

GENERAL MENTAL ABILITY, NARROWER COGNITIVE ABILITIES, AND JOB PERFORMANCE: THE PERSPECTIVE OF THE NESTED-FACTORS MODEL OF COGNITIVE ABILITIES

JONAS W. B. LANG

Department of Work and Social Psychology
Maastricht University

MARTIN KERSTING

Federal University of Applied Administrative Sciences

UTE R. HÜLSHEGER

Department of Work and Social Psychology
Maastricht University

JESSICA LANG

Institute of Occupational Medicine
RWTH Aachen University

The nested-factors model is a well-established structural model of cognitive abilities in cognitive ability research but has not yet been used to investigate the role of cognitive abilities in job performance. Core assumptions of the nested-factors model are that a broad general mental ability (GMA) exists besides narrower abilities and that this GMA differs from the narrower cognitive abilities in breadth but not in subordination. The authors of this article propose that a recently emerging statistical technique—relative importance analysis—corresponds to the assumptions of the nested-factors model. To empirically study the implications of using the nested-factors model, the authors applied relative importance analysis to a meta-analytic matrix linking measures of 7 narrower cognitive abilities from an established ability taxonomy (Thurstone's primary mental abilities), GMA, and job performance. Results revealed that GMA accounted for 10.9% to 28.6% of the total variance explained in job performance and that GMA was not consistently the most important predictor. The discussion focuses on potential theoretical, methodological, and practical implications of the nested-factors model for personnel psychology.

Researchers have accumulated considerable evidence that cognitive abilities are highly successful predictors of job performance (Bertua, Anderson, & Salgado, 2005; Salgado, Anderson, Moscoso, Bertua, & de Fruyt, 2003; Schmidt, Hunter, & Pearlman, 1981). One longstanding

Correspondence and requests for reprints should be addressed to Jonas W. B. Lang, Department of Work and Social Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; jonas.lang@maastrichtuniversity.nl.

question in personnel psychology, however, is the importance of the general cognitive ability construct—general mental ability (GMA)—relative to the importance of narrower cognitive abilities in job performance. In addressing this question, personnel psychologists frequently rely on assumptions of Charles Spearman's two-factor theory (Spearman, 1904) and its successor—the higher-order factor model (Gustafsson & Balke, 1993; Yung, Thissen, & McLeod, 1999). In this paper, we refer to these two theories as Spearmanian theories. As we detail in this paper, researchers commonly either explicitly draw on the assumptions of Spearmanian theories or implicitly rely on statistical methodology in a way that resembles the assumptions of the Spearmanian theories when they examine the role of GMA and narrower cognitive abilities in job performance. From the perspective of cognitive ability research, the sole reliance on the assumptions of Spearmanian theories may be somewhat surprising. An alternative view of the structure of cognitive abilities has emerged in the literature since the middle of the last century (Holzinger & Harman, 1938; Holzinger & Swineford, 1937; Thurstone, 1944). This alternative view is known as the nested-factors model, the general hierarchical model of cognitive abilities, or the bifactor model and has been endorsed by cognitive ability scholars like Holzinger (Holzinger & Harman, 1938; Holzinger & Swineford, 1937), Humphreys (1981), Gustafsson (Gustafsson & Balke, 1993), and Carroll (1995, 2003). Throughout this paper, we refer to this view as the nested-factors model. Although the nested-factors model shares many elements with Spearmanian theories, it differs from these theories in important aspects. As a result, there is currently a discrepancy between the prevalent theoretical models in cognitive ability research (both Spearmanian theories and the nested-factors model) and the theoretical perspective commonly used for investigating the role of cognitive abilities in the field of personnel psychology (only Spearmanian theories).

Measurement conceptualizations like Spearmanian theories and the nested-factors model can have important implications for substantive conclusions regarding relationships between a construct and other constructs. As Schwab (1980, see also Bacharach, 1989) has noted, substantive conclusions regarding a construct are only valid when the measurement theory for the construct is also valid. Consequently, the specification of an underlying measurement model (or theory) for a construct profoundly affects subsequent theoretical hypotheses, decisions on the use of statistical methods, and theoretical conclusions in theory-driven research. When alternative measurement conceptualizations exist, it is therefore critically important to investigate their implications. This paper consequently extends extant research on the relationship between cognitive abilities and job performance (relying predominantly on Spearmanian theories) and investigates the relationship between cognitive abilities and job

performance using the alternative perspective of the nested-factors model. In so doing, we contribute to the literature in two ways.

In the first part of this paper, we develop a conceptual and methodological framework that (a) clarifies the theoretical differences between the nested-factors model and Spearmanian theories and (b) addresses previous methodological difficulties in using the nested-factors model in research on the relation between cognitive abilities and criteria like job performance. The adequate statistical method to study the relation between cognitive abilities and job performance from the perspective of Spearman's two-factor theory and the higher-order factor model—incremental validity analysis—is readily available to personnel psychologists as part of standard statistical software packages and basic training in applied psychology. In contrast, researchers intending to rely on the nested-factors model in studying job performance (and other criterion variables) have faced difficulties in finding an adequate statistical technique to do so. To address these difficulties, we show how a recently emerging statistical technique in personnel psychology—relative importance analysis (e.g., LeBreton, Hargis, Griepentrog, Oswald, & Ployhart, 2007)—can be applied in a way that corresponds to the theoretical assumptions of the nested-factors model.

In the second part of this paper, we use the proposed conceptual and methodological framework to conduct the first empirical analysis of which we are aware that contrasts the importance of GMA and narrower cognitive abilities in job performance using the perspective of the nested-factors model. To conduct this empirical investigation, we first assemble a meta-analytic matrix linking measures for a broad range of narrower cognitive abilities, GMA, and job performance by meta-analyzing data from a structural intelligence battery. We then conduct a relative importance analysis on the resulting matrix. We contrast this analysis with an analysis of the same data that relies on Spearmanian assumptions and incremental validity analysis. In so doing, we also examine the implications of the two theoretical perspectives for personnel psychology.

Spearmanian Theories and Job Performance

Charles Spearman introduced two-factor theory in a seminal paper published in 1904 (Spearman, 1904). Spearman observed that matrices of a wide variety of different cognitive ability tests commonly show a positive correlation (so-called positive manifold phenomenon). He theorized that an underlying fundamental individual difference—GMA—caused the positive correlations in the matrices. Building on this idea, Spearman's two-factor theory proposes that two different types of abilities exist—GMA and specific abilities (Spearman, 1904, 1927). In two-factor theory,

GMA is the shared variance among a set of ability tests. Specific abilities correspond to the variance in ability tests that does not belong to GMA or error. Consequently, Spearman's two-factor theory assumes that there is only one fundamental source (or latent variable) for shared variance among cognitive ability tests—GMA. Because of this conceptualization, any shared variance between abilities is automatically credited to GMA.

A direct successor of Spearman's two-factor theory is the higher-order theory of cognitive abilities. Following Spearman's initial publication in 1904, researchers began to question two-factor theory on the grounds that some methods of factor analysis commonly found more than one factor in correlation matrices of ability tests (see Carroll, 1993, for an extensive review). These first-order factors, however, were commonly correlated to a notable degree. Researchers were consequently able to conduct second-order factor analyses on the correlation matrices of the first-order factors. These second-order factor analyses commonly found one factor. As a result, the higher-order factor model conceptualizes GMA as a higher-order factor explaining variance in narrower first-order ability factors (Gustafsson & Balke, 1993; Mulaik & Quaretti, 1997; Yung et al., 1999). Effectively, the difference between Spearman's two-factor theory and the higher-order factor model is the nature of the narrower cognitive abilities. In two-factor theory, the specific abilities are manifest variables (cognitive ability tests). In the higher-order factor model, in contrast, the narrower cognitive abilities have the status of latent variables and consequently may have more than one indicator variable. Important for studying the relation between cognitive abilities and job performance is that the higher-order factor model and Spearman's two-factor theory assume a similar relation between GMA and the narrower cognitive abilities. Both theories assume that shared variance between GMA measures and narrower cognitive ability measures solely originates from the GMA construct.

Figure 1A illustrates this core assumption using a very simple scenario with only one narrower cognitive ability measure. We use this simple example for illustrative purposes. Actually, analyses usually contain several measures of narrower cognitive abilities. As indicated by Figure 1A, Spearmanian theories assume that the shared variance between the GMA measure and the measure of the narrower cognitive ability completely belongs to GMA. Only the remaining variance in the narrower ability measure belongs to the narrower ability construct.

The assumptions of Spearmanian theories have direct consequences when one studies the importance of cognitive ability constructs in job performance. Figure 1B illustrates this point. As indicated by Figure 1B, the shared variance among the GMA measure, the narrower cognitive ability measure, and the job performance measure entirely belongs to GMA. A statistical method that maintains these assumptions in studying effects

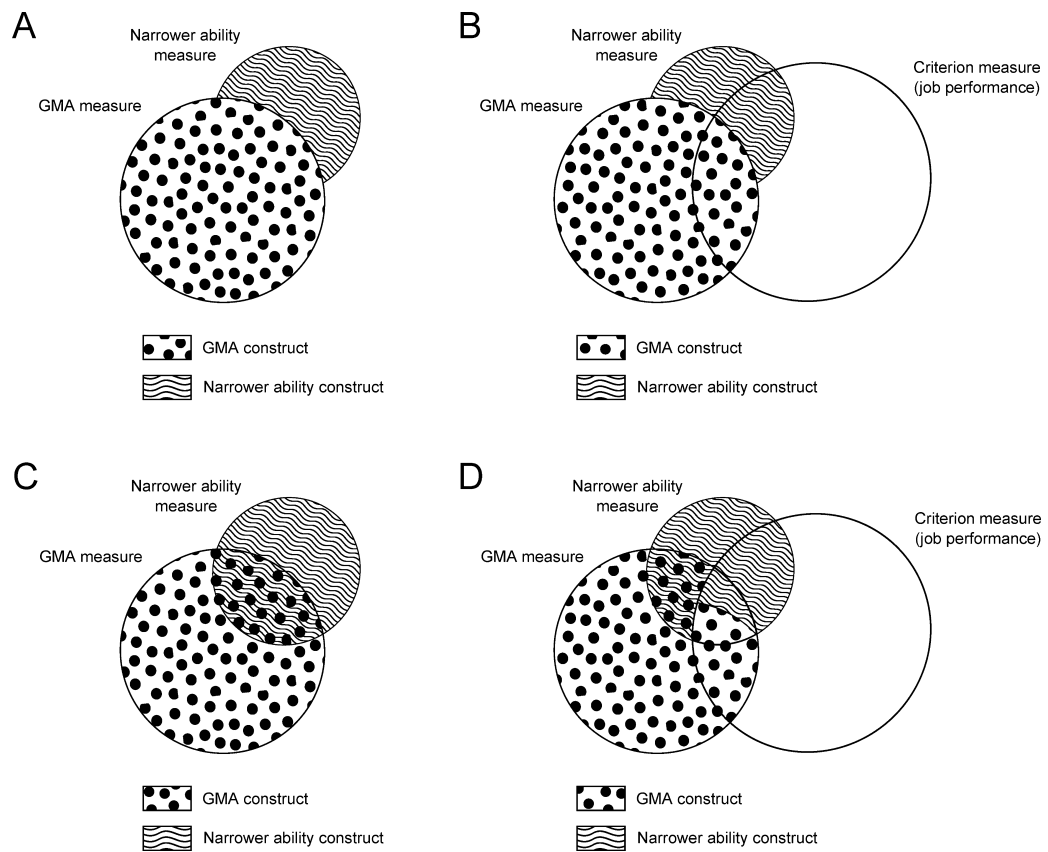


Figure 1: The Graph Illustrates the Implications of Spearmanian Theories, and the Nested-Factors Model for Job Performance Research.

For the sake of simplicity, the graph shows a very simple scenario with only one narrower cognitive ability measure. As displayed in the first panel (A), Spearmanian theories attribute all shared variance between the general mental ability (GMA) measure and narrower cognitive ability measure to the GMA construct. Consequently, criterion variance in job performance jointly explained by the GMA measure and the narrower cognitive ability measure is also attributed to GMA (B). The nested-factors model, in contrast, assumes that the shared variance between the GMA measure and the narrower cognitive ability measure is also shared by the constructs (C). Accordingly, criterion variance in job performance jointly explained by the GMA measure and the narrower cognitive ability measure may either belong to the GMA construct or to the narrower cognitive ability construct. To study the relative importance of the GMA construct and the narrower ability construct from the perspective of the nested-factors model, one consequently needs a methodological technique (i.e., relative importance analysis) to determine to which constructs jointly explained variance belongs. When this is successfully done, a situation like the situation illustrated in the last panel (D) results.

on a criterion variable is incremental validity analysis. In conducting this type of analysis, one first carries out a regression analysis with the GMA measure as the predictor and the job performance measure as the criterion. This analysis provides an estimate of the variance GMA explains in job performance from the perspective of Spearmanian theories. In the next

step, one adds the narrower cognitive ability to the regression analysis as a second predictor. The incremental change in R^2 then yields an estimate for the effect of the narrower cognitive ability construct on job performance according to the Spearmanian theories. The key point here is that any variance that the GMA measure, the narrower ability measure, and the job performance measure share is considered to be the result of GMA. In line with Spearmanian theories, only variance that the narrower ability measure and the job performance measure share (and thus is unique to the narrower ability measure) is considered to belong to the narrower ability construct.

Researchers in personnel psychology have long been interested in the relative importance of GMA and narrower cognitive abilities in explaining differences in job performance, and they have primarily adopted the framework of Spearmanian theories to investigate this question. Consistently, empirical studies conducted by personnel psychology researchers have found that narrower cognitive abilities explain only a small amount of incremental variance in job performance (see Ree & Carretta, 2002, for a thorough review). The majority of the reported studies used the General Ability Test Battery (GATB) and the Armed Forces Vocational Aptitude Battery (ASVAB) to assess cognitive abilities. Typically, these studies conclude that the narrower cognitive abilities from ASVAB or GATB batteries explain about 2% of job performance variance (cf. Ree, Earles, & Teachout, 1994). GMA, in contrast, typically explains more than 20% of the variance in job performance in these studies (Ree & Carretta, 2002; Ree et al., 1994). From the perspective of the Spearmanian theories, there is consequently “not much more than g ” (p. 518; Ree et al., 1994).

The Nested-Factors Model of Cognitive Abilities and Job Performance

The nested-factors model includes two major assumptions (Gustafsson & Balke, 1993; Holzinger & Harman, 1938; Holzinger & Swineford, 1937; Rindskopf & Rose, 1988; Thurstone, 1944; Yung et al., 1999). The first assumption is that two different types of factors directly explain variance in cognitive ability tests: GMA and narrower cognitive ability factors. Accordingly, each cognitive ability test has two-factor loadings: one loading on GMA, and one loading on a narrower cognitive ability factor. The second assumption of the nested-factors model is that GMA has no causal effect on narrower cognitive abilities and that GMA is thus not a higher-order factor. In the nested-factors model, GMA and narrower cognitive ability factors only differ in the number of manifest indicators (ability tests). This conceptualization confirms to Humphreys' (1981) view of the conceptual distinction between GMA and narrower cognitive abilities (Yung et al., 1999). Humphreys (1981) noted that “Breadth is the

key concept, not subordination.” (pp. 90–91). This conceptualization has important theoretical implications. As a result of the missing causal relation between GMA and narrower cognitive abilities in the nested-factors model, shared variance between GMA and narrower cognitive abilities does not necessarily belong to GMA. Instead, shared variance between GMA and narrower cognitive abilities can potentially be attributed to GMA, narrower cognitive abilities, or both.

Link to the Higher-Order Model

Theoretically, a nested-factors model can include all possible correlations between the cognitive abilities in the model (GMA and several narrower cognitive abilities). In factor analytic research, however, researchers commonly orthogonalize some or all factors to ease the interpretation of the factor loadings (Carroll, 1995) and to reduce the complexity of the model (Mulaik & Quaretti, 1997; Rindskopf & Rose, 1988; Yung et al., 1999). When all factors are orthogonalized, the factor loadings of the nested-factors model are fairly similar to factor loadings produced by the frequently used Schmid-Leiman solution (Schmid & Leiman, 1957). The Schmid-Leiman solution is an analysis based on a higher-order factor model and transforms the loadings of a higher-order model so that GMA and the narrower cognitive abilities have orthogonal loadings like in the orthogonalized version of the nested-factors model (Carroll, 1995; Schmid & Leiman, 1957; Yung et al., 1999).

The orthogonalized nested-factors model and the Schmid-Leiman solution are also formally related. Yung et al. (1999) showed that higher-order models and accordingly Schmid-Leiman solutions based on higher-order models are mathematically equivalent to a nested-factors model with orthogonal factors and additional constraints. Conversely, the link between the nested-factors and the higher-order factor model can also be investigated and interpreted from the perspective of the higher-order model. From this perspective, a nested-factors model with orthogonal factors is equivalent to a higher-order factor model with additional direct effects of GMA on the tests and some constraints on these direct effects (Yung et al., 1999).

Arguments Supporting the Nested-Factors Model

The relations between higher-order and nested-factors models shown by Yung et al. (1999) suggest that factor-analytic research cannot conclusively answer whether GMA is the underlying cause of narrower cognitive abilities (like in the higher-order model) or not (like in the nested-factors conceptualization). To evaluate the appropriateness of the

two conceptualizations, researchers consequently need to seek other evidence. Evidence favoring the nested-factors model over Spearmanian theories or vice versa has long been rare, and the two conceptualizations consequently have been coexisting in cognitive ability research for a considerable time. Recent theoretical work, however, questions some of the assumptions of the Spearmanian theories regarding the relation between GMA and the narrower cognitive abilities (Blair, 2006; van der Maas, Dolan, Grasman, Wicherts, Huizenga, & Raijmakers, 2006). As the nested-factors conceptualization makes fewer assumptions on the relation between GMA and narrower cognitive abilities, the nested-factors model can more readily account for these new insights.

Traditionally, researchers have explained the positive manifold phenomenon with the existence of a single latent factor, GMA. Spearman (1904) originally proposed that mental energy is the main cause or origin of GMA. More recent proposals of this type either argue that underlying cognitive factors (e.g., speed or efficiency of information processing, working memory) or underlying biological factors (e.g., brain size, neural efficiency, or neural plasticity) are the source of GMA (e.g., Deary, 2002; Duncan et al., 2000; Garlick, 2002; Gray & Thompson, 2004). There is evidence that most of these factors are considerably related to GMA and thus there is some evidence for a cognitive and biological basis of GMA (Deary, 2002; Duncan et al., 2000; Garlick, 2002; Gray & Thompson, 2004). However, none of these factors can fully explain the positive manifold phenomenon (Ackerman, Beier, & Boyle, 2005; Blair, 2006; Luciano et al., 2005).

Recently, researchers have developed theoretical models that offer alternatives to the single latent factor explanation of the positive manifold. Most notably, van der Maas et al. (2006, see also Dickens, 2007, for a similar model) have proposed a developmental theory suggesting that shared variance among cognitive abilities can also result from the interaction of narrower cognitive abilities. Theoretically, their theory can explain the existence of the positive manifold in adult populations (like working populations) without the existence of an underlying unitary cognitive or biological mechanism. The theory assumes that several specific cognitive processes reciprocally influence each other during human development. Initially, all processes are undeveloped and uncorrelated. During development, each process facilitates the development of other processes and these processes in turn influence the initial process and other processes. These dynamic interactions result in systems (i.e., people) with a generally positive development on many different processes and systems with less positive development. The explanation is that small random advantages in initial resources (genetic or environmental) for cognitive processes can exist. These random advantages may vary between persons—one person

may have high resources for certain processes (e.g., memory processes), and another person may have high resources for other processes (e.g., spatial processes). During development, these resource advantages can facilitate the development of other cognitive processes so that individuals with different types of initial advantages all develop high performance on a broad variety of cognitive processes and individuals without initial resource advantages do not. Accordingly, cognitive processes become correlated during development. Van der Maas et al.'s theory can consequently predict the positive manifold phenomenon without making the assumption that a single source of the positive manifold exists.

The existence of van der Maas et al.'s (2006) alternative explanation for the positive manifold phenomenon does not necessarily suggest that GMA is not a useful construct. As stated previously, empirical evidence suggests that general cognitive and biological mechanisms related to GMA measures exist (e.g., Duncan et al., 2000). There is thus some empirical support for the notion that GMA has a cognitive or biological basis. The important implication of van der Maas et al.'s theory, however, is that shared variance between GMA measures and narrower cognitive ability measures does not necessarily result from GMA. Instead, van der Maas et al.'s work suggests that shared variance between GMA measures and narrower cognitive ability measures can also result from specific cognitive processes. In fact, several authors have argued that some phenomena in the cognitive ability literature can more readily be explained by interactions between specific cognitive processes than by a single latent factor explanation (Blair, 2006; Dickens, 2007; van der Maas et al., 2006). The available evidence consequently supports several different sources of shared variance between cognitive ability measures—GMA, more specific mechanisms, or both—and consequently also supports the assumptions of the nested-factors model.

In sum, the reviewed literature provides some evidence (a) that both GMA and narrower cognitive abilities can explain empirical findings in the cognitive abilities literature and (b) that shared variance between GMA measures and narrower cognitive ability measures can either be explained by GMA or by narrower cognitive ability constructs. As Spearmanian theories assume that shared variance between GMA and narrower cognitive abilities entirely belongs to GMA, Spearmanian theories are not compatible with the idea that specific cognitive processes can explain shared variance between GMA and narrower abilities. The nested-factors model, in contrast, makes no assumptions on the origin of shared variance between GMA and narrower cognitive abilities. The model is consequently compatible both with research suggesting that a single latent factor like GMA exists and with research suggesting that specific cognitive mechanisms explain variance shared by GMA and narrower abilities. Figure 1C

illustrates this point. As indicated by Figure 1C, the nested-factors model allocates the shared variance between the GMA measure and the narrower cognitive ability measure in Figure 1C to both the GMA construct and the narrower cognitive ability construct. This conceptualization corresponds to the core assumption that the GMA construct and narrower cognitive ability constructs only differ in breadth and not in subordination.

Conceptual and Methodological Issues in Using the Nested-Factors Model in Job Performance Research

The nested-factors model has methodological implications when one wants to study the relation between cognitive abilities and job performance from its perspective. As shared variance between GMA measures and the narrower cognitive ability measures belongs to either GMA or narrower cognitive ability constructs in the model, incremental validity analysis does not correspond to the model's assumptions. Researchers consequently need a statistical technique that is able to determine the explained variance belonging to each ability construct even when the measures of the constructs are considerably correlated. A technique of this type should consequently enable researchers not only to determine the amount of unique variance each predictor explains in the criterion and the overall amount of explained variance, but it should also allow researchers to determine to which ability constructs jointly explained variance belongs.

The need to explore the role of GMA and narrower cognitive abilities in outcome criteria using the nested-factors model has been recognized by others. Gustafsson and Balke (1993) have studied the relative contribution of a GMA measure and narrower cognitive ability measures to school achievement and avoided the problem of ability correlations by orthogonalizing all ability measures using factor analytic methods. In so doing, Gustafsson and Balke comprehensively studied the predictive power of GMA and narrower cognitive abilities in correspondence with the nested-factors conceptualization. Despite these advantages, factorial orthogonalization approaches clearly have their drawbacks. Especially, there is a consensus in the literature that it is theoretically reasonable to assume that cognitive abilities are correlated (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Kline, 1991). As orthogonalized variables cannot be perfect approximations of the original collinear variables (Johnson & LeBreton, 2004; Mulaik & Quaretti, 1997), imposing constraints that require cognitive abilities to be uncorrelated likely diminishes the accuracy of ability-construct operationalizations in the statistical analyses.

In recent years, a new methodological technique—relative importance analysis—is gaining increased attention in personnel psychology (LeBreton et al., 2007). Relative importance analysis allows researchers to determine the specific contribution for each predictor even when the predictors are correlated and no foregoing assumptions on the order of predictors (like in Spearman's two-factor theory) exist. In the following section, we build on Gustafsson and Balke's (1993) conceptual work and aim to determine the role of GMA and narrower cognitive abilities in a criterion from the perspective of the nested-factors model. In conceptually extending Gustafsson and Balke's work, we do not use orthogonalization but directly study the contribution of correlated predictors—cognitive ability measures—to the criterion measure of interest—job performance—using relative importance analysis.

Relative Importance Analysis

The most widely accepted definition, by Johnson and LeBreton (2004), defines relative importance as “the proportionate contribution each predictor makes to R^2 , considering both its direct effect (i.e., its correlation with the criterion) and its effect when combined with the other variables in the regression equation” (p. 240). In accordance with this definition, modern relative importance methods (Budescu, 1993; Grömping, 2007; Johnson, 2000) portion an overall R^2 into variance shares for each predictor. Consequently, modern relative importance methods correspond to the assumptions of the nested-factors model.

Figure 1D conceptually illustrates the use of relative importance analysis in studying the effects of cognitive abilities on job performance from the perspective of the nested-factors model. As illustrated by Figure 1D, the shared variance between the GMA measure, the narrower cognitive ability's measure, and the job performance measure distributes among GMA's construct and the narrower cognitive ability's construct. Consequently, the narrower cognitive ability's construct explains more variance in this analysis than in the incremental validity analysis of the same data (compare Figure 1D to Figure 1B). Again, note that the example in Figure 1 is a very simple example for illustrative purposes—relative importance analyses using the nested-factors model will commonly include several narrower cognitive ability measures.

Traditional Measures of Relative Importance

Researchers have long been interested in the relative importance of predictors in multiple regression (e.g., Englehart, 1936). Traditionally,

researchers have used squared standardized regression coefficients and the product measure to determine the relative importance of variables.

Squared standardized regression coefficients are adequate and useful measures of relative importance when predictors are uncorrelated. In this case, squared betas are equal to squared zero-order correlations and sum to R^2 . When variables are correlated to a notable degree, however, squared standardized regression coefficients do not sum up to R^2 and do not adequately partition the variance shared among multiple correlated variables and the criterion (Budescu, 1993; Darlington, 1968; Johnson, 2000; Johnson & LeBreton, 2004). The reason is that standardized regression coefficients are designed to measure the effects of predictors for a specific combination with other predictors. When one compares two predictors using standardized regression coefficients, one effectively contrasts the effects of the two predictors under different and very specific circumstances (Budescu, 1993). Specifically, one compares the effect of Predictor 1 given Predictor 2 (and possibly other variables) with the effect of Predictor 2 given Predictor 1 (and possibly other variables). The contribution a predictor makes alone (its direct effect) and in subsets of predictors (partial effects) is not considered (LeBreton, Ployhart, & Ladd, 2004). In practice, squared standardized regression coefficients commonly overestimate the relative importance of the predictor variable most highly correlated with the criterion and underestimate the relative importance of the other variables in the model (Johnson, 2000).

The *product measure* is a popular early proposal for a measure of relative importance (Hoffman, 1960). The measure is calculated by multiplying the zero-order correlation and the standardized regression coefficient for each predictor. When predictors are uncorrelated, the product measure is an adequate measure of relative importance and equal to squared standardized regression coefficients and squared zero-order correlations. When predictors are correlated, the product measure has a number of desirable properties as a measure of relative importance (Johnson & LeBreton, 2004; Pratt, 1987). Most notably, the sum of the importance weights is equal to R^2 . The measure, however, also has several problems. Johnson and LeBreton (2004) noted that the product measure essentially ignores the magnitude of one of its components if the magnitude of the other component is very low. Another problem is that product measures can be negative when a predictor's zero-order correlation and its standardized regression coefficient have different signs. Especially when predictors are correlated, negative product measures are not an unusual phenomenon (Johnson & LeBreton, 2004). Consequently, the product measure can only be interpreted as a measure of relative importance when all zero-order correlations and the corresponding standardized regression coefficients have the same sign (Johnson & LeBreton, 2004; Pratt, 1987).

Modern Measures of Relative Importance

To address the problems of traditional measures of relative importance, researchers have proposed a variety of alternative measures of relative importance (Grömping, 2007; Johnson & LeBreton, 2004, for reviews). In this paper, we focus on two newer statistics of relative importance—general dominance weights and relative weights. Both general dominance weights and relative weights correspond to the definition of relative importance by Johnson and LeBreton (2004). Furthermore, several simulation studies have found that general dominance weights and relative weights are adequate measures of relative importance.

General dominance weights (Azen & Budescu, 2003; Budescu, 1993; Grömping, 2007) draw on a proposal by Lindeman, Merenda, and Gold (1980). Lindeman et al. proposed to use the mean of each predictor's squared semipartial correlation in all possible subset regression models as a measure of the predictor's relative importance. Azen and Budescu (2003) noted that the resulting general dominance weights can be interpreted as the average "usefulness" of a predictor across all subset regression models. General dominance weights are always positive and sum to the model R^2 . Consequently, they can be interpreted as effect size estimates that correspond to the amount of variance a predictor contributes to R^2 (LeBreton et al., 2007).

An advantage of general dominance weights is that researchers cannot only draw inferences from the overall results—the general dominance weights—but may also examine more specific patterns of dominance. Two meaningful specific patterns are complete and conditional dominance. Complete dominance of one predictor over another can be established if the predictor's additional contribution is greater than the other predictor's contribution for all subsets of models. Conditional dominance is a form of dominance that refers to a particular model size. To examine conditional dominance, one calculates the average contribution a predictor makes in each model size. For instance, for a model with three predictors, one calculates the average contribution for models with one, two, and three predictors. Conditional dominance of one predictor over another is established when the predictor has a larger average contribution within each model size (Azen & Budescu, 2003). Both complete and conditional dominance are more strict forms of dominance than general dominance and have the potential to provide relevant additional insights.

Relative weights (Johnson, 2000) address the problem of variable collinearity using a transformation approach. This approach includes several computational steps. First, one creates uncorrelated predictor variables (Z_k) that are maximally correlated with the original predictor variables (X_j). Second, one runs a regression analysis regressing the criterion

(Y) on the newly created uncorrelated predictor variables (Z_k) to get the standardized regression coefficients β_k . Third, one conducts a regression analysis regressing the original correlated predictor variables (X_j) on (Z_k) to get the standardized regression coefficients λ_{jk} . In the fourth and final step, relative weights can be calculated by combining the squared β_k and λ_{jk} values. An advantage of relative weights is that their estimation requires less computational effort than the estimation of general dominance weights. Especially when the number of predictors is large, the estimation procedure for relative weights is quicker and easier.

Relative weights and general dominance weights have similar characteristics. Like general dominance weights, relative weights sum up to R^2 , cannot be negative, and can be interpreted as effect size estimates. Furthermore, relative weights also consider both the direct effect and the effect of a variable in combination with other variables and consequently correspond to the definition of relative importance by Johnson and LeBreton (2004). Accordingly, results of empirical analyses relying on general dominance weights are frequently similar to results from corresponding analyses with relative weights. In a simulation study and several field studies, general dominance weights and relative weights yielded almost identical estimates of relative importance (LeBreton et al., 2004).

This Study

The first part of this paper developed a conceptual and methodological framework that (a) clarifies the theoretical differences between the nested-factors model and Spearmanian theories for studying the role of cognitive abilities in job performance and (b) allows researchers to use the nested-factors model in empirical research without making additional assumptions. In the second part of this paper, we apply the proposed conceptual and methodological framework in a meta-analytic investigation on the role of GMA and narrower cognitive abilities in job performance. We also conduct analyses using the perspective of the Spearmanian theories and incremental validity analyses. This allows us to compare the implications of the two theoretical perspectives.

In conceptualizing our meta-analytic investigation, we considered that several scholars have argued that empirical investigations on the importance of different cognitive abilities in a criterion like job performance are only valid when they cover an adequate range of narrower cognitive abilities (Carroll, 1982; Goldstein, Zedeck, & Goldstein, 2002; Wittmann & Süß, 1999). These scholars have recommended that researchers rely on established ability taxonomies to ensure that an adequate range of narrower abilities is covered. This study therefore focuses on meta-analyzing data from a battery designed to measure Thurstone's primary mental abilities

(PMAs; Thurstone, 1938a, 1938b; Thurstone & Thurstone, 1941). Thurstone's PMA taxonomy is one of the most established ability taxonomies (Carroll, 1993). The PMAs map the complexity of human abilities in terms of seven different factors: word fluency, verbal comprehension, spatial visualization, number facility, associative memory, reasoning, and perceptual speed. This basic structure has been supported in large-scale factor analytic studies using both the perspective of the higher-order model (Thurstone & Thurstone, 1941) and the nested-factors conceptualization (Holzinger & Harman, 1938).

On the basis of our core goal of investigating the role of cognitive abilities in job performance using the perspective of the nested-factors model and the PMA taxonomy, we formulated three more specific research questions to guide our investigation. These are as follows:

1. How much of the job performance criterion variance explained by cognitive ability measures (measures for GMA and all seven PMAs) belongs to GMA in a relative importance analysis?
2. Do PMAs exist that are relatively more or relatively equally important as GMA in explaining job performance variance?
3. How do the conclusions regarding the role of cognitive abilities in job performance reached by using the nested-factors model differ from those reached by relying on Spearmanian theories?

Method

Research on the role of predictor combinations in job performance requires not only predictor–criterion relationships but also predictor intercorrelations to reach viable conclusions (Viswesvaran & Ones, 2002). For the purpose of this study, we therefore assembled a meta-analytic intercorrelation matrix using data regarding the Wilde Intelligence Test (WIT; Jäger, 1963; Jäger & Althoff, 1983, 1994; WIT-2; Kersting, Althoff, & Jäger, 2008)—a German ability test battery designed to assess the PMAs.

Our reliance on meta-analysis allowed us to combine data for a broad range of different jobs and a broad range of narrower cognitive abilities. Restricting the scope of our analysis to just one battery of tests enabled us to avoid two typical methodological problems in assembling meta-analytic matrices (Bobko, Roth, & Potosky, 1999; Sackett, Lievens, Berry, & Landers, 2007; Viswesvaran & Ones, 1995). First, we were able to avoid imputation of missing predictor intercorrelations because correlations for all test–test combinations were readily available. Second, we did not face the issues associated with grouping more or less similar cognitive ability tests into categories based on a coding scheme (cf. Ackerman et al., 2005; Viswesvaran & Ones, 1995).

TABLE 1
*Sample-Weighted Average Range Restriction in Validity Studies and
 Sample-Weighted Average Test Reliability for Each Ability Test*

Test \bar{u}	Range restriction		Test reliability			
	k	N	\bar{r}_{xx}	k	N	
Proverbs (V)	.93	9	1,137	.83	7	773
Similar words (V, MT)	.90	13	1,451	.75	18	11,264
Verbal analogies (V, R)	.90	14	1,510	.77	19	10,643
Letter series (R, MT)	.95	4	559	.80	8	1,153
Number series (R, NF)	.92	7	751	.86	13	3,906
Estimating numbers (NF)	.89	11	1,360	.73	8	1,153
Embedded arithmetic problems (NF)	.94	10	1,233	.81	12	3,899
Basic arithmetic problems (NF, MT)	.94	6	875	.75	11	3,272
First and last letters (W, MT)	.99	3	417	.87	5	558
Faces (P)	.95	14	1,669	.83	7	773
Mirror images (S, MT)	.92	5	639	.90	11	2,759
Surface development (S)	.96	5	599	.82	12	4,130
Memory (M)	.92	10	1,063	.76	11	2,942

Note. Reliability coefficients were either internal consistencies (Cronbach's α , Spearman-Brown split-half reliability, or Flanagan-Rulon split-half reliability), test-retest reliabilities, or parallel-test reliabilities. \bar{u} = sample-weighted average range restriction, \bar{r}_{xx} = sample-weighted average test reliability, k = number of reported coefficients, N = total sample size, V = verbal comprehension composite of the Wilde Intelligence Test (WIT), R = reasoning composite of the WIT, NF = number facility composite of the WIT, W = word fluency, P = perceptual speed, S = spatial visualization, M = memory, MT = marker test for the respective primary mental ability (Thurstone & Thurstone, 1949).

The WIT-battery (WIT, Jäger, 1963; Jäger & Althoff, 1983, 1994; WIT-2, Kersting et al., 2008) was initially developed during the 1950s and is among the most frequently used cognitive ability test batteries in personnel selection in Germany today (Hülshager, Maier, Stumpp, & Muck, 2006). The WIT is based on Thurstone's PMA taxonomy and uses 13 ability tests to operationalize GMA and the PMAs. The tests and their corresponding PMA factors are provided in Table 1. Note that all of these tests use tasks that are identical or conceptually similar to those used and described by Thurstone to operationalize the PMAs (Thurstone, 1938a, 1938b; Thurstone & Thurstone, 1941, 1949).

Analytical Approach

In conducting this meta-analytic study, we followed recommendations of Viswesvaran and Ones (1995). Viswesvaran and Ones developed an analytical approach for conducting meta-analytic studies with meta-analytic correlation matrices. This approach includes two stages. In the first stage, researchers assemble a matrix of all available measures that operationalize

the constructs of interest. This matrix should be as specific as possible and all correlations should be corrected for possible artifacts. In the second stage, one can use the artifact-corrected meta-analytic matrix of measures to generate a smaller matrix of construct relationships. The advantage of this procedure is that constructs can be operationalized using the maximal available number of parallel measures. The operationalizations of constructs may include only one or several measures. When constructs are operationalized using multiple measures, it is possible that construct operationalizations include more measures than in all or most of the primary studies. A meta-analytic study of correlation matrices can consequently go beyond the scope of the primary studies. In line with Viswesvaran and Ones' recommendations, our meta-analysis focused on the 13 PMA tests of the WIT and linked these 13 tests to job performance criterion information. The major advantage of this procedure was that we could operationalize GMA using a composite of all 13 ability tests even though several primary studies only reported data on some of these tests.

Literature Search and Database

Studies for possible inclusion in the meta-analysis were identified by three different methods. First, we conducted literature searches in the PsycINFO (1872–2008) database and the German PSYINDEX (1977–2008) database using the search terms *WIT*, *Wilde*, and *Wilde-Test*. Second, we scanned the WIT test manuals for references to additional research using the WIT. Third, we sent out a request for unpublished WIT studies to the mailing lists of the industrial and organizational psychology and the psychological assessment divisions of the German psychological science association. All studies that contained test reliabilities, standard deviations, test–test correlations, or test–job performance correlations for the 13 WIT tests in our meta-analytic matrix were coded and added to a database. In articles with sample overlaps, the larger or more complete data were included in the meta-analysis, and the matching or overlapping samples were excluded.

The final database contained 54 studies. Forty studies contained test–job performance correlations for a total of 30 different jobs. These jobs covered six of the nine job categories (Categories 0/1, 2, 3, 5, 6, and 7) in the *Dictionary of Occupational Titles* (U.S. Department of Labor, 1991). Criterion information came from supervisory ratings of job performance and success in German apprenticeship (*Berufsausbildung*), certified foreperson (*Industriemeister*), or artisan foreperson (*Handwerksmeister*) programs (see Muller, 1999; Schmidt & Foster, 1999). Criterion information from German apprenticeship and foreperson programs has occasionally been referred to as training performance in the literature

(Hülshager, Maier, & Stumpp, 2007; Schuler, Funke, & Baron-Boldt, 1990). Yet, it is important to note that these types of performance criteria do not share much similarity with the concept of training performance in the international literature, where training performance is typically based on a test of job-specific knowledge and is not considered to be on the job (Schmidt & Hunter, 2004). In contrast, German apprenticeship and foreperson programs are on-the-job training programs. Apprenticeship programs last 3 years where trainees spend 70% to 80% of their time working on the jobs and only 20% to 30% of their time in formal training (Schmidt & Foster, 1999), whereas the design of foreperson programs is typically shorter but varies with occupation. In both types of programs, success criteria are a conglomerate of supervisory ratings, on-the-job performance assessments, formal written exams, and oral examinations. Thus, it may be more appropriate to think of success in German apprenticeship and foreperson programs as an overall performance criterion. We therefore combined correlations between WIT tests and success in apprenticeship and certified or artisan foreperson programs with correlations between WIT tests and supervisory ratings of job performance into one job performance criterion.

Stage 1: Meta-Analytic Matrix Linking Tests and Job Performance

In the first stage of Viswesvaran and Ones' analytical approach, we used Raju, Burke, Normand, and Langlois' (RBNL; Raju, Burke, Normand, & Langlois, 1991) meta-analytic procedure and its recent extensions (Burke & Landis, 2003; Raju, Lezotte, Fearing, & Oshima, 2006) to assemble a matrix linking the 13 WIT tests and job performance. The RBNL procedure uses a random-effects model and is designed to estimate the mean and the standard deviation of a set of artifact-corrected correlation coefficients. The RBNL procedure consists of two major steps. In the first step, the RBNL procedure uses all available sample-specific information to correct each correlation coefficient for artifacts and to estimate the sampling variance of each corrected correlation coefficient. Three different artifacts can be corrected: sampling error, unreliability, and range restriction. Although sampling error corrections are always applied, corrections for unreliability and range restriction are optional and depend on the research question. When sample-specific information is missing for unreliability and range restriction, Raju and colleagues recommend using the weighted mean of available estimates in the meta-analysis, estimates from other meta-analyses, or hypothetical estimates. In the second step, the RBNL approach uses the estimated corrected correlations and the corresponding sampling error estimates to calculate the estimated mean corrected correlation ($M_{\hat{\rho}}$) and the standard deviation of estimated

corrected correlations ($SD_{\hat{\rho}}$). Extensions of the original RBNL procedure also allow researchers to calculate standard errors for the estimated mean corrected correlation. These standard errors can be used to derive confidence intervals. In this analysis, we used a formula provided by Burke and Landis (2003; p. 295) to estimate standard errors for estimated mean corrected correlations. The advantage of this formula is that it not only accounts for subject-level sampling error but also for study-level sampling error.

Meta-Analytic Estimates of Test–Job Performance Relationships

To estimate mean corrected test–job performance correlations, we first obtained all raw correlations from the meta-analytic database linking one of the 13 ability tests to job performance criterion information. When a single study reported several job performance criteria, we used the most general criterion. When several criteria of the same level of generality were reported, we created an overall estimate of these criteria. As none of the studies that reported several criteria (and did not report an overall criterion) provided correlations between these criteria (which would have permitted the use of composite formulas, cf. Hunter & Schmidt, 2004), we averaged the available correlations in these cases and used the averaged correlation as the effect size.

In the next step, we corrected all test–job performance correlations for test unreliability, range restriction, and criterion unreliability. For test unreliability corrections, we used the reliability information provided in Table 1. The available reliability coefficients included three different types of reliability coefficients: internal consistencies (Cronbach's α , Spearman–Brown split-half reliability, or Flanagan–Rulon split-half reliability), test–retest reliabilities, and parallel-test reliabilities. In line with the RBNL approach, we used sample-specific reliability values when this information was available. When sample-specific artifact information was not available, we used a sample-weighted mean estimate based on the available reliability coefficients for the respective test.

For range-restriction corrections, we employed a similar procedure. Again, sample-specific values were used when available, and missing values were otherwise substituted with sample-weighted mean range restriction values. An overview of the available range restriction information for the 13 WIT tests is provided in Table 1.

Because almost none of the available studies reported criterion measure reliabilities, we substituted the missing criterion reliability values with an estimate from a comprehensive meta-analysis on the retest reliability of job performance evaluations (Sturman, Chermie, & Cashen, 2005). Sturman and colleagues found an average retest reliability of .776

for job performance evaluations. Past meta-analyses on job performance have frequently used estimates of interrater reliability from meta-analyses on the reliability of job performance evaluations to correct for unreliability in job performance. The reasoning behind this decision has commonly been that interrater reliabilities can be interpreted as measures of parallel test reliability. In principle, measurement theory suggests that parallel test reliabilities are the best choice for unreliability corrections (Brennan, 2001; Hunter & Schmidt, 2004). However, scholars have argued that interrater reliabilities from meta-analyses underestimate the parallel test reliability of job performance ratings (LeBreton, Burgess, Kaiser, Atchley, & James, 2003; Murphy & DeShon, 2000). Recently, Putka, Le, McCloy and Diaz (2008) extended the literature by showing that the designs and statistical methods used to estimate interrater reliability in primary studies are inadequate and likely underestimate the true parallel test reliability. Meta-analyses on interrater reliabilities consequently do not provide precise estimates. As a result, no adequate and precise meta-analytic estimate of parallel test reliability is currently available for job performance. We consequently followed the general advice in the measurement and meta-analysis literatures (Brennan, 2001; Hunter & Schmidt, 2004) to use retest-reliability estimates when parallel test reliability estimates are not available and relied on the meta-analytic retest-reliability estimate reported by Sturman et al. (2005).

Meta-Analytic Estimates of Test–Test Relationships

The intercorrelations between the 13 WIT tests were estimated and corrected using effect size, sample size, and reliability information from general-workforce applicant samples (samples of applicants applying for miscellaneous jobs or German high school populations near graduation) and were not corrected for range-restriction. We decided not to use correlations from selected incumbent samples because predictor intercorrelations from incumbent samples are typically distorted by a complex form of range restriction (Sackett et al., 2007). This type of range restriction arises from composite selection and can only be corrected using extensive information on the nature of the selection process from primary studies (Sackett et al., 2007). This information is rarely available in selection meta-analyses in general (Sackett et al., 2007), and it was not available from any of the incumbent samples in our database.

Stage 2: Meta-Analytic Matrices of Construct Relationships

In the second stage of Viswesvaran and Ones' analytical approach, one transforms the meta-analytic matrix of estimated mean corrected

correlations generated in the first stage of their approach to a meta-analytic matrix of construct relationships. When a construct has several measures in the matrix of estimated mean corrected correlations, Viswesvaran and Ones recommend forming unit-weighted composites of these measures. Therefore, they recommend using the following formula proposed by Nunnally and Bernstein (1994, p. 173; see also Viswesvaran & Ones, 1995, p. 873, Kuncel, Hezlett, & Ones, 2001, p. 174): $r_{wy} = \Sigma R_{wy} / \Sigma R_w^{1/2} \Sigma R_y^{1/2}$. This formula uses the correlation matrix of the measures for Composite W (R_w), the correlation matrix of the measures for Composite Y (R_y), and the correlation between the measures of W and Y (R_{wy}) to estimate the correlation between Composite W and Composite Y (r_{wy}). The formula consequently allows one to calculate correlations between either two unit-weighted composites or a unit-weighted composite and a single measure (In this case, the R_y or R_w for the single measure consists of only one correlation). An important advantage of using unit-weighted composites is that this method yields robust construct estimates (Ackerman & Beier, 2006; Cohen, 1990). Accordingly, researchers frequently use this method in meta-analytic research (e.g., Kuncel, Hezlett, & Ones, 2001) and in primary studies on cognitive abilities (e.g., Ackerman & Beier, 2006).

For the purpose of this investigation, we generated two different matrices of construct relationships. The first matrix of construct relationships was based on the original conceptualization of the WIT test and used varying numbers of tests to operationalize the PMAs like in the WIT test. The second matrix, in contrast, used a similar approach as Thurstone and Thurstone's (1949) PMA Test and used marker tests for the PMAs.

Matrix of Construct Relationships with PMA Composites

In the original conceptualization of the WIT, the 13 PMA tests are not equally distributed across the seven PMAs. Although some PMAs—word fluency, perceptual speed, and memory—are measured by one test, the other PMAs—verbal comprehension, reasoning, numerical facility, and spatial visualization—are operationalized using multiple tests. Table 1 shows the assignment of the WIT tests to the PMA composites in the WIT.

One noteworthy feature of the WIT is the composition of the reasoning composite. As indicated in Table 1, the reasoning composite consists of one specific reasoning test (letter series) and additionally uses two tests also included in the numerical-facility composite and the verbal-comprehension composite, respectively. The reason for this double assignment of tests is the nature of the reasoning concept in Thurstone's theory. Reasoning should assess the ability to find a rule or principle

(Thurstone, 1938a). Naturally, it is not possible to assess this ability without using content material. Tests assessing reasoning consequently use either verbal material, numerical material, or both, and commonly have considerable loadings on numerical or verbal factors (Carroll, 1993; Jäger, 1963; Thurstone, 1938a; Thurstone & Thurstone, 1941). In constructing the WIT, its authors have consequently decided to increase the economy of the test by using a numerical reasoning test and a verbal reasoning test for both the reasoning composite and the respective content composites. This approach is not uncommon in structural cognitive ability batteries. Many popular structural batteries use similar approaches to form composites (e.g., the GATB, Mellon, Daggett, MacManus, & Moritsch, 1996; and the Berlin Intelligence Test, Beauducel & Kersting, 2002, Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002; see also Carroll, 1993). We maintained the original conceptualization of the WIT in the first matrix of construct relationships, and accordingly we estimated a matrix linking the PMA operationalizations in the WIT, GMA, and job performance. We therefore used the meta-analytic matrix linking the 13 WIT tests and job performance we had assembled in the first stage of Viswesvaran and Ones's (1995) analytical approach. To estimate construct relationships between PMAs with multiple indicators and other variables in the matrix, we followed the recommendations of Viswesvaran and Ones (1995) and relied on the formula by Nunnally and Bernstein (1994, p. 173).

For GMA, we used a similar approach. GMA estimates are commonly nearly identical across a great variety of different estimation methods and factorial models (Jensen & Weng, 1994; Ree & Earles, 1991). Ree and Carretta (2002) have therefore suggested that unit weighting a considerable number of single tests is a particular easy and robust way to derive an adequate measure of GMA. We consequently used Nunnally and Bernstein's composite formula to estimate a unit-weighted GMA composite of the available 13 cognitive ability tests and used this composite as an indicator for GMA.

Matrix of Construct Relationships With PMA Marker Tests

A varying number of tests to estimate the PMAs may potentially alter estimates of relative importance. PMAs measured with multiple tests are potentially broader. Consequently, it is possible that these measures are more similar to GMA and that these differences influence the findings of relative importance analyses.

To investigate this possibility, we assembled a second matrix of construct relationships and conducted all analyses for both matrices. The second matrix was identical to the first matrix with the exception that the matrix was estimated using only one test for each of the seven

PMAs—even when multiple tests were available. This approach balanced the bandwidth of the PMAs in the matrix and is similar to the approach used by Thurstone and Thurstone (1949) in constructing the PMA Test. In conceptualizing this test, Thurstone and Thurstone selected one marker test to assess each of five PMAs (verbal comprehension, spatial visualization, reasoning, number facility, and word fluency) they considered to be particularly relevant in applied contexts.

An important question in the estimation of the second matrix of construct relationships was the selection of appropriate marker tests for PMAs with multiple WIT tests. In addressing this question, we build on the previous work by Thurstone and Thurstone (1949) and the fact that the WIT battery contains tests similar to the tests included in Thurstone and Thurstone's battery (see also Schaie, 1979). We accordingly used WIT tests similar to Thurstone and Thurstone's (1949) marker tests for PMAs with multiple WIT tests (see Table 1 for details).

Relative Importance Analyses

In conducting relative importance analyses, we focused on the two classic and the two modern measures of relative importance we described in the introduction. Squared standardized regression coefficients, the product measures, and general dominance weights were calculated using the *relaimpo* package (Grömping, 2008) in the R environment (R Development Core Team, 2009). To calculate relative weights, we used a script written by Johnson (2000).

Results

Meta-analytic estimates of the relations among cognitive ability tests and job performance are presented in Table 2. For each relationship, Table 2 contains mean uncorrected correlations (\bar{r}), and estimated mean corrected correlations ($M_{\hat{\rho}}$). As indicated by Table 2, estimated mean corrected correlations for test–test relationships ranged from $M_{\hat{\rho}} = .10$ to $M_{\hat{\rho}} = .68$, and estimated mean corrected correlations for test–job performance relations ranged from $M_{\hat{\rho}} = .18$ to $M_{\hat{\rho}} = .35$.

Table 3 presents the two matrices of construct relationships linking measures for the PMAs, GMA, and job performance, which we estimated using the estimated mean corrected correlations in Table 2. Note that the estimated predictor–criterion relations for GMA and the PMAs are somewhat similar to previous meta-analytic findings on cognitive ability–job performance relationships (Bertua et al., 2005; Schmidt et al., 1981). Table 3 also shows that the estimated mean corrected correlations between the cognitive ability measures are quite substantial. Mean estimates for

TABLE 2
Matrix of Meta-Analytic Relationships Between Cognitive Ability Tests and Job Performance

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Job performance	—	.24 (.12)	.24 (.09)	.24 (.12)	.25 (.09)	.19 (.12)	.20 (.12)	.24 (.15)	.21 (.13)	.26 (.15)	.16 (.14)	.16 (.16)	.14 (.16)	.17 (.11)
2. Proverbs	.32 (.10)	—	1,700/20	1,832/22	920/13	1,112/16	1,622/18	1,228/17	1,276/15	535/14	1,689/23	1,788/19	1,060/13	1,834/23
3. Similar words	(.21, .44)	.60 (.09)	—	2,330/7	2,081/6	2,081/6	2,081/6	2,081/6	1,185/5	1,976/5	2,330/7	2,081/6	2,081/6	1,917/5
4. Verbal analogies	.35 (.00)	(.27, .43)	.68 (.07)	—	2,461/7	4,695/10	2,619/8	4,543/11	2,907/7	1,976/5	2,461/7	4,081/9	4,198/9	2,984/6
5. Letter series	(.26, .43)	.60 (.14)	(.64, .73)	.51 (.09)	—	4,879/11	2,868/9	4,543/11	2,907/7	1,976/5	2,710/8	4,081/9	4,382/10	2,984/6
6. Number series	.34 (.00)	.39 (.05)	.25 (.04)	.51 (.09)	.43 (.09)	34 (.05)	38 (.04)	38 (.04)	18 (.03)	29 (.07)	30 (.07)	40 (.04)	36 (.05)	30 (.13)
7. Estimating numbers	(.24, .44)	.31 (.46)	(.18, .32)	(.43, .60)	—	2,461/7	2,461/7	2,461/7	1,185/5	1,976/5	2,461/7	2,081/6	2,081/6	1,668/4
8. Embedded arithmetic problems	.27 (.03)	.35 (.07)	.34 (.06)	.51 (.11)	.52 (.09)	—	.42 (.07)	.52 (.11)	.37 (.15)	.27 (.11)	.24 (.06)	.39 (.05)	.37 (.09)	.35 (.12)
9. Basic arithmetic problems	(.17, .37)	(.26, .43)	(.27, .40)	(.43, .58)	(.44, .60)	.54 (.06)	—	3,458/9	2,024/6	1,976/5	2,461/7	3,078/8	3,262/9	2,507/5
10. First and last letters	.29 (.12)	.27 (.08)	.27 (.03)	.40 (.07)	.45 (.01)	.48, .60)	—	.51 (.06)	.38 (.10)	.29 (.08)	.25 (.06)	.30 (.04)	.29 (.03)	.18 (.09)
11. Faces	(.19, .39)	(.18, .37)	(.20, .34)	(.33, .48)	(.39, .51)	(.48, .60)	.66 (.05)	2,868/9	1,185/5	1,976/5	2,710/8	2,239/7	2,239/7	1,668/4
12. Mirror images	.32 (.13)	.36 (.08)	.41 (.13)	.57 (.12)	.48 (.00)	.62 (.12)	.66 (.05)	—	.42 (.09)	.24 (.11)	.19 (.04)	.38 (.06)	.40 (.08)	.30 (.09)
13. Surface development	(.21, .42)	(.27, .45)	(.32, .50)	(.49, .65)	(.43, .53)	(.53, .71)	(.61, .71)	—	2,024/6	1,976/5	2,461/7	3,078/8	3,078/8	2,507/5
14. Memory	.30 (.13)	.10 (.10)	.20 (.08)	.25 (.11)	.23 (.00)	.44 (.17)	.50 (.11)	.52 (.08)	—	.14 (.04)	.26 (.03)	.23 (.05)	.17 (.08)	.22 (.13)
	(.18, .41)	(-.02, .22)	(.11, .28)	(.15, .35)	(.15, .31)	(.30, .59)	(.38, .62)	(.43, .61)	—	1,080/4	1,185/5	2,024/6	2,024/6	1,611/4
	.32 (.06)	.42 (.12)	.44 (.09)	.43 (.13)	.35 (.06)	.32 (.12)	.36 (.08)	.29 (.12)	.17 (.00)	—	.23 (.03)	.25 (.05)	.25 (.03)	.28 (.14)
	(.17, .47)	(.31, .54)	(.34, .54)	(.30, .57)	(.27, .43)	(.20, .44)	(.27, .46)	(.17, .41)	(.09, .25)	1,976/5	1,976/5	1,976/5	1,976/5	1,668/4
	.21 (.12)	.25 (.03)	.18 (.07)	.26 (.04)	.36 (.05)	.29 (.05)	.32 (.03)	.24 (.00)	.33 (.00)	.27 (.00)	—	.36 (.04)	.30 (.05)	.20 (.05)
	(.11, .31)	(.18, .31)	(.10, .26)	(.20, .32)	(.30, .43)	(.22, .36)	(.26, .38)	(.18, .30)	(.25, .40)	(.21, .33)	—	2,081/6	2,081/6	1,917/5
	.21 (.16)	.28 (.06)	.24 (.04)	.45 (.09)	.47 (.00)	.44 (.03)	.37 (.00)	.44 (.04)	.27 (.00)	.29 (.02)	.42 (.00)	—	.51 (.07)	.26 (.11)
	(.10, .32)	(.20, .36)	(.18, .29)	(.37, .52)	(.42, .52)	(.39, .49)	(.31, .43)	(.39, .49)	(.20, .33)	(.23, .35)	(.36, .47)	3,078/8	3,078/8	2,507/5
	.18 (.16)	.32 (.06)	.27 (.07)	.52 (.09)	.44 (.03)	.44 (.08)	.38 (.00)	.49 (.07)	.21 (.07)	.29 (.00)	.37 (.02)	.58 (.06)	—	.25 (.10)
	(.05, .32)	(.24, .40)	(.20, .34)	(.45, .59)	(.38, .50)	(.37, .51)	(.32, .44)	(.42, .55)	(.12, .29)	(.23, .35)	(.31, .43)	(.53, .64)	—	2,507/5
	.24 (.02)	.42 (.09)	.47 (.08)	.53 (.14)	.38 (.15)	.43 (.13)	.24 (.11)	.38 (.11)	.27 (.15)	.35 (.16)	.26 (.00)	.31 (.12)	.30 (.10)	—
	(.13, .35)	(.32, .52)	(.39, .56)	(.40, .65)	(.21, .55)	(.30, .56)	(.11, .37)	(.27, .49)	(.11, .44)	(.17, .52)	(.19, .32)	(.19, .44)	(.20, .41)	—

Note. Table entries below the diagonal are estimated mean corrected correlations (M_p) with standard deviations of the estimated corrected correlations (SD_p) provided in parentheses. Below each correlation appear 95% confidence intervals for estimated mean corrected correlations (CI_{M_p}). Table entries above the diagonal are mean observed correlations (\bar{r}) with standard deviations of the observed correlations (SD_r) provided in parentheses. Below each correlation appears the total sample size for all studies combined (N) followed by the number of correlations (k). The harmonic mean of the sample sizes across the cells of the matrix is $N = 2,015$.

TABLE 3
Meta-Analytic Matrix of Construct Relationships With Primary Mental Ability Composites (Below the Diagonal) and Meta-Analytic Matrix of Construct Relationships With Primary Mental Ability Marker Tests (Above the Diagonal)

Variable	1	2	3	4	5	6	7	8	9
1. Job performance	—	.44	.35	.30	.34	.21	.32	.21	.24
2. General mental ability	.44	—	.63	.53	.69	.65	.59	.53	.63
3. Verbal comprehension	.39	.79	—	.20	.25	.24	.44	.18	.47
4. Number facility	.36	.83	.48	—	.23	.27	.17	.33	.27
5. Reasoning	.39	.90	.72	.75	—	.47	.35	.36	.38
6. Spatial visualization	.22	.74	.45	.52	.63	—	.29	.42	.31
7. Word fluency	.32	.59	.50	.35	.45	.33	—	.27	.35
8. Perceptual speed	.21	.53	.26	.36	.37	.44	.27	—	.26
9. Memory	.24	.63	.55	.41	.54	.35	.35	.26	—

Note. Table entries are estimated mean corrected correlations ($M_{\hat{\rho}}$) and are based on the matrix of estimated mean corrected correlations ($M_{\hat{\rho}}$) in Table 2 (harmonic mean of sample sizes across cells: $N = 2,015$). For primary mental abilities with multiple WIT tests (verbal comprehension, number facility, reasoning, and spatial visualization), coefficients estimated using all available tests (composites) are provided below the diagonal (see Table 1 for details). Coefficients above the diagonal show the corresponding correlations when only one prototypical marker test for each of these primary mental abilities is used (see Table 1). GMA is a unit-weighted composite of all available ability tests (see Table 1).

corrected ability–ability correlations ranged from $M_{\hat{\rho}} = .26$ to $M_{\hat{\rho}} = .90$ in the matrix with composites, and from $M_{\hat{\rho}} = .26$ to $M_{\hat{\rho}} = .69$ in the matrix with marker tests. Although these estimates may appear to be quite high, their level is common in cognitive ability research and corresponds to the level found in large-scale studies (e.g., Beauducel & Kersting, 2002) and other cognitive ability meta-analyses (e.g., Ackerman et al., 2005).

Results for the relative importance analysis on the matrix linking the PMAs, GMA, and job performance are provided in Table 4.¹ A first examination of the estimated relative importance indices indicated that

¹An anonymous reviewer noted that the use of the same ability tests in estimating narrower cognitive ability measures and a GMA measure can potentially be viewed as a confounding factor in multiple regression analyses. We therefore reran all incremental validity analyses using GMA composites that excluded the ability for which we explored incremental validity in the respective analysis. We also ran relative importance analyses using a GMA composite that excluded the indicator variables for one of the seven PMAs and both modern relative importance measures. Both sets of analyses lead to identical conclusions. This is consistent with literature suggesting that GMA measures are highly stable once a considerable number of tests is included (Ree & Earles, 1991) and literature suggesting that modern relative importance measures are largely independent from a specific combination of other predictors (Budescu, 1993; Johnson, 2000). We consequently do not report the incremental validity estimates and relative importance measures for these 28 supplemental analyses. The results of these analyses can be obtained from the first author upon request.

TABLE 4
Relative Importance and Incremental Validity Analyses

Variable	β_j^2	$\beta_j M_{\hat{\rho}_{yj}}$ (%)	C_j (%)	ε_j (%)	R_{GMA}^2	ΔR^2
With composites for verbal comprehension, number facility, reasoning, and spatial visualization						
General mental ability	6.397	1.101 (471.0)	.058 (24.9)	.042 (18.1)	.190	
Verbal comprehension	.197	-.174 (-74.6)	.040 (17.3)	.045 (19.5)		.006
Number facility	.478	-.249 (-106.7)	.033 (13.9)	.038 (16.1)		.000
Reasoning	.096	-.120 (-51.4)	.034 (14.4)	.036 (15.2)		.000
Spatial visualization	.370	-.134 (-57.3)	.018 (7.9)	.015 (6.2)		.023
Word fluency	.033	-.058 (-25.0)	.026 (11.3)	.033 (14.1)		.007
Perceptual speed	.066	-.055 (-23.4)	.010 (4.4)	.013 (5.4)		.001
Memory	.100	-.077 (-32.8)	.014 (6.0)	.012 (5.3)		.002
All PMAs						.044
Sum		.234 (100)	.234 (100)	.234 (100)		
With marker tests for verbal comprehension, number facility, reasoning, and spatial visualization						
General mental ability	.001	.012 (5.1)	.059 (24.5)	.037 (15.3)	.190	
Verbal comprehension	.046	.075 (31.3)	.045 (18.7)	.052 (21.6)		.010
Number facility	.036	.057 (23.6)	.035 (14.6)	.042 (17.5)		.006
Reasoning	.042	.070 (29.1)	.039 (16.2)	.045 (18.7)		.003
Spatial visualization	.001	-.008 (-3.2)	.010 (4.1)	.008 (3.5)		.010
Word fluency	.017	.041 (17.2)	.031 (13.2)	.036 (15.0)		.007
Perceptual speed	.000	.003 (1.2)	.009 (3.7)	.010 (4.0)		.001
Memory	.002	.010 (-4.4)	.012 (5.0)	.011 (4.5)		.002
All PMAs						.050
Sum		.239 (100)	.239 (100)	.239 (100)		

Note. The analyses are based on the matrices of construct relationships in Table 3, which are in turn based on the meta-analytic matrix of estimated mean corrected correlations ($M_{\hat{\rho}}$) in Table 2 (harmonic mean of sample sizes across cells: $N = 2,015$). PMAs = primary mental abilities. β_j^2 = squared standardized regression coefficient, $\beta_j M_{\hat{\rho}_{uj}}$ = product of standardized regression coefficient and estimated mean corrected correlation (product measure), C_j = general dominance weight, ε_j = relative weight.

the two traditional measures—squared beta coefficients and product measure estimates—were difficult to interpret. In the analyses using the construct-relationship matrix with composites, GMA received very high estimates of relative importance and the narrower cognitive abilities received very low values. An opposite pattern emerged when we used marker tests to operationalize verbal comprehension, reasoning, number facility, and spatial visualization. In the analyses using the construct-relationship matrix with marker tests, GMA received very small values and narrower cognitive abilities had high relative importance values (see Table 4). For the product measures, we additionally found negative estimates in all analyses—a condition in which product measure estimates should not be interpreted (Johnson & LeBreton, 2004; Pratt,

1987). The likely reason for the traditional measures' difficulties is the level of collinearity in our matrices. Several authors argued that traditional measures of relative importance only work successfully when the level of collinearity is relatively low (e.g., Johnson & LeBreton, 2004). We consequently solely focused on the modern estimates of relative importance in addressing our three research questions.

We began by addressing Research Question 1 (i.e., How much of the job performance criterion variance explained by cognitive ability measures belongs to GMA in a relative importance analysis?) and examined the variance shares the general dominance analyses and the relative weight analyses assigned to GMA. In the analysis using the construct-relationship matrix with composites, GMA explained 5.8% of the criterion variance in the general dominance analysis and 4.2% of the criterion variance in the relative weight analysis (see the upper part of Table 4). These estimates contrast with an overall R^2 of .234 and consequently correspond to 24.9% and 18.1% of the total explained variance in the general dominance analysis and the relative weight analysis, respectively. Analyses using the construct-relationship matrix with marker tests yielded similar estimates (see the lower part of Table 4), suggesting that the use of composites or marker tests had only small influence on the estimated variance shares for GMA.

To address Research Question 2 (i.e., Do PMAs exist that are relatively more or relatively equally important as GMA in explaining job performance variance?), we contrasted the modern relative importance estimates (general dominance weights and relative weights) for GMA with the corresponding estimates for each of the PMAs. An interesting finding was that GMA was not the most important predictor in the relative weight analyses because verbal comprehension received a slightly stronger relative weight in both the relative weight analysis with composites and the relative weight analysis with marker tests (see Table 4).² Three other PMAs—numerical facility, reasoning, and word fluency—received relative weights that were close to the relative weight for GMA. Depending on the use of composites or marker tests, these abilities received relative weights that were slightly smaller, about equal, or slightly larger than the relative weights for GMA (see Table 4).

Although relative weight analyses commonly correspond with general dominance weights (Johnson & LeBreton, 2004; LeBreton et al., 2004), this was not the case in this study. In contrast to the findings from the relative weight analyses, the general dominance weight for GMA was higher than the general dominance weights for each of the PMAs in both the general dominance analysis using composites and the general dominance

²We thank an anonymous reviewer for his suggestion to examine relative weights.

analysis using marker tests (see Table 4). As in the relative weight analyses, however, verbal comprehension, number facility, reasoning, and word fluency received weights close to the weights assigned to GMA.

To further examine the relative importance of GMA and each of the PMAs in dominance analyses, we also examined the two more restrictive forms of dominance—conditional and complete dominance—proposed by Azen and Budescu (2003). For the analysis using the construct-relationship matrix with the PMA composites, results revealed that GMA dominated all PMAs in terms of complete and conditional dominance. A different pattern emerged in the analysis using marker tests. In this analysis, we found no evidence for complete or conditional dominance of GMA over all of the PMAs. A closer examination revealed that GMA was not conditionally (and thus also not completely) dominant because PMA measures occasionally received higher loadings in the subset models with few predictors.

We finally turned to Research Question 3 (i.e., How do the conclusions regarding the role of cognitive abilities in job performance reached by using the nested-factors model differ from those reached by relying on Spearmanian theories?). To address this research question, we used incremental validity analyses to estimate the amount of variance GMA explained in job performance and the incremental contribution of each PMA. We then contrasted these estimates with the modern relative importance measures. In so doing, we directly contrasted the implications of using a Spearmanian perspective with the perspective of the nested-factors model. Incremental validities are provided in the last column of Table 4. As indicated by Table 4, GMA explained 19% of the criterion variance in both analyses. The PMAs, in contrast, explained only a small proportion of the criterion variance. The contribution of each PMA was less than 3%. The incremental contribution of all seven PMAs was 4.4% in the analysis using PMA composites and 5.0% in the analysis using only marker tests (see Table 4). GMA consequently explained 81.2% of the total explained variance ($R^2 = .232$) when the matrix with composites was used and 79.5% of the total explained variance ($R^2 = .239$) when we used marker tests. Obviously, these findings differ dramatically from the variance shares we estimated using the modern relative importance measures. Our findings regarding Research Question 3 consequently provide evidence that theoretical assumptions (nested-factors model vs. Spearmanian perspective) have a strong influence on variance share estimates for cognitive abilities.

Supplementary Analysis

Several researchers have found that job complexity moderates cognitive ability–job performance relationships (Gutenberg, Arvey, Osburn, &

Jeanneret, 1983; Hunter & Hunter, 1984; Salgado, Anderson, Moscoso, Bertua, de Fruyt, & Rolland, 2003). Specifically, cognitive ability–job performance relationships for complex jobs are commonly higher than cognitive ability–job performance relationships for jobs with low levels of complexity. We accordingly examined how job complexity affected relative importance measures and incremental validity estimates. Like previous investigations (e.g., Salgado et al., 2003), we used the data dimension of the *Dictionary of Occupational Titles* (U.S. Department of Labor, 1991) to classify jobs into jobs of high (data dimension 0–3) and low complexity (data dimension 4–6) and reconducted our analyses for both job complexity categories.

Results are shown in Tables 5 and 6. Table 5 contrasts meta-analytic estimates of the relationships between the 13 WIT tests and job performance in jobs of high and low complexity. Table 6 then shows estimated mean corrected correlations between operationalizations of GMA, the PMAs, and job performance for jobs with high and low levels of complexity. As indicated by both tables, estimated predictor–criterion relationships tended to be higher for complex jobs than for less complex jobs. This finding is similar to findings from previous large-scale meta-analyses (Hunter & Hunter, 1984; Salgado et al., 2003).

We proceeded by examining how job complexity affected conclusions on the relative importance of GMA and narrower cognitive abilities. As indicated by Table 6, high job complexity lead to modestly stronger variance shares for GMA. The total criterion variance explained by GMA was 7.9% in the general dominance analysis using composites, 8.6% in the general dominance analysis using marker tests, 5.9% in the relative weight analysis using composites, and 5.6% in the relative weight analysis using marker tests (26.0%, 28.6%, 19.4%, and 18.5% of the total explained variance). As a result, GMA was also the relatively most important cognitive ability in three of the four analyses for high job complexity. (In the relative weight analysis using the matrix with marker tests, reasoning received a stronger weight than GMA, but the variance share for GMA was nevertheless larger than in the overall analysis reported in Table 4.)

The analysis for jobs with low job complexity showed an opposite pattern (see Table 6). GMA accounted for a smaller proportion of the criterion variance and the total explained variance in each analysis. Specifically, the criterion variance explained by GMA dropped to 3.7% in the general dominance analysis using composites, 3.4% in the general dominance analysis using marker tests, 2.6% in the relative weight analysis using composites, and 2.2% in the relative weight analysis using marker tests (20.1%, 16.7%, 14.3%, and 10.9% of the total explained variance). As a result, GMA was not the most important predictor in any of the four analyses. In all four analyses for low job complexity, relative weights and

TABLE 5
Test–Job Performance Relationships for Jobs With High and Low Job Complexity

Test	High job complexity					Low job complexity				
	M_{ρ} (SD_{ρ})	\bar{r} ($SD_{\bar{r}}$)	k	N	95% $CI_{M_{\rho}}$	M_{ρ} (SD_{ρ})	\bar{r} ($SD_{\bar{r}}$)	k	N	95% $CI_{M_{\rho}}$
Proverbs	.41 (.00)	.30 (.07)	9	587	(.28, .54)	.21 (.12)	.16 (.12)	4	479	(.03, .40)
Similar words	.37 (.00)	.25 (.10)	14	1,101	(.27, .47)	.32 (.00)	.22 (.08)	6	599	(.19, .45)
Verbal analogies	.36 (.01)	.25 (.11)	15	1,181	(.27, .45)	.32 (.14)	.23 (.14)	7	651	(.16, .48)
Letter series	.42 (.00)	.30 (.07)	11	578	(.31, .53)	.20 (.00)	.15 (.02)	2	342	(-.04, .43)
Number series	.37 (.00)	.26 (.08)	12	650	(.26, .48)	.13 (.00)	.09 (.09)	4	462	(-.06, .32)
Estimating numbers	.32 (.14)	.21 (.13)	13	1,083	(.19, .45)	.22 (.00)	.16 (.07)	5	539	(.07, .37)
Embedded arithmetic problems	.39 (.04)	.29 (.12)	11	670	(.28, .50)	.23 (.15)	.18 (.15)	6	558	(.06, .41)
Basic arithmetic problems	.32 (.07)	.23 (.12)	10	681	(.19, .45)	.28 (.17)	.19 (.13)	5	595	(.07, .48)
First and last letters	.31 (.00)	.25 (.17)	8	174	(.14, .48)	.32 (.11)	.27 (.15)	6	361	(.04, .61)
Faces	.26 (.07)	.20 (.12)	13	951	(.15, .37)	.15 (.14)	.12 (.15)	10	738	(-.02, .33)
Mirror images	.31 (.03)	.24 (.10)	12	1,074	(.21, .41)	.05 (.15)	.04 (.15)	7	714	(-.12, .22)
Surface development	.26 (.21)	.20 (.19)	9	581	(.07, .45)	.09 (.00)	.07 (.04)	4	479	(-.07, .25)
Memory	.25 (.00)	.17 (.08)	12	1,369	(.14, .36)	.22 (.13)	.14 (.16)	11	465	(.01, .43)

Note. M_{ρ} = estimated mean corrected correlation, SD_{ρ} = standard deviations of estimated corrected correlations, \bar{r} = mean observed correlation, $SD_{\bar{r}}$ = standard deviation of observed correlations, 95% $CI_{M_{\rho}}$ = 95% confidence interval for an estimated mean corrected correlation.

TABLE 6
Relative Importance Analyses and Incremental Validity Analyses for Jobs of High and Low Complexity

Variable	High job complexity				Low job complexity					
	M_ρ	C_j (%)	ϵ_j (%)	R_{GMA}^2	ΔR^2	M_ρ	C_j (%)	ϵ_j (%)	R_{GMA}^2	ΔR^2
With composites for verbal comprehension, number facility, reasoning, and spatial visualization										
General mental ability	.51	.079 (26.0)	.059 (19.4)	.262	.003	.32	.037 (20.1)	.026 (14.3)	.105	.014
Verbal comprehension	.44	.049 (16.0)	.056 (18.5)			.33	.033 (18.2)	.037 (20.0)		.000
Number facility	.43	.045 (14.9)	.052 (17.2)			.27	.020 (10.7)	.023 (12.5)		.004
Reasoning	.47	.050 (16.5)	.053 (17.5)			.26	.017 (9.0)	.016 (9.0)		.055
Spatial visualization	.32	.025 (8.2)	.025 (8.1)		.007	.08	.023 (12.4)	.015 (8.3)		.028
Word fluency	.31	.022 (7.2)	.026 (8.4)		.000	.32	.039 (21.1)	.047 (25.9)		.001
Perceptual speed	.26	.015 (5.0)	.018 (6.0)		.000	.15	.006 (3.0)	.008 (4.1)		.000
Memory	.25	.019 (6.3)	.015 (4.8)		.008	.22	.010 (5.5)	.011 (6.1)		.000
All PMAs					.042					.079
Sum		.304 (100)	.304 (100)				.184 (100)	.184 (100)		
With marker tests for verbal comprehension, number facility, reasoning, and spatial visualization										
General mental ability	.51	.086 (28.6)	.056 (18.5)	.262	.003	.32	.034 (16.7)	.022 (10.9)	.105	.022
Verbal comprehension	.37	.045 (15.0)	.054 (18.0)			.32	.042 (20.6)	.047 (22.8)		.015
Number facility	.32	.034 (11.2)	.042 (13.8)		.003	.28	.039 (19.3)	.045 (21.9)		.001
Reasoning	.42	.062 (20.7)	.072 (23.9)		.009	.20	.011 (5.2)	.013 (6.3)		.044
Spatial visualization	.31	.022 (7.3)	.024 (8.1)		.001	.05	.014 (7.1)	.008 (4.2)		.028
Word fluency	.31	.023 (7.7)	.027 (8.9)		.000	.32	.048 (23.6)	.052 (25.6)		.001
Perceptual speed	.26	.014 (4.5)	.015 (4.9)		.000	.15	.005 (2.5)	.006 (3.0)		.000
Memory	.25	.015 (5.0)	.011 (3.8)		.008	.22	.010 (5.1)	.011 (5.3)		.000
All PMAs					.039					.099
Sum		.301 (100)	.301 (100)				.204 (100)	.204 (100)		

Note. Relative importance and incremental validity analyses are based on matrices of construct relationships, which are in turn based on meta-analytic matrices of estimated mean corrected correlations (M_ρ). The meta-analytic matrices of estimated mean corrected correlations consist of the estimated mean corrected test–test correlations in Table 2 and the estimated mean corrected test–job performance correlations in Table 5. The harmonic means of sample sizes across the cells of these matrices are $N = 1,520$ for low job complexity and $N = 1,649$ for high job complexity. PMAs = primary mental abilities. M_ρ = estimated mean corrected correlations, C_j = general dominance weight, ϵ_j = relative weight.

general dominance weights for at least one PMA were larger than the corresponding estimates for GMA (see Table 6).

Discussion

This study contributes to the literature by introducing a popular and important theoretical perspective in the cognitive ability literature—the nested-factors model—into research on the role of cognitive abilities in job performance. The goal of the first part of this article was to build on previous work (Gustafsson & Balke, 1993) and develop a conceptual and methodological framework for studying the role of cognitive abilities in job performance from the perspective of nested-factors model. The proposed framework elaborates how a recently emerging statistical technique—relative importance analysis (LeBreton et al., 2007)—can be used to study the role of GMA and narrower cognitive abilities in job performance in accordance with the assumptions of the nested-factors model. A secondary goal of this paper was to use this approach to conduct the first empirical examination linking GMA and narrower cognitive abilities using the perspective of the nested-factors model. We begin by discussing the empirical findings.

GMA, Narrower Cognitive Abilities, and Job Performance

Our meta-analytic investigation revealed that GMA accounted for 15.3% to 24.9% of the total explained criterion variance in relative importance analyses using modern relative importance measures and an overall job performance measure as the criterion. These estimates suggest that GMA is an important predictor of job performance but that narrower cognitive abilities also have an important role from the perspective of the nested-factors model. These conclusions are further supported when one compares the variance share of GMA with the variance shares for the narrower cognitive abilities in our analyses. Interestingly, the two established measures of relative importance did not consistently suggest that GMA was the single most important predictor. Although dominance analyses suggested that GMA was the single most important predictor, relative weight analyses suggested that verbal comprehension was more important than GMA.

Some additional insights emerged from the supplementary analyses. In the supplementary analyses, we split the dataset into jobs with high and low complexity. For high job complexity, GMA received modestly stronger relative weights so that GMA was the most important cognitive ability in three of the four analyses. In the analyses for jobs with low complexity, in contrast, GMA received smaller relative weights so that both modern

measures of relative importance did not assign the largest variance share to GMA. The results of the supplementary analyses consequently suggest that the relative importance of GMA partly depended on the level of job complexity with GMA being relatively more important in high-complexity jobs and being relatively less important in low-complexity jobs.

Evidently, the findings from both the main and the supplementary analyses suggest that GMA is an important predictor of job performance. However, it is also obvious that these findings are not in line with much of the previous literature. In the existing literature, one frequently finds the general notion that GMA is clearly an overwhelmingly dominant predictor of job performance with narrow cognitive abilities contributing little to the explanation of individual differences in job performance.

The reason why our conclusions differ from the conclusions in the previous literature clearly lies in the differences between the underlying theoretical models (nested-factors model vs. Spearmanian theories) and the corresponding analytical strategies. This notion was further supported when we contrasted the findings we obtained using the theoretical perspective of the nested-factors model (and relative importance analyses) with corresponding analyses based on the commonly used Spearmanian perspective (and incremental validity analyses). The Spearmanian incremental validity analyses of our data suggested that GMA was an overwhelmingly dominant predictor of job performance and thus corresponded to the empirical findings and the conclusion in the existing literature.

Theoretical Implications

Our meta-analytic study showed that theoretical assumptions (nested-factors model vs. Spearmanian theories) have a dramatic impact on conclusions regarding the importance of GMA and narrower cognitive abilities. Consequently, an obvious question is. Which perspective is more appropriate? In cognitive ability research, the two theoretical perspectives have been coexisting for decades. As we detailed, factor-analytic research cannot conclusively answer the question whether GMA is the underlying cause of narrower cognitive abilities or not (Yung et al., 1999). Accordingly, the appropriateness of the two perspectives can only be based on the perspectives' ability to explain empirical phenomena in cognitive ability research.

As previously noted, the nested-factors model can account for recent theoretical work (Dickens, 2007; van der Maas et al., 2006) suggesting that not only a unitary GMA mechanism but also developmental interactions between specific cognitive processes can explain the positive manifold phenomenon (the finding that almost all ability tests are correlated). Spearmanian theories, in contrast, assume that GMA is the source of

all shared variance between GMA and narrower cognitive abilities and can consequently not account for the possibility that specific cognitive mechanisms explain variance GMA and narrower abilities share.

Researchers may use these arguments to generally adopt the perspective of the nested-factors model in job performance research. We believe, however, that it is commonly more appropriate to opt for a more cautious approach and contrast the conclusions reached using both theoretical perspectives. In so doing, researchers can ensure consistency with previous research and avoid potential misinterpretations. For instance, readers might get the impression that the data in a new investigation using the nested-factors model fundamentally differ from the data in an older study that relied on a Spearmanian theory. In reality, however, the two datasets may be highly similar, and the (likely) diverging conclusions reached by the two investigations are only a result of different theoretical assumptions on the structure of cognitive abilities and different analytical procedures corresponding to these assumptions.

Methodological Implications and Remaining Methodological Issues

Although analytical procedures to test models relying on Spearman's theory have long been available to researchers (incremental validity analysis), researchers intending to rely on the nested-factors model have long faced methodological difficulties. Consequently, an important goal of this paper was to propose an analytical approach that allows researchers to study the role of cognitive abilities in job performance from the perspective of the nested-factors model.

Fundamentally, the proposed use of relative importance analysis provides a unique way to study the role of cognitive abilities in job performance from the perspective of the nested-factors model. In contrast to previous approaches, relative importance analysis avoids problems inherent in alternative methodological approaches (i.e., orthogonalization). We note, however, that not all measures of relative importance seem to be suited to provide insights on the role of cognitive abilities in job performance. Only the modern relative importance measures—relative weights and general dominance weights—permitted meaningful conclusions. The two traditional measures of relative importance, in contrast, did not provide useful estimates of relative importance in our meta-analytic study. We consequently suggest that researchers rely on modern relative importance measures when they study the role of cognitive abilities in job performance.

An important limitation of the analytical approach we proposed is that the approach cannot be used when GMA is an exact linear composite of the narrower cognitive abilities. This situation occurs when the GMA

measure is a perfect linear function of the narrower cognitive abilities included in the analysis (e.g., T , X , Y , and Z are the narrower cognitive abilities, and GMA is a composite of these abilities, $GMA = T + X + Y + Z$). The matrix is then singular, and consequently relative importance analyses and other regression-based analyses cannot be run. The reason is that regression analyses cannot be conducted on singular matrices (Cohen, Cohen, West, & Aiken, 2003).

The analyses we report in this paper ran because GMA was not an exact linear composite of the PMAs in the two matrices of construct relationships we used. In the first matrix, the GMA measure closely resembled a linear composite of the PMA measures. However, the fact that the WIT uses two tests twice to estimate PMA composites made GMA not an exact composite of the PMA measures. In the second matrix, the use of one test for each PMA had the result that six tests were not used in operationalizing the PMAs. The GMA measure, in contrast, naturally used all available information and was consequently not a linear composite of the PMA measures.

The fact that the proposed analytical approach cannot be used when GMA is a perfect linear composite of the narrower cognitive abilities limits the applicability of the proposed approach in research. Researchers interested in studying the relative importance of GMA and narrower cognitive abilities from the perspective of the nested-factors model may face situations where it makes conceptual sense to use an exact linear composite of the narrower cognitive abilities to operationalize GMA. For instance, this problem will arise when an ability test battery uses each of its tests to operationalize one ability construct and simultaneously relies on all its tests to operationalize GMA.

To address this issue, researchers need to alter either the composition of the narrower cognitive ability measures or the GMA measure. One relatively easy solution is to include an additional short GMA test. This test can then be added to the GMA composite so the GMA measure is not a linear combination of the tests. Theoretically, a potential problem with a “statistical finesse” of this type is that its use alters the originally desired measurement model for either GMA or the narrower cognitive abilities, slightly. In practice, however, the difference between a slightly altered measurement approach for GMA and an exact linear composite might be negligible when a considerable number of tests are used. As several authors (Jensen & Weng, 1994; Ree & Carretta, 2002; Ree & Earles, 1991) have noted, the commonly considerable correlations between cognitive ability tests make GMA composites very robust. When a considerable number of tests is already used, adding additional tests or altering the weights for the tests usually has minor effects on GMA composite measures (Jensen & Weng, 1994; Ree & Carretta, 2002; Ree & Earles, 1991).

An analysis we ran on our meta-analytic data illustrates this point. As described previously, the battery we used in this study did not have a measurement model for the narrower cognitive abilities that made GMA a perfect linear composite of the narrower cognitive abilities (the PMA composites). This situation enabled us to estimate the mean corrected correlation between the GMA measure we used (a composite of the 13 WIT tests) and a perfect linear composite of the PMA operationalizations (using the formula by Nunnally and Bernstein, 1994, described in the Methods section of this article). Results revealed that the estimated mean corrected correlation between the GMA measure and an exact linear composite of the seven PMA measures was $M_{\hat{\rho}} = .99$ in the matrix using all WIT tests to operationalize the PMAs and $M_{\hat{\rho}} = .95$ in the matrix using only marker tests to operationalize the PMAs. This level of correspondence is similar to GMA–GMA correlations commonly found in studies comparing different approaches to estimate GMA composites (Jensen & Weng, 1994; Ree & Carretta, 2002; Ree & Earles, 1991). The findings accordingly suggest that the proposed strategies will likely enable researchers to effectively deal with the singular-matrix problem.

Nevertheless, we believe that the opportunity to compare operationalizations of narrower cognitive abilities with their exact linear combination would greatly increase the flexibility and the potential scope of applications for the proposed approach. Consequently, a useful avenue for future research would be to develop statistical methodology that addresses this limitation and allows researchers to estimate relative weights and/or general dominance weights even when the underlying matrix is singular.

Practical Implications

The empirical findings from our meta-analytic investigation have relevant practical implications for prediction. In prediction, personnel psychologists seek to choose a practically useful predictor or assemble a practically useful predictor combination (LeBreton et al., 2007). In so doing, personnel psychologists consider a variety of relevant criteria. Commonly, personnel psychologists consider statistical criteria and other criteria like costs, applicant reactions, fairness, organizational goals, managerial acceptance, and legal aspects when making a decision. LeBreton and colleagues proposed that the relative importance of a variable is a particularly useful statistical criterion in decisions of this type. In contrast to other regression statistics, relative importance estimates are also useful when predictors overlap and are consequently helpful in choosing an appropriate predictor from an overlapping set of potentially measurable predictors. Accordingly, the modern relative importance estimates from

our meta-analytic study provide estimates of relative importance that can help personnel psychologists in choosing cognitive ability measures for prediction.

When personnel psychologists rely on this information, an important implication will be that GMA may not always be the predictor they choose. Depending on the relative importance analysis approach and job complexity, GMA was not always the relatively most important predictor in our analyses. The most common finding was that GMA and a couple of narrower cognitive abilities were about equally important. Thus, the findings suggest that GMA frequently does not have a statistical advantage over several narrower cognitive abilities, and the findings consequently provide an empirical basis for using measures of narrower cognitive abilities instead of GMA measures in personnel selection. Personnel psychologists may use this added flexibility to additionally consider nonstatistical criteria in selecting cognitive ability predictors.

One nonstatistical criterion on which narrower cognitive abilities may potentially outperform GMA measures are applicant reactions to selection procedures. Applicant reactions have relevant implications for both applicants and organizations (Hülshager & Anderson, 2009). GMA tests are generally evaluated less favorably than other common selection procedures like interviews, work-sample tests, or resumés (Anderson & Witvliet, 2008; Steiner & Gilliland, 1996). We are not aware of any empirical studies directly comparing applicant reactions to GMA measures with applicant reactions to measures of narrower cognitive abilities. Several authors, however, have speculated that using narrower cognitive ability measures instead of GMA measures can potentially improve applicant reactions to cognitive ability measures (Murphy, 2009; Ones, Viswesvaran, & Dilchert, 2005). One argument is that narrower cognitive ability concepts are more in line with layperson theories about cognitive abilities (e.g., Sternberg & Kaufman, 1998) and are accordingly generally preferred by applicants (Ones et al., 2005). Another argument is that narrower cognitive abilities can be matched to what applicants and potentially also managers and stakeholders consider relevant and adequate (Murphy, 2009). Perceived characteristics of selection measures like face validity and job relatedness are key antecedents of positive applicant reactions (Gilliland, 1993; Hausknecht, Day, & Thomas, 2004). Accordingly, choosing predictor measures according to what applicants prefer is important for personnel psychologists (Murphy, 2009). In so doing, a useful strategy may be to use applicant evaluations of cognitive ability measures (and potentially also evaluations by managers or stakeholders) and combine this information with empirical findings from the present and potential future studies on the relative importance of corresponding cognitive abilities. This information can then be used to select predictors

(i.e., narrower cognitive abilities) with an optimal balance between both criteria.

Study Limitations

This meta-analytic investigation has some notable limitations. One limitation is that the PMAs do not map the full and wide range of all cognitive abilities with respect to both content and level of specificity (cf. Carroll, 1993). The PMAs are only one of the many models of narrower cognitive abilities (albeit a very successful one) that have been proposed in the literature.

Another limitation is that our study is restricted to a specific test battery applied in one nation so that the findings may not be generalizable to other tests in other nations. However, there are some reasons to be optimistic about the generalizability of the findings. The WIT features prominently in German (Hülsheger et al., 2006, 2007) and European meta-analyses (Salgado, Anderson, Moscoso, Bertua, & deFruyt, 2003; Salgado, Anderson, Moscoso, Bertua, deFruyt, & Roland, 2003) on the relation between cognitive abilities and job performance criteria. These meta-analyses have generally reported results similar to those reported by North American meta-analyses and no evidence that the WIT is different from other German/European tests.

Finally, a third limitation results from the meta-analytic approach of our study. Specifically, the primary studies in our database do not account for recent efforts to expand the criterion space of job performance (Sackett & Lievens, 2008). A variety of different extensions of the criterion space have been proposed. Particularly popular are proposals to extend the criterion space to include (a) organizational citizenship behavior (e.g., LePine & Van Dyne, 2001; Sackett & Lievens, 2008), (b) counterproductive work behavior (e.g., Dilchert, Ones, Davis, & Rostow, 2007; Sackett & Lievens, 2008), (c) adaptive performance (e.g., Lang & Bliese, 2009; Pulakos, Arad, Donovan, & Plamondon, 2000), (d) role-based performance (e.g., Griffin, Neal, & Parker, 2007; Welbourne, Johnson, & Erez, 1998), and (e) innovative performance (Hülsheger, Anderson, & Salgado, 2009; Welbourne et al., 1998). Not all of these criteria are necessarily related to cognitive abilities. For instance, research on organizational citizenship behavior has provided no evidence for a substantive relation between cognitive abilities and citizenship behavior ($M_{\rho} \leq .07$; Bartram, 2005; LePine & Van Dyne, 2001), and research on counterproductive work behavior has yielded both negative (Dilchert et al., 2007) and positive effects (Roberts, Harms, Caspi, & Moffitt, 2007). Nevertheless, it is likely fruitful to incorporate organizational citizenship behavior and counterproductive work

behavior as well as other relevant measures like adaptive performance, role-based performance, and innovative performance into future meta-analyses once a sufficient number of studies linking these dimensions to cognitive abilities become available.

Conclusion

The goal of this paper was to clarify the role of theoretical assumptions (nested-factors model vs. Spearmanian perspectives) and analytical strategies (relative importance vs. incremental validity analysis) in research on the importance of GMA and narrower cognitive abilities in explaining job performance. In our opinion, the nested-factors model is a very useful and theoretically sound structural conceptualization and provides important additional insights in job performance research. We consequently encourage other researchers to conduct future studies using the nested-factors conceptualization and the proposed methodological approach.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

- Ackerman PL, Beier ME. (2006). Determinants of domain knowledge and independent study learning in an adult sample. *Journal of Educational Psychology*, 98, 366–381.
- Ackerman PL, Beier ME, Boyle MO. (2005). Working memory and intelligence: The same or different constructs. *Psychological Bulletin*, 131, 30–60.
- *Althoff K. (1974). Kurzbericht über eine bewährungskontrolle der eignungsprognosen für bewerber der allgemeinen verwaltung in Nordrhein-Westfalen [Short report of a validity study for applicants to the public administration in North-Rhine Westphalia, Germany]. *DGP Informationen*, 12/74, 12–21.
- *Althoff K. (1975). Die vorhersage des ausbildungserfolges von aufstiegsbeamten der kriminalpolizei [Prediction of training success for promoted police officers in the criminal investigation department]. *DGP Informationen*, 04/75, 13–23.
- *Althoff K. (1977). Zusammenhänge zwischen ergebnissen von eignungstests und beruflicher bewährung: Dargestellt am beispiel der auswahl von bewerbern für den gehobenen und höheren polizeivollzugsdienst [Relationship between aptitude tests and professional probation: The example of assessment of applicants applying for upper grade prison officer positions]. *Schriftenreihe der Polizei-Führungsakademie*, 1, 6–27.
- Anderson N, Witvliet C. (2008). Fairness reactions to personnel selection methods: An international comparison between the Netherlands, the United States, France, Spain, Portugal, and Singapore. *International Journal of Selection and Assessment*, 16, 1–13.
- Azen R, Budescu DV. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, 8, 129–148.
- Bacharach SB. (1989). Organizational theories: Some criteria for evaluation. *Academy of Management Review*, 14, 496–515.
- Bartram D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90, 1185–1203.

- Beauducel A, Kersting M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment, 18*, 97–112.
- Bertua C, Anderson N, Salgado JF. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology, 78*, 387–409.
- Blair C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences, 29*, 109–160.
- Bobko P, Roth PL, Potosky D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *PERSONNEL PSYCHOLOGY, 52*, 561–589.
- Brennan RL. (2001). *Generalizability theory*. New York: Springer-Verlag.
- *Bretz E, Oldendörp H. (1992). Bewährungskontrolle: Vorhersage des Ausbildungserfolges im angestelltenlehrgang I [Validity study: Prediction of success in employee training course I]. *DGP Informationen, 03/29*, 75–84.
- Budescu DV. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin, 114*, 542–551.
- Burke MJ, Landis RS. (2003). Methodological and conceptual challenges in conducting and interpreting meta-analysis. In Murphy KR (Ed.), *Validity generalization: A critical review* (pp. 287–309). Mahwah, NJ: Erlbaum.
- Carroll JB. (1982). The measurement of intelligence. In R. Steinberg (Ed.), *Handbook of human intelligence* (pp. 29–120). London: Cambridge University Press.
- Carroll JB. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Carroll JB. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research, 30*, 429–452.
- Carroll JB. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In Nyborg H (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 1–22). Oxford, UK: Elsevier.
- Cohen J. (1990). Things I have learned (so far). *American Psychologist, 45*, 1304–1312.
- Cohen J, Cohen P, West SG, Aiken LS. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.
- Darlington RB. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69*, 161–182.
- Deary IJ. (2002). *g* and cognitive elements of information processing: An agnostic view. In Sternberg RJ, Grigorenko EL (Eds.), *The general factor of intelligence* (pp. 151–182). Mahwah, NJ: Erlbaum.
- Dickens WT. (2007, November 19). *What is g?* Paper presented at the second Thomas S. Schelling Symposium, University of Maryland, College Park, MD. Retrieved from <http://www.publicpolicy.umd.edu/symposium/speakers.htm>.
- *Diepenbrock W, Thielepape M. (1973). Untersuchung über die Vorhersage des Ausbildungserfolges von Meisteranwärtern in der Automobilproduktion: Eine Bewährungskontrolle [Prediction of success for master craftsmen trainees in the automobile production industry: A proficiency control]. *DGP Informationen, 09/37*, 1–14.
- Dilchert S, Ones DS, Davis RD, Rostow CD. (2007). Cognitive ability predicts objectively measured counterproductive work behaviors. *Journal of Applied Psychology, 92*, 616–627.

- Duncan J, Seitz RJ, Kolodny J, Bor D, Herzog H, Ahmed A, et al. (2000). A neural basis for general intelligence. *Science*, 289, 457–460.
- Englehart MD. (1936). The technique of path coefficients. *Psychometrika*, 1, 287–293.
- Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.
- Garlick D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in neural plasticity as an explanatory mechanism. *Psychological Review*, 109, 116–136.
- Gilliland SW. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18, 694–734.
- Goldstein HW, Zedeck S, Goldstein IL. (2002). *g*: Is this your final answer? *Human Performance*, 15, 123–142.
- *Görllich Y, Schuler H. (2007). AZUBI-TH: *Arbeitsprobe zur berufsbezogenen Intelligenz* [AZUBI-TH: Work sample for work-related intelligence]. Göttingen, Germany: Hogrefe. *Arbeitsprobe zur berufsbezogenen Intelligenz. Technische und handwerkliche Tätigkeiten (AZUBI-TH)*. [Work sample test on work-related intelligence — Technical and mechanical Jobs (AZUBI-TH)]. Göttingen, Germany: Hogrefe.
- *Graudenz H. (1982). Bewährungskontrolle 2 – Vorhersage des ausbildungserfolges von beamtenanwärtern des gehobenen dienstes beim RP Darmstadt [Proficiency control 2—Prediction of training success of applicants for higher grade civil services of RP Darmstadt]. *DGP Informationen*, 12/28, 32–50.
- *Graudenz H. (1984). Personalauswahl für den mittleren justizvollzugsdienst in Hessen [Employee selection for the middle grade of the judiciary civil service in Hessen, Germany]. *DGP Informationen*, 12/48, 1–17.
- Gray JR, Thompson PM. (2004). Neurobiology of intelligence: Science and ethics. *Nature Reviews Neuroscience*, 5, 471–482.
- *Greif S. (1972). *Gruppenintelligenztests: Untersuchungen am WIT, IST, LPS und AIT* [Group intelligence tests: Studies of WIT, IST, LPS and AIT]. Bern, Switzerland: Herbert Lang.
- Griffin M, Neal A, Parker SK. (2007). A new model of work role performance: Positive behavior in uncertain and interdependent contexts. *Academy of Management Journal*, 50, 327–347.
- Grömping U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *American Statistician*, 61, 139–147.
- Grömping U. (2008). *Relaimpo: Relative importance of regressors in linear models (Version 2.1)* [Computer software]. Vienna: R Foundation for Statistical Computing. Retrieved from <http://cran.r-project.org/>.
- Gustafsson J-E, Balke G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407–434.
- Gutentag RL, Arvey RD, Osburn HG, Jeanneret PR. (1983). Moderating effects of decision-making/information-processing job dimensions on test validities. *Journal of Applied Psychology*, 68, 602–608.
- Hausknecht JP, Day DV, Thomas SC. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *PERSONNEL PSYCHOLOGY*, 57, 639–683.
- *Haustedt T. (2001). *Prognostische Validität eines Auswahlverfahrens in einem Transportunternehmen* [Predictive validity of a selection procedure in a transport company] Unpublished diploma thesis, University of Bielefeld, Bielefeld, Germany.
- Hoffman PJ. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin*, 57, 116–131.

- Holzinger KJ, Harman HH. (1938). Comparison of two factorial analyses. *Psychometrika*, 1, 45–60.
- Holzinger KJ, Swineford F. (1937). The bi-factor method. *Psychometrika*, 2, 41–54.
- Hülshager UR, Anderson N. (2009). Applicant perspectives in selection: Going beyond preference reactions. *International Journal of Selection and Assessment*, 17, 335–345.
- Hülshager UR, Anderson N, Salgado JF. (2009). Team-level predictors of innovation at work: A meta-analysis spanning three decades of research. *Journal of Applied Psychology*, 94, 1128–1145.
- Hülshager UR, Maier GW, Stumpp T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. *International Journal of Selection and Assessment*, 15, 3–18.
- Hülshager UR, Maier GW, Stumpp T, Muck PM. (2006). Vergleich kriteriumsbezogener validitäten verschiedener intelligenztests zur vorhersage von ausbildungserfolg in Deutschland: Ergebnisse einer metaanalyse. [Comparing criterion-related validities of different intelligence tests for the prediction of training success in Germany: A meta-analysis]. *Zeitschrift für Personalpsychologie*, 5, 145–162.
- Humphreys LG. (1981). The primary mental ability. In Friedman MP, Das JR, O'Connor N (Eds.), *Intelligence and learning* (pp. 87–102). New York: Plenum.
- Hunter JE, Hunter RF. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72–98.
- Hunter JE, Schmidt FL. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- *Jäger AO. (1963). Der Wilde-Test: Ein neues intelligenzdiagnostikum. [The Wilde-Test: A new instrument to assess intelligence]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 10, 260–278.
- *Jäger AO, Althoff K. (1983). *Der Wilde-Intelligenz-Test (WIT): Ein strukturdiagnostikum* [The Wilde Intelligence Test (WIT): A structural assessment instrument]. Göttingen, Germany: Hogrefe.
- *Jäger AO, Althoff K. (1994). *Der Wilde-Intelligenz-Test (WIT): Ein strukturdiagnostikum* [The Wilde Intelligence Test (WIT): A structural assessment instrument]. (2nd ed.). Göttingen, Germany: Hogrefe.
- Jensen AR, Weng L-J. (1994). What is a good g? [Editorial]. *Intelligence*, 18, 231–258.
- Johnson JW. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35, 1–19.
- Johnson JW, LeBreton JM. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods*, 7, 238–257.
- *Kersting M, Althoff K, Jäger AO. (2008). *Der Wilde Intelligenztest 2 (WIT-2)*. [The Wilde Intelligence Test 2 (WIT-2)]. Göttingen, Germany: Hogrefe.
- *Kleinevoss R. (1983). Untersuchung zur vorhersage des ausbildungserfolges von anwärtern des gehobenen dienstes einer Bundesbehörde [Study on the prediction of training success of candidates for the upper grade civil services of a German federal agency]. *DGP Informationen*, 12/38, 41–72.
- Kline P. (1991). *Intelligence: The psychometric view*. New York: Routledge.
- *Knebelau M. (2002). *Evaluation des WILDE-Intelligenztests (WIT) und seiner funktion als eignungsdiagnostisches instrument* [Evaluation of the WILDE intelligence test (WIT) and its function as a personnel selection assessment instrument]. Unpublished diploma thesis, RWTH (Rheinisch-Westfälische Technische Hochschule) Aachen University, Aachen, Germany.
- Kuncel NR, Hezlett SA, Ones DS. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162–181.

- *Lang K. (1990). *Entwicklung und längsschnittliche validierung einer eignungsdiagnostischen testbatterie für auszubildende gewerblich-technischer berufe* [Development and longitudinal validation of an aptitude test battery for selection of trainees in commercial-technical occupations]. Unpublished diploma thesis, Justus-Liebig-University Gießen, Gießen, Germany.
- Lang JWB, Bliese PD. (2009). General mental ability and two types of adaptation to unforeseen change: Applying discontinuous mixed-effects models to the task-change paradigm. *Journal of Applied Psychology*, 92, 411–428.
- LeBreton JM, Burgess JRD, Kaiser RB, Atchley EK, James LR. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar. *Organizational Research Methods*, 6, 80–128.
- LeBreton JM, Hargis MB, Griepentrog B, Oswald FL, Ployhart RE. (2007). Multidimensional approach for evaluating variables in organizational research and practice. *PERSONNEL PSYCHOLOGY*, 60, 475–498.
- LeBreton JM, Ployhart RE, Ladd RT. (2004). A Monte Carlo comparison of relative importance methodologies. *Organizational Research Methods*, 7, 258–282.
- LePine JA, Van Dyne L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential effects of Big-Five personality characteristics and general cognitive ability. *Journal of Applied Psychology*, 86, 326–336.
- Lindeman RH, Merenda PF, Gold RZ. (1980). *Introduction to bivariate and multivariate analysis*. Glenview, IL: Scott, Foresman and Company.
- Luciano M, Posthuma D, Wright MJ, de Geus EJC, Smith GA, Geffen GM. (2005). Perceptual speed does not cause intelligence, and intelligence does not cause perceptual speed. *Biological Psychology*, 70, 1–8.
- Mellon SJ, Jr, Daggett M, McManus V, Moritsch B. (1996). *Development of General Aptitude Battery (GATB) Forms E and F*. Washington, DC: U. S. Department of Labor. Retrieved from http://www.onetcenter.org/dl_files/Develop_GATB_EF.pdf.
- Mulaik SA, Quaretti DA. (1997). First order or higher order general factor? *Structural Equation Modeling*, 4, 193–211.
- Muller M. (1999). Human resource management under institutional constraints: The case of Germany. *British Journal of Management*, 10, 31–44.
- Murphy KR. (2009). Content validation is useful for many things, but validity isn't one of them. *Industrial and Organizational Psychology*, 2, 453–464.
- Murphy KR, DeShon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *PERSONNEL PSYCHOLOGY*, 53, 873–900.
- Nunnally JC, Bernstein IH. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Ones DS, Viswesvaran C, Dilchert S. (2005). Cognitive ability in personnel selection decisions. In Evers A, Voskuil O, Anderson N (Eds.), *Handbook of personnel selection* (pp. 143–173). Oxford, UK: Blackwell.
- Pratt JW. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In Pukilla T, Duntaneu S (Eds.), *Proceedings of second Tampere conference in statistics* (pp. 245–260). Finland: University of Tampere.
- Pulakos ED, Arad S, Donovan MA, Plamondon KE. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology*, 85, 612–624.
- Putka DJ, Le H, McCloy RA, Diaz T. (2008). Ill-structured measurement designs in organizational research: Implications for estimating interrater reliability. *Journal of Applied Psychology*, 93, 959–981.

- R Development Core Team. (2009). *R: A language and environment for statistical computing (Version 2.8.1) [Computer software]*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://cran.r-project.org/>.
- Raju NS, Burke MJ, Normand J, Langlois GM. (1991). A new meta-analytic approach. *Journal of Applied Psychology, 76*, 432–446.
- Raju NS, Lezotte DV, Fearing BK, Oshima TC. (2006). Note on correlations corrected for unreliability and range restriction. *Applied Psychological Measurement, 30*, 145–149.
- Ree MJ, Carretta TR. (2002). g2K. *Human Performance, 15*, 3–23.
- Ree MJ, Earles JA. (1991). The stability of g across different methods of estimation. *Intelligence, 15*, 271–278.
- Ree MJ, Earles JA, Teachout MS. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology, 79*, 518–524.
- Rindskopf D, Rose T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research, 23*, 51–67.
- Roberts BW, Harms PD, Caspi A, Moffitt TE. (2007). Predicting the counterproductive employee in a child-to-adult prospective study. *Journal of Applied Psychology, 92*, 1427–1436.
- Sackett PR, Lievens F. (2008). Personnel selection. *Annual Review of Psychology, 59*, 419–450.
- Sackett PR, Lievens F, Berry CM, Landers RN. (2007). A cautionary note on the effects of range restriction on predictor intercorrelations. *Journal of Applied Psychology, 92*, 538–544.
- Salgado JF, Anderson N, Moscoso S, Bertua C, de Fruyt F. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *PERSONNEL PSYCHOLOGY, 56*, 573–605.
- Salgado JF, Anderson N, Moscoso S, Bertua C, de Fruyt F, Rolland JP. (2003). A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology, 88*, 1068–1081.
- Schaie KW. (1979). The primary mental abilities in adulthood: An exploration in the development of psychometric intelligence. In Baltes PB, Brim OG, Jr, (Eds.), *Life span development and behavior* (pp. 67–115). New York: Academic Press.
- Schmid J, Leiman JM. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 83–90.
- Schmidt K, Foster PN. (1999). Germany's dual vocational training system. *Tech Directions, 57*, 15–17.
- Schmidt FL, Hunter JE. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*, 162–173.
- Schmidt FL, Hunter JE, Pearlman K. (1981). Task differences as moderators of aptitude test validity in selection: A red herring. *Journal of Applied Psychology, 66*, 166–185.
- *Schmidt-Atzert L, Deter B, Jaeckel S. (2004). Prädiktion von Ausbildungserfolg: Allgemeine Intelligenz (g) oder spezifische kognitive Fähigkeiten? [Prediction of training success: General intelligence or specific cognitive abilities?]. *Zeitschrift für Personalpsychologie, 3*, 147–158.
- *Schmied H. (1971). *Aspekte des zeit-güte problems im Wilde-Intelligenz Test*. [Aspects of the time-quality problem in the Wilde-Intelligence Test.] Unpublished diploma thesis, Free University Berlin, Berlin, Germany.
- Schuler H, Funke U, Baron-Boldt J. (1990). Predictive validity of school grades: A meta-analysis. *Applied Psychology: An International Review, 39*, 89–103.

- Schwab DP. (1980). Construct validity in organizational behavior. *Research in Organizational Behavior*, 2, 3–43.
- *Seggebruch G. (1982). Zur Vorhersage des Berufserfolgs von Bewerbern für den mittleren Verwaltungsdienst in einer Kommunalverwaltung [Prediction of job success for applicants of the middle grade civil service in a local administration]. *DGP Informationen*, 12/82, 6–31.
- *Seggebruch G. (1984). Vergleich der ergebnisse aus eignungsuntersuchungen und zwischenprüfungen im gehobenen allgemeinen verwaltungsdienst—Eine Bewährungskontrolle [Comparison of the results of aptitude tests and intermediate examinations in middle grade general administration services—e proficiency control]. *DGP Informationen*, 12/84, 56–75.
- *Seidel K, Knebelau M, Kersting M. (2002, September). *Methoden-kritische analysen und aktuelle Befunde zur Kriteriumsvalidität des WILDE-Tests*. [Method-critic analyses and up-to-date findings on the criterion validity of the WILDE-test]. Paper presented at the 43th congress of the Deutsche Gesellschaft für Psychologie, Berlin, Germany.
- Spearman C. (1904). “General intelligence,” objectively determined and measured. *American Journal of Psychology*, 15, 201–292.
- Spearman C. (1927). *The abilities of man, their nature and measurement*. New York: Macmillan.
- Steiner DD, Gilliland SW. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology*, 81, 134–141.
- Sternberg RJ, Kaufman JC. (1998). Human abilities. *Annual Review of Psychology*, 49, 479–502.
- Sturman MC, Cheramie RA, Cashen LH. (2005). The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *Journal of Applied Psychology*, 90, 269–283.
- Süß H-M., Oberauer K, Wittmann WW, Wilhelm O, Schulze R. (2002). Working memory explains reasoning—and a little bit more. *Intelligence*, 30, 261–288.
- *Thielepape M. (1980). Untersuchung über die Vorhersage des Ausbildungserfolgs von Krankenschwestern und -Pflegerinnen — eine Bewährungskontrolle. [Prediction of training success for nurses and caregivers — a proficiency control]. *DGP Informationen*, 7/08, 21–43.
- Thurstone LL. (1938a). *Primary mental abilities*. Chicago: University of Chicago Press.
- Thurstone LL. (1938b). The perceptual factor. *Psychometrika*, 3, 1–17.
- Thurstone LL. (1944). Second-order factors. *Psychometrika*, 9, 71–100.
- Thurstone LL, Thurstone TG. (1941). *Factorial studies on intelligence*. Chicago: University of Chicago Press.
- Thurstone LL, Thurstone TG. (1949). *Examiner manual for the SRA Primary Mental Abilities Test*. Chicago: Science Research Associates.
- U.S. Department of Labor. (1991). *Dictionary of occupational titles* (4th ed.). Washington, DC: Author.
- van der Maas HLJ, Dolan CV, Grasman RPPP, Wicherts JM, Huizenga HM, Raijmakers MEJ. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Bulletin*, 113, 842–861.
- Viswesvaran C, Ones DS. (1995). Theory testing: Combining psychometric meta-analysis and structural equation modeling. *PERSONNEL PSYCHOLOGY*, 48, 865–885.
- Viswesvaran C, Ones DS. (2002). Agreements and disagreements on the role of general mental ability in industrial, work, and organizational psychology. *HUMAN PERFORMANCE*, 15, 211–231.

- *Weber W, Werner R. (1983). Auswahlverfahren und prüfungsleistungen bei anwärtern des mittleren und gehobenen nicht-technischen dienstes—bericht über zwei bewährungskontrollen [Selection procedure and achievement in examinations of applicants for the middle and upper grade nontechnical civil services—report on two proficiency controls]. *DGP Informationen*, 12/83, 1–40.
- Welbourne TM, Johnson DE, Erez A. (1998). The role-based performance scale: Validity analysis of a theory-based measure. *Academy of Management Journal*, 41, 540–555.
- Wittmann WW, Süß H-M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem solving performances via Brunswik Symmetry. In Ackerman PL, Kyllonen PC, Roberts RD (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 77–108). Washington, DC: APA.
- *Wolf B. (1990). *Zusammenhänge zwischen ergebnissen von eignungsdiagnostischen verfahren und ausbildungserfolg-dargestellt am beispiel der auswahl von bewerbern für den gehobenen polizeivollzugsdienst*. [Relationships between results of aptitude assessments and training success- the example of applicant selection for the higher grade prison police service]. Unpublished diploma thesis, Georg-August-University, Göttingen, Germany.
- *Wolff P, Voullaire C. (1968). Eignungsbegutachtung von körperbehinderten für einen verwaltungsberuf – eine bewährungskontrolle. [Proficiency assessment for physically handicapped persons for an administrative profession- a proficiency control]. *Diagnostica*, 14, 70–87.
- Yung Y-F, Thissen D, McLeod LD. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, 64, 113–128.
- *Zeissig HJ. (1989). *Berufsfindung und arbeitserprobung der berufsförderungswerke, möglichkeiten ihrer ergänzung hin zum assessment-center* [Career choice and work-related testing in German rehabilitation institutions, possibilities of supplementation with assessment center tasks]. Unpublished dissertation, University of Bielefeld, Bielefeld, Germany.