More Than *g*-Factors: Second-Stratum Factors Should Not Be Ignored

Serena Wee
Singapore Management University

Daniel A. Newman and Q. Chelsea Song
University of Illinois at Urbana–Champaign

Ree, Carretta, and Teachout (2015) outlined a compelling argument for the pervasiveness of dominant general factors (DGFs) in psychological measurement. We agree that DGFs are important and that they are found for various constructs (e.g., cognitive abilities, work withdrawal), especially when an "unrotated principal components" analysis is conducted (Ree et al., p. 8). When studying hierarchical constructs, however, a narrow emphasis on uncovering DGFs would be incomplete at best and detrimental at worst. This commentary largely echoes the arguments made by Wee, Newman, and Joseph (2014), and Schneider and Newman (2015), who provided reasons for considering second-stratum cognitive abilities. We believe these same arguments in favor of second-stratum factors in the ability domain can be applied to hierarchical constructs more generally.

Hierarchical Constructs: Modern Psychometric Analyses Reveal the Second Stratum

Hierarchical constructs are everywhere. Even in the domain of cognitive ability, where positive manifold and empirical evidence for DGF is perhaps the strongest of any content domain, Carroll's (1993) empirical review of over 400 datasets led him to conclude that cognitive ability was best described *not* by a unidimensional model but by a *hierarchical* three-stratum model (also see McGrew, 2009). According to the hierarchical factor model[1] (see, e.g., Figure 1), a set of cognitive tests (e.g., tests of reading comprehension, vocabulary, and grammar) reflects a more-specific intellectual ability—that is, *second-stratum ability* (e.g., reading–writing ability). In turn, a set of second-stratum ability factors (e.g., reading-writing, quantitative reasoning, visual-spatial processing) reflects Spearman's higher order *g* factor. Hierarchical factor models of cognitive ability typically fit the data better than do unidi-

Serena Wee, School of Social Sciences, Singapore Management University; Daniel A. Newman and Q. Chelsea Song, Department of Psychology, University of Illinois at Urbana–Champaign.

Correspondence concerning this article should be addressed to Serena Wee, Singapore Management University, School of Social Sciences, 90 Stamford Road, Singapore 178903. E-mail: serenawee@smu.edu.sg

[1] We use the term "hierarchical" to refer generically to both higher order and bifactor models (see Yung, Thissen, & McLeod, 1999).

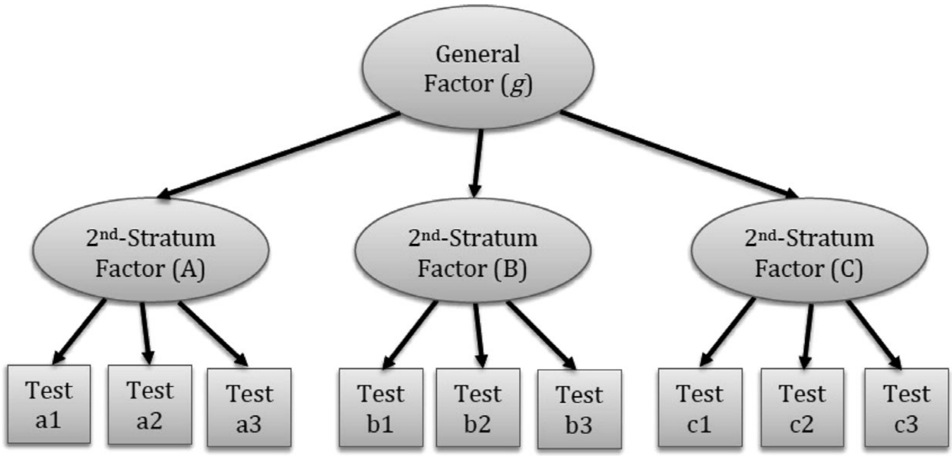

Figure 1. Hierarchical model with three strata (example). Examples of second-stratum cognitive abilities might include numerical ability, verbal ability, spatial ability, and clerical ability. Examples of tests that reflect numerical ability might include both arithmetic reasoning (word problems) and math knowledge (algebra-geometry-fractions-exponents). Examples of tests that reflect clerical ability/cognitive speed might include both numerical operations (a speeded test of simple math problems) and coding speed (a speeded test of recognizing arbitrary number strings; see Outtz & Newman, 2010).

mensional models (e.g., MacCann, Joseph, Newman, & Roberts, 2014; Outtz & Newman, 2010). The emerging consensus is thus that cognitive abilities are a set of hierarchically organized constructs: A higher order factor (i.e., *g*, the cognitive DGF) may be extracted from the positively correlated lower order factors (Schneider & Newman, 2015). We speculate that hierarchical factor models would also fare well in content domains such as job attitudes (Newman, Joseph, & Hulin, 2010) and work withdrawal (Hanisch, Hulin, & Roznowski, 1998), as well as in the many domains where Ree et al. identified DGFs; we assert that in all of these domains, psychometric models that include the second-stratum factors will tend to provide better fit to the data than do unidimensional models that include only a DGF.

As Ree et al. highlighted, the DGF often accounts for the majority of the test variance in a given psychological construct. When hierarchical factor analyses are conducted, each second-stratum ability factor accounts for less test variance than the DGF does. But focusing on the DGF while ignoring second-stratum ability factors may indicate a construct-deficient measurement model. In the cognitive ability domain for example, in addition to loading on the DGF, tests often also load substantially onto second-stratum ability factors. Mean loadings on the second-stratum factors were .42 for the Woodcock-Johnson Psycho-Educational Test Battery Manual sample (vs. .59

on *g*; Carroll, 2003) and also .42 for the 1960 Project TALENT sample (vs. .55 on *g*; Reeve, 2004). Although we concur with Ree et al. that DGFs are very important, we disagree with giving short shrift to second-stratum factors.

Some esteemed scholars have decried second-stratum factors as artifactual because it is plausible to attribute the appearance of second-stratum factors to *factor fractionation* or *swollen specifics* (Humphreys, 1962; Kelley, 1939). That is, any factor solution can be conditioned by adding tests from the same narrow domain until a lower order factor emerges. This argument is logically valid. By the same logic, however, any DGF (including *g* itself) might likewise be considered a swollen specific, which emerges because researchers have measured a given domain using relatively homogeneous instrumentation (Outtz & Newman, 2010). More specifically, the application of the cornerstone principle of convergent validity—in which a test is considered to measure cognitive ability only if it correlates highly with other cognitive ability tests—leads to *homometric reproduction* of constructs and instruments (Outtz & Newman, 2010). It is thus potentially inconsistent to claim that second-stratum factors emerge due to swollen specifics while simultaneously ignoring the possibility that DGFs are themselves swollen specifics, arising through the same process of specifying factor models on arbitrarily homogeneous indicators.

Specific Validity Depends on the Criterion Variable

Diversity outcomes. The use of cognitive tests in high-stakes selection typically results in adverse impact (i.e., the selection of disproportionately fewer minority applicants as compared with majority applicants), harming the diversity outcomes of a selection system. This is because the mean Black–White subgroup difference on a cognitive test composite (measuring *g*) is approximately 1 standard deviation in magnitude (Roth, Bevier, Bobko, Switzer, & Tyler, 2001). By contrast, second-stratum cognitive abilities can vary substantially in terms of the magnitude of their Black–White subgroup differences (Hough, Oswald, & Ployhart, 2001; Wee et al., 2014). By considering second-stratum cognitive abilities, rather than *g* alone, it is possible for specific cognitive ability factors to be differentially weighted so as to attenuate the trade-off between selection quality and organizational diversity. This is achieved by differentially weighting second-stratum abilities to achieve Pareto-optimal selection quality–diversity tradeoffs (De Corte, Lievens, & Sackett, 2007). For example, across two large samples comprising a total of 15 job families, Wee et al. (2014) showed it was possible to improve the proportion of hires from the minority group across all job families studied, with little to no decrement in selection quality compared with a unit-weighted cognitive test composite (essentially, compared with *g*). At least 8% diversity improvement was possible in all job families, and in four of the 15 job families, the adverse impact ratio more than doubled, greatly improving the proportion of job offers extended to minority candidates. Diversity improvement was typically achieved by assigning more weight in the selection system to second-stratum numerical ability and clerical ability and less weight to second-stratum verbal ability (Wee et al., 2014).

As is the case with other types of estimation, it is difficult to robustly estimate the weights assigned to second-stratum abilities at small sample sizes (e.g., $N < 50$), and more work examining the cross-validity of the technique remains to be conducted. Nonetheless, Wee et al. (2014) offer a "proof of concept" that organizational diversity may be improved without loss of selection quality, when compared with a unit-weighted *g* composite. For those interested in organizational diversity outcomes, Wee et al.'s results may augur a renewed interest in second-stratum abilities.

The compatibility principle. Beyond diversity outcomes, we acknowledge there has been only modest evidence for specific validity (i.e., incremental prediction of work performance criteria by second-stratum abilities, beyond *g*), especially for the criteria of training grades and work samples (see review by Ree & Carretta, 2002). Some scholars have noted that the meager results for the incremental validity of specific abilities might be due to how the performance criterion is measured (Reeve & Hakel, 2002; Viswesvaran & Ones, 2002). Echoing these authors and Ajzen and Fishbein (1977; also Fishbein & Ajzen, 1974), Schneider and Newman (2015) proposed an *ability–performance compatibility principle*: "General abilities predict general job performance, whereas specific abilities predict specific job performance." Schneider and Newman also note, "To our knowledge, only the first half of the ability–job performance compatibility principle (i.e., general ability predicts general job performance) has been rigorously evaluated to date" (p. 15). These authors then reviewed some suggestive results that are potentially relevant to the claim that specific abilities predict specific job performance criteria (Hogan & Holland, 2003; Joseph & Newman, 2010). However, research efforts are still hampered by a failure to evaluate *specific job performance criteria* (e.g., verbal job performance, spatial job performance). Until the job performance criterion is measured with compatible bandwidth to the cognitive ability construct, it will be difficult to draw unequivocal conclusions about the incremental validity of second-stratum abilities beyond *g*.

We expect the compatibility principle to generalize across relationships other than the cognitive ability–job performance relationship. That is, we should not expect second-stratum factors to predict general criteria. Instead, *second-stratum factors should predict second-stratum criteria*. Keeping this in mind should aid future researchers in designing potentially clearer tests of the specific validity hypothesis.

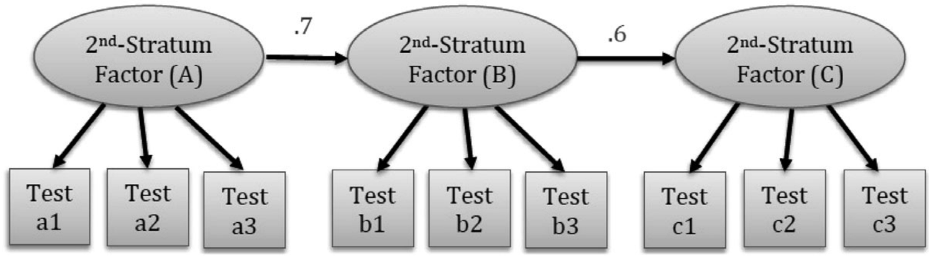

Figure 2A. Cascading model (example).

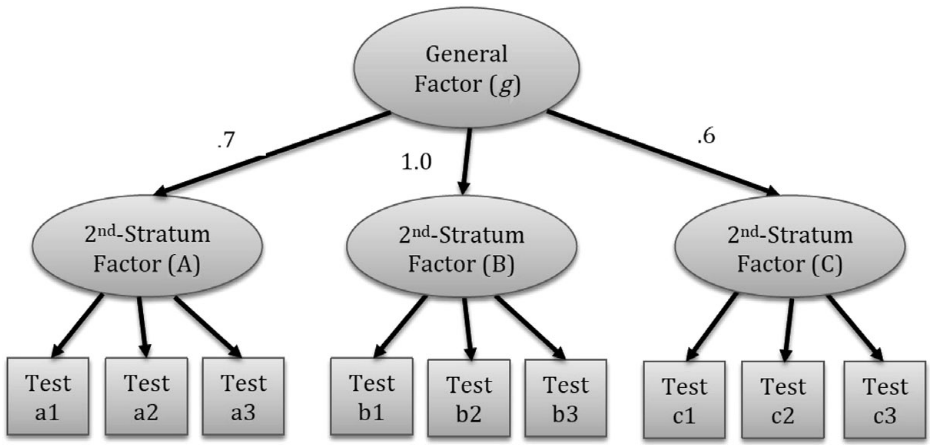

Figure 2B. Model with general factor (supported by same data as Figure 2A).

Positive Manifold Without *g*: Cascading Models

Ree and colleagues acknowledge that van der Maas et al.'s (2006) mutualism model could produce positive manifold in the absence of *g*, and we agree. Another, much simpler model—a cascading or mediation model—is also able to produce positive manifold in the absence of *g*. A *cascading model* is a type of mediation model that implies the sequential development of a set of related constructs over time, where development of one construct or skill enables the development of another construct or skill. For example, in Joseph and Newman's (2010) cascading model, emotional intelligence facets are connected in a developmental sequence: Emotion perception gives rise to emotion understanding, which in turn gives rise to emotion regulation. Yet, emotion perception, emotion understanding, and emotion regulation can also be treated as second-stratum factors of a higher order emotional intelligence construct (MacCann et al., 2014). It turns out that this is not uncommon: A cascading model (see Figure 2A) and a model containing a DGF (see Figure 2B) can often be fit equally well to the same data. So the data can often be equivocal as to whether a cascading model versus a DGF model is a more accurate theoretical specification. Whether the positive manifold is due to cascading/mediation versus a general higher order factor will likely require longitudinal data to resolve (see Cole & Maxwell, 2003, for a description of how longitudinal data might be useful in establishing mediation or cascading effects).

Summary

Empirically, we agree with Ree et al. that a positive manifold exists in many psychological constructs and that disregarding this positive manifold muddies the theoretical waters. Methodologically, as compared with an unrotated principal components analysis, more effective analytic strategies such as hierarchical and bifactor analyses exist to disentangle DGFs from second-stratum factors. We believe an emphasis on DGFs—at the expense of second-stratum factors—would prevent us from developing a deeper theoretical understanding of the nomological networks of psychological constructs. Both DGFs *and* second-stratum factors should be considered. We concur that ignoring DGFs is a chronic and widespread problem in many domains of organizational research, but overemphasizing DGFs to the disregard of second-stratum factors (as exemplified in the domain of cognitive ability research; Schneider & Newman, 2015; Wee et al., 2014) can also be a problem. The predictive value of second-stratum factors should be audited while keeping in mind *multiple criteria* (e.g., diversity outcomes), the *compatibility principle* (i.e., specific predictors lead to specific criteria), and the possibility that positive manifold can emerge from *cascading models* (i.e., mediation models).

References

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical literature. *Psychological Bulletin, 84*, 888–918.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. Cambridge, United Kingdom: Cambridge University Press.

Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen*. Oxford, United Kingdom: Elsevier Science Pergamon Press.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558–577.

De Corte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology, 92*, 1380–1393.

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*, 59–74.

Hanisch, K. A., Hulin, C. L., & Roznowski, M. (1998). The importance of individuals' repertoires of behaviors: The scientific appropriateness of studying multiple behaviors and general attitudes. *Journal of Organizational Behavior, 19*, 463–480.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job–performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*, 100–112.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9*, 152–194.

Humphreys, L. G. (1962). The organization of human abilities. *American Psychologist, 17*, 475–483.

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*, 54–78.

Kelley, T. L. (1939). Mental factors of no importance. *Journal of Educational Psychology, 30*, 139–143.

MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second-stratum factor of intelligence: Evidence from hierarchical and bifactor models. *Emotion*, *14,* 358–374.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10.

Newman, D. A., Joseph, D. L., & Hulin, C. L. (2010). Job attitudes and employee engagement: Considering the attitude "A-factor." In S. Albrecht (Ed.), *The handbook of employee engagement: Perspectives, issues, research, and practice* (pp. 43–61). Northampton, MA: Edward Elgar.

Outtz, J. L., & Newman, D. A. (2010). A theory of adverse impact. In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 53–94). New York, NY: Routledge/Taylor & Francis Group.

Ree, M. J., & Carretta, T. R. (2002). *g*2K. *Human Performance, 15*, 3–23.

Ree, M. J., Carretta, T. R., & Teachout, M. S. (2015). Pervasiveness of dominant general factors in organizational measurement. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *8*(3), 409–427.

Reeve, C. L. (2004). Differential ability antecedents of general and specific dimensions of declarative knowledge: More than *g*. *Intelligence*, *32*, 621–652.

Reeve, C. L., & Hakel, M. D. (2002). Asking the right questions about *g*. *Human Performance, 15*, 47–74.

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S. I., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54*, 297–330.

Schneider, W. J., & Newman, D. A. (2015). Intelligence is multidimensional: Theoretical review and implications of narrower cognitive abilities. *Human Resource Management Review, 25*, 12–27.

van der Maas, H. L., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*, 842–861.

Viswesvaran, C., & Ones, D. S. (2002). Agreements and disagreements on the role of general mental ability (GMA) in industrial, work, and organizational psychology. *Human Performance, 15*, 211–231.

Wee, S., Newman, D. A., & Joseph, D. L. (2014). More than *g*: Selection quality and adverse impact implications of considering second-stratum cognitive abilities. *Journal of Applied Psychology, 99*, 547–563.

Yung, Y. F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika, 64*, 113–128.