervisor-rating FFM validity estimates (i.e., bivariate correlations in Table 1) for attenuation using reliability of .52 and then multiplying the corrected validity by the square root of .93 estimates what the FFM operational validities would be if 12.44 supervisors had provided ratings. This step documents the validity gains to be expected through an increase in reliability alone (see Schmidt & Zimmerman, 2004, for more details). As can be seen in the last column of Table 3, the estimates of FFM operational validities with 12.44 supervisor ratings were considerably smaller than when 11.44 peer and subordinate ratings (6.75 peers and 4.69 subordinates) are added to the single supervisor rating (1). Thus, it is clear that our results are not simply a function of increased reliability of ratings.

An important caveat for our study is that our data did not come from a personnel selection setting. As is often the case with 360 degree performance ratings, ratings in our study were collected for the developmental benefit of incumbent managers. Thus, the applicability of our findings for personnel selection requires future research. However, we note that validation research commonly uses concurrent validity designs, such as in the present study. It is also considered crucial in validation research to use criterion ratings that were not collected for the purpose of organizational decision making (e.g., raises, promotions) because such administrative ratings usually lack variance, and our 360 degree developmental purpose ratings meet the standard of not having been used for administrative purposes. Also, because the 360 degree ratings were used only for developmental purposes, and not for administrative purposes, the chances are improved that respondents provided fairly honest ratings. Additionally, Ellingson, Sackett, and Connelly (2007) compared personality test responses in developmental versus selection contexts and found limited degrees of response distortion. Therefore, although we feel the above caveat regarding implications for selection is warranted, we also feel that the concurrent and developmental nature of our 360 degree ratings lent some strengths to our study and that our results have strong implications for choosing performance rating sources in personnel selection system validation research.

We feel we should mention that 360 degree ratings can be used for purposes other than criterion ratings. In fact, Mount, Barrick, and Strauss (1994) found that observer ratings, rather than selfreports, of personality were more valid when they measured personality using 360 degree ratings; their study did not measure job performance using 360 degree ratings. The present study measured job performance, but not personality, using 360 degree ratings. Thus, it would be informative to relate both personality traits and performance using 360 degree ratings (although common method variance would be an obvious concern in such a design). Such a study could address which source of personality traits is more predictive of which source of performance ratings and could show under what condition the maximum validity of personality is achieved. We think that this type of research will further enrich our understanding of the personality and performance linkage in a predictor-criterion-balanced way.

Finally, the present study used the same data set as was used in Berry et al. (2007), so we should discuss the implications of Berry et al.'s results for those of the present study. Essentially, Berry et al. found that when Self-Deceptive Enhancement scale scores (which reflect a relatively unconscious form of response distortion on self-report measures; Paulhus, 1998) were accounted for in regression equations, the criterion-related validity of Extraversion and Emotional Stability increased. Thus, Berry et al.'s results dealt most directly with concerns over response distortion in self-report personality. Put differently, Berry et al. made the case that response distortion issues with the predictors (i.e., self-report personality scales) affected conclusions about FFM personality validity. The present study, on the other hand, makes the case that criterion deficiency issues (resulting from the use of single-source supervisor performance ratings) affect conclusions about FFM personality validity. As Berry et al. dealt with the predictor domain and the present study dealt with the criterion domain, the results found in the present study are conceptually distinct from those of Berry et al.'s. Actually, from the point of view of optimal estimation of FFM personality validity, the results of Berry et al. are likely incremental to the results of the present study, such that if we also accounted for Self-Deceptive Enhancement, we might find even greater validity gains in the present study. Note, however, that this has no bearing on the present study's conclusion that part of the reason for low FFM validity may be due to the use of deficient criterion measures.

Conclusion

Both researchers and practitioners are interested in valid predictors of managerial performance. Personality has been studied as one of the predictors and at the same time has often encountered skepticism for its low validity. Our results demonstrate that estimates of the validity of personality for predicting managerial performance are greater when managerial performance is assessed from multiple perspectives by utilizing the 360 degree performance rating system. Far from being a source of nonmeaningful error variance, the discrepancies among ratings from multiple perspectives can in fact capture meaningful variance in multilevel managerial performance when they are analyzed with a methodological technique such as the present study's MTMM-based

structural equation modeling methods. In sum, applied researchers interested in the validity of personality should realize that criterion attributes (both positive and negative) can affect estimates of validity as much as predictor attributes do (Austin, 1964). Overall, our study contributes to the evidence that personality is useful in predicting managerial performance.

References

- Austin, A. W. (1964). Criterion-centered research. Educational and Psychology Measurement, 24, 807–822.
- 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., Mount, M. K., & Judge, T. A. (2001). Personality and job performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9–30.
- Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. (2001). Evaluation of 360 degree feedback ratings: Relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior*, 22, 775–788.
- Bentler, P. M. (2007). Can scientifically useful hypotheses can be tested with correlations. *American Psychologist*, 62, 772–782.
- Berry, C. M., Page, R. C., & Sackett, P. R. (2007). Effects of self-deceptive enhancement on personality-job performance relationships. *Interna*tional Journal of Selection and Assessment, 15, 94–109.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349–381). San Francisco, CA: Jossey-Bass.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187–215.
- Borman, W. C. (1997). 360° ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review*, 7, 299–315.
- Borman, W. C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6, 1–21.
- Campbell, J. P., Dunnette, M. D., Lawler, E. E., III., & Weick, K. E., Jr. (1970). Managerial behavior, performance, and effectiveness. New York, NY: McGraw-Hill.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel* selection in organizations (pp. 35–70). San Francisco, CA: Jossey-Bass.
- Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology*, 31, 187–212.
- Conway, J. M. (1999). Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 84, 3–13.
- Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10, 331–360.
- Conway, J. M., Lievens, F., Scullen, S. E., & Lance, C. E. (2004). Bias in the correlated uniqueness model for MTMM data. *Structural Equation Modeling*, 11, 535–559.
- Conway, J. M., Lombardo, K., & Sanders, K. C. (2001). A meta-analysis of incremental validity and nomological networks for subordinate and peer ratings. *Human Performance*, 14, 267–303.
- Craig, S. B., & Hannum, K. (2006). Research update: 360-degree performance assessment. Consulting Psychology Journal: Practice and Research, 58, 117–122.
- Dilchert, S., & Ones, D. S. (2008). Personality predictors of managerial salary. Zeitschrift für Personalpsychologie, 7, 1–23.

- Ellingson, J. E., Sackett, P. R., & Connelly, B. S. (2007). Personality assessment across selection and developmental contexts: Insights into response distortion. *Journal of Applied Psychology*, 92, 386–395.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer–supervisor ratings. *Personnel Psychology*, 41, 43–62.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88, 100–112.
- Hogan, R. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, 2nd ed., pp. 327–396). Palo Alto, CA: Consulting Psychologists Press.
- Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. Human Performance, 11, 129–144.
- Hough, L. M. (1992). The "big five" personality variables—construct confusion: Description versus prediction. *Human Performance*, 5, 139– 155.
- Johnson, J. W. (2003). Toward a better understanding of the relationship between personality and individual job performance. In M. R. Barrick & A. M. Ryan (Eds.), *Personality and work: Reconsidering the role of personality in organizations* (pp. 83–120). San Francisco, CA: Jossey-Bass.
- Judge, T. A., Bono, J. R., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87, 765–780.
- Lance, C. E., Baxter, D., & Mahan, R. P. (2006). Evaluation of alternative perspectives on source effects in multisource performance measures. In W. Bennett, C. E. Lance, & D. J. Woehr (Eds.), *Performance measurement: Current perspectives and future measurement* (pp. 49–76). Mahwah, NJ: Erlbaum.
- LeBreton, J. M., Burgess, J. R. D., Kaiser, R. B., Atchley, E. K., & James, L. R. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods*, 6, 80–128.
- LeBreton, J. M., & Senter, J. L. (2008). Answers to twenty questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11, 815–852.
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. Multivariate Behavioral Research, 39, 327–356.
- Lim, B.-C., & Ployhart, R. E. (2006). Assessing the convergent and discriminant validity of Goldberg's International Personality Item Pool. Organizational Research Methods, 9, 29–54.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93, 568–592.
- 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.
- Morgeson, F. P., Mumford, T. V., & Campion, M. A. (2005). Coming full circle: Using research and practice to address 27 questions about 360degree feedback programs. Consulting Psychology Journal: Practice and Research, 57, 196–209.
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality dimensions. *Journal of Applied Psychology*, 79, 272–280.
- Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology*, 51, 557–576.
- Murphy, K. R., & Cleveland, J. N. (1995). Understanding performance appraisal: Social, organizational, and goal-based perspectives. Thousand Oaks, CA: Sage.
- Murphy, K. R., & DeShon, R. P. (2000). Inter-rater correlations do not

estimate the reliability of job performance ratings. *Personnel Psychology*, 53, 873–900.

- Nilsen, D. (1995). An investigation of the relationship between personality and leadership performance. Unpublished doctoral dissertation, University of Minnesota, Twin Cities Campus.
- Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.).
 New York, NY: McGraw-Hill.
- Ones, D. S., Hough, L., & Viswesvaran, C. (2005, October). Personality of executives and managers: Career interests, choice, progression and success. Paper presented at the consortium of the Society for Industrial and Organizational Psychology Fall, St. Louis, MO.
- Page, R. C. (2007). Work Behavior Inventory user's manual. Minnetonka, MN: Human Resource Consultants.
- Paulhus, D. L. (1998). Paulhus Deception Scales: The Balanced Inventory of Desirable Responding–7 user's manual. North Tonawanda, NY: Multi-Health Systems.
- Pedhazur, E. J. (1982). Multiple regression in behavioral research: Explanation and prediction (2nd ed.). New York, NY: Holt, Rinehart & Winston.
- Ployhart, R. E., Weekly, J., & Ramsey, J. (in press). The consequences of human resource stocks and flows: A longitudinal examination of unit service orientation and unit effectiveness. Academy of Management Journal.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003).
 Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879–903.
- Sala, F., & Dwight, S. A. (2002). Predicting executive performance with

- multirater surveys: Whom you ask makes a difference. Consulting Psychology Journal: Practice and Research, 54, 166–172.
- Saucier, G. (2002). Orthogonal markers for orthogonal factors: The case of the Big Five. *Journal of Research in Personality*, 36, 1–31.
- Schmidt, F. L., & Zimmerman, R. D. (2004). A counterintuitive hypothesis about employment interview validity and some supporting evidence. *Journal of Applied Psychology*, 89, 553–561.
- Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85, 956–970.
- Tett, P. R., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, 88, 500–517.
- Tett, R. P., Guterman, H. A., Bleier, A., & Murphy, P. J. (2000). Development and content validation of a "hyperdimensional" taxonomy of managerial competence. *Human Performance*, 13, 205–251.
- Thorndike, R. L. (1949). Personnel selection. New York, NY: Wiley.
- Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, 81, 525–531.
- Viswesvaran, C., & Ones, D. S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment*, 8, 216–226.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574.
- Widaman, K. F. (1985). Hierarchically nested covariance structures models for multitrait–multimethod data. Applied Psychological Measurement, 9, 1–26.

Appendix

The WBI Scales and Definitions

Emotional Stability

Emotional control. Tendency to maintain composure and avoid emotional outbursts, not show anger or lose one's temper, respond calmly when provoked, and avoid aggressive behavior in difficult situations.

Self-confidence. Tendency to show confidence in one's ability to be successful, be self-assured and optimistic in new or challenging situations, be confident of one's future, and not be intimidated by people or situations.

Resilience. Tendency to work calmly and effectively in stressful situations, accept criticism and not have feelings easily hurt, not worry about problems, and quickly recover from disappointments.

Extraversion

Persuasiveness. Tendency to persuade and change the opinions of others, sell ideas or products, negotiate and debate issues, and deploy strategies to influence individuals or groups.

Leadership. Tendency to seek opportunities to lead and direct others, assert oneself over others, inspire others to achieve goals, and get others to do more than they expected to do.

Sociability. Tendency to build work relationships with a wide range of individuals, be approachable and easy to communicate with, quickly build rapport with individuals, and create a network of contacts.

Energy. Tendency to work quickly and energetically, handle simultaneous tasks, prefer to work in a fast-paced environment, and sustain a rapid mental or physical pace over extended periods of time.

Openness

Flexibility. Tendency to adjust quickly to new goals and work procedures, embrace new ways of doing things, adjust to constant change, and work effectively where there is a great deal of uncertainty.

Innovation. Tendency to generate new and creative ideas, approach problems from a fresh perspective, have a vivid imagination, and offer original thought to arrive at inventive solutions.

Analytical thinking. Tendency to gather facts and information, systematically use logic and analysis, anticipate problems and potential solutions, and identify connections or patterns in complex issues.

Independence. Tendency to make decisions without the guidance of others, figure out things for oneself, work independently with little or no supervision, and make decisions on one's own and accept the consequences.

Agreeableness

Teamwork. Tendency to collaborate with others, assist and support team members and goals, help build group morale, and actively work to resolve conflict within teams.

Diplomacy. Tendency to interact with individuals with tact and courtesy, not offend others, trust new acquaintances, and be polite and respectful with difficult people.

Customer service. Tendency to be sensitive to others' needs and feelings, be understanding and helpful to others, be compassionate in providing advice, and provide high levels of service to others.

Conscientiousness

Attention to detail. Tendency to be careful and thorough in the details of work, carefully check work products for accuracy and quality, be organized and tidy, and maintain careful records.

Achievement drive. Tendency to set challenging goals, define standards and measure performance against them, constantly work to improve performance, and strive to exceed standards of performance.

Initiative and persistence. Tendency to take action and start things without being asked, do more than is typically required, volunteer to take on new assignments, and assume the risks that go with additional responsibilities.

Dependability. Tendency to be reliable and dependable, meet attendance standards and deadlines, avoid risks that may lead to accidents, and refrain from impulsiveness that may hinder reliability or performance.

Note. From "Effects of Self-Deceptive Enhancement on Personality–Job Performance Relationships," by C. M. Berry, R. C. Page, and P. R. Sackett, 2007, International Journal of Selection and Assessment, 15, p. 100. Copyright 2007 by Ronald Page. Adapted with permission.

Received March 10, 2008
Revision received June 15, 2009
Accepted July 15, 2009