MULTISOURCE PERFORMANCE APPRAISAL AND PERSONALITY: A VIEW FROM THE DARK SIDE

Blaine Gaddis, Ph.D.
Jeff Foster, Ph.D.
Hogan Assessment Systems

We examine the effects of "dark side" personality dimensions on various dimensions of job performance as rated by supervisors, peers, direct reports, and "others" (e.g., customers). Results illustrate the complex effects of personality on performance and reveal both consistencies and striking differences in perceptions of performance across rater groups.

The literature on multisource performance appraisal, or "360 performance appraisals," (Conway & Huffcutt, 1997) is well established. These methods are used for a variety of applications such as predicting and assessing individual performance (Atwater & Yammarino, 1992), employee development (Dalesio, 1998), and predicting and assessing team performance (Reilly & McGourty, 1998). However, some still question the utility of 360's because of issues concerning rater errors and biases (Hauenstein, 1998; Kane & Lawler, 1978; Pulakos, 1997) and a perceived lack of difference between ratings provided by different parties (Yammarino & Atwater, 1997). Nevertheless, if used prudently, multisource ratings provide organizations with a means of assessing performance on job-relevant competencies for a variety of purposes (Catano, Darr & Campbell, 2007; Stoker & Van der Heijden, 2001).

Considering the value of multisource performance appraisal, researchers have made efforts to predict performance on job-relevant competencies by identifying constructs bearing significant relationships with those ratings. Such constructs include cognitive processes (Landy & Farr, 1980; Yammarino & Atwater, 1997), attitudes and beliefs (Tziner, Murphy & Cleveland, 2002), and personality measures such as conscientiousness (Roberts & Hogan, 2001) and agreeableness (Ones & Viswesvaran, 2001). Although efforts at linking personality measures to 360 ratings have traditionally focused on Five-Factor Model personality factors (McCrae & Costa, 1987), other more specific personality facets may also show significant relationships with multisource ratings of performance. As an alternative, we offer a different perspective on the personality, focusing on facets of personality that may derail or otherwise inhibit an individual's career progression (Dotlich & Cairo, 2003). Specifically, we use the Hogan Development Survey (R. Hogan & J. Hogan, 2009) to measure these potentially negative personality dimensions.

The Hogan Development Survey

At a fundamental level, the HDS is rooted in socioanalytic theory. Grounded in both interpersonal (Carson, 1969; Leary, 1957; Wiggins, 1979) and evolutionary (Barrett, Dunbar, & Lycett, 2002; Dawkins, 1976) psychology, socioanalytic theory argues that, as group-living animals, humans evolved strategies for maximizing individual and group survival. All groups exist in status hierarchies, and interaction within groups involves negotiating for acceptance and status. As such, people are motivated both to "get along" (maximize popularity) and to "get ahead," (maximize status relative to other members of the group). Even when individuals

are satisfied with their position in the status hierarchy, they still must "get along" with others in order to maintain their position within the group. In this context, personality characteristics are dysfunctional if they interfere with an individual's capacity to get along or get ahead. These dysfunctional dispositions impairing a person's ability to get along or get ahead reflect the "dark side" of personality (cf. Conger, 1990).

The HDS fills a unique need in psychological assessment because, despite technical competence in work-related skills, talented people sometimes fail. Various factors can limit peoples' ability to perform consistently at a high level. Then, difficulties at work appear as problematic behaviors in the interpersonal context that defines the workplace. These problematic behaviors arise from any of a number of common dysfunctional dispositions. These dispositions (a) reflect people's distorted beliefs about themselves, how others will treat them, and the best means to achieve their personal goals, and (b) negatively influence people's careers and life satisfaction. The Hogan Development Survey (HDS; R. Hogan & J. Hogan, 2009) assesses eleven of these dysfunctional personality dimensions. The HDS can be used both as an element of the personnel selection process and as a source of developmental information for helping employees gain "strategic self-awareness" about how to improve their job performance and relationships with others at work. It is important to emphasize that the HDS is not designed to measure personality disorders—the personality disorders are manifestations of mental disorder, and we are assessing self-defeating expressions of normal personality.

The HDS contains 168 items in the form of statements to which a respondent indicates "agree" or "disagree" (0 = disagree, 1 = agree). Each scale contains 14 items rationally derived from its distinguishing syndrome feature; scale scores range from 0 to 14. Items are scored so that higher scores represent more dysfunctional tendencies. There is no item overlap among the 11 scales. The items were screened for content that might seem offensive or to invade privacy. There are no items concerning sexual preferences, religious beliefs, criminal or illegal behavior, racial/ethnic attitudes, or attitudes about disabled individuals. Table 1 presents the 11 HDS scales, their definitions, and an example item for each scale.

Table 1. HDS Scales, Definitions, and Example Items

Scale	Definition	Example Item
Excitable	Being moody and inconsistent, showing enthusiasm about new persons or projects and then becoming disappointed with them	My mood can change quickly
Skeptical	Being cynical, distrustful, overly sensitive to criticism, and questioning others' true intentions	There are few people I can really trust
Cautious	Being resistant to change and reluctant to take even reasonable chances for fear of being evaluated negatively	It is difficult for me to be assertive
Reserved	Being socially withdrawn and lacking interest in or awareness of the feelings of others	I prefer spending time by myself
Leisurely	Being autonomous, indifferent to other people's requests, and becoming irritable when they persist	I ignore people who don't show respect
Bold	Being unusually self-confident and, as a result, unwilling to admit mistakes or listen to advice, and unable to learn from experience	I do most things well
Mischievous	Enjoying taking risks and testing the limits	I have few regrets

Table 1. HDS Scales, Definitions, and Example Items (Continued)

Scale	Definition	Example Item
Colorful	Being expressive, dramatic, and wanting to be noticed	Other people pay attention to me
Imaginative	Acting and thinking in creative and sometimes unusual ways	I am creative about my appearance
Diligent	Being careful, precise, and critical of the performance of others	I take pride in organizing my work
Dutiful	Being eager to please, reliant on others for support, and reluctant to take independent action	I leave the big decisions up to others

In the 1997 HDS manual (R. Hogan & J. Hogan, 1997), a principal components analysis of the 11-scale correlation matrix resulted in the extraction of three "global" components which accounted for 62% of the variance in the matrix. The Excitable, Skeptical, Cautious, Reserved, and Leisurely scales defined the first component, which resembled the theme of "moving away from people" in Horney's (1950) model of flawed interpersonal tendencies. The Bold, Mischievous, Colorful, and Imaginative scales comprised the second component, corresponding to Horney's (1950) theme of "moving against people." Finally, the third component, corresponding to Horney's (1950) theme of "moving toward people," included the Diligent and Dutiful scales. As such, these three components define the HDS global factors presented in Table 2.

Table 2. HDS Global Factors, Definitions, and Associated Scales

Global Factor	Definition	Associated Scales
Moving Away	Avoiding others	Excitable, Skeptical, Cautious, Reserved, Leisurely
Moving Against	Dominating and intimidating others	Bold, Mischievous, Colorful, Imaginative
Moving Toward	Building alliances to minimize threats	Diligent, Dutiful

As previously noted, efforts combining personality measures and multisource performance appraisal traditionally focus on Five-Factor Model personality factors (McCrae & Costa, 1987). As a goal, these efforts focus on identifying the personality characteristics most beneficial to performance as rated by supervisors, peers, direct reports, and other parties. As an alternative, we decided to analyze the impact of "dark side" personality characteristics on 360 performance ratings to identify the personality characteristics most adverse to performance as rated by these parties. Moreover, we compared these other ratings to self ratings of performance to investigate whether individuals and those around them agree about the impact of these potentially negative personality characteristics on performance. Our findings show that, in some areas, consensus exists about the impact of these personality dimensions on performance. More interestingly, however, is our finding that disagreement also exists about the impact of some of these personality characteristics on performance. Specifically, while the individual may find a particular personality dimension beneficial to performance, others around him or her may not.

Methods

To examine relationships between negative aspects of personality and multisource performance ratings, we identified studies in the Hogan archive containing both Hogan Development Survey data (HDS; R. Hogan & J. Hogan, 2009) and 360 performance appraisal ratings. The Hogan archive contains data from over 280 criterion studies examining relationships between personality measures and job performance. The Hogan Development Survey is comprised of 11 scales defining areas where, under periods of heavy stress, behaviors emerge that interfere with a person's ability to build relationships and create cohesive, goal-oriented teams (Hogan, Hogan & Warrenfeltz, 2007). With HDS and multisource performance appraisal data in hand, we examined correlations between HDS scale scores and self, supervisor, peer, direct report, and "other" ratings provided for each job-relevant competency.

First, we identified 29 independent samples (total N = 3,146) from published articles, chapters, technical reports, and dissertations between 1997 and 2008 that were catalogued in the Hogan archive. The studies met the following criteria: (a) they used job analysis to estimate personality-based job requirements; (b) they used a concurrent (k = 14) or predictive (k = 15) validation strategy with working adults; (c) the criteria were content explicit, not just overall job performance, and these were classified reliably by Subject Matter Experts (SMEs); and (d) HDS scales served as predictor variables. We excluded studies using: (a) clinical patients and therapists; (b) undergraduate or graduate students; (c) self-reported performance criteria; (d) performance criteria other than ratings; (e) only an overall performance criterion; (f) laboratory or assessment center studies; and (g) studies unrelated to work contexts. Studies compiled for the meta-analysis included occupational samples of managers, executives, and professionals. The studies included are empirical validation studies with either supervisor ratings as criteria or 360 degree feedback ratings as criteria. All data were collected in the English language, although some studies were completed using data collected outside of the US. Before conducting meta-analyses, we categorized performance ratings criteria using dimensions in the Competency Evaluation Tool (CET; Hogan Assessment Systems, 2008) as an organizing framework.

We used meta-analytic procedures specified by Hunter and Schmidt (2004) to analyze results across studies and assess effect sizes. All studies used zero-order product-moment correlations, which eliminated the need to convert alternative statistics to values of r. Corrections were made for sampling error, criterion unreliability, and range restriction. Although some researchers (e.g., Murphy & De Shon, 2000) argue against the use of rater-based reliability estimates, we followed procedures outlined by Barrick and Mount (1991) and Tett et al. (1991), and used the .508 reliability coefficient proposed by Rothstein (1990) as the estimate of the reliability of supervisory ratings of job performance. Also, we computed a range restriction index for HDS scales. Following procedures described by Hunter and Schmidt (2004), we divided each HDS scale's within-study standard deviation by the scale's standard deviation reported in the HDS manual (R. Hogan & J. Hogan, 2009). This procedure produced an index of range restriction for each HDS scale within each study, and we used these values to correct each predictor scale for range restriction.

Hunter and Schmidt (2004) point out that meta-analytic results can be biased unless each sample contributes about the same number of correlations to the total. To eliminate such bias,

we averaged correlations within studies so that each sample contributed only one point estimate per predictor scale. For example, if more than one criterion from any study was classified as "Leadership," the correlations between each predictor scale and those criteria were averaged to derive a single point estimate of the predictor-criterion relationship. Note that this procedure uses both negative and positive correlations rather than mean absolute values for averaging correlations. This is the major computational difference between the current analyses and those presented by Tett et al. (1991, p. 712). We did not correct correlation coefficients to estimate validity at the construct level. Although some (e.g., Mount & Barrick, 1995; Ones, Schmidt, & Viswesvaran, 1994) argue this is a relevant artifact that can be corrected, we believe it is premature to estimate the validity of the perfect construct when there is no firm agreement on the definition of the perfect construct.

In meta-analyzing the data, we examined correlations between HDS scales and each competency by rater source (self, supervisor, peer, direct report, other). Following recommendations made by Scherbaum (2005), we then combined results across studies using meta-analysis to provide indicators of relationships between HDS scales and competencies across organizations and jobs. Specifically, we examined each scale-by-competency-*by-rater* grouping. These meta-analyses, following procedures outlined by Hunter and Schmidt (2004), provide stable estimates of the relationships between each of the 11 HDS scales and performance ratings.

Results

We used meta-analytic techniques to determine stable relationships between HDS scales and 48 competencies from the CET where sufficient data were available from the 29 studies included. The results of our meta-analyses provide some interesting results reflecting both consistencies and inconsistencies in how various rater parties view the impact of various negative personality characteristics on performance outcomes. In some instances, raters share the same perception of personality characteristics and their impacts on performance. However, in other cases, their perceptions are completely contradictory.

For example, of the studies included in this research, between two and ten studies (varying by rater group) contained at least one performance rating mapped to Stress Tolerance. These results appear in Table 3. As this table shows, all raters viewed the HDS Excitable scale as having a detrimental effect on one's ability to effectively handle stress. More specifically, results for this scale emerged as significant across all raters except for direct reports, with these data constrained by a small sample size. One can logically assume that additional data for this rater group would have helped this result gain statistical significance as well.

Table 3. Meta-Analyzed HDS Excitable – Stress Tolerance Correlations by Rater Group

Rater	*K*	*N*	EXC	SKE	CAU	RES	LEI	BOL	MIS	COL	IMA	DIL	DUT
SUP	10	1096	***-.21****	-.14*	-.08	-.08*	-.07	-.04	-.07	-.08*	-.15*	-.02	.10
PEER	3	357	***-.34****	-.08	-.07*	-.06*	.04*	.00	-.04	-.10*	-.16*	-.06	.01
DR	2	73	***-.13***	-.03	.03	.11	-.05	.12	.05	.02	.16	-.20	-.02
OTHER	4	219	***-.24****	-.14*	.10	-.01	.07	-.04	-.07	-.13	-.19*	-.01	.17*
SELF	5	565	***-.50****	-.08	-.40*	-.12*	-.10	.20*	.14*	.04	-.02	.05	-.09*

Note. * Correlation is significant at 0.05 level; *K* = Number of studies; *N* = Number of participants across *K* studies; SUP = Supervisor; DR = Direct Report; EXC = Excitable; SKE = Skeptical; CAU = Cautious; RES = Reserved; LEI = Leisurely; BOL = Bold; MIS = Mischievous; COL = Colorful; IMA = Imaginative; DIL = Diligent; DUT = Dutiful.

The results presented in Table 3 indicate that, regardless of perspective, raters view individuals who are moody, tense, and easily annoyed (elevated HDS Excitable) as deficient in their ability to effectively handle stress. This same pattern of results repeats across nearly all of the other 47 competencies included in this research, with raters across perspectives viewing the Excitable scale as a detrimental influence on performance criteria.

Another example of consistent results across rater groups appears for ratings of Leadership. Depending on rater group, between six and 19 studies contained at least one performance rating mapped to this performance outcome. These results appear in Table 4. As this table shows, all raters viewed the HDS Cautious scale as having a significantly detrimental effect on one's ability to effectively lead others. Note that the same results appear for the Excitable scale as well.

Table 4. Meta-Analyzed HDS Cautious – Leadership Correlations by Rater Group

Rater	*K*	*N*	EXC	SKE	CAU	RES	LEI	BOL	MIS	COL	IMA	DIL	DUT
SUP	19	2437	-.07*	-.13*	***-.07****	-.04	-.10*	.00	-.06*	.00	-.10*	.03	-.03
PEER	8	1451	-.09*	.00	***-.06****	-.13*	-.04	.04	-.01	.07	.02	-.01	.03
DR	7	1328	-.08*	-.02	***-.09****	-.15*	-.06*	.12*	.02	.12*	.02	.03	-.04
OTHER	6	305	-.16*	.04	***-.15****	-.09	.01	.10	.11	.15	-.04	.01	.01
SELF	9	1549	-.24*	.08*	***-.29****	-.19*	-.03	.30*	.12*	.21*	.09*	.15*	-.03

Note. * Correlation is significant at 0.05 level; *K* = Number of studies; *N* = Number of participants across *K* studies; SUP = Supervisor; DR = Direct Report; EXC = Excitable; SKE = Skeptical; CAU = Cautious; RES = Reserved; LEI = Leisurely; BOL = Bold; MIS = Mischievous; COL = Colorful; IMA = Imaginative; DIL = Diligent; DUT = Dutiful.

The results presented in Table 4 indicate that, regardless of perspective, raters view individuals who are reluctant to take chances, unassertive, and resistant to change (elevated HDS Cautious) as deficient in their ability to effectively lead others. Again, note that raters also view highly Excitable individuals as deficient in their leadership capabilities. For many of the other 47 competencies included in this research, this same pattern of results emerges with raters across perspectives viewing the Cautious scale as a detrimental influence on performance criteria.

Interestingly, raters also viewed some HDS scales as beneficial to performance along some competencies. Although the HDS assesses personality characteristics that typically derail career progression, raters viewed these tendencies as positive influences on performance for some well-aligned competencies. For example, of the studies included in this research, between two and six studies (varying by rater group) contained at least one performance rating mapped to Initiative. These results appear in Table 5. As this table shows, raters viewed the HDS Bold scale as having a positive influence on one's ability to take proactive action without being told what to do. More specifically, results for this scale emerged as significant across all raters except for "others," where insufficient data were available to conduct these analyses.

Table 5. Meta-Analyzed HDS Bold – Initiative Correlations by Rater Group

Rater	*K*	*N*	EXC	SKE	CAU	RES	LEI	BOL	MIS	COL	IMA	DIL	DUT
SUP	6	1469	-.03	-.06	-.10	.00	-.11*	***.08****	-.04	.00	-.03	.09	-.04
PEER	3	1165	.02	.07	-.07	-.12*	-.07	***.13****	.02	.11*	.14*	.07*	-.07
DR	2	1060	-.03	-.01	-.11	-.12*	-.12	***.12****	.00	.10*	-.01	.07*	-.12
SELF	2	1081	-.14*	.07	-.26*	-.17*	-.07*	***.28****	.13*	.15	.10	.22*	-.06*

Note. * Correlation is significant at 0.05 level; *K* = Number of studies; *N* = Number of participants across *K* studies; SUP = Supervisor; DR = Direct Report; EXC = Excitable; SKE = Skeptical; CAU = Cautious; RES = Reserved; LEI = Leisurely; BOL = Bold; MIS = Mischievous; COL = Colorful; IMA = Imaginative; DIL = Diligent; DUT = Dutiful.

The results presented in Table 5 indicate that, regardless of perspective, raters view individuals who are assertive, demanding, and aggressive (elevated HDS Bold) as able to take initiative without direction from others. Interestingly, this pattern of results also repeats across several other competencies included in this research, with raters across perspectives viewing elevations in the Bold scale as beneficial to performance.

Although the above results provide interesting examples of instances where raters across perspectives provided similar ratings of the target individual, in other instances raters viewed the influence of certain HDS scales on performance competencies in vastly different lights. In fact, for some competencies, results across rater groups appear nearly contradictory. For example, of the studies included in this research, between two and three studies (varying by rater group) contained at least one performance rating mapped to Goal Setting. These results appear in Table 6. For this competency, results for the Colorful scale emerged as significant across all raters except for "others," where insufficient data were available to conduct these analyses. However, significant results appear in opposite directions across raters. Individuals viewed their own Colorful behaviors as positive influences on their ability to identify objectives and take steps to achieve them. Direct reports of these individuals also viewed these behaviors as beneficial to setting and pursuing goals. However, colleagues and supervisors of these individuals provide negative ratings, viewing Colorful tendencies as detrimental influences on the target's ability to set goals.

Table 6. Meta-Analyzed HDS Colorful – Goal Setting Correlations by Rater Group

Rater	*K*	*N*	EXC	SKE	CAU	RES	LEI	BOL	MIS	COL	IMA	DIL	DUT
SUP	2	225	.01	-.18*	.09	.01	-.03	-.10	-.09*	***-.17****	-.13	-.01	-.11
PEER	3	277	-.04	.01	.02	.02	.04	-.01	-.14	***-.08****	-.20*	.17	.05
DR	2	210	-.06	-.01	-.09	-.10*	.10*	.14	.06*	***.14****	.00	-.04	.06
SELF	2	226	-.15*	.20*	-.23*	-.21*	-.04	.29*	.05	***.27****	.04	.11	-.03

Note. * Correlation is significant at 0.05 level; *K* = Number of studies; *N* = Number of participants across *K* studies; SUP = Supervisor; DR = Direct Report; EXC = Excitable; SKE = Skeptical; CAU = Cautious; RES = Reserved; LEI = Leisurely; BOL = Bold; MIS = Mischievous; COL = Colorful; IMA = Imaginative; DIL = Diligent; DUT = Dutiful.

Contrary to the first three results provided above, the results presented in Table 6 indicate that raters view the ability of individuals who are dramatic, self-promoting, and easily distracted (elevated HDS Colorful) to set and pursue goals in vastly different lights depending on their perspective. As these individuals typically confuse activity with productivity, they and their subordinates may view these behaviors in a positive light. However, peers and supervisors of these individuals, viewing these Colorful behaviors from different perspectives, rate these tendencies as detrimental to effective goal setting. Also unlike results previously discussed, raters across perspectives view the Colorful scale in different lights depending on the performance dimension under consideration. For some competencies, raters consistently view

Colorful behaviors as beneficial to performance. In fewer cases, raters consistently view these tendencies as detrimental to performance. Finally, for other competencies, results appear consistent with those presented in Table 6, with raters viewing Colorful behaviors in different lights depending on their perspective to the target individual.

As a second example of results differing by rater, consider the results for Negotiation presented in Table 7. Of the studies included in this research, between two and five studies (varying by rater group) contained at least one performance rating mapped to this competency. For Negotiation, results for the Diligent scale emerged as significant across self- and supervisor-ratings. Peer ratings did not emerge as statistically significant and insufficient data existed for direct reports and others to conduct these analyses. However, where individuals viewed their own Diligent behavior as beneficial to their ability to negotiate, their supervisors rated these behaviors as negative influences on performance along this competency. Also note that, while not statistically significant, peer ratings appear in the same direction as supervisory ratings.

Table 7. Meta-Analyzed HDS Diligent – Negotiation Correlations by Rater Group

Rater	*K*	*N*	EXC	SKE	CAU	RES	LEI	BOL	MIS	COL	IMA	DIL	DUT
SUP	5	581	-.17	-.08	-.16*	-.11	.00	.13*	.04	.03	-.05	***-.10****	-.12
PEER	3	433	-.13	-.09	-.04	-.03	.05	.01	.10*	.01	.00	***-.03***	.01
SELF	2	448	-.24*	.10*	-.33*	-.10*	.09	.34*	.24*	.29*	.13*	***.14****	-.06

Note. * Correlation is significant at 0.05 level; *K* = Number of studies; *N* = Number of participants across *K* studies; SUP = Supervisor; DR = Direct Report; EXC = Excitable; SKE = Skeptical; CAU = Cautious; RES = Reserved; LEI = Leisurely; BOL = Bold; MIS = Mischievous; COL = Colorful; IMA = Imaginative; DIL = Diligent; DUT = Dutiful.

Results presented in Table 7 indicate that perfectionistic, micro-managing, and inflexible individuals (elevated HDS Diligent) view those tendencies as positive influences on performance in negotiation. However, supervisors of these individuals view these same behaviors as detrimental to negotiation, perhaps because they see the difficulties these individuals may experience in the imperfect, give-something-to-get-something environment characterizing negotiations. Across other competencies included in this research, raters frequently view Diligent behaviors as beneficial to performance. However, for many other competencies, results appear consistent with those presented in Table 7, with raters viewing Diligent behaviors in different lights depending on their perspective to the target individual.

In still other cases, individuals may believe that certain personality characteristics have no influence on performance along job-relevant competencies. However, others around the target individual may find these characteristics highly influential, but disagree about the direction of this influence. As an example, consider the results for Interpersonal Skills presented in Table 8. Of the studies included in this research, between two and four studies (varying by rater group) contained at least one performance rating mapped to this competency. For Interpersonal Skills, results for the Dutiful scale emerged as significant across supervisor- and direct report-ratings. Peer and other ratings did not emerge as statistically significant. From these results, note first that individuals found nearly no relationship between their own Dutiful scores and their Interpersonal Skills. Supervisors of these individuals rated these behaviors as beneficial to performance along this competency. However, direct reports of these individuals viewed these same behaviors as detrimental influences on interpersonal skill.

Table 8. Meta-Analyzed HDS Dutiful – Interpersonal Skills Correlations by Rater Group

Rater	*K*	*N*	EXC	SKE	CAU	RES	LEI	BOL	MIS	COL	IMA	DIL	DUT
SUP	4	233	-.19*	-.17	-.17*	-.24	-.18*	-.02	.06	.04	-.12	.07	***.27****
PEER	2	108	-.04	-.30*	-.17*	-.18*	-.22*	.07*	-.05	.16	-.27*	-.02	***.03***
DR	2	90	-.03	-.05	-.19	-.16*	-.17	.04	-.05*	.00	-.37*	-.15*	***-.15****
OTHER	3	197	-.34	-.16	-.02	-.15	-.03	.07	.07	.09*	-.04	-.10	***.09***
SELF	4	288	-.27*	.03	-.34*	-.44*	-.07	.26*	.13*	.22	-.01	.16*	***-.02***

Note. * Correlation is significant at 0.05 level; *K* = Number of studies; *N* = Number of participants across *K* studies; SUP = Supervisor; DR = Direct Report; EXC = Excitable; SKE = Skeptical; CAU = Cautious; RES = Reserved; LEI = Leisurely; BOL = Bold; MIS = Mischievous; COL = Colorful; IMA = Imaginative; DIL = Diligent; DUT = Dutiful.

Results presented in Table 8 indicate that raters view the Interpersonal Skills of individuals who are conforming, reliant on others, and eager to please (elevated HDS Dutiful) in different lights depending on their perspective. As these individuals are comfortable following others and work hard to gain approval, their supervisors may hold positive perceptions about their Interpersonal Skills. However, as these same individuals may also be reluctant to defend their subordinates, direct reports rate these tendencies as detrimental to Interpersonal Skills. Across other competencies included in this research, raters frequently view Dutiful behaviors as beneficial to performance. However, for some other competencies, results appear consistent with those presented in Table 8, with raters viewing Dutiful behaviors in different lights depending on whether they are the recipient of that obliging behavior.

Discussion

This research represents a new and innovative foray into multisource performance appraisal, and provides new insights about the impact of personality on performance ratings from different sources. Unlike earlier research examining the effects of "bright side" Five-Factor Model personality constructs on 360 performance appraisal ratings (Ones & Viswesvaran, 2001; Roberts & Hogan, 2001), the current study focused on "dark side" personality dimensions. With typical "bright side" personality measures, most generally consider increased levels of dimensions better for their effects on job performance constructs. In contrast, examining "dark side" personality constructs in the current effort reveals more complex relationships between personality and performance. Specifically, for some HDS scales, results are consistent with the above notion that higher construct levels positively relate to job performance. However, for other HDS scales, increased levels of the personality dimension *negatively* associate with competency-based job performance. As these results indicate, *more* personality is not always *better*.

Following from these observations, the current research bears some interesting implications for personnel development applications. Because of the positive orientation of FFM scales and typical job performance rating dimensions, developmental feedback based on the interplay of these dimensions is likely to take a similarly positive tone. As such, personnel may gain insights about which positively-oriented personality dimensions they currently "under-do" and should try to apply to work more often. In other words, this developmental feedback will inform personnel about performance deficiencies where they should "start" certain behaviors. However, these initiatives provide no information about the potentially negative personality dimensions they may currently "over-do" and from which they should refrain. In contrast to the nature of positively-oriented personality, assessing multisource performance appraisal

ratings against negatively-oriented personality dimensions provides insights about how excesses in "dark side" personality characteristics negatively affect job performance. In turn, developmental feedback based on this information will inform personnel not only about behaviors they should "start" and "keep" doing, but other behaviors they should "stop" as well. With regard to multisource performance appraisal, combining "bright side" and "dark side" personality constructs provides a more comprehensive view of how the individual may enhance job performance through efforts to resolve both behavioral deficiencies and excesses.

Another important conclusion to draw from this research is an apparent disconnect between rater groups. Specifically, from self ratings it appears that individuals believe that elevations in certain HDS scales (i.e., Skeptical) will prove beneficial to certain aspects of job performance (i.e., Achievement Orientation). However, examining the same scale-by-competency pairings for other rater groups (i.e., supervisor) it appears that others view these same HDS scale elevations as either unrelated or detrimental to specific facets of job performance. These findings stand in contrast to the perceived lack of differences between rater groups noted by Yammarino and Atwater (1997). However, the more important issue is *why* different groups provide such discrepant ratings concerning the impact of HDS scales on performance competencies.

Perhaps, as a function of impression management, individuals provide different behavioral samples to each rater group. In this context, the performance ratings provided by each group would each be considered accurate, though each group is essentially rating different behavioral samples provided by the same target individual. This seems logical, as employees do not behave the same around their supervisors as they do around their peers, direct reports, or customers.

As an alternative, it may be the case that supervisors, peers, direct reports, and other parties hold different expectations about what they consider acceptable performance in a target individual. These differences may reflect differing frames of reference held by different rater groups about what they consider to be acceptable performance in a target individual. In this context, the target individual may behave consistently across rater groups. However, these same behaviors may match one rater's frame of reference for performance more closely than another's, resulting in different performance ratings across raters.

These potential explanations behind the differences noted between rater groups provide hypotheses for future research efforts to investigate. Although it does appear that rating differences do exist between rater groups, the exact nature of these differences remains unclear. What the current research does make clear, however, is the value of examining multisource performance appraisal ratings in the context of individual personality. Regardless of whether organizations intend to use results for personnel selection or development applications, examining performance in light of *both* "bright side" and "dark side" personality dimensions can provide valuable insights about the effects of individual personality characteristics on performance as rated by supervisors, peers, direct reports, and other parties.

References

Atwater, L. E., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology, 45*, 141 – 164.

Barrett, L., Dunbar, R., & Lycett, J. (2002). *Human evolutionary psychology.* London: Palgrave.

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

Carson, R. C. (1969). *Interaction concepts of personality.* Chicago: Aldine.

Catano, V. M., Darr, W., & Campbell, C. A. (2007). Performance appraisal of behavior-based competencies: A reliable and valid procedure. *Personnel Psychology, 60*, 201 – 230.

Conger, J. A. (1990). The dark side of leadership. *Organizational Dynamics, 19,* 44-55.

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.

Dalesio, A. (1998). Using multi-source feedback for employee development and personnel decisions. In J. W. Smither (Ed.), *Performance Appraisal: State of the Art in Practice* (pp. 277 – 330). San Francisco, CA: Jossey-Bass Publishers.

Dawkins, R. (1976). *The selfish gene.* Oxford: Oxford University Press.

Dotlich, D. L., & Cairo, P. C. (2003). *Why CEOs Fail: The 11 Behaviors that Can Derail Your Climb to the Top – and How to Manage Them.* San Francisco, CA: Jossey-Bass.

Hauenstein, N. M. A. (1998). Training raters to increase the accuracy of appraisals and the usefulness of feedback. In J. W. Smither (Ed.), *Performance Appraisal: State of the Art in Practice* (pp. 404 – 444). San Francisco, CA: Jossey-Bass Publishers.

Hogan Assessment Systems (2008). *Competency Model Development* [Unpublished technical manual]. Tulsa, OK: Author.

Hogan, R., & Hogan, J. (1997). *Hogan Development Survey Manual*. Tulsa, OK: Hogan Assessment Systems.

Hogan, R., & Hogan, J. (2009). *Hogan Development Survey manual* (2nd ed.). Tulsa, OK: Hogan Press.

Hogan, R., Hogan, J., & Warrenfeltz, R. (2007). *The Hogan Guide: Interpretation and Use of Hogan Inventories*. Tulsa, OK: Hogan Assessment Systems.

Horney, K. (1950). *Neurosis and human growth.* New York: Norton.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks, CA: Sage Publications, Inc.

Kane, J. S., & Lawler, E. E. (1978). Methods of peer assessment. *Psychological Bulletin, 85*, 555 – 586.

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87*, 72 – 107.

Leary, T. F. (1957). *Interpersonal diagnosis of personality.* New York: Ronald Press.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validity of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81-89.

Mount, M. K., & Barrick, M. R. (1995). The big five personality dimensions: Implications for research and practice in human resources management. In G. R. Ferris (Ed.), Research in Personnel and Human Resources Management (vol. 13, pp. 153-200). JAI Press.

Murphy, K. R., & DeShon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology*, 53(4), 873-900.

Ones, D. S., Schmidt, F. L., & Viswesvaran, C. (1994, April). *Examination of construct validity with linear composites and generalizability coefficient corrected correlations*. Paper presented at the annual conference of the Society for industrial and Organizational Psychology, Nashville, TN.

Ones, D. S., & Viswesvaran, C. (2001). Personality at work: Criterion-focused occupational personality scales used in personnel selection. In B. W. Roberts and R. Hogan (Eds.), *Personality Psychology in the Workplace*. Washington, D.C.: American Psychological Association.

Pulakos, E. D. (1997). Ratings of job performance. In D. Whetzel and G. R. Wheaton (Eds.), *Applied Measurement Methods in Industrial Psychology* (pp. 291 – 318). Palo Alto, CA: Davies-Black Publishing.

Reilly, R. R., & McGourty, J. (1998). Performance appraisal in team settings. In J. W. Smither (Ed.), *Performance Appraisal: State of the Art in Practice* (pp. 244 – 277). San Francisco, CA: Jossey-Bass Publishers.

Roberts, B. W., & Hogan, R. (Eds.) (2001). *Personality Psychology in the Workplace*. Washington, D.C.: American Psychological Association.

Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology*, 75, 322–327.

Scherbaum, C. A. (2005). Synthetic validity: Past, present, and future. *Personnel Psychology, 58,* 481-515.

Stoker, J. I., & Van der Heijden, B. I. J. M. (2001). Competence development and appraisal in organizations. *Journal of Career Development, 28*, 97 – 113.

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

Tziner, A., Murphy, K. R., & Cleveland, J. N. (2002). Does conscientiousness moderate the relationship between attitudes and beliefs regarding performance appraisal and rating behavior? *International Journal of Selection and Assessment, 10*, 218 – 224.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 295-412.

Yammarino, F. J., & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human resource management. *Organizational Dynamics*, 35 – 44.