A Conceptual and Empirical Review of the Structure of Assessment Center Dimensions

Article *in* Journal of Management · June 2014

Impact Factor: 6.86 · DOI: 10.1177/0149206314522299

3 authors, including:

Brian James Hoffman
University of Georgia
46 PUBLICATIONS **1,637** CITATIONS
SEE PROFILE

David J. Woehr
University of North Carolina at …
81 PUBLICATIONS **2,160** CITATIONS
SEE PROFILE

Journal of Management
Vol. 40 No. 5, July 2014 1269–1296
DOI: 10.1177/0149206314522299

A Conceptual and Empirical Review of the Structure of Assessment Center Dimensions

John P. Meriac
University of Missouri–St. Louis
Brian J. Hoffman
University of Georgia
David J. Woehr
University of North Carolina-Charlotte

Although the design, scoring, and interpretation of assessment centers (ACs) commonly focuses on job-relevant dimensions, over three decades of past studies have questioned the evidentiary basis underlying dimension-based interpretations of ACs. This review combines multiple approaches to examine the structure of AC dimensions. First, we consulted the AC, job performance, leadership, and personality literatures to articulate competing models of the dimensions underlying AC ratings. Next, meta-analytic confirmatory factor analysis (CFA) was used to compare the fit of these models to existing AC data. The results supported a model including administrative skills, relational skills, and drive. Third, socioanalytic theory was used as a basis to examine the nomological network of these three broad factors, specifically their relationships with general mental ability and the five factor model of personality. The analyses supported the nomological network of drive and administrative skills but less so for relational skills. These findings are discussed with regard to the construct-related validity of AC dimensions, the fidelity of ACs to the broader criterion domain, and the value of applying generalizable models to the analysis of AC ratings.

Keywords: *assessment center; job performance; leadership; personnel selection*

Authors' Note: References for the primary studies included in the meta-analysis are available at http://jom.sagepub.com/supplemental.

Acknowledgments: Portions of this article are based on the first author's doctoral dissertation, which was directed by the third author, and were presented at the 27th meeting of the Society for Industrial and Organizational Psychology. We would like to thank Matthew Fleisher, Lauren Wood, Matthew Milunski, and John Slifka for assistance in study coding and conducting the literature search.

Corresponding author: John P. Meriac, Department of Psychology, University of Missouri–St. Louis, 425 Stadler Hall, One University Boulevard, St. Louis, MO 63121-4499, USA.

E-mail: meriacj@umsl.edu

Since the inception of the assessment center method (AC; Bray & Grant, 1966), dimension scores have played a central role in the design, scoring, and interpretation of ACs. Despite extensive inquiry over the past 30 years (Sackett & Dreher, 1982), there remains considerable disagreement as to whether ACs can effectively measure dimensions at all, with some going so far as to propose that dimensions be removed entirely from the scoring and interpretation of ACs (e.g., Lance, 2008). Given that dimensions provide practitioners with a link between ACs and other HR functions (e.g., employee development, succession planning; Thornton & Gibbons, 2009) and that dimension-based ACs are the most common approach to scoring ACs in research and applied settings, clarifying the presence, nature, and psychometric properties of dimensions is central to the effective application of the AC method. The purpose of this review was to explore generalizable models of AC performance and to empirically examine their applicability to ACs. Specifically, we consulted the AC, job performance, and leadership literatures to derive competing models applicable to the AC context. Next, we compared these alternative models in a meta-analytic confirmatory factor analysis of of the internal structure of dimensions in ACs as represented in the extant literature. Finally, we use socioanalytic theory (Hogan & Holland, 2003) to articulate and test a nomological network of AC dimensions.

Despite over 50 years of AC research, no theory of AC dimensions has been offered (Arthur, 2012; Arthur, Day, & Woehr, 2008). Instead, dimensions are usually based on a job analysis or competency model, with limited consideration of ongoing theoretical developments in the measurement and classification of leader skills and behaviors (Austin & Crespin, 2006). Although a great strength of the AC method is its flexibility (Howard, 2008) and thus, applicability across jobs, idiosyncrasies across ACs also pose a significant challenge to the development of generalizable inferences about the AC method. Indeed, the diversity in dimension labels precluded direct comparisons across different ACs (Arthur, Day, McNelly, & Edens, 2003). Similarly, the diversity and sheer number of dimensions assessed in operational ACs might be partly responsible for persistent questions as to the construct-related validity of AC dimensions, and the failure to develop and test generalizable models of AC performance has posed a substantial roadblock to progress in the scientific literature. In contrast, the job performance and leadership literatures have made substantial progress in developing and validating taxonomic, generalizable models. Although ACs are advertised as moderate to high fidelity simulations of the criterion domain, the AC literature and the literature on the targeted criterion space (usually managerial performance) have persisted relatively independently, with minimal attempt at theoretical and conceptual integration (Arthur, 2012; Hoffman, Melchers, Blair, Kleinmann, & Ladd, 2011). Accordingly, we draw from other areas of management to inform the structure underlying AC ratings.

In the AC literature, Arthur et al. (2003) introduced a framework to organize observed AC dimensions into a broader and more generalizable model. Arthur et al.'s framework has been widely adopted, in both meta-analytic (e.g., Bowler & Woehr, 2006; Meriac, Hoffman, Woehr, & Fleisher, 2008) and primary studies (e.g., Dilchert & Ones, 2009). Despite its popularity, Arthur et al.'s framework was not based on an a priori model of the managerial skills and performance domain, empirical research has not investigated its fit to the available data, and this model has never benefited from a large-scale comparison to alternative models. In light of the popularity of framework, these are critical research needs.

Accordingly, this review advances the AC literature by forwarding a set of generalizable, commonly defined, and empirically supported factors of AC performance that were drawn

from, and thus generalize to, models of the criterion domain. By applying multiple approaches to evaluating the construct-related validity of ACs (Woehr, Meriac, & Bowler, 2012), the present study seeks to summarize the current state of the measurement of AC dimensions. These findings can potentially inform practice by clarifying expectations for underlying dimension structures and providing a unified framework for organizing dimension ratings when communicating results to researchers and practitioners.

Searching for Assessment Center Dimensions

The AC method has remained popular in the hiring, promotion, and development of managers (Thornton & Rupp, 2006) and has seen an increasing presence in global strategic talent management (Povah & Thornton, 2011). Research has consistently indicated that ACs demonstrate sound criterion-related validity evidence (Gaugler, Rosenthal, Thornton, & Bentson, 1987; Hermelin, Lievens, & Robertson, 2007; Meriac et al., 2008). Also, ACs are generally assumed to possess high levels of content-related validity and result in favorable applicant reactions in comparison with cognitive ability tests (Macan, Avedon, Paese, & Smith, 1994). Finally, although limited, evidence for developmental applications is relatively favorable (Gibbons, Rupp, Snyder, Hollub, & Woo, 2006; Rupp et al., 2006).

Despite years of research on the topic, considerable debate persists on AC construct-related validity evidence, specifically whether ACs actually measure the dimensions they are designed to measure (Arthur et al., 2008; Lance, 2008; Lievens & Christiansen, 2012). Based on the methods used in Sackett and Dreher's (1982) classic analysis of AC construct-related validity, the vast majority of these studies have applied multitrait–multimethod matrices (MTMM; Campbell & Fiske, 1959) to dimensions rated in each exercise, or within-exercise dimension ratings (WEDRs). When applied to ACs, the general expectation has been that correlations between the same dimensions across exercises should be stronger relative to correlations of different dimensions within exercises. To date, the MTMM approach to construct validation has been the subject of numerous primary studies and three large scale reviews (Bowler & Woehr, 2006; Lance, Lambert, Gewin, Lievens, & Conway, 2004; Lievens & Conway, 2001). On the basis of these reviews, when commonly accepted approaches are used to model the MTMM structure of AC WEDRs, researchers have generally concluded that there is limited evidence for the measurement of dimensions in ACs. This pattern of results has led some to the controversial suggestion that dimensions should be removed entirely in favor of exercise-based design, scoring, and interpretation (Lance, 2008).

However, several MTMM analyses have questioned this conclusion. Specifically, these studies have shown that by specifying fewer dimensions (Gaugler & Thornton, 1989), broader dimensions (Hoffman et al., 2011), and modeling dimensions within exercises (Hoffman & Meade, 2011; Putka & Hoffman, 2013) as many as three distinguishable dimensions can be empirically supported in the analysis of WEDRs. Despite preliminarily supportive evidence, such broad-factor frameworks have only been applied in a few samples and typically only using WEDRs. Given the considerable variability that exists across ACs, it is important to ensure that such broad dimension models are applicable to the broader literature. In addition, although WEDRs are considered to be a building block to scoring ACs and thus a meaningful unit of analysis, it is also important to evaluate the construct-related validity of other units of scoring (Rupp, Thornton, & Gibbons, 2008).

Construct-Related Validity of Across-Exercise Dimension Ratings

AC scores that summarize dimensional performance across all exercises, or across-exercise dimension ratings (AEDRs), are perhaps the most commonly used metric for the summary and interpretation of AC performance (Rupp et al., 2008). Although we have learned much about the psychometric qualities of WEDRs over the past 30 years, there has been far less consideration of the psychometric quality of AEDRs. Given the centrality of AEDRs to AC practice, the limited attention to this unit of scoring is a notable deficiency in the literature. A particularly glaring omission is the dearth of theory pertaining to the evaluation of the quality of AC dimensions (Arthur, 2012). Instead, AC dimensions are typically developed to meet organizational objectives and ideally, on the basis of a job analysis or competency model (International Task Force on Assessment Center Guidelines, 2009). Although the ability to tailor the AC to the focal job context is a key advantage of the method, this flexibility is not without cost.

First, the diversity in AC dimensions across local administrations presents a challenge to drawing generalizable conclusions. For instance, in their review of the criterion-related validity of AC dimensions, Arthur et al. (2003) identified 168 different dimension labels across 34 different ACs. Similar to personality research prior to the emergence of the five factor model (FFM; Digman, 1997; Goldberg, 1990) and job performance research prior to the current focus on developing a generalizable performance framework (Austin & Villanova, 1992), generalizing knowledge has been a challenge in light of the diversity in AC dimensions. Similarly, the application of generalizable performance models to ACs offers the potential for coherence in a relatively fragmented literature (Arthur et al., 2003).

Second, to the extent that operational objectives and demands drive the inclusion of dimensions, without consideration of accumulated knowledge on performance measurement and performance theory, conceptual overlap in content is a concern. A great deal of overlap is evident among dimensions within individual ACs, where operational ACs typically include multiple dimensions that capture similar behavioral content (Arthur, 2012; Hoffman et al., 2011) In light of this overlap, it is not surprising that the analysis of observed dimensions yields minimal support for their construct-related validity. As stated by Arthur et al. (2008: 108),

> The fundamental issue here is one of construct validity and an emphasis on the fact that merely labeling data as reflecting a particular construct (espoused construct) does not mean that is the construct that is being assessed (actual construct). Yet for some unexplainable reason, this practice appears to be the norm in AC research and practice where statements about what exercises measure (e.g., stress tolerance, social competence, factual argumentation, activity, imaginativeness) are by self-proclamation with rarely any systematic psychometric test development evidence presented to support these assertions.

Consistent with these observations, ACs as reported in the literature measures an average of 11 (with a range of 3 to 25) different performance dimensions (Meriac et al., 2008; Woehr & Arthur, 2003), far more dimensions than are traditionally supported in other areas of inquiry (e.g., ratings of job performance and leadership). Similarly, evaluations of AC design show that convergent and discriminant validity of ACs is improved when fewer, more conceptually distinct dimensions are evaluated in a given AC (Gaugler & Thornton, 1989;

Guenole, Chernyshenko, Stark, Cockerill, & Drasgow, 2011; Woehr & Arthur, 2003). As noted above, content overlap among observed dimensions is one explanation for the consistent failure of MTMM analyses to support dimensions (Hoffman et al., 2011).

In fact, as early as the Management Progress Studies, Bray and Grant (1966) recognized this issue. Subsequently, they used exploratory factor analysis to reduce observed dimensions to a more generalizable and empirically plausible number. Over the past several decades, a few primary studies have evaluated the factor structure underlying AC dimensions (e.g., Schmitt, 1977; Shore, Thornton, & Shore, 1990). However, such analyses are rare, the majority have been exploratory, and past studies have focused on single ACs. Thus, available research has neither the theoretical, empirical, or evidentiary foundation to forward generalizable inferences about the structure of AC dimensions (Woehr et al., 2012). Hence, much can be gained by applying theoretically relevant and generalizable performance models to understand the structure of ACs.

Generalizable Models of AC Dimensions

In light of these issues, Arthur et al. (2003) forwarded a framework that organizes manifest AC dimensions into seven overarching categories. Arthur et al. first coded primary study dimensions reported in the literature into Thornton and Byham's (1982) list of commonly used dimension labels. These labels were next grouped into seven categories by subject matter experts based on content similarity, and the criterion-related validity of the resulting categories was examined. Specifically, the categories they proposed were (a) *problem solving*, (b) *tolerance for stress/uncertainty*, (c) *influencing others*, (d) *consideration/awareness of others*, (e) c*ommunication*, (f) *organizing and planning*, and (g) *drive*. However, Arthur et al. subsequently removed tolerance for stress/uncertainty due to heterogeneity and ambiguity among the associated dimensions, resulting in a six-dimension framework.

In the decade since this study's publication, Arthur et al.'s framework has seen widespread adoption in both quantitative reviews (e.g., Bowler & Woehr, 2006; Lievens, Chasteen, Day, & Christiansen, 2006; Meriac et al., 2008) and primary studies (e.g., Connelly, Ones, Ramesh, & Goff, 2008; Dilchert & Ones, 2009). However, as noted by Arthur et al. (2003), this model was intended to be a first step in the development of a general structure of AC dimensions, and the authors encouraged researchers and practitioners to refine this framework. However, thus far no large-scale examinations of this sort have been undertaken. Although Arthur et al.'s framework represents a critical step forward in organizing this literature, more work is needed to guide researchers and practitioners seeking a generalizable and empirically supported taxonomy of AC dimensions (Hoffman et al., 2011). Accordingly, a key contribution of this study is to empirically test the adequacy of this model in describing the dimension structure from past AC research. In light of the popularity of this framework, these analyses have the potential to inform multiple areas of AC research and practice. In contrast to the AC literature, job performance and leadership research has seen substantial progress in the development of generalizable performance taxonomies (Austin & Crespin, 2006; Bartram, 2005; Borman & Brush, 1993; Conway, 1999; Hogan & Shelton, 1998; Mintzberg, 1973). However, this approach has been rarely undertaken in the AC literature.

Given that ACs are proposed to reflect moderate to high fidelity simulations of the criterion domain, a similarity in the structure underlying ACs and that underlying performance on

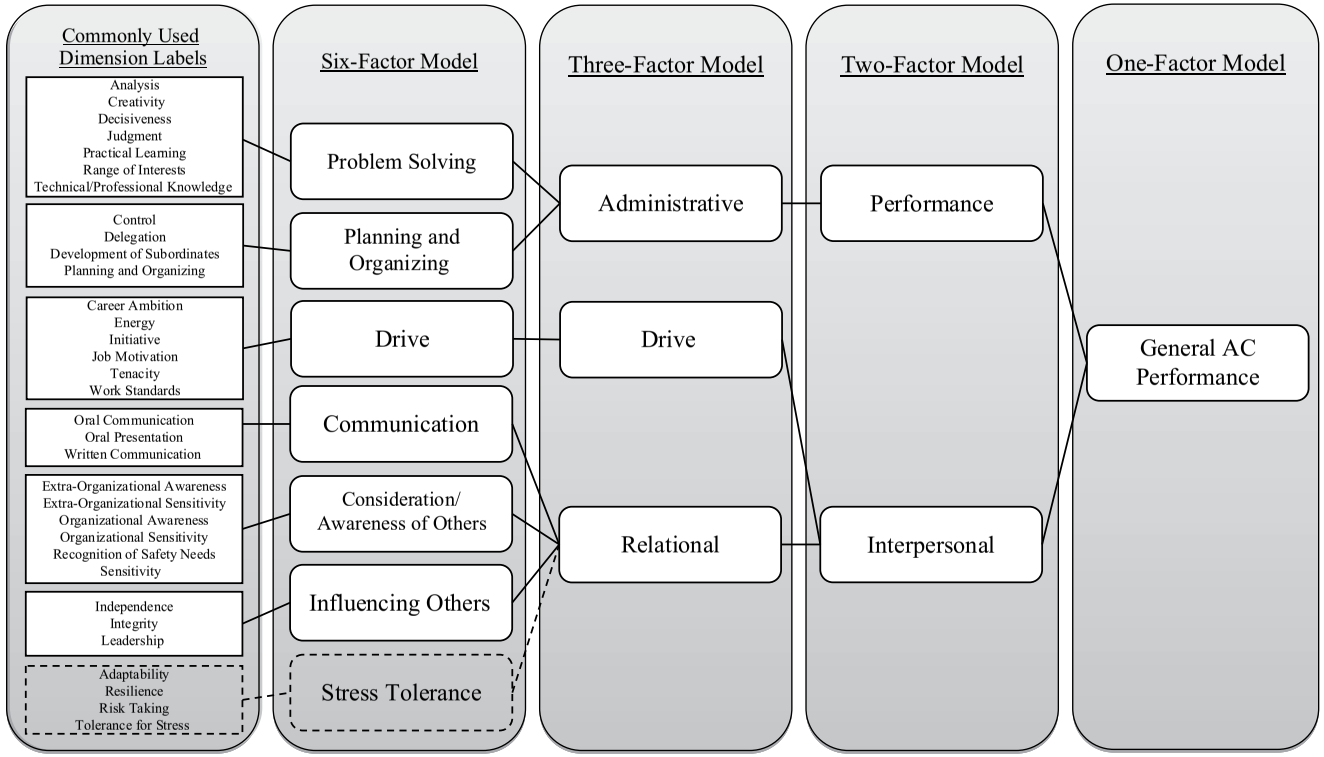
the job is both desirable and expected. Furthermore, the rich history of developing and testing generalizable models in related literature areas can provide a useful, but thus far untapped, resource in articulating models of the structure underlying AC ratings. In addition, such an approach can be informative in evaluating the nomological network and theoretical underpinnings of the constructs underlying ACs. Finally, integrating models from the broader job performance literature is useful, because it can clarify the linkages between AC ratings and performance behaviors of importance to organizations by putting both in the same language.

Given the dearth of attention to the theoretical and empirical structure underlying AC dimensions, we draw from the aforementioned literatures in articulating competing models of the constructs underlying ACs dimensions (Hoffman et al., 2011). Figure 1 presents a summary of the guiding frameworks used to organize manifest AC dimensions into competing models. Because of the widespread adoption in the AC literature, we specified Arthur et al.'s (2003) framework. Arthur et al. had initially started with seven dimensions and dropped stress tolerance because the primary study dimensions in this broader dimension were heterogeneous in nature, especially when compared with the other six dimensions. However, to examine the veracity of the decision to exclude stress tolerance and for the sake of completeness, we tested the seven-factor model as well. Next, we examine alternative models from the broader literature and discuss their representation in existing AC research. The extant literature suggests potential alternative conceptualizations to Arthur et al.'s model that include three-, two-, and one-factor models. The following sections elaborate on the conceptual underpinnings of each of these alternative models.

Three-factor models. Three-factor models of performance have emerged in the job performance, leadership, and AC literatures. For instance, Katz's (1955) model of managerial skills delineated technical, conceptual, and human skills and has recently been adopted (Mount, Judge, Scullen, Sytsma, & Hezlett, 1998; Scullen, Mount, & Goff, 2000) and expanded (Mumford, Zaccaro, Harding, Jacobs, & Fleishman, 2000) to describe the managerial performance domain . However, this taxonomy has subsequently been critiqued for difficulty in distinguishing conceptual and technical skills (Hoffman, Lance, Bynum, & Gentry, 2010). Borman and Brush (1993) proposed a similar three-factor model to describe managerial skill requirements, including interpersonal dealings and communication, leadership and supervision, and technical activities and the mechanics of management. This model has also been supported in the broader managerial performance literature (Conway, 1999; Hoffman & Woehr, 2009). A similar structure has achieved widespread adoption in the job performance literature, with factors corresponding to task performance (similar to technical skills), individually directed citizenship behaviors (similar to relational skills), and organizationally directed citizenship behaviors (similar to drive/motivation; Borman & Motowidlo, 1997; Williams & Anderson, 1991).

Three-factor models have also been identified in the AC literature. In fact, Thornton and Byham's (1982) early review of available factor analytic evidence speculated about the existence of three overarching dimensions of AC performance (i.e., administrative skills, interpersonal skills, and activity or drive). This model was supported by an exploratory factor analysis of dimension ratings in two operational ACs (Hinrichs, 1969; Schmitt, 1977), and a similar model was supported in a recent study using MTMM-based analyses to evaluate AC dimensions (Hoffman et al., 2011). Kolk, Born, and van der Flier (2004) proposed a similar

Figure 1
Summary of Models Examined and Nesting of Assessment Center Dimensions



Note: Stress tolerance is included in the figure in dashed lines to denote that it was omitted from Arthur, Day, McNelly, and Edens's (2003) final framework.

"triadic" model that naturally underlies ratings and comprises feeling, thinking, and power factors. Summarizing these key findings from the available literature, an underlying model specifying technical, drive, and relationship-maintenance components has consistently emerged. Based on prominent models from the job performance literature, and convergence across other literature bases, we test a three-factor model comprising *administrative skills, relational skills*, and *drive.*

Hogan's (1983) socioanalytic theory (Hogan & Shelton, 1998) provides additional theoretical precedent for this model and proposes that two broad motive patterns govern interpersonal interactions in organizations: "behavior designed to *get along* with other members of the group and behaviors to *get ahead* or achieve status vis-à-vis other members of the group" (Hogan & Holland, 2003: 100). Whereas administrative skills are clearly related to cognitive ability, relational skills and drive should be more strongly related to noncognitive variables. In the context of taxonomies of leader and managerial behaviors and by extension, ACs, getting along maps closely onto the relational skills dimension and getting ahead maps onto drive (Hogan & Kaiser, 2005). Individuals high on trait markers of getting along are "evaluated by others as good team players, organizational citizens, and service" (Hogan & Holland, 2003: 101). Individuals high on trait markers of getting ahead "take initiative, seek responsibility, compete, and try to be recognized" (Hogan & Holland, 2003: 101). Past studies have found that the FFM can be meaningfully grouped into these two broad factors and that these factors evince the expected pattern of associations with related behaviors in work settings (Bartram, 2005; Chiaburu, Oh, Berry, Li, & Gardner, 2011). Given that interpersonal interactions, or at a minimum hypothetical depictions of interpersonal interactions, are evident in virtually all AC exercises and that dimensions corresponding to both sets of behaviors are commonly measured in ACs, we expect both sets of interpersonally oriented behavior styles to emerge in AC ratings.

Two-factor models. An overarching two-factor model of performance has also been supported in both the leadership and job performance domains (Organ, 1988). These models proposed two broad aspects of performance: those that are more task-oriented and those that are more interpersonally oriented. This model simplifies the three-factor model by collapsing relational skills and drive into a broader interpersonal skills category. This model suggests that the two interpersonal styles specified by socioanalytic theory are not distinct in ACs but, combined, are distinguishable from the administrative skills factor.

A broad two-factor distinction between more technically oriented performance and more interpersonally oriented performance is evident across various domains of work behavior. For instance, classic research on leader behavior style distinguished initiating structure from consideration leader behaviors (Bowers & Seashore, 1966; Fleishman, 1957; Stogdill, 1974). Similarly, the job performance literature has also simplified the distinction between organizational citizenship behavior (OCB) directed toward individuals and organizations into an overall OCB factor (Hoffman, Blair, Meriac, & Woehr, 2007; LePine, Erez, & Johnson, 2002) that is described as affiliative (Mackenzie, Podsakoff, & Podsakoff, 2011) and is distinguished from more technically oriented aspects of task performance. A similar distinction is evident in descriptions of managerial performance, where the task and social context are commonly distinguished (Dierdorff, Rubin, & Bachrach, 2012; Dierdorff, Rubin, & Morgeson, 2009). Finally, this model has also emerged in AC research. Specifically, Shore et al. (1990) supported a similar model using exploratory factor analysis of dimensions from

an operational AC and supported two factors corresponding to *interpersonal-style* and *performance-style* factors. On the basis of past theoretical and empirical findings, this model collapses drive and relational skills to reflect a broader category of interpersonal style that is distinct from more administratively oriented aspects of performance.

General factor model. Given the positive manifold across the categories contained in Arthur et al.'s (2003) framework, a unidimensional model may provide the best representation of AC performance. Certainly, previous research has frequently indicated relatively strong correlations among categories assessed in ACs, and a general performance factor has been supported in recent MTMM applications (Hoffman et al., 2011; Lance et al., 2004). Kuncel and Sackett (2014) used composite reliability analyses to justify support for composite dimensions derived on the basis of WEDRs and then subjected the composite-based dimensions to factor analysis, supporting the presence of a general factor of AC dimension ratings relative to specific dimensions. On the basis of these analyses, they concluded that investigations of the construct validity of WEDRs using MTMM analyses have been misguided and instead, research should focus on reducing the magnitude of the general factor and better distinguishing between AEDRs. The present study directly examines the plausibility of a general factor model relative to multidimensional models to directly examine this matter.

Some evidence for a general factor would not be surprising, as Sackett and Dreher (1982) also noted substantial overlap among different dimensions, and research on the criterion domain substantiates evidence for a general performance factor in job performance ratings (Viswesvaran, Schmidt, & Ones, 2005). Practically, a general factor model is consistent with the frequent use of an overall assessment rating (OAR) as a composite indicator of AC performance in both research and practice, where an OAR represents the aggregation of ratings on separate dimensions (and/or exercises) into a single overall score. In sum, the primary goal of this study was to meta-analytically summarize and provide a direct comparison of alternative models of the structure underlying AC ratings.

AC Factors and External Correlates

Consistent with recent suggestions to use a nomological network approach to evaluating the construct-related validity of ACs (Hoffman et al., 2011; Rupp et al., 2008; Thornton & Rupp, 2012), we supplement this review of the structural validity of AC dimensions with a meta-analysis of the correlations between the empirically supported dimensions and a nomological network of individual difference constructs. Whereas the factor analytic approach described above is based on the correlations among indicators within the same method of measurement (i.e., the AC), the nomological network strategy to construct validation hinges on the focal construct evincing relationships with theoretically relevant *external* variables (Cronbach & Meehl, 1955; Woehr et al., 2012). The application of nomological network analyses in conjunction with the confirmatory factor analyses (CFAs) will afford multiple perspectives toward evaluating the construct-related validity of AC dimensions.

The present nomological network analysis focuses on general mental ability (GMA) and the FFM and in particular uses socioanalytic theory as a framework to understand the relationships between the FFM and AC dimensions. The FFM is useful because it has been explicitly used as a framework to propose relationships between the FFM and AC

dimensions (Lievens et al., 2006). However, we do not offer specific hypotheses for these relationships, because the underlying model must be discerned from the CFAs, and providing competing hypotheses based on multiple dimensions models is impractical. However, below we briefly review past studies linking individual differences to AC dimensions and general expectations concerning the nature of these relationships on the basis of socioanalytic theory.

General mental ability. ACs are typically administered for complex roles, usually supervisory/managerial positions. GMA has been proposed to be a key antecedent across diverse facets of job performance. In fact, GMA has been proposed to underlie findings of a general factor of job performance (Schmidt & Hunter, 1998; Viswesvaran et al., 2005), and it is expected that GMA will be correlated with AC factors in all models. Thus, although we expect GMA to be associated with performance on all dimensions, we do not expect these relationships to be uniform. Specifically, we expect that GMA should be more strongly related to problem-solving and administratively oriented factors.

As noted by Thornton and Rupp (2012), some dimensions are expected to be more "cognitively loaded" than others. In the present context, those dimensions specified to load on Arthur et al.'s (2003) problem solving and the broader technical skills factors are expected to reflect intelligence to a greater extent than other dimensions. Consistent with this suggestion, Shore et al. (1990) found that cognitive variables were more strongly correlated with a performance-style factor than they were with an interpersonal-style factor. Finally, although their focus was on incremental validity, Meriac et al.'s (2008) meta-analysis reported correlations between GMA and Arthur et al.'s (2003) broad dimensions, and GMA tended to evidence stronger associations with more problem-solving-oriented dimensions. Accordingly, we expect GMA to be especially strongly related to administrative/technically oriented factors across the various performance models.

Five factor model of personality. The expected relationships between AC factors and personality are somewhat more complex. Scholars have proposed relationships between the FFM and AC dimensions, but evidence for associations between personality and AC dimensions has been mixed. For instance, Lievens et al. (2006) proposed several expected relationships between Arthur et al.'s (2003) framework and FFM factors. Specifically, they offered that communication and influencing others should be more strongly related to extraversion, consideration/awareness of others should be related to agreeableness, drive and organizing and planning should be related to conscientiousness, stress tolerance should be related to emotional stability, and problem solving should be related to openness to experience. Dilchert and Ones (2009) provided support for some but not all of these SME-based associations between personality and AC dimensions. Other scholars have proposed broader associations. For instance, Shore et al. (1990) found that their interpersonal-style factor was more strongly related to several conceptually similar personality variables, compared to their performance-style factor. Thus, some studies have supported the hypothesized nomological network of AC dimensions whereas others have supported a few broad dimensions.

Drawing from socioanalytic theory, we expect meaningful patterns of results to emerge, although they are less nuanced given the parsimony of this framework relative to existing AC models. Furthermore, the FFM does not directly measure getting along and getting ahead, few if any personality inventories do. Instead, the value of socioanalytic theory is in its

applicability across personality measures and constructs (Hogan & Shelton, 1998). Previous research has successfully mapped the FFM onto getting along and getting ahead and documented expected patterns of relationships between these traits and on-the-job behaviors (Bartram, 2005; Chiaburu et al., 2011; Hogan & Holland, 2003).

Moreover, the applicability of this interpretation is not limited to organizational settings; a similar two-factor model based on the FFM has been substantiated across the psychological sciences on the basis of both factor analytic and nomological network analyses (DeYoung, 2006; Digman, 1997; Markon, Krueger, & Watson, 2005). Although the labels differ across disciplines, extraversion and openness are grouped to form a factor corresponding to getting ahead from socioanalytic theory and conscientiousness, agreeableness, and emotional stability are grouped to form a factor indicative of getting along (Bartram, 2005; Hogan & Holland, 2003). Thus, there is considerable interdisciplinary support for this interpretation of the FFM. Finally, because socioanalytic theory is a theory of the ways in which personality traits predict different interpersonal styles at work (Hogan & Shelton, 1998) and has been explicitly linked to leadership styles (Hogan & Kaiser, 2005) it provides a particularly useful framework for evaluating the nomological network of interpersonal simulations, such as ACs. On the basis of socioanalytic theory, we expected that conscientiousness and agreeableness would be more strongly associated with dimensions indicative of getting along (e.g., consideration/awareness of others) and extraversion and openness would be more strongly associated with getting ahead (e.g., drive; Hogan & Holland, 2003). In short, the present review combined CFA and nomological network approaches to construct validation to provide a comprehensive analysis of the construct-related validity of AC dimensions. The goals of these analyses were, first, to test a generalizable model of AC dimensions and, second, to test a theoretically derived nomological network of the dimension structure through the lens of socioanalytic theory.

Method

Meta-Analytic Procedure

Dimension intercorrelations reported in primary studies were synthesized using meta-analytic procedures recommended by Hunter and Schmidt (2004). To locate studies for inclusion, the following databases were searched: PsycINFO, Web of Science, and Business Source Premier. The following search terms were used in these databases to identify studies: *assessment center, AC*, and *dimension ratings*. Also, the reference lists of previous AC meta-analyses were examined. The initial search resulted in a total of 664 studies that were further reviewed for inclusion.

Once relevant primary studies were identified, they were evaluated based on several inclusion criteria. Specifically, to be included, (a) a study must have reported dimension intercorrelations or values that could be converted into correlations, (b) dimension intercorrelations must have been reported as across-exercise or postconsensus dimension ratings (i.e., they could not be WEDRs), (c) dimension labels must have been provided, and (d) the study must have reported the size of the sample on which dimension intercorrelations were computed. Authors of studies that presented only partial data were contacted to request the necessary information from their studies. After evaluating whether located studies met the inclusion criteria, a total of 68 independent samples were identified (references for these

studies are available as an online supplement). The ACs in the studies retained for the final analyses evaluated an average of 12.92 dimensions (*Mdn* = 11 using an average of 5.78 exercises; *Mdn* = 5). The mean number of participants (*n*) in each study was 346.53 (*Mdn* = 153, total *N* = 23,564).

Coding of Primary Studies

Dimension intercorrelations from primary studies were used as input for the meta-analysis and subsequent CFAs. Consistent with prior reviews (e.g., Arthur et al., 2003), dimension labels from primary studies were coded into Thornton and Byham's (1982) list of 33 commonly used AC dimension labels. This list served as a common framework for classifying dimensions reported in the primary studies, and has been utilized by previous meta-analyses (e.g., Arthur et al., 2003) to provide a common approach of grouping of primary AC dimensions. Arthur et al. found that the majority of dimensions listed in the primary studies they identified could be classified using one of these labels. Hence, primary dimensions that were categorized into the dimensions in this list served as indicators for the models tested in this study.

Primary study information (i.e., correlations, dimension labels, sample size) was coded by the first and second authors with the help of two industrial-organizational psychology doctoral students. Definitions of dimensions were reviewed, and dimensions were then coded by the first and second authors and a doctoral student into Thornton and Byham's (1982) dimension labels. The coders initially agreed on 98% of the dimension classifications, and the remaining discrepancies were resolved by discussion. When two dimensions (e.g., planning, organizing) were identified that fit into the same Thornton and Byham label (e.g., planning and organizing), the average of the two dimensions was retained for inclusion in the meta-analyses to ensure that studies were not double counted (i.e., to ensure independence).

Analyses

Once primary dimensions were sorted into Thornton and Byham's (1982) dimension labels, the meta-analysis procedures developed by Hunter and Schmidt (2004) were employed, and sample-weighted mean correlations were computed using the SAS PROC MEANS syntax developed by Arthur, Bennett, and Huffcutt (2001). The meta-analytically derived correlation coefficients were used to construct a correlation matrix among the AC dimensions mentioned above (see Table 1). In several instances, no primary study dimensions could be identified that represented a Thornton and Byham dimension. Specifically, we could not locate sufficient primary study information to include the following nine dimensions in the analyses: (a) *extra-organizational awareness*, (b) *extra-organizational sensitivity*, (c) *recognition of employee safety needs*, (d) *integrity*, (e) *practical learning*, (f) *technical and professional knowledge*, (g) *resilience*, (h) *development of subordinates*, and (i) *range of interests*. These dimensions either were not reported by any studies or had very few correlations with other dimensions. Four dimensions had very few missing correlations with other variables, but their inclusion as measured variables in the analyses would have resulted in empty cells/missing correlations. However, these variables were conceptually similar, and fell into the same categories of Arthur et al.'s (2003) model. As a result, the following four composite variables were created: (a) *oral communication/oral presentation*, (b) *organizational awareness/organizational sensitivity*, (c) *delegation/control*, and (d) *adaptability/risk*

Table 1
Dimension Intercorrelations, Sample Size (*N*), and Number of Studies (*k*)

	1	2	3	4	5	6	7	8	9	10
1. OC	—	17,162 (34)	4,984 (17)	11,556 (44)	2,073 (11)	6,210 (24)	2,457 (14)	4,767 (9)	1,190 (4)	6,660 (20)
2. WC	.38	—	3,630 (11)	8,223 (28)	700 (4)	3,697 (11)	923 (4)	4,673 (8)	554 (2)	4,541 (10)
3. OS	.37	.24	—	1,763 (12)	1,629 (8)	4,857 (16)	190 (2)	172 (1)	372 (2)	2,694 (15)
4. SN	.45	.37	.33	—	1,025 (6)	2,904 (18)	2,457 (14)	4,608 (8)	1,190 (4)	5,607 (16)
5. CA	.37	.31	.22	.26	—	1,629 (8)	172 (1)	266 (2)	172 (1)	1,457 (7)
6. EN	.52	.28	.39	.43	.48	—	1,440 (8)	263 (2)	608 (2)	3,168 (17)
7. IT	.44	.17	.58	.31	.64	.52	—	172 (1)	608 (2)	674 (3)
8. JM	.49	.37	.68	.44	.64	.53	.64	—	554 (2)	3,292 (2)
9. TN	.36	.27	.36	.32	.65	.41	.60	.40	—	1,018 (3)
10. WS	.41	.38	.35	.46	.38	.59	.60	.51	.50	—
11. ID	.42	.13	.29	.23	.29	.55	.57	.55	.67	.43
12. LD	.52	.32	.44	.45	.33	.64	.56	.41	.36	.44
13. DL	.26	.35	.41	.22	.62	.48	.34	.61	.42	.32
14. PO	.45	.42	.34	.38	.36	.47	.42	.43	.47	.47
15. AN	.41	.36	.40	.40	.34	.44	.34	.47	.60	.47
16. CR	.41	.28	.20	.40	.50	.47	.38	.57	.46	.52
17. DS	.40	.33	.21	.36	.34	.34	.51	.44	.59	.42
18. JG	.44	.38	.37	.43	.37	.44	.43	.44	.35	.47
19. AD	.45	.23	.52	.52	.35	.52	.40	.35	.32	.37
20. ST	.55	.35	.31	.43	.32	.50	.36	.54	.55	.44

(continued)

Table 1 (continued)

	11	12	13	14	15	16	17	18	19	20
1. OC	4,697 (15)	20,858 (52)	2,132 (13)	20,731 (46)	16,306 (32)	6,833 (21)	9,604 (28)	19,107 (45)	6,873 (28)	7,504 (26)
2. WC	2,902 (7)	16,450 (31)	1,737 (10)	17,077 (33)	14,246 (20)	1,095 (4)	7,736 (18)	16,261 (31)	5,194 (19)	6,064 (18)
3. OS	3,230 (6)	5,075 (18)	570 (4)	4,908 (17)	3,682 (9)	1,365 (9)	3,664 (8)	5,075 (8)	5,075 (18)	2,666 (15)
4. SN	1,217 (7)	11,645 (44)	2,132 (13)	10,956 (39)	6,467 (22)	2,817 (14)	5,708 (19)	9,618 (36)	3,311 (20)	6,380 (21)
5. CA	1,048 (5)	1,924 (10)	172 (1)	2,073 (11)	1,421 (7)	675 (4)	1,421 (7)	2,073 (11)	1,723 (9)	1,872 (10)
6. EN	3,751 (8)	6,301 (25)	352 (2)	5,266 (17)	3,902 (10)	1,894 (11)	3,804 (9)	5,173 (19)	4,997 (17)	2,757 (16)
7. IT	708 (4)	2,457 (14)	1,118 (6)	1,600 (7)	1,427 (9)	1,933 (8)	584 (5)	1,329 (8)	963 (4)	1,006 (4)
8. JM	250 (2)	4,460 (9)	172 (1)	4,767 (9)	4,385 (8)	263 (2)	3,987 (7)	4,460 (9)	804 (4)	3,922 (7)
9. TN	436 (1)	1,190 (4)	372 (2)	1,190 (4)	172 (1)	608 (2)	172 (1)	754 (3)	754 (3)	172 (1)
10. WS	1,475 (6)	6,660 (20)	380 (2)	6,323 (18)	4,449 (9)	1,177 (6)	4,211 (7)	5,986 (17)	3,076 (16)	5,404 (15)
11. ID	—	4,697 (15)	201 (2)	4,541 (13)	4,037 (11)	436 (1)	4,091 (12)	4,261 (14)	3,480 (8)	1,388 (9)
12. LD	.55	—	2,057 (12)	19,749 (43)	15,833 (30)	2,669 (13)	9,529 (27)	19,393 (46)	6,873 (28)	7,446 (26)
13. DL	.22	.39	—	2,038 (12)	1,503 (11)	805 (2)	758 (8)	2,057 (12)	1,333 (8)	1,186 (5)
14. PO	.35	.50	.47	—	15,736 (27)	1,886 (7)	9,366 (26)	19,034 (42)	6,706 (27)	7,337 (25)
15. AN	.33	.46	.43	.59	—	1,043 (4)	8,982 (25)	15,702 (29)	4,816 (16)	5,966 (15)
16. CR	.53	.49	.21	.45	.34	—	172 (1)	1,541 (7)	2,140 (11)	1,748 (8)
17. DS	.28	.40	.64	.55	.48	.67	—	9,273 (27)	4,305 (15)	5,387 (15)
18. JG	.33	.52	.39	.65	.61	.53	.57	—	6,873 (28)	7,595 (27)
19. AD	.43	.55	.29	.37	.38	.31	.29	.41	—	3,535 (18)
20. ST	.40	.49	.26	.45	.46	.32	.50	.49	.50	—

Note: Sample-weighted mean correlations are beneath the diagonal. Cell *N* and *k* are above the diagonal, with *k* values in parentheses. AD = risk taking/adaptability; AN = analysis; CA = career ambition; CR = creativity; DL = delegation/control; DS = decisiveness; EN = energy; ID = independence; IT = initiative; JG = judgment; JM = job motivation; LD = leadership; OC = oral communication/oral presentation; OS = organizational sensitivity/organizational awareness; PO = planning and organizing; SN = sensitivity; ST = stress tolerance; TN = tenacity; WC = written communication; WS = work standards.

taking. As a result, the final input matrix contained 20 manifest variables, representing 24 of Thornton and Byham's (1982) list of commonly used dimensions. In total, this matrix was composed of 190 separate meta-analytically derived correlation coefficients. These values served as input for the CFAs.

As described by Viswesvaran and Ones (1995), the combination of meta-analysis with covariance structure analysis offers the opportunity to conduct a CFA or evaluate a structural equation model with the data. Once meta-analytic estimates were derived, the sample-weighted meta-analytic correlation matrix was subjected to a set of CFAs to investigate the structure underlying AC performance. These analyses were conducted using LISREL 8.70 (Jöreskog & Sörbom, 2004), and models were compared using several model-data fit indices. As recommended by Viswesvaran and Ones, the harmonic mean (M = 972) of the sample sizes for the mean correlations was used as the sample size for the subsequent analyses.

Overall model fit was examined by comparing the relative fit across models. Four goodness-of-fit indices were examined, including the chi-square (χ^2) model fit test statistic, Steiger's (1990) root mean square error of approximation (RMSEA), the comparative fit index (CFI; Bentler, 1990), and Akaike's (1987) information criterion (AIC). The χ^2 test for goodness of fit is rarely used in isolation, since it is prone to reject anything other than perfect fit, especially when the sample size is large (Brown, 2006). However, χ^2 allows for a comparison of models when they are arranged in a parameter-nested sequence. RMSEA (Steiger, 1990) provides a test that makes an adjustment for model complexity (i.e., impacted by degrees of freedom). Values of .06 or less indicate a close fit to the data, and values above .08 are out of acceptable range (Hu & Bentler, 1999). The CFI is an incremental measure of fit relative to a null model; values can range from 0 to 1.0, where values of .95 or greater indicate an acceptable level of fit (Hu & Bentler, 1999). AIC represents a model fit index that can be used to compare non-nested models, where smaller values indicate better fit. Models were compared by evaluating this set of fit indices to determine which model best explains the structure of AC dimensions.

Results

Factor Structure of AC Dimensions

The first objective of this study was to evaluate the appropriateness of Arthur et al.'s (2003) framework. Each of the Thornton and Byham (1982) dimensions loaded onto one of Arthur et al.'s (2003) AC categories as proposed in their initial study. The only constraints placed on the model were setting factor variances to 1.0 for starting values so the models could converge. We began by testing the six-factor framework corresponding to Arthur et al.'s original framework but without stress tolerance. Based on rational grounds, the meta-analytic correlations, and past AC research (e.g., Schmitt, 1977), written communication was specified as loading on administrative and problem-solving dimensions in the six-factor model. The six-factor model fit the data very well in absolute terms, as all indices were within acceptable rules-of-thumb, $\chi^2(120) = 463.35$; RMSEA = .054; CFI = .95; AIC = 565.35 (see Table 2).

Because researchers in previous studies included stress tolerance as a seventh factor, we also examined the fit of this model. The seven-factor model also fit the data well in absolute terms, $\chi^2(149) = 579.88$; RMSEA = .052; CFI = .95; AIC = 701.88. However, this model also

Table 2
Model-Data Fit Indices for Confirmatory Factor Analysis Models

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	RMSEA	CFI	AIC
Six-factor	463.345	120			.054	.947	563.345
Three-factor	468.642	132	5.297^{ns}	12	.051	.947	546.642
Two-factor	474.913	134	6.271*	2	.051	.947	548.913
One-factor	543.804	135	68.891***	1	.056	.940	615.804

Note: AIC = Akaike information criterion; CFI = comparative fit index; RMSEA = root mean square error of approximation.
*$\Delta\chi^2$ between models was significant at $p < .05$.
***$\Delta\chi^2$ between models was significant at $p < .001$.

contained inadmissible parameter estimates. Specifically, a latent factor correlation was greater than 1.0, between consideration/awareness of others and stress tolerance ($\varphi = 1.09$) and two other correlations were very strong ($\varphi = .99$), indicating redundancy among the factors. In contrast to the seven-factor model, all of the parameter estimates were admissible in the six-factor model. Having said that, this solution also yielded substantial overlap among the factors (i.e., $\varphi = 1.0$ in the standardized solution between communication and consideration/awareness of others), and the correlation between planning and organizing and problem solving was $\varphi = .99$, indicating potential redundancy in the factors. Nevertheless, given that this model provided an acceptable fit and did so with admissible parameter estimates, we retained the six-factor model as a baseline for the remaining model tests.

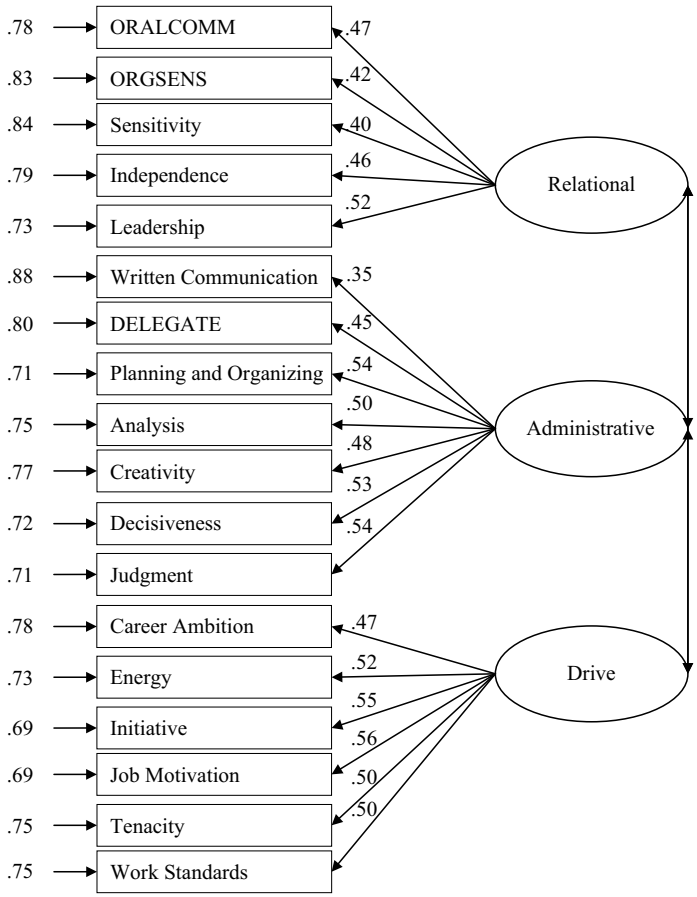
The three-factor model also provided a close fit to the data in absolute terms, $\chi^2(132) = 468.64$; RMSEA = .051; CFI = .95; AIC = 546.64 (see Figure 2), did not differ significantly from the six-factor model, $\Delta\chi^2(12) = 5.29$, $p = .95$, and provided a more parsimonious accounting of the data. Thus, given that the more parsimonious three-factor model did not yield a significant decrement in model fit and multiple latent factor correlations met or approached unity in the six-factor model, the three-factor model was retained as the more appropriate model. However, drive and relational skills overlapped substantially ($\varphi = .94$). Thus, the three-factor model was compared with the two-factor model to more directly test their distinguishability.

We tested a two-factor model that collapsed drive and relational skills to form a broader interpersonal skills factor. This model also provided a close fit to the data, $\chi^2(134) = 474.91$; RMSEA = .051; CFI = .95; AIC = 548.91. However, the two-factor model yielded a modest but significant decrement in fit relative to the three-factor model, $\Delta\chi^2(2) = 6.27$, $p < .05$. Thus, although relational skills and drive were very strongly correlated in the three-factor model, the significant decrement in fit suggests that, to some extent, they were empirically distinguishable.

For the sake of comprehensiveness, we also tested a general factor model by collapsing all manifest dimensions into a single factor. The one-factor model resulted in a significant and practically meaningful decrement in model fit relative to the two-factor model, $\chi^2(135) = 543.80$; RMSEA = .056; CFI = .94; AIC = 615.80; $\Delta\chi^2(1) = 68.89$, $p < .001$, suggesting a clear distinction between administrative skills and the broader interpersonal skills factor.

Together, the three-factor model specifying administrative skills, relational skills, and drive provided the most appropriate approximation of the data on the basis of parsimony

Figure 2
Path Diagram for Three-Factor Model With Standardized Estimates



Note: Three of the manifest indicators were pairs of conceptually similar variables that were collapsed to form a full matrix. DELEGATE = delegation/control; ORALCOMM = oral communication/oral presentation; ORGSENS = organizational sensitivity/organizational awareness.

relative to the six-factor model, and the modest but significant decrement in fit for the two-factor model. However, the value of distinguishing drive from relational skills was equivocal based on their strong factor correlations and the small differences in fit between the two- and three-factor models. Thus, we turned our attention to the construct-related validity of these three broad factors on the basis of their overlap with a nomological network composed of the FFM and GMA.

Relationships With GMA and FFM Traits

To evaluate the nomological network of the three-factor model, we meta-analytically summarized the correlations between the FFM and GMA and dimensions from each of the

models (see Table 3). First, administrative skills were most strongly correlated with GMA (ρ = .30). To test for significant differences between estimates, we used the formula for testing dependent meta-analytic correlations provided by Riketta and Van Dick (2005). Results indicated that all three factors were significantly correlated with GMA, but tests for significant differences between correlations revealed that cognitive ability was more strongly correlated with administrative skills, compared with relational skills (ρ = .21; $Z = 7.38$, $p < .001$) or drive (ρ = .23; $Z = 3.27$, $p < .001$). Results also indicated that all three AC factors were significantly correlated with extraversion, but extraversion was more strongly correlated with drive (ρ = .25) than with administrative skills (ρ = .09; $Z = 3.00$, $p < .01$) or relational skills (ρ = .14; $Z = 2.01$, $p < .05$). Relational skills and adminstrative skills did not differ in their relationship with extraversion ($Z = 1.49$, $p = .14$). Relational skills were weakly but significantly related to agreeableness (ρ = .06, $p < .05$) but not conscientiousness or neuroticism.

Discussion

The notable absence of theory in articulating the factors measured by ACs has hindered comparisons across studies and possibly resulted in unrealistic expectations regarding the number of dimensions that we expect to be measured in ACs. By drawing from multiple literatures to test theoretically derived models and relationships with relevant variables in the nomological network analysis, this review offers an unprecedented understanding of the status of dimension measurement in ACs. The results of this theoretical and empirical analysis suggest that ACs, as typically conducted in the extant literature, distinguish between two or three broad factors roughly corresponding to administrative skills, drive, and, potentially, relational skills. Implications of this pattern of results for modern conceptualizations of the construct-related validity of ACs are discussed below.

Main Findings

By comparing the factor structure of Arthur et al.'s (2003) model relative to theoretically derived alternative models using a large-scale summary of the literature, this study advances the development of a generalizable model of AC performance (Arthur & Villado, 2008; Hoffman et al., 2011). Similarly, doing so with a large scale meta-analysis helps to bolster confidence in the relevance of the results to various applications of the AC method. We should note from the outset that although generalizability to typical ACs is a strength of a large-scale review of the literature, there is considerable variability in AC methods and even the behaviors associated with a given dimension label across local ACs. Thus, it is very likely that the present findings, although informative to the myriad applications of the AC method, do not describe every AC. We return to this issue where relevant throughout our discussion.

Next, despite the increasing use of Arthur et al.'s (2003) model, until now the empirical structure of this model has never been directly tested relative to alternatives, either in primary studies of AEDRs or meta-analytically. Our results show that three broad factors more parsimoniously account for Arthur et al.'s six factors and their associated subdimensions. In addition, these inferences are bolstered by converging evidence from both internal (Campbell & Fiske, 1959) and external (Cronbach & Meehl, 1955) approaches to construct validation. These main findings are indicative of a considerably more parsimonious interpretation of

Table 3
Assessment Center Factor Correlations With External Variables

Dimension	GMA	N	E	O	A	C
Six-factor						
Communication	.29* (.01) 15,402/34	–.09* (.04) 777/6	.11* (.03) 1,333/8	.12* (.02) 1,703/8	.10* (.03) 1,333/8	.05^{ns} (.03) 1,200/7
Consideration/awareness	.17* (.01) 8,582/30	–.06^{ns} (.04) 574/5	.06^{ns} (.03) 1,315/9	.06* (.03) 1,199/7	.07* (.03) 1,247/8	.13* (.03) 1,065/7
Drive	.23* (.01) 7,408/29	–.04^{ns} (.04) 788/5	.25* (.03) 1,159/8	.08^{ns} (.04) 794/6	.13* (.03) 1,091/7	.09^{ns} (.03) 909/6
Influencing others	.22* (.01) 14,906/34	.02^{ns} (.03) 1,020/7	.15* (.02) 2,134/11	.07* (.03) 1,645/9	.04^{ns} (.02) 1,693/10	.01^{ns} (.03) 1,443/8
Organizing and planning	.29* (.01) 14,549/27	–.06^{ns} (.04) 777/6	.11* (.03) 1,447/11	.10* (.02) 1,820/9	.08* (.03) 1,379/8	.06^{ns} (.03) 1,197/7
Problem solving	.29* (.01) 15,263/37	–.05^{ns} (.04) 777/6	.09* (.03) 1,450/9	.12* (.02) 1,927/10	.06^{ns} (.03) 1,557/10	.08* (.03) 1,200/7
Thee-factor						
Relational skills	.21*^{a} (.01) 14,860/34	–.02^{ns} (.03) 1,004/7	.14*^{a} (.02) 2,186/12	.10* (.02) 2,070/10	.06* (.02) 1,677/10	.04^{ns} (.03) 1,495/9
Administrative skills	.30*^{b} (.01) 15,107/36	–.06^{ns} (.04) 777/6	.09*^{a} (.03) 1,518/10	.12* (.02) 1,927/10	.06* (.03) 1,557/10	.07^{ns} (.03) 1,268/8
Drive	.23*^{a} (.01) 7,408/29	–.04^{ns} (.04) 788/5	.25*^{b} (.03) 1,159/8	.08^{ns} (.04) 794/6	.13* (.03) 1,091/7	.09^{ns} (.03) 909/6
Two-factor						
Interpersonal-style	.21*^{a} (.01) 14,979/35	–.02^{ns} (.03) 1,445/8	.15* (.03) 2,186/12	.10* (.02) 2,070/10	.09* (.02) 2,118/11	.02^{ns} (.02) 1,936/10
Performance-style	.30*^{b} (.01) 15,107/36	–.06^{ns} (.04) 777/6	.09* (.03) 1,518/10	.12* (.02) 1,927/10	.06* (.03) 1,557/10	.07^{ns} (.03) 1,268/8
One-factor						
General performance	.26* (.01) 15,180/37	–.03^{ns} (.03) 1,445/8	.14* (.02) 2,186/12	.10* (.02) 2,177/11	.09* (.02) 2,225/12	.04^{ns} (.02) 1,936/10

Note: Each cell contains the AC factor's composite correlation with nomological network variables. Each cell contains the meta-analytic correlation, standard error (in parentheses), sample size, and number of studies on which the values are based. Different superscripts indicate that correlations between AC factors and external correlates are significantly different at $p < .05$. Reliability information from primary studies was used to correct the nomological network variables as follows: cognitive ability (GMA; .90), neuroticism (N; .90), extraversion (E; .76), openness (O; .74), agreeableness (A; .73), and conscientiousness (C; .74).
*$p < .05$.

ACs than is typical based on the average number of dimensions measured locally (i.e., 13 in the present review) and even relative to more parsimonious frameworks (e.g., Arthur et al., 2003). As we articulate in the implications, this central finding offers a host of implications to AC design, interpretation, and even formatting.

Finally, the present support for two to three overarching factors is remarkably consistent with more recent MTMM-type analyses of WEDRs (Hoffman et al., 2011; Hoffman & Meade, 2011). Together, the convergence in (a) the analysis of WEDRs, (b) the results of the analyses of AEDRs presented here, and (c) various theoretical models indicate consistent findings on approximately three broader factors in ACs. In this way, despite substantial disagreement in recent years, it seems that the AC literature, regardless of analytic method or scoring approach, is beginning to converge on a dimension structure composed of two to three overarching performance dimensions that correspond to behavioral styles identified across different domains.

This is important because AC MTMM and more dimension-based research (usually using AEDRs) has occurred in parallel and been characterized by conflicting results. Proponents of the dimension-based perspective have argued that WEDRs are an inappropriate unit of scoring and analysis (Arthur, 2012; Kuncel & Sackett, 2014; Rupp et al., 2008; Thornton & Rupp, 2012) because they are rarely used in applied settings. Similarly, task-based AC advocates have argued that AEDRs are an inappropriate unit of analysis (Jackson, 2012; Lance, 2012) because they omit exercises. In conjunction with recent MTMM research, these results indicate that there may be more in common across the results of these seemingly disparate approaches, at least in terms of the dimension structure underlying ACs. Of course these two scoring and analytic approaches continue to provide somewhat unique perspectives, with AEDRs providing a scoring unit that is more consistent with current practice and WEDRs providing a more molecular analysis of the inner workings of ACs. Nevertheless, the consistency across approaches in dimension structure is encouraging for the integration and resolution of these seemingly competing perspectives.

Implications

The present results are discussed in light of three overarching implications for AC research and practice. First, these findings have implications for the ongoing debate as to whether ACs measure dimensions (Lance, 2008). Second, these results provide a preliminary indication of a potential deficiency in the AC method. Third, these findings have implications for how AC dimensions are conceptualized in AC research and practice.

Do ACs measure dimensions? First, the present findings are important in light of recent claims that ACs do not measure dimensions at all and that dimensions should be removed entirely from the AC process (Lance, 2008). Bolstered by the consistency across internal structure and nomological network analyses, our results indicate that it would be a mistake to remove dimensions entirely, given that ACs generally appear to capture up to three factors (administrative skills, relational skills, and drive). The categories underlying these broad factors seem to correspond with models stemming from the work performance (Borman & Brush, 1993), personality (Hogan & Shelton, 1998), leader behavior (Fleishman, 1957), and managerial context literatures (Dierdorff et al., 2009). Furthermore, as outlined

above, support for two to three broad factors has also emerged in MTMM-based analyses of WEDRs (Hoffman et al., 2011). Thus, in contrast to recent assertions that ACs measure a single general performance factor (Kuncel & Sackett, 2014) or multiple exercises, it appears that AC dimensions are alive and well.

Having said that, far fewer dimensions were supported than are commonly measured, the fit did not differ drastically between the models, and factors were strongly correlated within each of the models. Administrative skills were clearly supported based on the distinguishability from models across factors. However, the distinguishability of drive and relational skills, the two more noncognitive dimensions, was less clear. The nomological network approach yielded similar conclusions. GMA was expected to overlap with administrative skills, and the FFM factors were expected to associate with drive and relational skills in the directions proposed by socioanalytic theory. Notably, most of the correlations between the FFM and GMA and the AC dimensions were weak. However, the pattern of relationships is key in nomological network analyses (Cronbach & Meehl, 1955).

Given that informational complexity is a defining characteristic of the managerial role (Dierdorff et al., 2009; Mintzberg, 1973), the moderate associations between GMA and all three AC factors was not surprising. However, administrative skills reflect more cognitively oriented constructs. Thus, the stronger overlap between GMA and administrative skills is consistent with the expected nomological network of this factor. In addition, the clear drop in fit when administrative skills was collapsed with relational skills to form a general factor model further supports their distinguishability. In short, both sets of analyses supported the construct-related validity of the administrative skills factor.

Overall, few of the nomological network correlations based on the FFM corresponded to the trends predicted by socioanalytic theory. The only clear support based on this theory was for extraversion and drive. This pattern of results provides initial support for the construct-related validity of the second broad factor corresponding to getting ahead. In contrast with the predictions of socioanalytic theory and based on the available nomological network provided by past studies, it seems that relational skills as typically measured in ACs are more strongly associated with extraversion than with predictors of prosocial and relationship-oriented behaviors (i.e., conscientiousness, agreeableness, and emotional stability). Although drive is typically conceptualized as a motivational construct (Lievens et al., 2006) expected to be associated with conscientiousness, it seems to reflect social motivation (extraversion) to a greater extent than task motivation (conscientiousness). Together, although the present review does support the notion that the ACs measure dimensions, only a broad distinction between administrative skills and drive received consistent support across internal structure and nomological network analyses. On the other hand, relational skills were strongly associated with drive, and both were associated with known markers of getting ahead to a greater extent than known markers of getting along.

A potential deficiency? Three potential explanations exist for the limited distinction between drive and relational skills. The first possibility is that typical ACs do not capture both interpersonal styles described by socioanalytic theory and instead, on the basis of correlations with extraversion, primarily capture a style indicative of getting ahead. Neither the structural nor the nomological network analyses supported more nuanced dimension models from the AC literature (Lievens et al., 2006) or socioanalytic theory (Hogan & Kaiser, 2005).

On the basis of socioanalytic theory, one possible explanation for this pattern of results is that ACs do not adequately distinguish relational styles associated with getting along from those associated with getting ahead and instead measure a single broader interpersonal construct that is primarily associated with a marker of getting ahead.

Consistent with this possibility, narcissism, thought to be a direct indicator of more destructive manifestations of getting ahead (O'Boyle, Forsyth, Banks, & McDaniel, 2012), has been linked to higher ratings in leaderless group discussions (Brunell, Campbell, Gentry, Hoffman, & Khunert, 2008). Although extraversion is certainly a relevant predictor in leadership roles, it is a stronger predictor of leader emergence than leader effectiveness (Judge, Bono, Ilies, & Gerhardt, 2002). In addition, more surgent leader behavior styles may be less effective in the management of the highly skilled, interdependent, and autonomous modern workforce (Grant, Gino, & Hofmann, 2011; Wood & Hoffman, 2012). Instead, researchers have proposed that more prosocial and relational styles of management are needed (Dierdorff et al., 2009).

Importantly, the FFM factors we labeled as getting along are also predictive of prosocial and relational behaviors across organizational roles (Bartram, 2005). For instance, FFM domains commonly labeled as getting along, such as agreeableness, conscientiousness, and, to a lesser extent, emotional stability, have been meta-analytically linked to consideration-oriented leader behaviors (DeRue, Nahrgang, Wellman, & Humphrey, 2011), transformational leadership (Judge & Piccolo, 2004), leadership effectiveness (Judge et al., 2002), and more relational forms of OCB (Chiaburu et al., 2011). However, the constructs measured in ACs did not correlate with these established predictors of more prosocial or relational behaviors. Hence, this multimethod, construct-oriented (Arthur & Villado, 2008) evaluation points to the possibility that ACs may not capture the more prosocial or relational styles of leadership.

A second potential explanation for the limited support for the relational skills factor is that the nomological network was insufficient. For instance, the FFM does not directly measure getting along and getting ahead. However, the FFM traits have routinely been linked to relevant work behaviors in ways predicted by socioanalytic theory (Bartram, 2005). Similarly, past AC researchers have explicitly proposed linkages between the FFM and each of Arthur et al.'s (2003) broad dimensions in patterns consistent with socioanalytic theory (Lievens et al., 2006).

Thus, additional research is needed to further expand the nomological network of AC dimensions, especially research examining the distinction between drive and relational skills. Studies using the FFM facets, rather than factors, might yield stronger support for ACs and their capacity to measure relational skills. For instance, it would be interesting to know whether the strong correlations with extraversion emerge for both sociability and dominance facets. Similarly, although emotional intelligence has strong ties to conscientiousness and emotional stability from the FFM (Joseph & Newman, 2010), including constructs such as emotional intelligence would be informative to determining the efficacy of ACs to measure relational skills.

A third possible explanation for the limited support for these findings is that some ACs measure relational skills more effectively than others. As emphasized throughout the article, our present results summarize ACs in general, as reported in the available literature to date. However, ACs vary considerably by design and purpose (Thornton & Rupp, 2006). Thus, it

is possible that some ACs measure relational skills better than others. Clearly, future research on the nature and assessment of relational skills is needed.

How can researchers and practitioners incorporate broad factors? Broad-factor models have been utilized for years in psychological assessment for paper-and-pencil measures of constructs. For example, the FFM model of personality operates such that facet-level information offers more specific information about an individual's personality beyond the five broader factors (Barrick & Mount, 1991). The FFM was never intended to explain every aspect of personality, although it has served as a useful framework for categorizing personality information and allowing for broad comparisons across studies, even when using different inventories. This three-factor AC model may operate in a similar manner. Because we demonstrate the assignment of more narrow dimensions into broader factors, these dimensions are applicable to ACs in their existing form.

Importantly, these results should in no way be interpreted as evidence that practitioners should only measure three dimensions in a given AC; this would be impractical. However, the expectation that large numbers of primary study dimensions should be distinct seems unrealistic (Hoffman et al., 2011; Woehr & Arthur, 2003). Although we do not advocate the abandonment of narrow dimensions, we propose that incorporating narrow dimensions into more generalizable models is useful as a framework that can inform the multiple phases of the AC design, interpretation, and research process.

There is evidence, however that practitioners have been reducing the number of dimensions measured in ACs. For instance, 13 dimensions were measured on average in the ACs included in this review, but recent global surveys suggest that 6 to 7 dimensions have become the more popular number measured (Povah, 2011). Thus, although we were unable to examine only modern and arguably improved ACs in this review, it seems that our core finding of a reduced dimension structure has, to some extent, already been embraced by modern AC designers. Having said that, and on the basis of the present results, one alternative to reducing the absolute number of dimensions in ACs might be to continue to rate the number of dimensions required by operational demands, but to use a hierarchical approach to organize the manifest dimensions into a more parsimonious and empirically supported set of broad dimensions.

For instance, in AC design, generalizable theoretical models can serve as a guide when selecting dimensions in designing a new AC. Practitioners could use these types of models to ensure that facets of the broader dimensions are measured. Doing so can enhance coverage of each of the overarching performance constructs. Similarly, narrow dimensions used in assessor training and on rating forms might be organized around broad factors. This approach could potentially yield increased reliability within broad factors and increased distinction between broad factors. This approach is common in other measures in similar domains, such as Center for Creative Leadership's BENCHMARKS® multisource feedback tool.

With respect to administering feedback in developmental ACs, practitioners could take operational AC dimension ratings and use this framework for grouping existing dimensions for the purpose of conveying assessment results to feedback recipients, the assessee's supervisor, and other organizational decision makers. Viewing primary dimensions as belonging to higher-order categories may allow for the flexibility and specificity needed for effective feedback while enhancing the construct-related validity of subsequent recommendations.

This in turn could yield more specific feedback as well as a more parsimonious set of constructs for other purposes, such as conducting validation studies or succession planning.

Limitations and Future Directions

First, it is possible that other models exist that might better explain the structure of AC dimensions. We tested a subset of plausible models that were identified in our review, drawing from the job performance, leadership, and AC literatures. However, it is important to note that the dimension structure in local ACs will vary to some extent. For instance, it is possible that other dimensions could be added or existing dimension structures could be revised consistent with modern leadership theories. Furthermore, it is possible that using a different framework of manifest indicators (i.e., primary study dimensions), an alternate framework might have seen stronger empirical support. We encourage continued exploration into generalizable models.

Importantly, we forward this model as a framework, not as a description of every AC. Given the differences in how dimensions are operationalized across ACs, it is likely that some flexibility in the underlying dimension structure is needed. For instance, we set conflict and influence-oriented behaviors to load on relational skills. We did so because drive provided the closest approximation of getting ahead and on the basis of past taxonomies (Borman & Brush, 1993). However, clearly, these types of behaviors have a strong element of directiveness and thus, could reasonably be specified as a facet of drive. Depending on the nature of the exercises in the AC and the specific definition of a given narrow dimension, the broader loading could differ from the framework presented here. Thus, although we encourage the incorporation of this framework in future research and practice, we also urge flexibility in the dimension classification on the basis of operational differences across ACs.

Similarly, there is substantial variability in operational ACs, and it is likely that AC design characteristics impact the structure and nomological network results (Woehr & Arthur, 2003). For instance, assessor training, the use of behavioral rating scales, and the nature of the exercises could impact the dimension structure and network. Future research should examine the extent that AC design characteristics can result in more distinct factors and the ability to better distinguish relational skills from drive. Although studies have attended to the influence of AC design on MTMM results, design characteristics have not been considered with respect to AEDRs. This is an important consideration for AC researchers, given that dimensions are typically the focus of practical applications of the method.

Another limitation of the present study is the small number of studies available for several of the FFM correlations with the AC factors and among several of the AC subdimensions. Continued research exploring the FFM, GMA, and other relevant individual differences is warranted. Although Thornton and Byham's (1982) list of commonly used labels serves as a useful starting point for categorizing AC dimensions from operational ACs, several of these categories were not represented in any of the studies identified, and several dimension intercorrelations were represented by relatively few studies. Future research might consider updating Thornton and Byham's taxonomy to incorporate any dimensions that have become more common in recent years and remove any less common dimensions.

General Conclusions

When compared to most mainstream predictors in psychological research, ACs have been deficient with respect to describing the constructs that are measured (Arthur et al., 2008). This quantitative review used multiple methods to arrive at a model of AC AEDRs consisting of administrative skills, drive, and possibly, relational skills. These results point to a considerably more parsimonious accounting of AC data than has been previously proposed. The results of this study support the continued use of dimensions in ACs and offer a potentially useful framework to guide future AC research and practice. However, these results also draw attention to a possible challenge in measuring relational skills, which warrants additional empirical attention.

References

Akaike, H. 1987. Factor analysis and AIC. *Psychometrika*, 52: 317-332.

Arthur, W., Jr. 2012. Dimension based assessment centers: Theoretical perspectives. In D. J. Jackson, B. J. Hoffman, & C. E. Lance (Eds.), *The psychology of assessment centers*: 96-120. New York: Routledge.

Arthur, W., Jr., Bennett, W., Jr., & Huffcutt, A. I. 2001. *Conducting meta-analysis using SAS*. Mahwah, NJ: Lawrence Erlbaum.

Arthur, W., Jr., Day, E. A., McNelly, T. L., & Edens, P. S. 2003. A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 56: 125-154.

Arthur, W., Jr., Day, E. A., & Woehr, D. J. 2008. Mend it, don't end it: An alternate view of assessment center construct-related validity evidence. *Industrial and Organizational Psychology*, 1: 105-111.

Arthur, W., Jr., & Villado, A. J. 2008. The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practice. *Journal of Applied Psychology*, 93: 435-442.

Austin, J. T., & Crespin, T. R. 2006. From "criterion problem" to problems of criteria in industrial and organizational psychology: Progress, pitfalls, and prospects. In W. Bennett, Jr., C. E. Lance, & D. J. Woehr (Eds.), *Performance measurement: Current perspectives and future challenges*: 9-48. Mahwah, NJ: Lawrence Erlbaum.

Austin, J. T., & Villanova, P. D. 1992. The criterion problem: 1917-1992. *Journal of Applied Psychology*, 77: 836-874.

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

Bartram, D. 2005. The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90: 1185-1203.

Bentler, P. M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin*, 107: 238-246.

Borman, W. C., & Brush, D. H. 1993. More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6: 1-21.

Borman, W. C., & Motowidlo, S. J. 1997. Task performance and contextual performance: The meaning for personnel selection research. *Human Performance*, 10: 99-109.

Bowers, D. G., & Seashore, S. E. 1966. Predicting organizational effectiveness with a four-factor theory of leadership. *Administrative Science Quarterly*, 11: 238-263.

Bowler, M. C., & Woehr, D. J. 2006. A meta-analytic evaluation of the impact of dimension and exercise factors on assessment center ratings. *Journal of Applied Psychology*, 91: 1114-1124.

Bray, D. W., & Grant, D. L. 1966. The assessment center in the measurement of potential for business management. *Psychological Monographs: General and Applied*, 80: 1-27.

Brown, T. A. 2006. *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Brunell, A., Campbell, K., Gentry, B., Hoffman, B., & Khunert, K. 2008. Emergent leadership: The case of narcissism. *Personality and Social Psychology Bulletin*, 34: 1663-1676.

Campbell, J. P., & Fiske, D. W. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56: 81-105.

Chiaburu, D. S., Oh, I., Berry, C. M., Li, N., & Gardner, R. G. 2011. The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, 96: 1140-1166.

Connelly, B. S., Ones, D. S., Ramesh, A., & Goff, M. 2008. A pragmatic view of assessment center exercises and dimensions. *Industrial and Organizational Psychology*, 1: 121-124.

Conway, J. M. 1999. Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 84: 3-13.

Cronbach, L., & Meehl, P. 1955. Construct validity in psychological tests. *Psychological Bulletin*, 52: 281-302.

DeRue, D. S., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. 2011. Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, 64: 7-52.

DeYoung, C. G. 2006. Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91: 1138-1151.

Dierdorff, E. C., Rubin, R. S., & Bachrach, D. G. 2012. Role expectations as antecedents of citizenship and the moderating effects of work context. *Journal of Management*, 38: 573-598.

Dierdorff, E. C., Rubin, R. S., & Morgeson, F. P. 2009. The milieu of managerial work: An integrative framework linking work context to role requirements. *Journal of Applied Psychology*, 94: 972-988.

Digman, J. M. 1997. Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73: 1246-1256.

Dilchert, S., & Ones, D. S. 2009. Assessment center dimensions: Individual differences correlates and meta-analytic incremental validity. *International Journal of Selection and Assessment*, 17: 254-270.

Fleishman, E. A. 1957. A leader behavior description for industry. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement*: 103-119. Columbus: Ohio State University, Bureau of Business Research.

Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., III, & Bentson, B. 1987. Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72: 493-511.

Gaugler, B. B., & Thornton, G. C., III. 1989. Number of assessment center dimensions as a determinant of assessor accuracy. *Journal of Applied Psychology*, 74: 611-618.

Gibbons, A. M., Rupp, D. E., Snyder, L. A., Hollub, A. S., & Woo, S. E. 2006. A preliminary investigation of developable dimensions. *Psychologist-Manager Journal*, 9: 99-123.

Goldberg, L. R. 1990. An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59: 1216-1229.

Grant, A. M., Gino, F., & Hofmann, D. A. 2011. Reversing the extraverted leadership advantage: The role of employee proactivity. *Academy of Management Journal*, 54: 528-550.

Guenole, N., Chernyshenko, O., Stark, S., Cockerill, T., & Drasgow, F. 2011. We're doing better than you might think: A large-scale demonstration of assessment centre convergent and discriminant validity. In N. Povah & G. C. Thornton III (Eds.), *Assessment centres and global talent management*: 15-32. Surrey, UK: Gower.

Hermelin, E., Lievens, F., & Robertson, I. T. 2007. The validity of assessment centres for the prediction of supervisory performance ratings: A meta-analysis. *International Journal of Selection and Assessment*, 15: 405-411.

Hinrichs, J. R. 1969. Comparison of "real life" assessments of management potential with situational exercises, paper-and-pencil ability tests, and personality inventories. *Journal of Applied Psychology*, 53: 425-432.

Hoffman, B. J. 2012. Exercises, dimensions, and the great battle of Lilliput: Evidence for a mixed model interpretation of AC performance. In D. J. Jackson, C. L. Lance, & B. J. Hoffman (Eds.), *The psychology of assessment centers*: 281-308. New York: Routledge.

Hoffman, B. J., Blair, C. A., Meriac, J. P., & Woehr, D. J. 2007. Expanding the criterion domain? A quantitative review of the OCB literature. *Journal of Applied Psychology*, 92: 555-566.

Hoffman, B. J., Lance, C. E., Bynum, B. H., & Gentry, W. A. 2010. Rater source effects are alive and well after all. *Personnel Psychology*, 63: 119-151.

Hoffman, B. J., & Meade, A. 2011. Alternate approaches to understanding the psychometric properties of assessment centers: An analysis of the structure and equivalence of exercise ratings. *International Journal of Selection and Assessment*, 20: 82-97.

Hoffman, B. J., Melchers, K., Blair, C. A., Kleinmann, M., & Ladd, R. T. 2011. Exercises and dimensions are the currency of assessment centers. *Personnel Psychology*, 64: 351-395.

Hoffman, B. J., & Woehr, D. J. 2009. Disentangling the meaning of multisource performance rating source and dimension factors. *Personnel Psychology*, 62: 735-765.

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. 1983. A socioanalytic theory of personality. In M. M. Page (Ed.), *1982 Nebraska symposium on motivation*: 55-89. Lincoln: University of Nebraska Press.

Hogan, R., & Kaiser, R. B. 2005. What we know about leadership. *Review of General Psychology*, 9: 169-180.
Hogan, R., & Shelton, D. 1998. A socioanalytic perspective of job performance. *Human Performance*, 11: 129-144.
Howard, A. 2008. Making assessment centers work the way they are supposed to. *Industrial and Organizational Psychology*, 1: 98-104.
Hu, L., & Bentler, P. M. 1999. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6: 1-55.
Hunter, J. E., & Schmidt, F. L. 2004. *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
International Task Force on Assessment Center Guidelines. 2009. Guidelines and ethical considerations for assessment center operations. *International Journal of Selection and Assessment*, 17: 243-253.
Jackson, D. J. R. 2012. Task-based assessment centers: Theoretical perspectives. In D. J. R. Jackson, C. E. Lance, & B. J. Hoffman (Eds.), *The psychology of assessment centers*: 173-189. New York: Routledge.
Jöreskog, K., & Sörbom, D. 2004. *LISREL 8.70*. Chicago: SSI.
Joseph, D. L., & Newman, D. A. 2010. Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology*, 95: 54-78.
Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. 2002. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87: 765-780.
Judge, T. A., & Piccolo, R. 2004. Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89: 755-768.
Katz, R. L. 1955. Skills of an effective administrator. *Harvard Business Review*, 33: 33-42.
Kolk, N. J., Born, M. P., & van der Flier, H. 2004. A triadic approach to the construct validity of the assessment center. *European Journal of Psychological Assessment*, 20: 149-156.
Kuncel, N. R., & Sackett, P. S. 2014. Resolving the assessment center construct validity problem (as we know it). *Journal of Applied Psychology*, 99: 38-47.
Lance, C. E. 2008. Why assessment centers do not work the way they are supposed to. *Industrial and Organizational Psychology*, 1: 84-97.
Lance, C. E. 2012. Research into task-based assessment centers. In D. J. R. Jackson, C. E. Lance, & B. J. Hoffman (Eds.), *The psychology of assessment centers*: 218-233. New York: Routledge.
Lance, C. E., Lambert, T. A., Gewin, A. G., Lievens, F., & Conway, J. M. 2004. Revised estimates of dimension and exercise variance components in assessment center postexercise dimension ratings. *Journal of Applied Psychology*, 89: 377-385.
LePine, J. A., Erez, A., & Johnson, D. E. 2002. The nature and dimensionality of organizational citizenship behavior: A critical review and meta-analysis. *Journal of Applied Psychology*, 87: 52-65.
Lievens, F., Chasteen, C. S., Day, E. A., & Christiansen, N. D. 2006. Large-scale investigation of the role of trait activation theory for understanding assessment center convergent and discriminant validity. *Journal of Applied Psychology*, 91: 247-258.
Lievens, F., & Christiansen, N. 2012. Core debates in assessment center research: Dimensions versus exercises. In D. J. Jackson, B. J. Hoffman, & C. E. Lance (Eds.), *The psychology of assessment centers*: 68-91. New York: Routledge.
Lievens, F., & Conway, J. M. 2001. Dimension and exercise variance in assessment center scores: A large-scale evaluation of multitrait-multimethod studies. *Journal of Applied Psychology*, 86: 1202-1222.
Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. 1994. The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology*, 47: 715-738.
Mackenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. 2011. Challenge-oriented organizational citizenship behaviors and organizational effectiveness: Do challenge-oriented behaviors really have an impact on the organization's bottom line? *Personnel Psychology*, 64: 559-592.
Markon, K. E., Krueger, R. F., & Watson, D. 2005. Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, 88: 139-157.
Meriac, J. P., Hoffman, B. H., Woehr, D. J., & Fleisher, M. S. 2008. Further evidence for the validity of assessment center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings. *Journal of Applied Psychology*, 93: 1042-1052.
Mintzberg, H. 1973. *The nature of managerial work*. New York: Harper & Row.
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.

Mumford, M. D., Zaccaro, S. J., Harding, F. D., Jacobs, T. O., & Fleishman, E. A. 2000. Leadership skills for a changing world: Solving complex social problems. *Leadership Quarterly*, 11: 23-35.

O'Boyle, E. H., Jr., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. 2012. A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, 97: 557-579.

Organ, D. W. 1988. *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.

Povah, N., 2011. A review of recent international surveys into assessment centre practices. In N. Povah & G. C. Thornton, III (Eds.), *Assessment centres and global talent management*: 329-350. Surrey, UK: Gower.

Povah, N. & Thornton, G. C., III. (Eds.). 2011. *Assessment centres and global talent management*. Surrey, UK: Gower.

Putka, D., & Hoffman, B. J. 2013. A closer look at reliable and unreliable sources of variance in assessment center ratings. *Journal of Applied Psychology*, 98: 114-133.

Riketta, M., & Van Dick, R. 2005. Foci of attachment in organizations: A meta-analytic comparison of the strength and correlates of workgroup versus organizational identification and commitment. *Journal of Vocational Behavior*, 67: 490-510.

Rupp, D. E., Gibbons, A. M., Baldwin, A. M., Snyder, L. A., Spain, S. M., Woo, S. E., Brummel, B. J., Sims, C. S., & Kim, M. 2006. An initial validation of assessment centers as accurate assessments for effective training interventions. *Psychologist-Manager Journal*, 9: 171-200.

Rupp, D. E., Thornton, G. C., III, & Gibbons, A. M. 2008. The construct validity of the assessment center method and usefulness of dimensions as focal constructs. *Industrial and Organizational Psychology*, 1: 116-120.

Sackett, P. R., & Dreher, G. F. 1982. Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology*, 67: 401-410.

Schmidt, F. L., & Hunter, J. E. 1998. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124: 262-274.

Schmitt, N. 1977. Interrater agreement in dimensionality and combination of assessment center judgments. *Journal of Applied Psychology*, 62: 171-176.

Scullen, S. E., Mount, M. K., & Goff, M. 2000. Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85: 956-970.

Shore, T. H., Thornton, G. C., III, & Shore, L. M. 1990. Construct validity of two categories of assessment center dimension ratings. *Personnel Psychology*, 43: 101-116.

Steiger, J. H. 1990. Structural model evaluation and modification: In interval estimation approach. *Multivariate Behavioral Research*, 25: 173-180.

Stogdill, R. M. 1974. *Handbook of leadership: A survey of theory and research*. New York: Free Press.

Thornton, G. C., III, & Byham, W. C. 1982. *Assessment centers and managerial performance*. New York: Academic Press.

Thornton, G. C., III, & Gibbons, A. 2009. Validity of assessment center for personnel selection. *Human Resource Management Review*, 19: 169-187.

Thornton, G. C., III, & Rupp, D. E. 2006. *Assessment centers in human resource management*. Mahwah, NJ: Lawrence Erlbaum.

Thornton, G. C., III, & Rupp, D. E. 2012. Research into dimension-based assessment centers. In D. J. Jackson, B. J. Hoffman, & C. E. Lance (Eds.), *The psychology of assessment centers*: 141-170. New York: Routledge.

Viswesvaran, C., & Ones, D. S. 1995. Theory testing: Combining psychometric meta-analysis and structural equation modeling. *Personnel Psychology*, 48: 865-885.

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. 2005. Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology*, 90: 108-131.

Williams, L. J., & Anderson, S. E. 1991. Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, 17: 601-617.

Woehr, D. J., & Arthur, W., Jr. 2003. The construct-related validity of assessment center ratings: A review and meta-analysis of the role of methodological factors. *Journal of Management*, 29: 231-258.

Woehr, D. J., Meriac, J. P., & Bowler, M. C. 2012. A review of methodological approaches in assessment center research. In D. J. Jackson, B. J. Hoffman, & C. E. Lance (Eds.), *The psychology of assessment centers*: 45-67. New York: Routledge.

Wood, L., & Hoffman, B. H. 2012. *Job characteristics: The times, they are a-changin'*. Poster presented at the 27th annual meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.