Industrial and Organizational Psychology, **1** (2008), 320–322.
Copyright © 2008 Society for Industrial and Organizational Psychology. 1754-9426/08

An Argument for Context-Specific Personality Assessments

ERIC D. HEGGESTAD AND HEATHER L. GORDON
University of North Carolina at Charlotte

Some scholars within our field (e.g., Morgeson et al., 2007; Murphy & Dzieweczynski, 2005) have begun to lament as disappointing the criterion-related validity evidence for personality assessments in the prediction of job performance. In meta-analytic work, for example, Barrick, Mount, and Judge (2001), Dudley, Orvis, Lebiecki, and Cortina (2006), Hurtz and Donovan (2000), and Tett, Jackson, and Rothstein (1991) have reported rather small (.12, .15, .13, and .12, respectively) uncorrected mean validity coefficients between Conscientiousness, the most predictive of the Big Five traits, and job performance. In their article, Hough and Oswald (2008) argued that a more careful consideration of situational demands on behavior is essential to understanding the personality–performance relationship and to the identification of conditions under which larger criterion-related validities can be expected. We could not agree more.

In his now classic book *Personality and Assessment*, Mischel (1968) reviewed the personality literature and made two key observations: (1) that self-report measures of personality rarely, if ever, correlate with behavior at a level above .30 and (2) that correlations of people's behaviors across situations are typically small in magnitude. Based on these observations, Mischel suggested that "the concept of personality traits as broad predispositions is thus untenable" (p. 140), a conclusion that would cripple personality psychology for some time. The defense made by personality psychologists to Mischel's strong conclusion came in the form of the principle of aggregation (Epstein & O'Brien, 1985). Stated simply, the principle of aggregation suggests (a) that the accumulation of behavioral information from a large number of different situations can result in a reliable indication of trait standing and (b) that trait assessments can be expected to predict (well in excess of .30) behavior when that behavior has been aggregated over a number of specific instances. Fundamentally, therefore, the principle of aggregation suggests that personality assessments are poor predictors of specific instances of behavior (because these behaviors are strongly influenced by the demands of the particular situation) but can be reasonably good predictors of lawful tendencies in behavior that emerge across different situations.

The disappointingly low criterion-related validities observed for personality assessments are likely a result, at least in part, of our lack of understanding of the principle of aggregation and its application to the context of selection. Many personality assessments used in personnel decision making have not been created specifically for use in selection settings but rather as general assessments of the respondent's personality. As such, the

Correspondence concerning this article should be addressed to Eric D. Heggestad. E-mail: edhegges@email.uncc.edu

Address: Department of Psychology, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001

Eric D. Heggestad, Department of Psychology, University of North Carolina at Charlotte; Heather L. Gordon, Organizational Science, University of North Carolina at Charlotte.

https://doi.org/10.1111/j.1754-9434.2008.00056.x Published online by Cambridge University Press

items and instruction sets for these measures are general and do not focus on a particular situational context. In other words, these assessments ask people to report their trait-related behaviors on average across the variety of situations they typically encounter. Working from the principle of aggregation, scores on such assessments should best predict behavioral tendencies that arise when behavior is aggregated across a wide variety of situations and settings. As noted by Hough and Oswald, and consistent with this notion, personality trait measures have been found to be good predictors of major life outcomes, such as health behaviors and occupational attainment (Bogg & Roberts, 2004; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007).

We, as industrial–organizational psychologists, however, are generally more interested in making predictions about a person's behavior at work than about his or her behavior in general. As such, our focus should be on the prediction of behavior aggregated across work-related situations. Thus, it can be reasoned that assessing an applicant with a general personality measure results in a mismatch between what our predictor is designed to predict (i.e., aggregated levels of general behavior) and the scope of the criterion we are interested in predicting (i.e., aggregated levels of work behavior). By using general personality measures in selection contexts, we are taking a predictor-centric approach rather than a more appropriate criterion-centric approach (Bartram, 2005; Guion, 1998).

As mentioned briefly by Hough and Oswald, several researchers (Bing, Whanger, Davison, & VanHook, 2004; DeGroot & Kluemper, 2007; Hunthausen, Truxillo, Bauer, & Hammer, 2003; Schmit, Ryan, Stierwalt, & Powell, 1995) have investigated contextualized personality assessments as predictors of context-specific outcomes (such as job performance). Contextualized personality assessments provide information about behavioral tendencies within a specific context, such as a workplace context. As such, the use of these measures represents a tighter predictor–criterion linkage than does the use of general personality assessments. Most often, contextualization is accomplished by instructing respondents to consider only their behavior within the work context when responding to the items or by adding contextual information, such as the phrase "at work," directly to the items. Research has demonstrated superior predictive capability of contextualized measures over standard, noncontextualized versions of the same measure when the criterion of interest is context specific (Bing et al., 2004; DeGroot & Kluemper, 2007; Hunthausen et al., 2003; Schmit et al., 1995). For example, Schmit et al. compared the criterion-related validities for a school-specific contextualized measure of Conscientiousness and a noncontextualized version of the same measure in the prediction of GPA. The criterion-related validity was found to be notably higher for the contextualized measure ($r = .41$) than for the noncontextualized measure ($r = .25$; see Bing et al., 2004, for similar findings). As another example, Hunthausen et al., who varied the degree of contextualization through an instructional manipulation, found that contextualization was associated with an increase in concurrent criterion-related validity (from $r = .10$ to .31) for a measure of Conscientiousness in the prediction of performance ratings among customer service employees in the airline industry.

Consideration of the principle of aggregation in the selection context suggests that the use of contextualized measures will create a tighter linkage between the predictor and our criteria of interest and that, consequently, the criterion-related validity of those assessments should be stronger than for similar, noncontextualized assessments. Although the research base is still rather limited, existing empirical evidence does support the superiority of contextualized assessments in the prediction of work-specific outcomes. Although we would like to see additional research in this area (e.g., how much contextualization is required in the items and within the measure; a clear evaluation of construct equivalence; the impact of faking), we believe that researchers and practitioners alike should

https://doi.org/10.1111/j.1754-9434.2008.00056.x Published online by Cambridge University Press

begin consistently using contextualized assessments of personality. Researchers investigating the relationships between personality and organizationally relevant outcomes are likely to find stronger associations when contextualized measures are used. Likewise, practitioners are likely to observe increased criterion-related validities and make better quality decisions when contextualized measures are used in employee selection systems. In fact, with some additional empirical support, we believe that the use of contextualized measures within selection contexts could soon become a best practice in our field.

References

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and 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.

Bartram, D. (2005). The Great Eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90*, 1185–1203.

Bing, M. N., Whanger, J. C., Davison, H. K., & VanHook, J. B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology, 89*, 150–157.

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin, 130*, 887–919.

DeGroot, T., & Kluemper, D. (2007). Evidence of predictive and incremental validity of personality factors, vocal attractiveness and the situational interview. *International Journal of Selection and Assessment, 15*, 30–39.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*, 40–57.

Epstein, S., & O'Brien, E. J. (1985). The person-situation debate in historical and current perspective. *Psychological Bulletin, 98*, 513–537.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

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

Hough, L. M., & Oswald, F. L. (2008). Personality testing and industrial–organizational psychology: Reflections, progress, and prospects. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 272–290.

Hunthausen, J. M., Truxillo, D. M., Bauer, T. N., & Hammer, L. B. (2003). A field study of frame-of-reference effects on personality test validity. *Journal of Applied Psychology, 88*, 545–551.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869–879.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology, 60*, 1029–1049.

Murphy, K., & Dzieweczynski, J. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18*, 343–357.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313–345.

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology, 80*, 607–620.

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.

https://doi.org/10.1111/j.1754-9434.2008.00056.x Published online by Cambridge University Press