Psychology Science, Volume 49, 2007 (3), p. 187-207

Working memory and fluid intelligence are both identical to *g*?! Reanalyses and critical evaluation

GILLES E. GIGNAC[1]

Abstract

In this investigation, two previously published confirmatory factor analytic studies that separately reported working memory and fluid intelligence higher-order loadings so large as to suggest isomorphism with *g* were evaluated critically within the context of internal consistency reliability. Specifically, based on two data analytic approaches, the previously reported higher-order loadings which suggested isomorphism with *g* were demonstrated to have been achieved via the substantial disattenuation effects observed within structural equation modeling, when the latent variable corresponding composite scores are associated with low levels of reliability. The two approaches were: (1) the obverse of the disattenuation procedure for imperfect reliability, and (2) the implied correlation between a corresponding phantom composite variable and a higher-order *g* factor. The results derived from the two approaches were found to correspond very closely. To allow for a more informative evaluation, researchers are encouraged to report the internal consistency reliabilities associated with the composite scores which correspond to their latent variables, as well as to report both the disattenuated and attenuated higher-order loadings within their multi-factor models.

Key words: intelligence; higher-order modeling; reliability; working memory; fluid intelligence

[1] Gilles E. Gignac, School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, WA, 6009, Australia; email: gilles@psy.uwa.edu.au

In practice, the application of factor analysis frequently involves the identification of a factor's most highly loading variable, which generally facilitates the labelling and interpretation of the factor. With respect to the general factor of intelligence, considerable effort has been devoted to the quest of identifying the most highly *g* loading type of subtest and/or factor, presumably because this may help uncover the most fundamental nature of the construct. Spearman (1927) described the fundamental cognitive process underlying the general factor based on three noegenetic laws: apprehension of experience; eduction of relations; and eduction of correlates. However, individual differences in general intelligence were considered by Spearman (1927) to be principally due to 'mental energy', an idea that continues to attract some attention in the theoretical literature (e.g., Lykken, 2005). From a more psychometric perspective, individual differences in the performance of cognitive ability tests are widely considered to yield a 'positive manifold' (i.e., positive correlations between all subtests; Jensen, 1998). Further, the vast majority of the shared variance between subtests has been argued to be reducible to a single factor, the *g* factor (Jensen, 1998). However, it is also widely acknowledged that a single *g* factor does not usually account for all of the inter-subtest covariance, and, consequently, that meaningful group-factors can be extracted from a matrix of inter-correlated subtests (Jensen, 1998). The practical significance of these group-factors remains an active area of research, with some investigators claiming a substantial amount of unique construct validity and others arguing that there is little more than *g* (see Sternberg & Grigorenko, 2002).

In conjunction with the unique construct validity research, some studies have focused upon examining the relative importance of group-factors as contributors of *g* factor variance using confirmatory factor analysis (CFA), with some rather surprising results. Specifically, several authors of CFA investigations have contended to have observed identity (i.e., 1.0) or near identity (i.e., >.90) associations between a first-order (group-level) factor and a *g* factor, which has prompted various schools of thought to claim that the fundamental nature of general intelligence resides within a particular cognitive domain.

It would appear that two of the most frequently cited lower-order factors claimed to be isomorphic with *g* are Working Memory (WM) and fluid intelligence (Gf), both of which have been reported to be associated with empirical support. Colom, Rebollo, Palacios, Juan-Espinosa, and Kyllonen (2004) and Gustafsson (1984) are perhaps two of the better known investigations that have used CFA and higher-order modeling to support contentions that a lower-order group-factor is isomorphic with *g*. In the case of Colom et al. (2004), WM was reported to have an association with *g* equal to 1.04 (in sample 1), while in the case of Gustafsson (1984), Gf was reported to be associated with *g* at 1.0. Obviously, it is doubtful that two apparently distinct lower-order group-factors can both be identical to *g*. The logical implications with the above disparate contentions are that if both Gf and WM are identical to *g*, then both Gf and WM are also identical to each other.[2] Clearly, the current status of the literature, which includes conflicting claims for the status of two lower-order group-factors as indicators of *g*, is in need of some work to help clarify the situation.

[2] Blair (2006) is a rare example where it has been suggested that Gf and WM are the same constructs. However, there is substantial theoretical and empirical evidence to suggest that Gf and WM do possess a non-negligible amount of unique construct validity. See, for example, several responses to Blair's (2006) target article. Also, consider that, in Colom et al.'s (2004) higher-order models, both WM and Gf were modeled as lower-order factors but only WM exhibited isomorphic-like loadings with *g*.

Consequently, the purpose of this investigation was to potentially resolve the above conflict in the literature. To this effect, the CFA investigation by Colom et al. (2004) and Gustafsson (1984) will be reviewed critically and the published correlation matrices re-analysed. An alternative methodological explanation for the observation of isomorphic (or near isomorphic) associations will be provided, based on internal consistency reliability and the disattenuation effects within structural equation modeling (SEM). Prior to evaluating and re-analysing the results and data reported in Colom et al. (2004) and Gustafsson (1984), a non-technical introduction and demonstration of internal consistency reliability within the context of latent variable modeling will be provided.

SEM and internal consistency reliability

It is well established that the association between two variables will be attenuated when the measured variable scores are associated with imperfect levels of reliability (Muchinsky, 1996). The degree of attenuation between the two variables has been shown to be a function of the square root of the product of the reliabilities of the correlated variables (Spearman, 1904). This product is commonly referred to as r_{max} and formulated as $\sqrt{(r_{xx})(r_{yy})}$, where r_{xx} and r_{yy} are equal to the reliabilities of the two correlated variables (Nunnally & Bernstein, 1994). Thus, to disattenuate a correlation for imperfect reliability, one divides the observed correlation by r_{max}. Fan (2003a) demonstrated the close correspondence between the disattenuation effects achieved via Spearman's (1904) Classical Test Theory approach and the disattenuation effects achieved by correlating latent variables within SEM.

The word 'composite' will be used in this paper to describe a variable or score that is based on the sum of two or more variables (e.g., items or subtests). It is important to make clear the distinction between the reliability of a latent variable's corresponding composite score and the reliability of the scores associated with the indicators (say, subtests) used to define a latent variable (and, correspondingly, used to create a composite sum score). The reliability of the indicator (subtest) scores is not directly relevant to the disattenuation effects observed within a latent variable framework. Rather, it is the reliability of the corresponding composite scores derived from an aggregation of the subtests that defines the latent variable that is of direct relevance (note that, technically, latent variables are devoid of measurement error).

Since the popularization of SEM, a small number of publications have cautioned against the application of SEM in cases where the observed variables chosen to define a latent variable (usually subtests in intelligence research) are correlated so weakly as to be associated with unacceptably low levels of corresponding composite score internal consistency reliability. For example, Cohen, Cohen, Teresi, Marchi, and Velez (1990) specifically recommended researchers to calculate and report the reliabilities associated with the composite scores which correspond to the latent variables within their models, based on the same principle for doing so when analysing data using conventional statistical analyses (e.g., Pearson correlation, multiple regression, etc.). Given the low frequency with which the Cohen et al.

(1990) paper has been cited[3], it would appear that few researchers have taken into consideration the issue of corresponding composite score reliability when applying SEM to their data.

To facilitate a clear understanding of the effects of the reliability of corresponding composite scores within latent variable modeling, a demonstration will be conducted based on a higher-order factor model with one second-order general factor and three first-order factors. Prior to the description and presentation of the results associated with the demonstration, some discussion of the methods used to estimate the internal consistency reliability of composite scores within a latent variable modeling framework will be presented.

Internal consistency reliability within SEM: Omega (Ω) and Phantom Omega (Ω_{Ph})

Internal consistency reliability is generally conceptualized as the ratio of true score variance to total variance (Crocker & Algina, 1986). While Cronbach's alpha (α) is the most popular method of estimating internal consistency reliability (Peterson, 1994), it may be argued to be a limited method of estimating reliability, as it assumes that each variable included in the composite is associated with an equal amount of true score variance (i.e., tau-equivalence), as well as that no two indicators share variance independently of the common dimension all indicators are expected to measure (i.e., independence of error terms) (McDonald, 1999). For this reason, the estimation of internal consistency reliability within a latent variable framework may be considered an advantage, as formulae have been established to accommodate model solutions associated with unequal factor loadings and correlated error terms. A well established method of estimating internal consistency reliability within a latent factor framework is known as omega (Ω; Hancock & Mueller, 2001; McDonald, 1999; Raykov, 2001), which can be formulated as:

$$\Omega = \frac{(\sum_{i=1}^{k} \ell_i)^2}{(\sum_{i=1}^{k} \ell_i)^2 + 2\sum_{1 \le i < j \le k} \delta_{ij} + \sum_{i=1}^{k} \delta_{ii}}$$

where ℓ_i = standardized factor loading, and δ_{ij} = correlated error term, and δ_{ii} = error variance (i.e., $1-\ell_i^2$). In the event that the factor model is not associated with correlated error terms, the $2\sum_{1 \le i < j \le k} \delta_{ij}$ portion of the Ω equation would be equal to zero, effectively dropping out of the equation. Note that the equation above is based on the standardized factor solution, which may be considered preferable to unstandardized formulations of Ω (e.g., Raykov, 2001), as the unstandardized formulations assume equal variances across all indicators (Hancock & Mueller, 2001).

[3] As of August 31st, 2007, PsycINFO reported that the Cohen et al. (1990) paper had been cited 22 times. However, the majority of the papers that have cited the Cohen et al. (1990) publication refer to aspects of the paper unrelated to the disattenuation effects within SEM (e.g., latent variables vs. emergent variables).

An alternative conceptualization of internal consistency reliability is based on the notion of the correlation between observed scores and true scores, which is referred to as the 'reliability index' (Nunnally & Bernstein, 1994). The estimation of internal consistency reliability via the reliability index has been established with popular SEM programs based on the utility of 'phantom variables' (see Fan, 2003b; Raykov, 1997). Phantom variables within SEM are neither observed variables nor latent variables. Rather, within the context composite score reliability estimation, they are simply the representation of sum scores from which implied correlations can be estimated. Thus, the correlation between a sum score (i.e., composite) and its corresponding latent variable (i.e., true score) is equal to the reliability index. When squared, the reliability index is equal to Ω (Fan, 2003b). For the purposes of clarity, the Ω estimated from the phantom variable modeling approach will be referred to as 'phantom Ω' (Ω_{Ph}).

In addition to the estimation of internal consistency reliability, the utility of phantom variables within the context of higher-order modeling pertains to the fact that implied correlations can be also estimated between group-factor corresponding composites and a *g* factor (r_{Ph*g}). Note that r_{Ph*g} correlations are not disattenuated for imperfect reliability, despite the fact that they are estimated within a SEM program. Consequently, r_{Ph*g} estimates would be expected to correspond closely to the higher-order loadings attenuated (or re-attenuated) via Ω based on the obverse of Spearman's (1904) classic disattenuation procedure. These attenuation procedures will now be described in more detail.

The attenuation procedure via Omega (Ω) and Phantom Omega (Ω_{Ph})

In order to demonstrate the disattenuation effects achieved within higher-order modeling, two correlation matrices consistent with three lower-order factors (three indicators each) and one higher-order general factor were generated: (1) based on indicators associated with "respectable" levels of composite score internal consistency reliability (i.e., .80), and (2) based on indicators associated with unacceptably low levels of composite score reliability (i.e., <.70). A graphical depiction of the model is presented in Figure 1 (Model 1). It can be observed that the higher-order model includes a phantom variable associated with each group-factor for the purposes of estimating Ω_{Ph} and r_{Ph*g}. The primary purpose of the demonstration was to reveal the very substantial disattenuation effects observed within CFA when the latent variable's corresponding composite scores are associated with very low levels of internal consistency.

For the purposes of this paper, the correlations between indicators designed to load onto the same lower-order group-factor will be referred to as 'intra-group correlations'. In contrast, the term 'extra-group correlations' will be used to refer to correlations between subtests from different lower-order group-factors. The mean intra-group correlations would be expected to be closely associated with the internal consistency of the reliability of the composite scores corresponding to that narrow group of aggregated subtests, while the mean extra-group correlations would reflect some indication of association with a *g* factor (assuming there was one). Within the first generated correlation matrix, the intra-group correlations were specified at .58, which corresponded to an equally weighted composite score reliability (Ω) of .806, i.e., a level which is generally considered acceptable for basic research in psy-

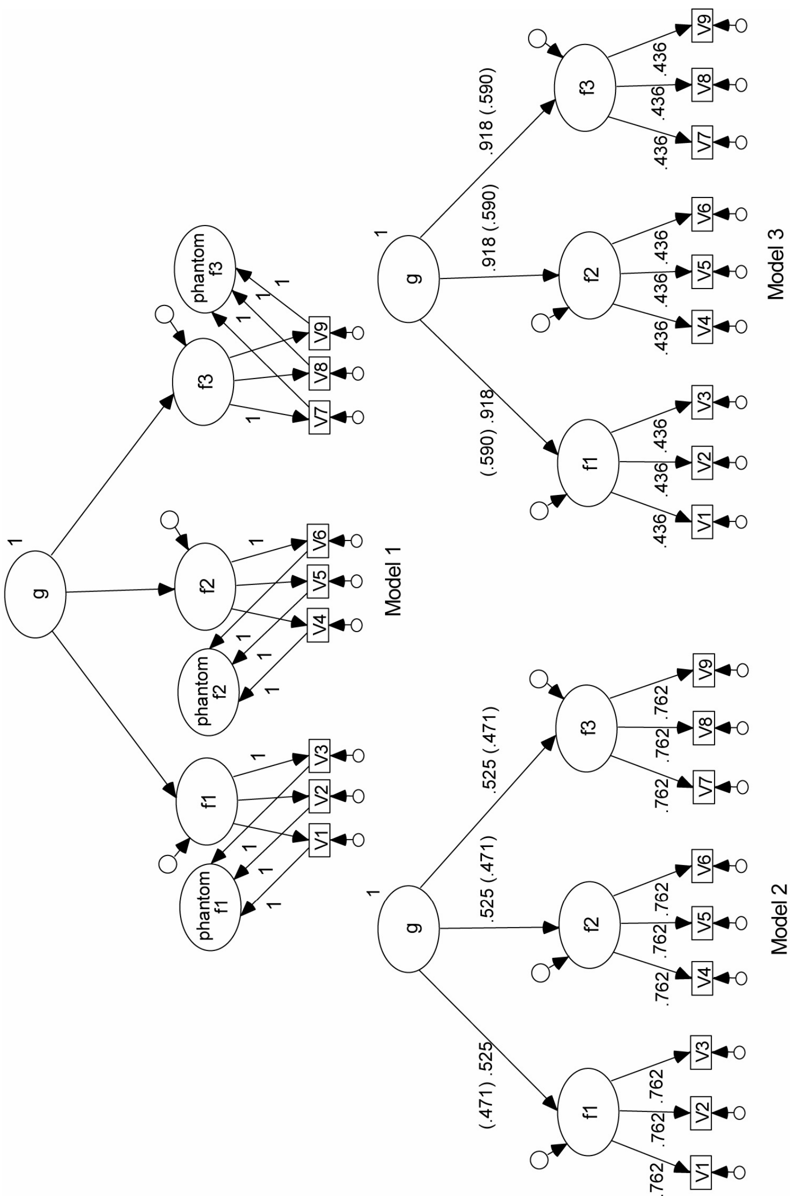

Figure 1:
Model 1 = higher-order model with phantom variables and corresponding constraints applied for the purposes of identification/scaling; Models 2 and 3 = high reliability/low reliability higher-order models with completely standardized maximum likelihood estimation parameter estimates (attenuated parameter estimates within parentheses).

chology (Nunnally & Bernstein, 1994). In contrast, the correlations between the indicators designed to load onto the two latent variables (i.e., extra-group correlations) were all specified to be equal to .16 (i.e., quite low; see Table 1). As can be seen in Figure 1 (Model 2), the first-order factor loadings were equal to .762 and the second-order factor loadings were equal to .525 (all model solutions were estimated via AMOS 5.0).

One may wonder to what degree the higher-order factor loadings may have been increased due to the dissattenuation effects achieved within SEM. Such a question would be in the spirit of Cohen et al.'s (1990) recommendation that researchers consider the effect of reliability on the estimation of the parameter estimates between latent variables within their SEM models. To estimate the attenuated higher-order factor loadings associated with the three lower-order factors, the disattenuated loadings were multiplied by the square root of their respective corresponding composite score reliabilities. Such a procedure is simply the obverse of the Spearman (1904) disattenuation for imperfect reliability procedure discussed above. First, however, the internal consistency reliabilities were verified to be .806 using the Ω formula:

$$\Omega = \frac{(.762+.762+.762)^2}{(.762+.762+.762)^2 + ([1-.581]+[1-.581]+[1-.581])} = \frac{5.226}{4.226+1.257} = .806$$

Further, based on the phantom variable modeling approach, the implied correlation between the two phantom variable composites and their corresponding latent variables was equal to .898, which corresponded to a Ω_{Ph} of $.898^2 = .806$ (i.e., identical to the Ω estimate of .806). To estimate the corresponding attenuated higher-order loadings, the second-order factor loading estimates of .525 were multiplied by the square root of their respective underlying composite score reliabilities (i.e., Ω), which in this case was equal to $.806^{1/2} = .898$. Thus, the corresponding attenuated factor loadings were equal to .525*.898 = .471. Correspondingly, the implied correlations between each phantom variable and the higher-order *g* factor (i.e., r_{Ph*g}) were also equal to .471.

Within the second generated correlation matrix, the intra-group correlations and the extra-group correlations were specified at .19 and .16, respectively, which corresponded to an equally weighted composite score reliability of .413 (i.e., very low), based on both the Ω formula

$$\Omega = \frac{(.436+.436+.436)^2}{(.436+.436+.436)^2 + ([1-.190]+[1-.190]+[1-.190])} = \frac{1.711}{1.711+2.430} = .413$$

and the implied phantom variable correlation the internal consistency reliability was estimated at .413 (i.e., $\Omega_{Ph} = .643^2 = .413$).

To estimate the corresponding attenuated higher-order loadings, the factor loading estimates of .918 were multiplied by the square root of their respective corresponding composite score reliabilities (Ω), which in this case was equal to $.413^{1/2} = .643$. Thus, the corresponding attenuated factor loadings were equal to .918*.643 = .590, which does not suggest isomorphism. As would be expected, the implied correlation between each phantom composite variable and the higher-order *g* factor (i.e., r_{Ph*g}) was also equal to .590.

The results of the above analyses have demonstrated four effects: (1) the estimation of composite score reliability (Ω) based on latent variable standardized factor loadings and errors; (2) the estimation of composite score reliability (Ω_{Ph}) based on the squared implied correlation between a phantom variable and its corresponding latent variable; (3) the convergence of Ω and Ω_{Ph}; and (4) the substantial effects of a latent variables' low levels of corresponding composite score internal consistency reliability on the estimation of higher-order loadings within a latent variable modeling framework. Based on the above information, it remained to be determined whether the latent variable corresponding composite scores associated with the Colom et al. and Gustafsson higher-order model latent variables were based on respectable levels of internal consistency reliability. Prior to analysing the Gustafsson and Colom et al. data, a brief review of the higher-order models they tested will be provided.

Brief review of Gustafsson (1984) and Colom et al. (2004)

Perhaps the most frequently cited study to claim a lower-order group-factor to be identical with *g* is that of Gustafsson (1984), where it was reported that Gf was identical to *g*, based on a higher-order model of several cognitive ability subtests in a sample of 981 sixth grade children. More specifically, Gustafsson (1984) tested a higher-order factor model defined by 10 first-order factors, three second-order factors, and one third-order *g* factor. The second-order factors were visual/spatial intelligence (Gv), crystallized intelligence (Gc), and fluid intelligence (Gf). Gustafsson (1984) reported that the Gf second-order factor had a standardized loading of 1.00 on the third-order *g* factor, while Gv and Gc had *g* loadings of .80 and .76, respectively. Thus, Gustafsson (1984) concluded that "Gf seems identical with '*g*'" (p. 193). A graphical depiction of the higher-order model endorsed by Gustafsson has been re-tested by the present author and reproduced in Figure 2 (Model 1).[4]

[4] Note that the higher-order model endorsed by Gustafsson (1984) included a large number of correlated residuals/errors and cross-loadings. Gustafsson was led to add a large number of freely estimated parameters to the model because he employed an exact-fitting (χ^2 p>.05) approach to model fit evaluation, which was common at the time.

It will also be noted that the Gf loading of 1.05 reported in this paper (Figure 2, Model 1) is larger than the loading of 1.00 reported by Gustafsson (1984, p. 192). It is highly unlikely that a second-order factor would be related to *g* at exactly 1.00. As described by Dillon, Kumar, and Mulani (1987), in many cases where a factor model yields a very large factor loading estimate, the loading is actually above 1.0 (i.e., Heywood case), which necessarily implies that the residual variance associated with the predicted observed/latent variable is negative. The observation of a Heywood case generally compromises interpretations of the CFA model, as all variances within an acceptable SEM model estimated via MLE must be non-negative (Joreskog & Sorbom, 1988). When a negative variance is observed, it is usually due to model misspecification and/or over-parameterization (Joreskog & Sorbom, 1988). Because the Gf standardized loading onto the third-order general factor was estimated at 1.05, it necessarily implied that the Gf latent variable residual variance was negative (in fact, it was estimated at -.01). Finally, and perhaps equally important, the solution was also associated with a non-positive definite matrix, which would render interpretations of the model questionable (Joreskog & Sorbom, 1988).

Perhaps the only SEM technique that could be used justifiably to deal with the negative residual variance would be to constrain the Gf factor's residual variance to a boundary value (i.e., .001), which would necessarily result in a Gf loading of 1.00. However, Joreskog and Sorbom (1988, p. 215) asserted that constraining variances to non-negative values should not be viewed as valid way to deal with this type of problem. Instead, it should be acknowledged that the "root of the problem is that the model is empirically over pa-

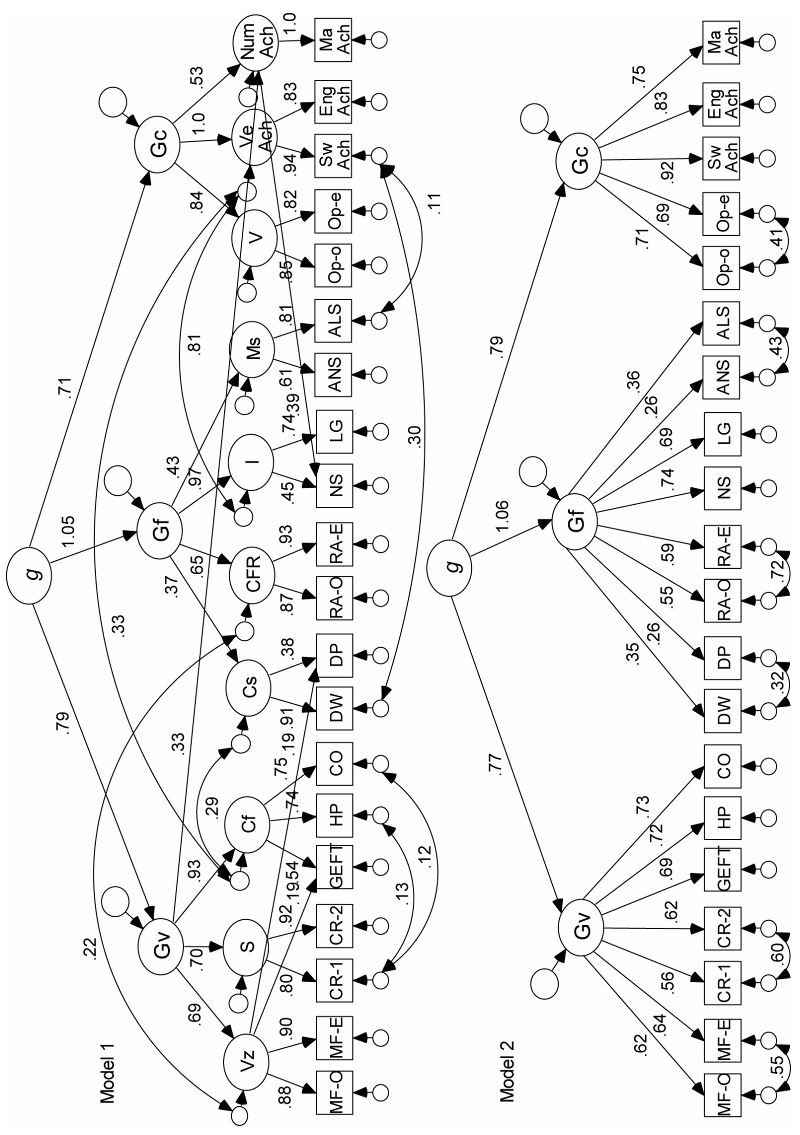

Figure 2:
Model 1 = Gustafsson (1984) endorsed higher-order model and completely standardized estimates; Model 2 = equivalent higher-order model with only two-orders and correlated residuals between indicators derived from the same subtest; for full subtest names see Gustafsson.

rameterised" (Joreskog & Sorbom, 1988, p. 215). Despite Joreskog and Sorbom's (1983) recommendation, the present author re-tested the higher-order model endorsed by Gustafsson (1984) with the added .001 constraint to Gf's residual variance. The chi-square difference test was found to be statistically significant ($\Delta\chi^2_1 = 4.26$, $p<.05$), which implied that the negative residual variance estimate was likely not due to sampling fluctuations (Chen, Bollen, Paxton, Curran, & Kirby, 2001). Further, the model solution was found to remain associated with a non-positive definite matrix, which suggested that the model was either wrong or that the sample size was too small (Joreskog & Sorbom, 1988). Given that the data were reported to be based on 1224 subjects, it is probably more plausible to suggest that the model was unacceptable. For the purposes of this investigation, however, the issue of the acceptability of the Gustafsson higher-order model will not be considered further.

Prior to the more recent WM and *g* CFA investigations, Kyllonen and Christal (1990) reported large (.80+) correlations between WM and general reasoning across several samples based on oblique factor models, which suggested that reasoning ability could be reduced to WM capacity. Inspired by the Kyllonen and Christal (1990) study, Colom, Rebollo, Palacios, Juan-Espinosa, and Kyllonen (2004) sought to test the hypothesis that WM was identical to general intelligence, using a CFA higher-order modeling strategy across three adult samples with a mean $N = 198$. Based on the first data set, Colom et al. modeled a higher-order model with four first-order factors (defined by three subtests each), which corresponded to Working Memory (WM), Processing Speed (PS), Crystallized Intelligence (Gc) and Fluid Intelligence (Gf). The second and third data sets included an additional three subtests used to form a fifth group-factor (psychometric speed, Gs). Based on the three data sets, Colom et al. (2004) found that a first-order WM factor exhibited a mean second-order factor loading of .96 onto the *g* factor. A graphical depiction of the models and parameter estimates associated with the Colom et al. investigation are presented in Figure 3. It can be observed that, in one case (sample 1), WM exhibited a second-order standardized factor loading of 1.04 onto *g*. Samples two and three yielded WM higher-order loadings of .89 and .93, respectively. Consequently, Colom et al. (2004) concluded that WM may possibly be considered the "essence of *g*" (p. 288).

In light of the above demonstrated effects relevant to internal consistency reliability and the disattenuation effects observed within SEM, the Colom et al. (2004) and Gustafsson (1984) data were re-analysed. It was hypothesized that the isomorphic higher-order loadings reported in Colom et al. and Gustafsson would be the result of substantial disattenuation effects due to unacceptably low levels of corresponding composite score reliability associated with the lower-order factors.

Method

Correlation matrices and measures

All analyses were based on the three correlation matrices reported in Colom et al. (2004) and the single correlation matrix reported in Gustafsson (1984). The Colom et al. correlation matrices were converted into covariance matrices based on the standard deviation information reported in Table 1 of Colom et al. (2004). As reported in Colom et al., 12 cognitive ability tests were administered to the first sample (N = 198) and 15 tests were administered to the second (N = 203) and third samples (N = 193). Sample 1 was based on undergraduates and samples 2 and 3 were based on US Air Force recruits. For further details, see Colom et al.

In the case of Gustafsson (1984), the data were reported to be based on a sample of 981 sixth grade children. No standard deviation information was reported, which precluded the possibility of converting the data into a covariance matrix. A total of 20 cognitive ability variables were included in the Gustafsson correlation matrix. Further details can be found in Gustafsson (1984).

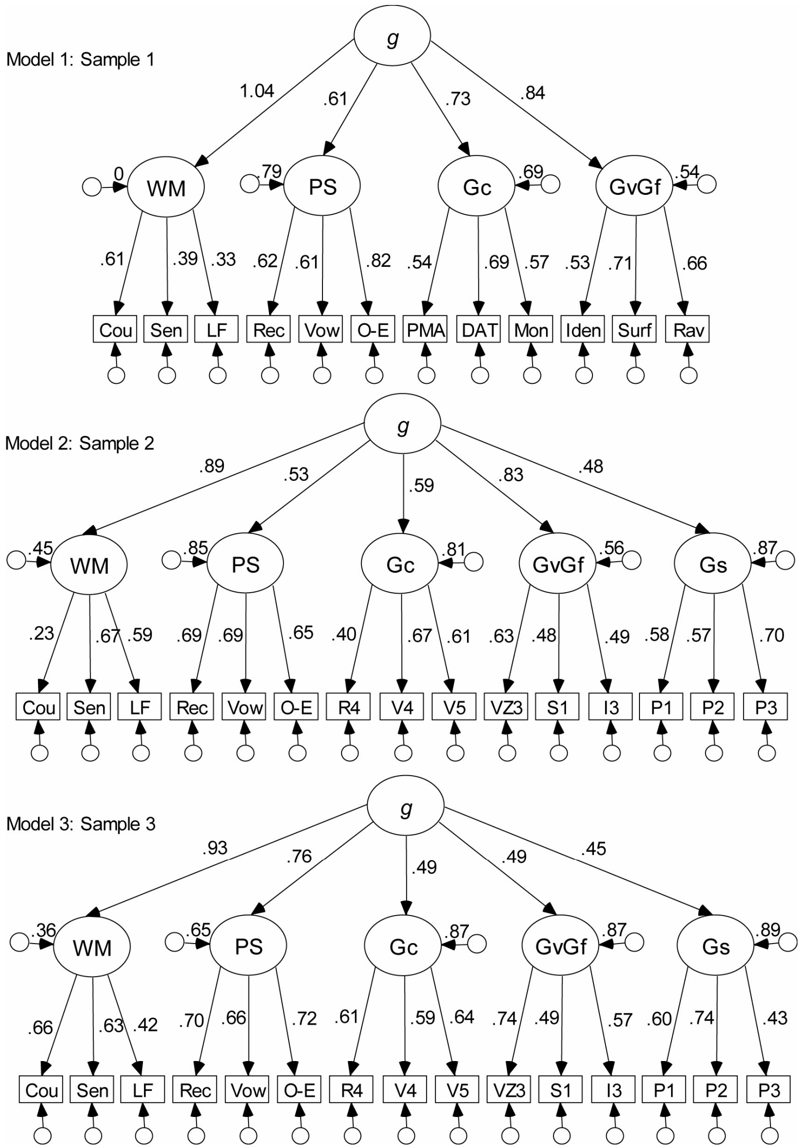

Figure 3:

Colom et al. (2004) higher-order models and standardized parameter estimates; Cou = Counter; Sen = Sentence Verification; LF = Line Formation; Rec = Rectangle-Triangle; Vow = Vowel-Consonant; O-E = Odd-Even; PMA = PMA-V; DAT = DAT-VR; Mon = Monedas; Iden = Identical Figures = Surf = Surface Development; Rav = Raven; R4 = Necessary Arithmetic Operations; V4 = Advanced Vocabulary; V5 = Vocabulary; VZ3 = Surface Development; S1 = Card Rotations; I3 = Figure Classifications; P1 = Finding A's; P2 = Number Comparison; P3 = Identical Pictures.

Table 1:
Generated correlation matrices consistent with adequate (below diagonal) and inadequate (above diagonal) levels of underlying composite score reliability

	V1	V2	V3	V4	V5	V6	V7	V8	V9
V1	1.0	**.19**	**.19**	.16	.16	.16	.16	.16	.16
V2	**.58**	1.0	**.19**	.16	.16	.16	.16	.16	.16
V3	**.58**	**.58**	1.0	.16	.16	.16	.16	.16	.16
V4	.16	.16	.16	1.0	**.19**	**.19**	.16	.16	.16
V5	.16	.16	.16	**.58**	1.0	**.19**	.16	.16	.16
V6	.16	.16	.16	**.58**	**.58**	1.0	.16	.16	.16
V7	.16	.16	.16	.16	.16	.16	1.0	**.19**	**.19**
V8	.16	.16	.16	.16	.16	.16	**.58**	1.0	**.19**
V9	.16	.16	.16	.16	.16	.16	**.58**	**.58**	1.0

Note. Intra-group correlations are in bold.

Data analytic strategy

Across all three of the Colom et al. samples and the single Gustafsson sample, the internal consistency reliabilities associated with the composite variables corresponding to the first-order factors were estimated using Cronbach's α, omega (Ω) and phantom omega (Ω_{Ph}). While Cronbach's α was not especially relevant to the purpose of this investigation, it was reported for the purposes of comparison, as most researchers would be expected to be familiar with α.[5] The Ω estimates were derived from the standardized factor loadings and errors associated with the first-order factors within the higher-order models. Finally, the internal consistency reliability estimates based on the phantom variable modeling approach (Ω_{Ph}) were obtained by including phantom composite variables corresponding to each lower-order factor (i.e., in the same manner as depicted in Figure 1, Model 1; see Fan, 2003b, for a full description of the method). The implied correlation between each phantom variable and corresponding latent variable was squared to derive the Ω_{Ph} estimates.[6]

Although the higher-order model solutions reported in Colom et al. (2004) do not appear to be associated with any errors, the higher-order models were re-tested and reported in this investigation to facilitate an understanding of the results. In the case of Gustafsson's (1984) higher-order model, a modified version which consisted of only two-orders (rather than three-orders) was evaluated for several technical reasons.

First, one of the first-order factors (i.e., Numerical Achievement; see Model 1, Figure 2) was defined by only one observed variable (i.e., Ma. Ach.) with its error variance constrained to zero. This procedure has been argued to be an invalid practice, because most any measured variable would be associated with some level of imperfect reliability (Joreskog &

[5] Again, it is not the Cronbach's α of the individual subtest scores that is referred to here; rather, it is the Cronbach's α associated with the composite scores based on the sum of the subtests used to define a particular lower-order factor.

[6] Note that in order to obtain the appropriate standardized implied correlations, it was necessary to specify all observed variable variances to 1.0 within the SPSS correlation matrix.

Sorbom, 1982). Secondly, Gustafsson's (1984) model also included eight first-order factors defined by only two observed variables, four of which consisted of different halves of the same test (i.e., odd-even split-halves). Given the recommendation that one should define latent variables by a minimum of three or preferably four indicators (Mulaik & Millsap, 2000), it could be argued that only three (i.e., 10%) of Gustafsson's (1984) first-order factors were defined adequately. Although a latent variable can technically be created with two indicators (Bollen, 1989), the effect of specifying a first-order factor based on an odd-even split of the items of a single subtest should be regarded critically, in comparison to creating a two-indicator latent variable defined by two separate subtests. That is, the effect of specifying a two-indicator latent variable based on an odd-even split of the items of a single test is to simply partition error variance from true score variance within one specific cognitive ability test (i.e., method), not a cognitive ability factor, per se.

Consider that the mean inter-correlation between the indicators within the Gustafsson (1984) correlation matrix was .33, while the mean inter-correlation between the odd-even split indictors was .71. It is plausible to suggest that the inclusion of correlations of this magnitude into a substantive portion of a CFA model would be expected to distort the results to some degree. A CFA multitrait-multimethod strategy could model the variance shared by methods into method factors (or correlated residuals) and the substantive covariance these indicators may share with other indicators (from different methods or subtests) into substantive factors (see Conway, 2002, for a non-technical review). Such a model would be expected to be more justifiable than the model endorsed by Gustafsson (1984), although there is probably not a compelling case to suggest that such a model would lead to substantially different higher-order factor loadings or underlying composite score reliabilities.[7]

Nonetheless, with respect to the Gustafsson data, a modified higher-order model which consisted of only two-orders was endorsed in this investigation for two reasons: (1) technically, it was considered a more justifiable model; and (2) it allowed for a straightforward application of the Ω equation described above. The modified higher-order model is depicted in Figure 2 (Model 2) and was associated with adequate levels of model close-fit ($\chi^2(161, N = 981) = 575.15$, $p < .01$, CFI = .958, TLI = .951, RMSEA = .051, SRMR = .041; Schermelleh-Engel, Moosbrugger, & Muller, 2003). Further, the higher-order factor loadings were very comparable to the Gustafsson endorsed higher-order model (e.g., Gf = 1.06 vs. Gf = 1.05).

Model identification was achieved by constraining one factor loading from each of the first-order latent variables to 1.0, as well as constraining the second-order general factor's variance to 1.0. With respect to the phantom composite variables, each indicator was fixed to 1.0 to reflect an equally weighted composite (see Fan, 2003b). Based on the Ω estimates, the higher-order factor loadings estimated within the Colom et al. and the Gustafsson higher-order models were attenuated using the attenuation procedure described above (i.e., each factor loading was multiplied by its corresponding $\Omega^{1/2}$). Consequently, the disattenuated (λ_g) and attenuated (λ'_g) higher-order factor loadings were then compared. As well, the implied correlations between the phantom variables and the higher-order *g* factor (r_{Ph*g}) were also estimated, as these correlations would be expected to reflect the unattenuated association

[7] In fact, the difference in the Ω_{Ph} estimates associated with the Gf second-order factor (Gustafsson model) and the Gf first-order factor (modified model) amounted to only .01, which suggested that the Gustafsson and modified Gustafsson model were highly comparable with respect to reliability.

between the group-level equally weighted composites and *g*. Thus, the λ'_g and r_{Ph*g} estimates were expected to correspond to each other very closely.

Finally, the mean intra-group and extra-group correlations across all three of the Colom et al. samples and the Gustafsson sample were calculated and reported to help understand the effects from a less technical perspective. That is, a lower-order factor with a large *g* loading would be expected to be based on subtests associated with relatively large extra-group correlations, in comparison to other subtests, rather than relatively small intra-group correlations.

All analyses were based on Maximum Likelihood Estimation via AMOS 5.0. In the case of Gustafsson (1984), only a correlation matrix was available rather than a covariance matrix. Although it is well established that CFA based maximum likelihood estimation (MLE) can pose some problems when analysing a correlation matrix (Cudek, 1989), these problems are largely restricted to multi-group factor analyses and/or models that place constraints on one or more of the parameter estimates (other than those constraints used to identify the model). In fact, one of the primary advantages associated with MLE is that it is scale invariant and scale free (Bollen, 1989). It will be noted, however, that despite its scale invariance/freeness, the standard errors associated with a MLE solution derived from a correlated matrix would not be expected to be accurate (Cudek, 1989). Consequently, to estimate the standard errors associated with the standardized parameter estimates, a Monte Carlo simulation of 2000 random samples of data based on the elements within the inputted correlation matrix was performed with the Monte Carlo utility within AMOS 5.0. Thus, the CFA results reported in this investigation were considered to be accurate, despite the fact that the Gustafsson re-analyses were based on a correlation matrix.

As the purpose of this investigation was to estimate composite score reliability and the corresponding attenuated factor loadings of previously published higher-order models, an evaluation of model-fit was not considered particularly relevant. Consequently, no model-fit indexes and/or criteria were reported in the results section.

Results

Reliability estimates

Table 2 includes the internal consistency reliability estimates associated with all of the first-order latent variable corresponding composites based on the Colom et al. data. It can be observed that the Ω reliability estimates associated with the WM equally weighted composite scores were .44, .50 and .61 for samples one, two, and three respectively, which corresponded very closely to the Ω_{Ph} estimates (maximum discrepancy = .02). The Ω estimates associated with the PS, Gc, GvGf, and Gs equally weighted composites were higher, but none were above .75.

In the case of the Gustafsson (1984) data, the Ω estimates were .76, .62, and .84 for Gv, Gf, and Gc, respectively, which, again, corresponded closely to the Ω_{Ph} estimates (see Table 3). Thus, the equally weighted WM composites from the Colom et al. data and the equally weighted Gf composite from the Gustafsson data were associated with very low levels of internal consistency reliability.

Table 2:
Summary of internal consistency reliability estimates (α , Ω and ΩPh), attenuated higher-order loadings (λ'g), and implied correlations between phantom composites and g (rPh*g) based on Colom et al.'s data

	Sample 1					Sample 2					Sample 3					Combined				
	α	Ω	Ω_{Ph}	λ'_g	r_{Ph*g}	α	Ω	Ω_{Ph}	λ'_g	r_{Ph*g}	α	Ω	Ω_{Ph}	λ'_g	r_{Ph*g}	α	Ω	Ω_{Ph}	λ'_g	r_{Ph*g}
WM	.41	.44	.43	.69	.68	.45	.50	.50	.63	.63	.59	.61	.59	.72	.71	.48	.52	.51	.68	.67
PS	.72	.73	.72	.52	.52	.72	.71	.72	.45	.45	.73	.73	.73	.65	.65	.72	.72	.72	.54	.54
Gc	.61	.63	.63	.58	.58	.55	.58	.59	.45	.45	.64	.64	.64	.40	.39	.60	.62	.62	.48	.47
GvGf	.66	.67	.67	.69	.69	.53	.55	.54	.61	.61	.61	.63	.63	.39	.39	.60	.62	.61	.56	.56
Gs	--	--	--	--	--	.65	.65	.65	.39	.39	.60	.63	.62	.36	.35	.63	.64	.64	.38	.37

Note. α = Cronbach's alpha; Ω = omega; Ω_{Ph} = phantom omega; λ'_g = higher-order factor loading attenuated via Ω ; r_{Ph*g} = correlation between phantom variable and higher-order *g* factor.

Table 3:
Summary of internal consistency reliability estimates (α , Ω and Ω_{Ph}), attenuated higher-order loadings (λ'_g), and implied correlations between phantom composites and g (r_{Ph*g}) based on Gustafsson's data.

	α	Ω	Ω_{Ph}	λ'_g	r_{Ph*g}
Gv	.86	.76	.79	.67	.69
Gf	.74	.62	.64	.84	.85
Gc	.89	.84	.86	.73	.73

Note. α = Cronbach's alpha; Ω = omega; Ω_{Ph} = phantom omega; λ'_g = higher-order factor loading attenuated via Ω ; r_{Ph*g} = correlation between phantom variable and higher-order *g* factor.

For the purposes of comparison, it will be noted that the corresponding Cronbach's α estimates were roughly comparable to the Ω and Ω_{Ph} estimates (see Tables 2 and 3), with the exception of the Gf composite *α* of .74 within the Gustafsson (1984) data. The discrepancy in the *α* versus Ω and Ω_{Ph} estimates in the Gustafsson Gf case was due to the fact that Cronbach's *α* can not take into consideration the large amount of variance associated with the correlated errors between the indicators derived from the same subtests used to define the Gf latent variable (see correlated error estimates in Figure 2, Model 2).

Attenuated higher-order loadings

With respect to Colom et al.'s sample one data, where the WM Ω estimate was equal to .44, the WM higher-order loading estimate of 1.04 was attenuated by multiplying it by the square root of .44, which yielded an attenuated factor loading (λ'_g) of .69 (i.e., $1.04*.44^{1/2} = 1.04*.66 = .69$). The same procedure was applied to all second-order factor loadings, across all three of the Colom et al. samples. Table 2 includes a summary of all the attenuated (λ'_g) higher-order factor loadings associated with all three of the Colom et al. higher-order model solutions. It can be observed that the mean attenuated WM factor loading was equal to .68, which was very substantially lower than the mean disattenuated WM factor loading of .95. The implied correlations between the phantom composites and the *g* factor (r_{Ph*g}; see Table 2) were all very comparable to the attenuated higher-order loadings (maximum discrepancy = .01), which supported further the hypothesis that the higher-order loadings of approximately 1.0 reported by Colom et al. were due to the substantial disattenuation effects achieved by modeling latent variables associated with low levels of latent variable corresponding composite score reliabilities.

With respect to the Gustafsson data, it can be observed that the Gf higher-order loading of 1.06 was attenuated to .84 (based on the Ω estimate of .62), which was comparable to the corresponding r_{Ph*g} correlation of .85. The remaining higher-order loading estimates are displayed in Table 3.

Intra-group versus extra-group correlations

The Colom et al. and the Gustafsson correlation matrices were also examined with respect to mean intra-group and extra-group correlations. As can be seen in Table 4, in the case of the Colom et al. data (sample one), the mean intra-group correlation associated with WM subtests was the lowest (.22) amongst all five subtests groupings. In contrast, the mean WM extra-group correlation (.18) was quite comparable to the other mean extra-group correlations. In the third column of the Table devoted to the Colom et al. data, the differences between the respective mean intra- and extra-group correlations are reported. It can be observed that the WM subtests were associated with only a small mean intra-group/extra-group correlation difference (i.e., .04). In the case of Gustafsson (1984), the Gf type subtests were also associated with the lowest mean intra-group correlation. In fact, the difference was quite substantial (.26 vs. .43 and .61). Further, the mean extra-group correlations associated with Gustafsson's (1984) data were all very comparable (i.e., range .28 to .29). As can be seen in Table 4, the Gf subtests were actually associated with a somewhat numerically higher level of extra-group correlation than intra-group correlation (i.e., Δr_{dif} = -.02).[8]

Table 4:
Intra-group and extra-group correlations associated with the Colom et al. (2004) and Gustafsson (1984) correlation matrices analysed in this investigation.

	Colom et al. (2004)				Gustafsson (1984)		
	Intra-*r*	Extra-*r*	Δr_{dif}		Intra-*r*	Extra-*r*	Δr_{dif}
WM	.22	.18	.04	Gv	.43	.29	.14
PS	.49	.15	.34	Gf	.26	.28	-.02
Gc	.29	.15	.14	Gc	.61	.29	.32
GvGf	.27	.17	.10	$\bar{X}$	.43	.29	
Gs	.38	.12	.26				
$\bar{X}$	.33	.15					

Note. Intra-*r* = mean intra-group correlation; Extra-*r* = mean extra-group correlation; Δr_{dif} = difference between mean intra-group correlation and mean extra-group correlation.

Discussion

The results reported in this investigation demonstrated that the lower-order group-factors reported by Colom et al. (2004) and Gustafsson (1984) to be isomorphic with *g* were associated with very low levels of latent variable corresponding composite score internal consistency reliability. Further, the observations of isomorphic (or near isomorphic) higher-order factor loadings were demonstrated to be achieved to a large degree by the disattenuation effects observed within a latent variable modeling framework. When the Colom et al. (2004)

[8] See footnote 4 for some discussion relevant to negative residual variances and the unacceptability of the Gustafsson higher-order model.

and Gustafsson (1984) data were examined from the mean intra/extra-group correlation perspective, both the WM intra-group correlations (Colom) and the Gf intra-group correlations (Gustafsson) were found to be relatively small, while the WM and Gf extra-group correlations were found to be quite comparable in size to that of the other subtest groupings.

Internal consistency reliability and SEM

As reported in Perterson (1994), Nunnally (1967; 1978) is the most frequently cited reference in support of a minimum reliability criterion of .70. However, as pointed out by Lance, Butts, Michels (2006), Nunnally (1978) recommended .70 only for early stage research. For basic research, Nunnally (1978) recommended a criterion of .80. Given the long history of factor analytic research in the area of intelligence (see Jensen, 1998), the data reanalysed in this investigation should be characterized as basic research, not early stage research. Consequently, a minimum level of reliability of .80 should be probably acknowledged, although it should be appreciated that any demarcation criterion of this nature will be to a large degree arbitrary and/or subjective. In both the Colom et al. and Gustafsson cases, the lower-order group-factors claimed to be isomorphic with *g* were demonstrated to be associated with latent variable corresponding composite score internal consistency reliabilities well below .80. In the case of Colom et al., the WM corresponding composite score reliabilities averaged .52 (Ω) and .51 (Ω_{Ph}) across all three samples: substantially lower than even the early stage research recommendation of .70 (Nunnally, 1978). While the Gustafsson data were generally more respectable, the Gf composite score reliability of Ω = .62 was nonetheless below the .70 recommendation for early stage research.

As demonstrated with the attenuation procedure, the higher-order factor loadings associated with the Colom et al. and Gustafsson higher-order models were reduced very substantially because of the low levels of internal consistency reliability. Consequently, the Colom et al. and Gustafsson observations of isomorphic or near isomorphic higher-order loadings are argued here to have occurred largely due to the very substantial disattenuation effects applied within a SEM framework, when the latent variable corresponding composites are associated with low levels of internal consistency reliability. It should be emphasized that the disattenuation effects observed within SEM are not unique to higher-order modeling. Rather, they would be expected to be observed within any approach based on latent variables, such as oblique-factor modeling (Fan, 2003a). Thus, for example, the SEM results reported in Kyllonen & Christal (1990), where working memory and reasoning ability were suggested to be effectively isomorphic, should be interpreted within the context of the SEM disattenuation effects described in this paper.

The large magnitude of the difference between the attenuated and the disattenuated factor loadings reported in this investigation raises the question as to whether the disattenuation procedure should be considered acceptable. Burt (1940) noted that some well-known researchers disattenuated correlations, while others did not. Burt (1940) clearly did not agree with the practice of disattenuating correlations and concluded: “In my view, if the reliability coefficient is so low that the raw correlation requires correction, that is a reason, not for factoring corrected correlations but for improving the experimental technique.” (Burt, 1940, p. 287). Burt’s (1940) view may or may not be shared by many contemporary researchers. Ultimately, the issue of whether the disattenuation procedure should or should not be considered appropriate in factor analytic research, either in the context of inter-subtest correlations or inter-latent variable correlations, is not expected to be resolved here. Consequently, it is recommended that, at the very least, researchers should be required to report the internal consistency reliability associated with the composite scores which correspond to their latent variables, as recommended previously by Cohen et al. (1990). Further, it would also be beneficial if researchers reported both the disattenuated and the attenuated factor loadings within a CFA model, in much the same way a researcher would be expected to report both the attenuated and disattenuated correlations within the more traditional Classical Test Theory approach to disattenuating a bivariate correlation between two observed variables (Muchinsky, 1996).

Intra-group and extra-group correlations

It is perhaps justifiable to recommend that any group of subtests purported to be identical with *g* should exhibit substantial levels of extra-group association, in comparison to other subtest groupings. In the case of both Colom et al. (2004) and Gustafsson (1984), the observation of a higher-order group-factor loading which suggested identity with *g* is argued here to have been achieved by the clustering of a group of subtests with poor intra-group correlations, rather than strong extra-group correlations. Thus, what the Gf subtests from the Gustafsson (1984) study and the WM subtests from the Colom et al. (2004) study have in common is that they were associated with the smallest differences between the respective intra- and extra-group correlations, not because they exhibited very strong extra-group correlations, but because they exhibited low levels of intra-group correlations.

Concluding considerations

In a previous investigation, Gignac (2006) found that well-fitting CFA models of several popular intelligence assessment batteries tended to be associated with general factors with largely equal factor saturation levels across all specific abilities. Thus, the results of this investigation coincide well with Gignac (2006), which suggests further that no single subtest or lower-order factor can totally account for the *g* factor within a properly specified model. Even if it were demonstrated that a lower-order group-factor were associated with *g* at a level to suggest isomorphism, it would nonetheless remain to be established whether that group-factor could predict theoretically relevant external criteria to the same degree as the *g* factor. That is, suggestive evidence of isomorphism based solely upon CFA (i.e., in the absence of predictive validity) should not be considered especially compelling, given the problem of equivalent models in SEM research (MacCallum, Wegener, Uchino, & Fabrigar, 1993). Ultimately, predictive validity studies should be relied upon to help support claims that a particular lower-order factor encompasses all of the validity associated with the *g* factor. Thus, while CFA studies should have a place in assessing hypotheses relevant to whether a lower-order domain is identical to *g*, it should not be expected to solve the problem of testing the hypothesis of 'identity with *g*' in a totally convincing way.

Finally, it will be noted that the Gustafsson higher-order model was associated with a statistically significant negative residual variance (see footnote 3), which should be viewed as evidence that the model was unacceptable. Further, the Colom et al. higher-order models were not associated with acceptable levels of model-fit based on conventional standards (e.g., Hu & Bentler, 1999). Thus, there remains the possibility that alternative models could be found to better represent the Gustafsson and Colom et al. correlation matrices. One particular model neglected by both Colom et al. (2004) and Gustafsson (1984) is the direct hierarchical model (a.k.a., bi-factor model or nested factor model) endorsed by Jensen and Weng (1994). Empirical support for the direct hierarchical model over a comparable higher-order model would have implications for contentions that a lower-order group-factor is isomorphic with *g*, because typical direct hierarchical models specify the association between group-factors and *g* to zero. Consequently, a potentially useful extension of the present investigation would consist of an evaluation of the plausibility of the direct hierarchical model based on the Gustafsson (1984) and Colom et al. (2004) data.

References

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.

Bollen, K.A. (1989). Structural equations with latent variables. New York: Wiley.

Burt, C. (1940). *The factors of the mind*. London: University of London Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge, UK: Cambridge University Press.

Chen, F., K.A. Bollen, P. Paxton, P. Curran, and J. Kirby. 2001. Improper Solutions in Structural Equation Models: Causes, Consequences, and Strategies. *Sociological Methods and Research*, *29*, 468-508.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal modeling. *Applied Psychological Measurement*, *14*, 183-196.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence*, *32*, 277-296.

Conway, J. M. (2002). Method variance and method bias in I/O Psychology. In S. G. Rogelberg (Eds.), *Handbook of Research Methods in Industrial and Organizational Psychology* (pp. 344-365). Oxford, UK: Blackwell Publishers.

Crocker, L., & Algina, J. (1986*). Introduction to classical test theory*. New York: Holt, Rinehart and Winston.

Cudek, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, *2*, 317-327.

Fan, X. (2003a). Two approaches for correcting correlation attenuation caused by measurement error: Implications for research practice. *Educational and Psychological Measurement*, *63*, 915-930.

Fan, X. (2003b). Using commonly available software for bootstrapping in both substantive and measurement analyses. *Educational and Psychological Measurement*, *63*(1), 24-50.

Gignac, G. E. (2006). Evaluating subtest 'g' saturation levels via the Single Trait-Correlated Uniqueness (STCU) SEM approach: Evidence in favour of crystallized subtests as the best indicators of 'g'. *Intelligence*, *34*, 29-46.

Gustafsson, J-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179-203.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit,& D. Sorebom (Eds.), *Structural equation modeling: Present and future—A festschrift in honor of Karl Jöreskog* (pp. 195-216). Lincolnwood, IL: Scientific Software International.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Jensen, A. R. (1998). The *g factor: The science of mental ability*. Westport: Praeger.

Jensen, A. R., & Weng, L. (1994). What is a good *g*? *Intelligence*, 18, 231-258.

Joreskog, K.G. & Sorbom, D. (1982). Recent developments in structural equation modeling. Journal of Marketing Research, *19*, 404-416.

Kyllonen, P., & Christal, R. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, *14*, 389-433.

Lance, C. E., Butts, M. M., Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, *9*, 202-220.

Lykken, D. T. (2005). Mental energy. *Intelligence*, *33*(4), 331-335.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R.(1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, *114*, 185-199.

McDonald, R. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum Associates.

Muchinsky, P. M. (1996). The correction for attenuation. *Educational and Psychological Measurement*, *56*(1), 63-75.

Mulaik, S. A., & Millsap, R. E. (2000). Doing the four-step right. *Structural Equation Modeling*, *7*, 36-73.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory*. New York: McGraw-Hill

Peterson, R. A. (1994). A meta-analysis of Cronbach's alpha. *Journal of Consumer Research*, *21*, 381-391.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*, 173-184.

Raykov, T. (2001). Bias in coefficient alpha for fixed congeneric measures with correlated errors. *Applied psychological measurement, 25*(1), 69-76.

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*, 23-74.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72-101.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Sternberg, R. J., & Grigorenko, E. L. (2002). *The General Factor of Intelligence: How general is it?* Mahwah, NJ: Lawrence Erlbaum Associates.