Intelligence 55 (2016) 57–68

Contents lists available at ScienceDirect

Intelligence

The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better

Gilles E. Gignac

School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia, 6009, Australia

ARTICLE INFO

Article history:
Received 14 July 2015
Received in revised form 14 December 2015
Accepted 18 January 2016
Available online xxxx

Keywords:
Higher-order model
Bifactor model
Model-fit

ABSTRACT

Recent model-fit simulation work relevant to the higher-order model and the bifactor model has neglected to focus upon the key statistical distinction between the two models: the proportionality constraint imposed by the higher-order model. Consequently, the purpose of this investigation was to demonstrate, in an accessible manner, the nature of the proportionality constraint. Secondly, it was hypothesized that the degree to which the proportionality constraint is violated will correlate perfectly with the degree to which the bifactor model fits better than the higher-order model. Finally, it was hypothesized that the bifactor model will not necessarily be favoured based on fit-indices which incorporate a penalty for model complexity (TLI, AIC, BIC). To test the hypothesis, a series of 12 correlation matrices were simulated such that the proportionality constraint violation increased progressively. The magnitude of the violation was observed to correlate essentially perfectly with the degree of model fit difference between the higher-order and the bifactor model. At moderate to high levels of proportionality constraint violation, the TLI favoured the bifactor model. Importantly, however, at low levels of proportionality constraint violation, the TLI favoured the higher-order consistently across all three sample sizes. The results associated with the chi-square difference test, AIC, and BIC were much more greatly affected by sample size. The results are interpreted to suggest that the proportionality constraint needs to be taken into consideration when comparing the higher-order model and the bifactor model. Additionally, fit indices which incorporate a penalty for model complexity do not necessarily favour the bifactor model. Consequently, researchers are encouraged to consider the bifactor model in appropriate research scenarios.

The majority (57.6%) of recent confirmatory factor analysis papers in the area of intelligence have used a higher-order modeling strategy (Reeve & Blacksmith, 2009). However, a non-negligible percentage of papers (22.7%) have used a bifactor modeling strategy (Reeve & Blacksmith, 2009). Reise (2012) suggested that the bifactor model is gaining in popularity. In the area of intelligence, it may be suggested that there are strong views about which model, higher-order or bifactor, is to be preferred from a theoretical perspective (Gignac, 2008). Under what circumstances one model may be expected to fit better than the other also remains a contentious and relatively unclear issue (Morgan, Hodge, Wells, & Watkins, 2015; Murray & Johnson, 2013). Consequently, the purpose of this investigation was to focus upon a relatively unknown and poorly understood key distinction between the two models: the proportionality constraint imposed by the higher-order model (Yung, McLeod, & Thissen, 1999). Specifically, the nature of the proportionality constraint will be described in a relatively accessible manner. Furthermore, the hypothesis that the bifactor model will be observed to fit better than the higher-order model to the extent that the data are inconsistent with the proportionality constraint will be tested.

1. Higher-order models and bifactor models

Prior to describing the nature of the proportionality constraint imposed by the higher-order model, it is important to understand the various terms associated with both the higher-order model and the bifactor model, as several of these terms are used to calculate ratios relevant to the proportionality constraint. Within the context of confirmatory factor analysis, intelligence researchers have, essentially, two options to represent a general factor in the presence of group-level factors (Rindskopf & Rose, 1988): (1) the higher-order model; and (2) the bifactor model. As can be seen in Fig. 1 (Model 1), in the typical higher-order model, each indicator/subtest (a1, a2, a2, b1, b2, b3, c1, c2, c3) is specified to load onto a first-order factor (FA, FB, or FC). Such first-order factors are known as group-level factors. The higher-order model depicted in Fig. 1 includes three group-level factors. In practice, higher-order models which include four or more group-level factors can also be specified. A higher-order model defined by only two group-level factors requires an equality constraint for identification purposes (Zinbarg, Revelle, & Yovel, 2007).

In the typical higher-order model, each group-level factor is specified to load onto the second-order general factor (g). Theoretically, the higher-order model implies that the shared variance between the

E-mail address: gilles.gignac@uwa.edu.au.

http://dx.doi.org/10.1016/j.intell.2016.01.006
0160-2896/

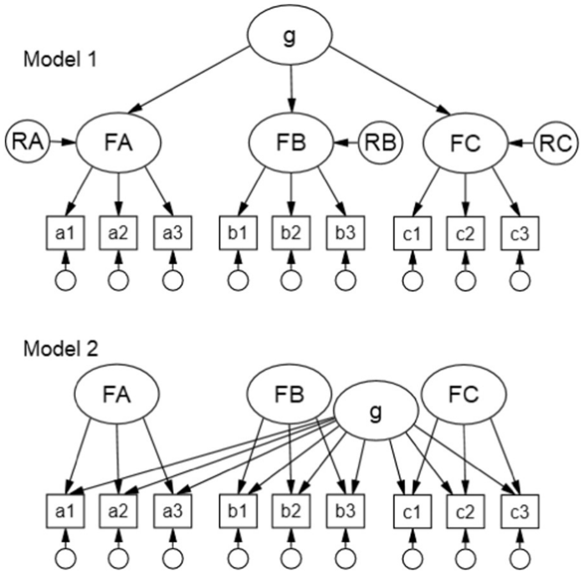

Fig. 1. The two competing models tested in this investigation: Model 1 = higher-order model; Model 2 = bifactor model.

group-level factors is caused by *g*: A key feature of the model for those who endorse it (Keith, 2005; Weiss, Keith, Zhu, & Chen, 2013). The higher-order model of intelligence also implies that the positive manifold (positive correlation across all subtest types) is caused by the shared variance between the group-level factors. Thus, only indirectly does the general factor in the higher-order model influence the indicators or subtests (Gignac, 2008; McDonald, 1999). Higher-order models with three or more orders are possible (e.g., Gustafsson, 1984); however, typically, second-order factor models are specified in practice.

Often neglected in the literature is the fact that the higher-order model has orthogonal latent variable terms that can be interpreted substantively, as well as used to predict various dependent variables (e.g., Crawford, Deary, Allan, & Gustafsson, 1998). These terms have been referred to as residualized group factors (Jensen & Weng, 1994) or lower-order factor residuals (Gignac, 2008). As can be seen in Fig. 1 (Model 1), the depicted higher-order model is associated with three lower-order factor residual terms: RA, RB, and RC. In the case of the FA first-order factor, for example, RA represents the shared variance between the a1, a2, and 3 subtests that is not accounted for by the general factor. In the area of intelligence, for example, RA might represent crystallised intelligence (g_c), independent of *g*. Within the context of intelligence, the perspective of this paper is that a lower-order factor within a higher-order model is considered an amalgamation of *g* and a specific or broad ability (e.g., crystallised intelligence, fluid intelligence, etc.), rather than a clear representation of a specific or broad ability. By contrast, the lower-order factor residual term is considered an unambiguous representation of a specific or broad ability, if it is a source of variance that is found to be statistically significant. Stated alternatively, the lower-order factor residual term is considered the broad ability, cleansed of *g*.

Thus, in a typical higher-order model, the shared variance between subtests is hypothesized to arise due to unique effects associated with the general factor, as well as unique effects associated with a broad ability, which is represented by one of the lower-order factor residual terms (McDonald, 1999). From this perspective, the first-order factor loadings are arguably difficult to interpret, as they are imbued with two sources of variance: latent general factor variance and latent group-level factor residual variance. For this reason, some researchers recommend decomposing the first-order factor loadings into their respective unique parts (Gignac, 2014a; Humphreys, 1962). Such a procedure is known as a Schmid–Leiman decomposition (or Schmid–Leiman transformation; Schmid & Leiman, 1957).[1] The Schmid–Leiman decomposition procedure is relatively simple to execute and very much akin to the procedure of calculating indirect effects in a typical mediation analysis. Specifically, in a mediation analysis, the relevant regression path coefficients are multiplied together (i.e., independent variable and intervening variable coefficient ⁎ intervening variable and dependent variable coefficient) to estimate the indirect effect between an independent variable and a dependent variable (Alwin & Hauser, 1975). With respect to the higher-order model, the general factor is the independent variable, the subtests are the dependent variables, and the first-order factors are the intervening variables, in this context (Gignac, 2008).

Typically, there are two indirect effects calculated from a higher-order model solution via the Schmid–Leiman decomposition procedure: (1) the indirect effect of the general factor on each subtest; and (2) the indirect effect of a residual lower-order factor term on each subtest associated with a particular group-level factor. In practical terms, to execute a Schmid–Leiman decomposition, each subtests' first-order factor loading is multiplied twice: (1) once by the corresponding second-order general factor loading; and (2) once by the corresponding group-level factor residual term coefficient (Gignac, 2007; Jensen & Weng, 1994). The first product represents the indirect association between the general factor and a particular subtest. The second product represents the indirect association between a unique group-level factor (residual) term and a particular subtest. It is important to note that, in the higher-order model, there are generally no direct effects between *g* and the subtests. In this sense, a well-fitting typical higher-order model may be viewed as evidence to suggest that the association between *g* and the subtests is mediated fully by the group-level factors (Gignac, 2008). That is, mediation is implied to occur through the group-level factors, as the direct effects between *g* and the subtests are constrained to zero.

In contrast to the higher-order model is the bifactor model (Holzinger & Swineford, 1937). The bifactor model has also been referred to as a nested factor model (Gustafsson & Balke, 1993), a general hierarchical factor model (Yung et al., 1999), and a direct hierarchical model (Gignac, 2008). As can be seen in Fig. 1 (Model 2), all of the latent variables in the bifactor model are defined by the subtests directly. Consequently, the general factor is a first-order factor in the bifactor model. The lower-order factor residuals in the higher-order model (RA, RB, and RC) are specified as orthogonal group-level first-order latent variables in the bifactor model (FA, FB, and FC), each defined directly by the relevant subtests. Thus, substantively, FA, FB, and FC within the higher-order model have more in common with RA, RB, and RC in the higher-order model. Unlike the higher-order model, there is no need to decompose the bifactor model solution to obtain the unique associations between each subtest and each latent variable term. However, a trade-off is that it is not possible to estimate the uncleansed effect of the broad factors on the subtests in the bifactor model. Yung et al. (1999) demonstrated that the higher-order model is nested within the bifactor model, consequently, the two competing models can be tested for differences in model fit via the chi-square difference test (Yung et al., 1999). As the higher-order model is associated with fewer parameters than a corresponding bifactor model, the higher-order model will never be found to fit better than a corresponding bifactor model from an implied model chi-square statistic perspective (Yung et al., 1999). However, whether the higher-order model can be found to fit better than the bifactor under certain conditions based on fit indexes which incorporate a penalty for model complexity (e.g., TLI, BIC, AIC) remains to be determined conclusively.

[1] Yung et al. (1999) refer to the Schmid–Leiman transformation as a "Schmid–Leiman hierarchical factor model". However, it may be somewhat misleading to refer to the Schmid–Leiman procedure as a model. The higher-order model is a model, for example, as it can be specified and tested statistically for plausibility. The Schmid–Leiman transformation, however, cannot be specified and/or tested for plausibility. It is simply used to calculate indirect effects, as described further below.

2. Proportionality constraint

As there are only direct effects in the bifactor model, all of the loadings are estimated freely. By contrast, in the higher-order model, the first-order loadings are not estimated freely. Instead, a "hidden" (Yung et al., 1999, p. 115) constraint is imposed by the higher-order model. The constraint is known as the proportionality constraint (Schmiedek & Li, 2004; Yung et al., 1999).

Essentially, the proportionality constraint restricts the partitioning of general and specific factor variance associated with indicators within a group-level factor to be perfectly proportional. Thus, the general factor (*g*) and specific factor (*s*) loading ratios (*g*/*s*) are constrained to equality across all subtests within a particular group-level factor in a higher-order model (Yung et al., 1999). Consequently, no subtest within a group-level factor can be observed to be associated relatively strongly with g, unless it is also observed to be associated relatively strongly with the corresponding first-order group level specific factor. The bifactor model, by contrast, does not impose any such constraint (Yung et al., 1999). Consequently, the *g*/*s* ratios can vary across all indicators in the bifactor model (Beaujean, Parkin, & Parker, 2014).

Based on a series of empirical analyses of previously published intelligence battery correlation matrices, Gignac (2008) found that the bifactor model tended to fit better than the higher-order model. Gignac (2008) suggested that the reason the bifactor model tends to fit better than the higher-order model is because researchers are not adept enough to create (or choose) subtests in a manner such that each subtests' *g* related variance would be mediated completely by the first-order factor they were specified to load upon. Arguably, Gignac's (2008) explanation is a rephrasing of the proportionality constraint imposed by the higher-order model described by Yung et al. (1999). Reynolds and Keith (2013) simulated two correlation matrices to demonstrate that the higher-order model (S–L decomposed) and the bifactor model will yield identical solutions when the proportionality constraint is satisfied and different solutions when the proportionality constraint is violated. However, perhaps surprisingly, to-date, violation of the proportionality constraint assumption imposed by the higher-order model has not been investigated systematically via simulation to help understand the effects on model fit.

3. Previous simulation research

Mulaik and Quartetti's (1997) investigation appears to be the first quantitative investigation into the differences between the higher-order and bifactor models. First, they demonstrated that a higher-order model and a bifactor model can be observed to yield the same solutions, once the higher-order model solution is decomposed via the Schmid–Leiman decomposition procedure, if both models fit perfectly. Next, in an attempt to evaluate the differences between the models, Mulaik and Quartetti (1997) "perturbed" (p. 199) the Schmid–Leiman decomposed solution such that some of the loadings were modified by a maximum of |.05|. Thus, the correlation matrix was assumed to be no longer perfectly represented by a higher-order model. A correlation matrix was then derived from the "perturbed" Schmid–Leiman solution. Finally, the competing higher-order and bifactor models were fitted to the modified correlation matrix. Mulaik and Quartetti (1997) reported that the bifactor model fit the data perfectly ($\chi^2 = .00$), and the higher-order model nearly did so, as well, $\chi^2 = 3.59$, $df = 101$. Based on a sample size of 1000, Mulaik and Quartetti (1997) estimated the power of rejecting the null hypothesis in this case at only .18. Based on several other similar simulations, Mulaik and Quartetti (1997) concluded that, although there may be small differences between the higher-order model and the bifactor model, it is likely that typical investigations would lack the statistical power to identify those differences as statistically significant, even with sample sizes as large as 1000.

Mulaik and Quartetti's (1997) investigation may be criticised on the basis that the perturbations of the Schmid–Leiman decomposed solution were applied in a non-systematic fashion. Thus, it is possible that the various increases and decreases in the Schmid–Leiman transformed solution loadings used to create the modified correlation matrices simply cancelled each other out in such a way as to largely maintain the proportionality constraint imposed by the higher-order model. Despite Mulaik and Quartetti's (1997) contention about a lack of statistical power, a number of applied empirical investigations appeared in the literature that reported better fitting bifactor models, in comparison to higher-order models, across a variety of sample sizes (e.g., Bechtel, Verdugo, & de Queiroz Pinheiro, 1999; Gignac, 2005; Gignac, Palmer, Manocha, & Stough, 2005; Mohlman & Zinbarg, 2000; Smith, Parrott, Diener, Hoyle, & Kim, 1999; Watkins & Kush, 2002).

Chen, West, and Sousa (2006) criticised Mulaik and Quartetti's (1997) investigation on the basis that the "perturbations" specified in the correlation matrices were so trivial in magnitude (i.e., factor loading differences of |.05| or less) that it should have come as no surprise that the chi-square difference test was underpowered. Consequently, Chen et al. (2006) conducted a follow-up simulation and found that with moderate disturbances in the Schmid–Leiman decomposed solution, power as high as .99 can be expected in many higher-order model versus bifactor model comparisons, even with very moderate sample sizes ($N = 200$). Although the Chen et al. (2006) helped clarify the issue of power in the context of comparisons between the higher-order model and the bifactor model, they did not focus upon the proportionality constraint imposed by the higher-order model in their simulations. Instead, the modifications applied to the correlation matrices were, again, unsystematic, although larger than those applied by Mulaik and Quartetti (1997).

More recently, Murray and Johnson (2013) conducted a simulation study to investigate whether commonly used fit indexes which incorporate a penalty for model complexity (TLI, AIC, and BIC) can be used validly to select which model is more plausible, the higher-order model or the bifactor model. To this effect, Murray and Johnson (2013) assumed that the higher-order model representation of cognitive abilities was true and, consequently, specified seven data sets that were contended to be consistent with a higher-order model. However, six of the seven data sets included various levels of model complexity. Specifically, low to high correlated residuals and/or low to high cross-loadings were included in the modified correlation matrices. Low correlated residuals were specified at .10 and high correlated residuals were specified at .20. Correspondingly, low cross-loadings were specified at .10 and high cross-loadings were specified at .20. Thus, in the "true" high-order model data set with the highest level of model complexity, there was a total of 10 correlated residuals, six cross loadings, in addition to the residuals and cross-loadings identified in the original test battery solution.

Across all levels of model complexity, Murray and Johnson (2013) found that the bifactor model fitted the data better than the higher-order model, even though the data were specified to be consistent with a higher-order model representation of cognitive abilities. Consequently, Murray and Johnson (2013) concluded that model-fit indexes, even those which incorporate a penalty for model complexity (TLI, AIC, and BIC), are biased in favour of the bifactor model. Thus, the recommended practical differences of TLI > .10 (Gignac, 2007) and/or AIC/BIC < 10 (Raftery, 1995) to indicate practical improvements in model fit should not necessarily be interpreted as evidence in favour of the bifactor model over the higher-order model. Murray and Johnson's (2013) interpretation of their results suggested that the main reason the bifactor model tends to fit better than the higher-order model is because it is better at accounting for "non-substantive sources" (p. 419) of variance than the higher-order model. However, as per Chen et al. (2006); Murray and Johnson (2013) did not take into account the proportionality constraint imposed by the higher-order model. Thus, it is possible that the complexity they added to the simulated correlation matrices had the unintended effect of violating the proportionality constraint imposed by the higher-order model.

Given the relatively small number of simulated data sets examined in Murray and Johnson (2013); Morgan et al. (2015) suggested that the influence of sampling error may be an important factor in the model-fit evaluation of the competing higher-order and bifactor models. Consequently, Morgan et al. (2015) conducted a Monte Carlo simulation across three types of models: the correlated factor model, the higher-order model, and the bifactor model. Furthermore, they examined these three types of models across two levels of factor definition: (1) a four factor model, where two of the factors were defined by three indicators and the other two factors were defined by only two indicators; (2) an additional four factor model, where all four factors were defined by three indicators. Two sample sizes were investigated: 200 and 800. For each model/condition, 1000 replications were performed.

As would be expected, Morgan et al. (2015) found that when the data were simulated to be consistent with a bifactor model structure, the fit indexes (RMSEA, SRMR, CFI, TLI, AIC, BIC) very typically tended to suggest that the bifactor model was the preferred model. However, when the data were specified to be consistent with a higher-order model structure, Morgan et al. (2015) reported much less consistency in the results. In many cases, the fit indexes erroneously suggested that the bifactor model should be favoured. However, it should be noted that Morgan et al. examined a number of fit indexes that incorporate only a minor penalty (e.g., CFI) or no penalty at all (e.g., SRMR) for model complexity (see Marsh, Hau, & Grayson, 2005). Thus, based on such comparisons, the bifactor model would be expected to fit better than the higher-order model in many resamples, as the bifactor model is never associated with a larger implied model chi-square value than the corresponding higher-order model (Yung et al., 1999). For this reason, comparisons between the higher-order model and the bifactor model should be based on indices that incorporate a relatively substantial penalty for model complexity such as TLI, AIC, and BIC (Gignac & Watkins, 2013; Marsh et al., 2005). Irrespective of which fit index is considered, it is very difficult to evaluate the results of Morgan et al. (2015), as the re-sampling of a 1000 samples from the true higher-order model correlation matrix would have resulted in many samples that deviated from the true higher-order model correlation matrix simply by chance. In fact, it is unlikely that any one of the 1000 re-samples corresponded precisely to the original higher-order model population correlation matrix. Some correlation matrices may have been inconsistent with the proportionality constraint imposed by the higher-order model.

Arguably, previous simulation research relevant to the higher-order model and the bifactor model has neglected to focus upon the key statistical distinction between the higher-order model and the bifactor model: the proportionality constraint imposed by the higher-order model (Yung et al., 1999). Consequently, the first purpose of this investigation was to demonstrate, in an accessible manner, the nature of the proportionality constraint. Secondly, the following hypothesis was tested empirically: the degree to which the proportionality constraint is violated will be correlated perfectly with the degree to which the bifactor model fits better than the higher-order model. Finally, the possibility that fit indices which incorporate a penalty for model complexity (e.g., TLI, AIC, BIC) will not necessarily favour the bifactor model will be evaluated across three sample sizes: 500, 1000, and 2000.

4. Method

4.1. Data generation

In order to demonstrate the nature of the proportionality constraint, as well as to test the hypothesis associated with this investigation, a total of 12 correlation matrices were simulated (see Tables S1 to S12). Additionally, the data were generated across three sample size conditions: $N = 500$, $N = 1000$ and $N = 2000$. Each correlation matrix consisted of nine observed variables: a1, a2, a3, b1, b2, b3, c1, c2, and c3. All of the data were simulated to be consistent with the presence of one general factor (*g*) and three group-level specific (*s*) factors (FA, FB, FC).[2]

The first correlation matrix was simulated such that the proportionality constraint was met perfectly (Satisfied 1; see Table S1). Thus, all of the *g*/*s* ratios were expected to be equal within each group-level factor. The remaining 11 correlation matrices were manipulated systematically in such a way that the variability in the *g*/*s* ratios was progressively larger. Thus, in the second correlation matrix (Violated 1; see Table S2), the assumption of proportional variance contributions from each of the a1, a2, and a3 indicators was violated only very minimally. Specifically, the a1 by a2 correlation was increased in magnitude by 5%, and the a2 by a3 correlation was decreased in magnitude by 5%. In the third correlation matrix (Violated 2; see Table S3), the a1 correlations with the b1 to c3 indicators were decreased in magnitude by 5% and the a3 correlations with the b1 to c3 indicators were increased in magnitude by 5%. Thus, in correlation matrices Violated 2 to Violated 3, the a1 indicator's contribution to the group-level FA factor residual was expected to increase (Violated 2) and its contribution to the general factor was expected to decrease (Violated 3). Such a manipulation may be said to be consistent with a violation of the proportionality constraint imposed by the higher-order model. Correspondingly, the a3 indicator's variance contribution to the group-level FA factor was expected to decrease and its contribution to the general factor was expected to increase: again, an effect that may be said to consistent with a violation of the proportionality constraint. A series of such 5% increases and decreases was applied to the a1 and 3 indicators in five steps (correlation matrices Violated 1 to Violated 5; see Tables S2 to S6). Then, the b1 and b3 indicators associated with FB group-level factor were additionally manipulated similarly in three steps (Violated 6 to Violated 8; see Tables S7 to S9). Finally, the c1 and c3 indicators associated with the FC group-level factor were additionally manipulated similarly in three steps (Violated 9 to Violated 11; see Tables S10 to S12). Thus, in the final three correlation matrices (Violated 9 to Violated 11), all three group-level factors were expected to be associated with some within group-level factor variability in the *g*/*s* ratios.

4.2. Data analysis

Corresponding higher-order and a bifactor models were fitted across all 12 correlation matrices (see Fig. 1). All solutions were estimated in Amos 21 via maximum likelihood estimation (Arbuckle, 2012). To demonstrate the nature of the proportionality constraint imposed by the higher-order model, all of the indirect effects associated with the higher-order model solutions were first calculated via the Schmid–Leiman decomposition procedure (Gignac, 2007; Schmid & Leiman, 1957). Then, each subtest's decomposed general factor loading (*g*) was divided by its corresponding Schmid–Leiman decomposed group-level specific factor loading (*s*). When the proportionality constraint is satisfied completely, the ratio of *g*/*s* is the same value for all subtests specified to load onto a particular group-level factor (Yung et al., 1999). Thus, *within* a particular group-level factor, the ratios must be equal. However, the *g*/*s* ratios do not have to be equal to a particular value across group-level factors, in order for the proportionality constraint to hold. In order to represent the degree of proportionality constraint violation, the coefficient of variation (CoV = *SD*/*M*) in *g*/*s* ratios was calculated for each respective group-level factor and then summed

[2] As will be reported below, this statement was verified by the fact that the bifactor model was associated with perfect model fit across all correlation matrices, as well as statistically significant latent variable variances, all defined by statistically significant indicator loadings.

across all three group-level factors. If the proportionality constraint were satisfied perfectly, the CoV would equal .00. By contrast, larger CoV values represented increasingly larger violations of the proportionality constraint. Note well that the higher-order model imposes a proportionately constraint (Yung et al., 1999). Consequently, the CoV values associated with a higher-order model solution will always be equal to zero. By contrast, the bifactor model does not impose a proportionality constraint (Yung et al., 1999). Consequently, the *g*/*s* ratios are free to vary when calculated from the corresponding bifactor model solution. Thus, the CoV values used in this investigation were calculated from the bifactor model solutions.

As parsimony is widely considered a virtue in science, and the bifactor model is a more complex model than the higher-order model in pure degrees of freedom terms (i.e., bifactor estimates more parameters; Rindskopf & Rose, 1988), it has been suggested that, in order to evaluate competing higher-order and bifactor model solutions, fit indexes which incorporate a penalty for model complexity should be preferentially evaluated (Gignac, 2008; Gignac & Watkins, 2013). Consequently, in order to estimate the level of fit associated with the higher-order and bifactor models in this investigation, particular focus was placed upon an evaluation of the Akaike Information Criterion (AIC; Akaike, 1973), the Bayesian Information Criterion (Schwarz, 1978), and the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973). The BIC is known to incorporate a greater penalty for model complexity than the AIC (Burnham & Anderson, 2004). Although the TLI is known to incorporate a greater penalty for model complexity than the CFI (Marsh et al., 2005), it is not known whether the penalty is comparable to the AIC or the BIC. Smaller AIC and BIC values are indicative of a better fitting model. Raftery (1995) suggested that an AIC difference of −6 was indicative of "strong" evidence in favour of a model, and a difference of −10 or greater was indicative of "very strong" evidence in favour of a particular model. In practice, the same guidelines are applied to BIC. As the AIC and the BIC are heavily influenced by sample size (Burnham & Anderson, 2004), Gignac (2007) suggested a TLI difference of .010 may be interpreted as a practical improvement in model fit, as the TLI is only minimally affected by sample size (Marsh et al., 2005). Thus, in order to test the hypothesis that there is a very strong correlation between the degree to which the proportionality constraint is violated and the degree to which the bifactor model fits better than the higher-order model, Pearson correlations were estimated between the CoV values and the difference (Δ) in AIC, BIC, and TLI index values associated with the higher-order and bifactor models. The corresponding *p* values were estimated via randomization tests, as random sampling was clearly not associated with these data. For the purposes of thoroughness, the differences in implied model chi-square values between the higher-order and bifactor models ($\Delta\chi^2$) were also reported.

As fit statistics and indexes are known to interact with sample size (Marsh et al., 2005), the simulation was repeated across three sample sizes: 500, 1000, and 2000. The model fit statistics/indexes were reported for all three sample sizes. However, the factor solution results (i.e., factor loadings) were reported only for the $N = 1000$ for the purposes of brevity (the point estimates were identical across all three sample sizes).

5. Results

As can be seen in Table 1, the first correlation matrix (Satisfied 1; $N = 1000$) was associated with higher-order model and bifactor model chi-square values of .00, which implied that both models fit perfectly, as expected. Additionally, the .00 chi-square value associated with the higher-order model implied that the proportionality constraint imposed by the higher-order model was satisfied perfectly. As can be seen in Table 2, the higher-order model (Schmid–Leiman decomposed) and the bifactor model solutions were identical. It can also be observed in Table 2 that all of the *g*/*s* ratios were equal to the same value within each group-level factor (FA: 1.24. FB: 1.53, and FC: 1.53). It will be noted, however, that the higher-order model was associated with smaller AIC and BIC values (ΔAIC = 12; ΔBIC = 41.45)[3] than the bifactor model. Such a result suggested that the higher-order model should be preferred over the bifactor model on grounds of parsimony, in this case (Raftery, 1995). By contrast, the TLI index failed to distinguish the two models numerically (i.e., ΔTLI = .000). However, based on Gignac's (2007) TLI criterion (ΔTLI = .010 or greater), the higher-order model was considered favoured.

Next, the higher-order and bifactor models were fitted against the 11 correlation matrices whereby the proportionality constraint was violated increasingly in magnitude (i.e., Violated 1 to Violated 11). As can be seen in Table 1, the CoV values were observed to increase in magnitude across all of the Violated correlation matrices, as expected (range: .13 to 1.90). It can also be seen in Table 1 that the bifactor model was observed to fit all correlation matrices perfectly. By contrast, the higher-order model was observed to be associated with increasingly poor levels of model fit. The degree to which the bifactor model fit better than the higher-order model across the correlation matrices can be observed in Table 1 (right side). Specifically, it can be observed that the chi-square difference values ($\Delta\chi^2$) increased in value from .00 to 120.29 across the correlation matrices. Based on 6 degrees of freedom, the chi-square differences were observed to be statistically significant ($p < .05$) at a CoV value of .37 or greater (from Violated 3 onwards).

In contrast to the chi-square difference values, the AIC values decreased in magnitude from 12 (favouring the higher-order model) to −108.29 (favouring the bifactor model). Based on Raftery's (1995) guidelines, the more complex bifactor model was observed to fit better than the higher-order model as evaluated via the AIC at a CoV value of .37 (ΔAIC = −6.26) or greater (Violated 3 onwards). Correspondingly, the BIC values decreased in magnitude from 41.45 (favouring the higher-order model) to −78.85 (favouring the bifactor model). Based on Raftery's (1995) guidelines, the more complex bifactor model was observed to fit better than the higher-order model based on BIC at a CoV value of .61 (ΔBIC = −9.75) or greater (Violated 5 onwards). Finally, it can be observed that the ΔTLI values increased from .000 up to a maximum value of .062. Based on Gignac's (2007) practical difference of .010 criterion for TLI, the more complex bifactor model was observed to fit better than the higher-order model at a CoV value of .37 or greater (Violate 3 onwards). Thus, very similar to AIC and chi-square.

As can be seen in Fig. 2 (Panel A), there was a near perfect, positive association between the degree to which the proportionality constraint was violated (CoV) and the degree to which the higher-order model was associated with a lack of model fit via chi-square ($r = .99$, $p < .001$). Thus, the hypothesis that the degree to which the proportionality constraint was violated would be correlated perfectly with the degree to which the bifactor model fits better than the higher-order model was considered supported. It will be noted, however, that the association was not completely linear. In part, the non-linearity was due to switching from violating the proportionality constraint only for the FA group-level factor to then both the FA and FB group level factors (Violated 5 to Violated 6); and then again from violating the proportionality constraint for the FA and FB group-level factors only to all three group-level factors (Violated 8 to Violated 9). That is, adding the first proportionality constraint violation to the second and third group-level factors did not result in as significant a change in model fit as would be expected based on the results of the preceding correlation matrix. Secondly, and correspondingly, it did appear that progressively larger changes in the violation of the proportionality constraint within

[3] I subtracted the bifactor model AIC, BIC, and TLI values from the higher-order model values, as a larger number of bifactor models were found to be associated with smaller AIC and BIC values, and, correspondingly, larger TLI values. Thus, in this investigation, positive ΔAIC and ΔBIC values were indicative of a better fitting higher-order model. Conversely, negative ΔAIC and ΔBIC values were indicative of a better fitting bifactor model. Positive ΔTLI values were indicative of a better fitting bifactor model. There were no negative ΔTLI values.

Table 1
Model fit statistics and indices associated with the higher-order and bifactor models: N = 1000.

		Higher-order model						Bifactor model						Difference (Δ)			
	CoV	χ^2	*df*	*p*	TLI	AIC	BIC	χ^2	*df*	*p*	TLI	AIC	BIC	χ^2	TLI	AIC	BIC
Satisfied 1	.00	.00	24	1.000	1.013	42.00	145.06	.00	18	1.00	1.013	54.00	186.51	.00	.000	12.00	41.45
Violated 1	.13	2.04	24	1.000	1.012	44.04	147.10	.00	18	1.00	1.013	54.00	186.51	2.04	.001	9.96	39.41
Violated 2	.25	8.33	24	.999	1.009	50.33	153.40	.00	18	1.00	1.013	54.00	186.51	8.33	.004	3.67	33.11
Violated 3	.37	18.26	24	.790	1.003	60.26	163.32	.00	18	1.00	1.013	54.00	186.51	18.26	.010	−6.26	23.19
Violated 4	.49	32.81	24	.108	.995	74.81	177.87	.00	18	1.00	1.013	54.00	186.51	32.81	.018	−20.81	8.64
Violated 5	.61	51.20	24	.001	.986	93.20	196.26	.00	18	1.00	1.013	54.00	186.51	51.20	.027	−39.20	−9.75
Violated 6	.94	60.06	24	.000	.981	102.06	205.12	.00	18	1.00	1.013	54.00	186.51	60.06	.032	−48.06	−18.61
Violated 7	1.10	70.82	24	.000	.975	112.83	215.89	.00	18	1.00	1.013	54.00	186.51	70.82	.038	−58.83	−29.38
Violated 8	1.25	86.09	24	.000	.968	128.09	231.15	.00	18	1.00	1.013	54.00	186.51	86.09	.045	−74.09	−44.64
Violated 9	1.58	94.80	24	.000	.964	136.80	239.86	.00	18	1.00	1.013	54.00	186.51	94.80	.049	−82.80	−53.35
Violated 10	1.74	105.35	24	.000	.958	147.35	250.42	.00	18	1.00	1.013	54.00	186.51	105.35	.055	−93.35	−63.91
Violated 11	1.90	120.29	24	.000	.951	162.29	265.36	.00	18	1.00	1.013	54.00	186.51	120.29	.062	−108.29	−78.85

Note. Satisfied = proportionality constraint satisfied; Violated = proportionality constraint violated; CoV = coefficient of variation in *g*/*s* ratios (corresponds to the degree of violation in the proportionality constraint); $\Delta\chi^2$ values equal to or greater than 12.60 were statistically significant, $p < .05$.

Table 2
Completely standardised solutions associated with the higher-order and bifactor models: Satisfied 1 and Violated 1.

	Proportionality constraint: Satisfied 1										Proportionality constraint: Violated 1									
	Higher-order (S–L)					Bifactor					Higher-order (S–L)					Bifactor				
	g	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*
A1	.58	.47			1.24	.58	.47			1.24	.59	.49			1.22	.58	.53			1.09
A2	.55	.44			1.24	.55	.44			1.24	.55	.45			1.22	.55	.44			1.25
A3	.52	.42			1.24	.52	.42			1.24	.50	.41			1.22	.52	.37			1.41
B1	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
B2	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
B3	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
C1	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C2	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C3	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53

Note. S–L = Schmid–Leiman decomposed; *g*/*s* = *g* loading divided by corresponding group-level specific factor loading; note that the *g*/*s* ratios reported above were calculated based on *g* and *s* values at the fifth decimal place. Consequently, some minor discrepancies based on *g* and *s* values at the second decimal place, as reported in the table above, may be observed.

Table 3
Completely standardised solutions associated with the higher-order and bifactor models: Violated 2 and Violated 3.

	Proportionality constraint: Violated 2										Proportionality constraint: Violated 3									
	Higher-order (S–L)					Bifactor					Higher-order (S–L)					Bifactor				
	g	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*
A1	.58	.48			1.20	.55	.58			.95	.58	.50			1.16	.55	.66			.83
A2	.55	.46			1.20	.55	.43			1.27	.55	.47			1.16	.55	.42			1.32
A3	.51	.43			1.20	.55	.35			1.59	.49	.42			1.16	.55	.30			1.82
B1	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
B2	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
B3	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
C1	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C2	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C3	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53

Note. S–L = Schmid–Leiman decomposed; *g*/*s* = *g* loading divided by corresponding group-level specific factor loading; note that the *g*/*s* ratios reported above were calculated based on *g* and *s* values at the fifth decimal place. Consequently, some minor discrepancies based on *g* and *s* values at the second decimal place, as reported in the table above, may be observed.

Table 4
Completely Standardised solutions associated with the higher-order and bifactor models: Violated 4 and Violated 5.

	Proportionality constraint: Violated 4										Proportionality constraint: Violated 5									
	Higher-order (S–L)					Bifactor					Higher-order (S–L)					Bifactor				
	g	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*	*g*	FA	FB	FC	*g*/*s*
A1	.57	.49			1.16	.52	.72			.72	.58	.53			1.10	.52	.84			.62
A2	.55	.47			1.16	.55	.40			1.37	.54	.49			1.10	.55	.38			1.45
A3	.50	.43			1.16	.58	.28			2.08	.47	.43			1.10	.58	.24			2.44
B1	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
B2	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
B3	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53	.59		.39		1.53
C1	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C2	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C3	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53

Note. S–L = Schmid–Leiman decomposed; *g*/*s* = *g* loading divided by corresponding group-level specific factor loading; note that the *g*/*s* ratios reported above were calculated based on *g* and *s* values at the fifth decimal place. Consequently, some minor discrepancies based on *g* and *s* values at the second decimal place, as reported in the table above, may be observed.

Table 5
Completely standardised solutions associated with the higher-order and bifactor models: Violated 6 and Violated 7.

	Proportionality constraint: Violated 6										Proportionality constraint: Violated 7									
	Higher-order (S–L)					Bifactor					Higher-order (S–L)					Bifactor				
	g	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*
A1	.58	.53			1.10	.52	.84			.62	.58	.53			1.10	.52	.84			.62
A2	.54	.49			1.10	.55	.38			1.45	.54	.49			1.10	.55	.38			1.45
A3	.47	.43			1.10	.58	.24			2.44	.47	.43			1.10	.58	.24			2.44
B1	.62		.42		1.48	.59		.55		1.09	.64		.47		1.35	.59		.83		0.72
B2	.59		.40		1.48	.59		.37		1.61	.59		.44		1.35	.59		.31		1.92
B3	.55		.37		1.48	.59		.28		2.15	.50		.37		1.35	.59		.18		3.26
C1	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C2	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53
C3	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53	.59			.39	1.53

Note. S–L = Schmid–Leiman decomposed; *g/s* = *g* loading divided by corresponding group-level specific factor loading; note that the *g/s* ratios reported above were calculated based on *g* and *s* values at the fifth decimal place. Consequently, some minor discrepancies based on *g* and *s* values at the second decimal place, as reported in the table above, may be observed.

Table 6
Completely Standardised solutions associated with the higher-order and bifactor models: Violated 8 and Violated 9.

	Proportionality constraint: Violated 8										Proportionality constraint: Violated 9									
	Higher-order (S–L)					Bifactor					Higher-order (S–L)					Bifactor				
	g	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*
A1	.58	.53			1.10	.52	.84			.62	.58	.53			1.10	.52	.84			.62
A2	.54	.49			1.10	.55	.38			1.45	.54	.49			1.10	.55	.38			1.45
A3	.47	.43			1.10	.58	.24			2.44	.47	.43			1.10	.58	.24			2.44
B1	.64		.47		1.35	.59		.83		.72	.64		.47		1.35	.59		.83		0.72
B2	.59		.44		1.35	.59		.31		1.92	.59		.44		1.35	.59		.31		1.92
B3	.50		.37		1.35	.59		.18		3.26	.50		.37		1.35	.59		.18		3.26
C1	.62			.42	1.48	.59			.55	1.09	.63			.44	1.42	.59			.66	0.90
C2	.59			.40	1.48	.59			.37	1.61	.59			.41	1.42	.59			.34	1.73
C3	.55			.37	1.48	.59			.28	2.15	.53			.37	1.42	.59			.23	2.60

Note. S–L = Schmid–Leiman decomposed; *g/s* = *g* loading divided by corresponding group-level specific factor loading; note that the *g/s* ratios reported above were calculated based on *g* and *s* values at the fifth decimal place. Consequently, some minor discrepancies based on *g* and *s* values at the second decimal place, as reported in the table above, may be observed.

Table 7
Completely standardised solutions associated with the higher-order and bifactor models: Violated 10 and Violated 11.

	Proportionality constraint: Violated 10										Proportionality constraint: Violated 11									
	Higher-order (S–L)					Bifactor					Higher-order (S–L)					Bifactor				
	g	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*	*g*	FA	FB	FC	*g/s*
A1	.58	.53			1.10	.52	.84			0.62	.58	.53			1.10	.52	.84			.62
A2	.54	.49			1.10	.55	.38			1.45	.54	.49			1.10	.55	.38			1.45
A3	.47	.43			1.10	.58	.24			2.44	.47	.43			1.10	.58	.24			2.44
B1	.64		.47		1.35	.59		.83		0.72	.64		.47		1.34	.59		.83		.72
B2	.59		.44		1.35	.59		.31		1.92	.59		.44		1.34	.59		.31		1.92
B3	.50		.37		1.35	.59		.18		3.26	.50		.37		1.34	.59		.18		3.26
C1	.63			.44	1.42	.59			.66	0.90	.64			.47	1.34	.59			.83	.72
C2	.59			.41	1.42	.59			.34	1.73	.59			.44	1.34	.59			.31	1.92
C3	.53			.37	1.42	.59			.23	2.60	.50			.37	1.34	.59			.18	3.26

Note. S–L = Schmid–Leiman decomposed; *g/s* = *g* loading divided by corresponding group-level specific factor loading; note that the *g/s* ratios reported above were calculated based on *g* and *s* values at the fifth decimal place. Consequently, some minor discrepancies based on *g* and *s* values at the second decimal place, as reported in the table above, may be observed.

Table 8
Model fit statistics and indices associated with the higher-order and bifactor models: N = 500.

		Higher-order model						Bifactor model						Difference (Δ)			
	CoV	χ^2	*df*	*p*	TLI	AIC	BIC	χ^2	*df*	*p*	TLI	AIC	BIC	χ^2	TLI	AIC	BIC
Satisfied 1	.00	.00	24	1.000	1.027	42.00	131.51	.00	18	1.00	1.027	54.00	167.79	.00	.000	12.00	36.28
Violated 1	.13	1.02	24	1.000	1.025	43.02	131.53	.00	18	1.00	1.026	54.00	167.79	1.02	.001	10.98	36.26
Violated 2	.25	4.16	24	1.000	1.022	46.16	134.67	.00	18	1.00	1.026	54.00	167.79	4.16	.004	7.84	33.12
Violated 3	.37	9.12	24	.997	1.016	51.12	139.63	.00	18	1.00	1.026	54.00	167.79	9.12	.010	2.88	28.16
Violated 4	.49	16.39	24	.874	1.008	58.39	146.89	.00	18	1.00	1.026	54.00	167.79	16.39	.018	−4.39	20.90
Violated 5	.61	25.57	24	.375	.998	67.57	156.08	.00	18	1.00	1.026	54.00	167.79	25.57	.028	−13.57	11.71
Violated 6	.94	30.00	24	.185	.994	72.00	160.51	.00	18	1.00	1.026	54.00	167.79	30.00	.032	−18.00	7.28
Violated 7	1.10	35.38	24	.063	.988	77.38	165.89	.00	18	1.00	1.026	54.00	167.79	35.38	.038	−23.38	1.90
Violated 8	1.25	43.00	24	.010	.980	85.00	173.51	.00	18	1.00	1.025	54.00	167.79	43.00	.045	−31.00	−5.72
Violated 9	1.58	47.35	24	.003	.976	89.35	177.86	.00	18	1.00	1.025	54.00	167.79	47.35	.049	−35.35	−10.07
Violated 10	1.74	52.62	24	.001	.970	94.62	183.13	.00	18	1.00	1.025	54.00	167.79	52.62	.055	−40.62	−15.34
Violated 11	1.90	60.09	24	.000	.963	102.09	190.59	.00	18	1.00	1.025	54.00	167.79	60.09	.062	−48.09	−22.80

Note. Satisfy = proportionality constraint satisfied; Violated = proportionality constraint violated; CoV = coefficient of variation in *g/s* ratios (corresponds to the degree of violation in the proportionality constraint); $\Delta\chi^2$ values equal to or greater than 12.60 were statistically significant, $p < .05$.

a group-level factor was associated with even larger changes in model fit in favour of the bifactor model than would be expected based purely upon a linear effect. Such non-linearity was most evident in the differences in the higher-order model and bifactor model fit levels associated with the Violated 4 to Violated 5 correlation matrices.

There was also a near perfect, negative, linear association ($r = -.99$, $p < .001$) between the degree to which the proportionality constraint was violated (CoV) and the magnitude with which the AIC favoured the bifactor model (ΔAIC; see Fig. 2, Panel B). Similarly, the correlation between the CoV values and the magnitude with which the BIC favoured the bifactor model (ΔBIC) was also very large and negative ($r = -.99$, $p < .001$; see Fig. 2, Panel C). Finally, as can be seen in Fig. 2 (Panel D), there was a strong, positive, and mostly linear association ($r = .99$, $p < .001$) between the degree of with which the proportionality constraint was violated (COV) and the magnitude with which the TLI favoured the bifactor model (ΔTLI).

With respect to the Schmid–Leiman decomposed loadings and the bifactor loadings, it can be observed across Tables 2 to 4 that the a1 and a2 indicator loadings on the FA group-level factor were progressively different. In particular, by correlation matrix Violated 3, the point at which ΔTLI .010 and ΔAIC = −6.20, the higher-order model suggested that the a1 and a2 indicators were associated with the FA factor at .50 and .42, respectively. By contrast, the bifactor model suggested that the corresponding a1 and 3 indicator associations with the FA factor were .66 and .30, respectively. By correlation matrix Violated 5, the point at which ΔBIC = −9.75, the higher-order model suggested that the a1 and a2 indicators were associated with the FA factor at .49 and .43. By contrast, the bifactor model suggested that the corresponding a1 and a3 associations with the FA factor were .72 and .28, respectively. A similar effect was observed for the b1 and b2 indicator loadings on the FB group-level factor across Tables 5 and 6. Finally, with respect to the correlation matrix with the greatest degree of proportionality constraint violation (Violated 11; CoV = 1.20), it can be observed that there were very substantial differences between the higher-order model solution S–L decomposed loadings and the bifactor model solution loadings across all three group-level factors (see Table 7; right side). Again, the higher-order model S–L decomposed solution suggested much greater homogeneity in the specific loadings, in comparison to the bifactor model solution.

Finally, as can be seen in Tables 8 and 9, the fit statistic/index results were consistent across the sample sizes of 500 and 2000. However, it was noted that the $\Delta\chi^2$, AIC and BIC values were much more substantially affected by sample size than the TLI. In particular, it was noted that the TLI suggested the bifactor model should be favoured at CoV = .37 (Violated 3) and above across all three sample sizes. However, the $\Delta\chi^2$ and the AIC suggested favourability for the bifactor model at CoV = .25 (Violated 2) and above at $N = 2000$, but CoV = .49 (Violated 4) and above at $N = 500$. Correspondingly, the BIC suggested favourability for the bifactor model at CoV = .49 (Violated 4) and above at $N = 2000$, but CoV = 1.25 (Violated 8) and above at $N = 500$.

6. Discussion

The results of this investigation suggest clearly that the bifactor model will fit better than the corresponding higher-order model to the degree that the data are inconsistent with the proportionality constraint imposed by the higher-order model. Additionally, the TLI indicated favourability for the higher-order model at low levels of proportionality constraint violation consistently across all three sample sizes ($N = 500$, 1000, and 2000). Although the chi-square difference test, the AIC and the BIC also tended to favour the higher-order model at lower levels of proportionality constraint violation, their behaviour was substantially affected by sample size. At moderate to high levels

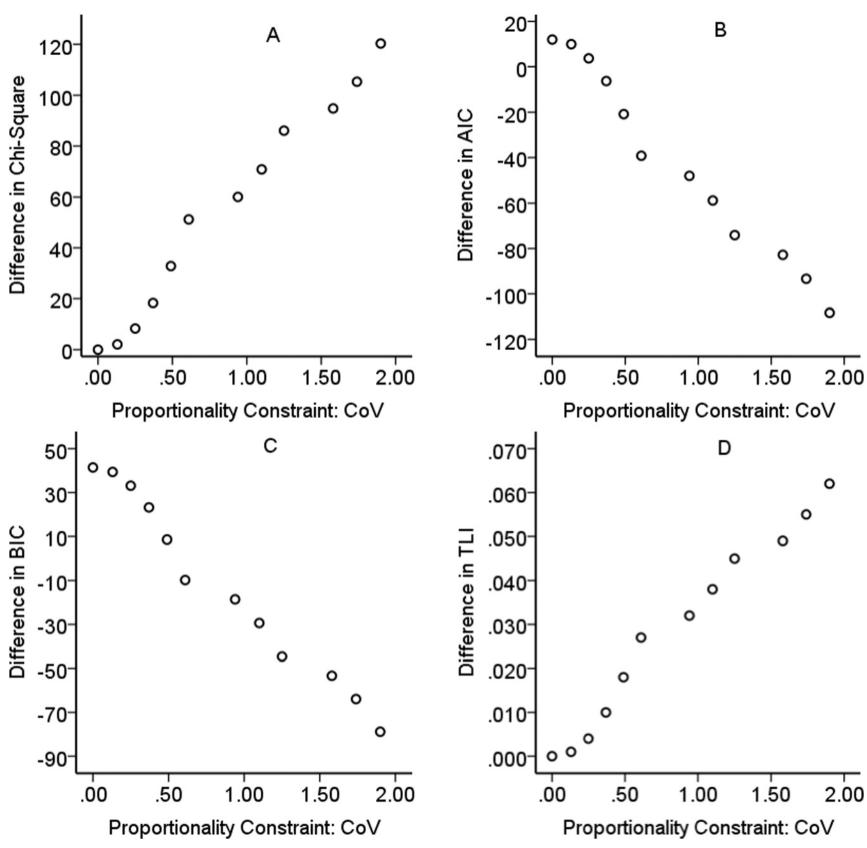

Fig. 2. Scatter plots depicting the association between the degree of violation in the proportionality constraint imposed by the higher-order model and the degree to which the bifactor model is associated with better model fit (AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; TLI = Tucker–Lewis Index).

Table 9
Model fit statistics and indices associated with the higher-order and bifactor models: N = 2000.

	CoV	Higher-order model χ^2	*df*	*p*	TLI	AIC	BIC	Bifactor model χ^2	*df*	*p*	TLI	AIC	BIC	Difference (Δ) χ^2	TLI	AIC	BIC
Satisfied 1	.00	.000	24	1.00	1.007	42.00	159.62	.000	18	1.00	1.007	54.00	205.22	.000	.000	12	45.6
Violated 1	.13	4.08	24	1.00	1.005	46.08	163.70	.000	18	1.00	1.006	54.00	205.22	4.08	.001	7.92	41.52
Violated 2	.25	16.68	24	.862	1.002	58.68	176.30	.000	18	1.00	1.006	54.00	205.22	16.68	.004	−4.68	28.92
Violated 3	.37	36.53	24	.049	.997	78.53	196.15	.000	18	1.00	1.006	54.00	205.22	36.53	.009	−24.53	9.07
Violated 4	.49	65.64	24	.000	.989	107.64	225.26	.000	18	1.00	1.006	54.00	205.22	65.64	.017	−53.64	−20.04
Violated 5	.61	102.44	24	.000	.979	144.44	262.06	.000	18	1.00	1.006	54.00	205.22	102.44	.027	−90.44	−56.84
Violated 6	.94	120.18	24	.000	.975	162.18	279.80	.000	18	1.00	1.006	54.00	205.22	120.18	.031	−108.18	−74.58
Violated 7	1.10	141.73	24	.000	.969	183.73	301.35	.000	18	1.00	1.006	54.00	205.22	141.73	.037	−129.73	−96.13
Violated 8	1.25	172.26	24	.000	.962	214.26	331.88	.000	18	1.00	1.006	54.00	205.22	172.26	.044	−160.26	−126.66
Violated 9	1.58	189.69	24	.000	.958	231.69	349.31	.000	18	1.00	1.006	54.00	205.22	189.69	.048	−177.69	−144.09
Violated 10	1.74	210.81	24	.000	.953	252.81	370.43	.000	18	1.00	1.006	54.00	205.22	210.81	.053	−198.81	−165.21
Violated 11	1.90	240.71	24	.000	.946	282.71	400.32	.000	18	1.00	1.006	54.00	205.22	240.71	.060	−228.71	−195.1

Note. Satisfied = proportionality constraint satisfied; Violated = proportionality constraint violated; CoV = coefficient of variation in *g*/*s* ratios (corresponds to the degree of violation in the proportionality constraint); $\Delta\chi^2$ values equal to or greater than 12.60 were statistically significant, $p < .05$.

of proportionality constraint violation, the TLI suggested favourability for the bifactor model consistently across all sample sizes.

In contrast to previous empirical treatments of the difference between higher-order and bifactor models in relation to model fit (Chen et al., 2006; Morgan et al., 2015; Mulaik & Quartetti, 1997; Murray & Johnson, 2013), the results of this investigation offer, arguably, a very clear explanation for when the bifactor model will be observed to fit better than the higher-order model and vice versa. Each increase in the degree of within group-level factor variability in the *g*/*s* ratios (i.e., CoV) was observed to be associated with an increase in level of fit progressively in favour of the bifactor model. Although the Pearson correlations between the CoV values and the degree of fit difference between the higher-order model and the bifactor model were nearly perfect ($r = .99$), there was some degree of curvilinearity associated with the effect. Based on the scatter plots, it would appear that minor to moderate increases in the violation of the proportionality constraint at lower levels of violation in the proportionality constraint are not as impactful on fit indices as the same minor to moderate increases in the violation of the proportionality constraint at higher levels of violation in the proportionality constraint.

Precisely why the effect is not perfectly linear remains to be determined. It is possible that the CoV is not a perfectly valid indicator of violation of the proportionality constraint across all levels of proportionality constraint violation. Alternatively, the minor non-linearity effect may be a simple consequence of the possibility that the maximum likelihood estimation fit function is affected by discrepancies between the observed covariance matrix and the implied model covariance matrix in such a way that larger discrepancies have an increasingly disproportional impact, as clearly observed in the ordinary least squares based discrepancy function (Browne, MacCallum, Kim, Andersen, & Glaser, 2002). Irrespective of the explanation for the non-linearity effect, it is arguably assuring to know that the higher-order model is not necessarily discriminated against in favour of the bifactor model, as relatively minor violations of the proportionality constraint were associated with index favourability for the higher-order model. In particular, TLI favoured consistently the higher-order model up to a CoV value of .37 (Violated 3), across all three sample sizes ($N = 500$, 1000, 2000). Although the chi-square difference test, the AIC, and the BIC also tended to favour the higher-order model at lower levels of proportionality constraint violation, their behaviour was observed to be affected greatly by sample size. Such an effect has been observed in other CFA simulations (e.g. Marsh & Balla, 1994). Thus, the TLI may be considered more attractive for model comparisons, in this context.

Given that TLI favoured the bifactor model only at moderate to high levels of proportionality constraint violation, the results reported in this investigation are not consistent with Murray and Johnson's (2013) conclusion that fit indices such as TLI and AIC are "biased" in favour of the bifactor model. Based on the results of this investigation, it may be suggested that the correlated residuals and cross-loadings ("unmodelled complexity") Murray and Johnson (2013) added to their correlation matrices may have had the unintended consequence of violating the proportionality constraint imposed by the higher-order model. Thus, instead of evidencing bias, the fit indices may have simply been responsive to this effect. Unfortunately, Murray and Johnson (2013) did not report a sufficient amount of information (correlation matrices; factor solutions) in their article for the reader to evaluate this possibility.

It is well documented that the higher-order model is associated with more degrees of freedom than a corresponding bifactor model (Rindskopf & Rose, 1988). In this sense, the higher-order model may be regarded as the simpler, preferable model over the bifactor model. However, based on the results of this investigation, the model simplicity associated with the higher-order model may be suggested to be achieved at the expense of having to incorporate an assumption that may prove difficult to satisfy in practice: the proportionality constraint. It is likely for this reason that the bifactor model has been observed to fit better than the higher-order model across a wide variety of intelligence batteries (e.g., Brunner & Süß, 2005; Canivez, in press; Gignac, 2005, 2006a, 2006b, 2008; Golay & Lecerf, 2011; Murray & Johnson, 2013).

From a theoretical perspective, it would seem "unjust" that the statistical evaluation of whether *g* is best considered a superordinate factor (higher-order) or a breadth factor (first-order) comes down to testing a statistical assumption of sorts: the proportionality constraint imposed by the higher-order model. Evidently, additional empirical and theoretical considerations will have to be provided, in order to evaluate the issue convincingly, as argued by others (Murray & Johnson, 2013). It should be acknowledged, nonetheless, that those who endorse a higher-order conceptualisation of intelligence should be required to explain why the nature of intelligence is such that any indicator of intelligence can contribute variance to *g* and a group-level factor residual in a manner that is perfectly proportional (within sampling fluctuations) across all indicators within a group-level factor. By contrast, those who endorse a breadth conceptualisation of intelligence (i.e., bifactor) do not have to provide any such explanation, as each subtest is allowed to contribute variance to *g* and a group-level factor freely in the bifactor model. From this perspective, perhaps the bifactor model may be considered simpler than the higher-order model.

However, it should also be acknowledged that both the correlated factors model and the higher-order model are more consistent with conventional views of intelligence, most all of which posit that the group-level factors of intelligence are inter-related with each other positively (e.g., Carroll, 1993; Gustafsson, 1984; Van Der Maas et al., 2006). Theoretical work which attempts to reconcile the bifactor model with current theories of intelligence is encouraged. Finally, it will be noted that hybrid bifactor models can be specified such that some of the nested group-level factors are allowed to correlate (e.g., Rindskopf &

Rose, 1988). Additionally, higher-order models can be specified such that there are direct effects between a second-order *g* factor and some of the subtests (e.g., Gignac, 2014b; Yung et al., 1999). The true nature of intelligence may be possibly consistent with a hybrid bifactor or hybrid higher-order model.

Whether the differences in model fit between the higher-order model and the bifactor model may be expected to be sufficiently large in practice to yield meaningful differences in the factor solutions is an important question. Ultimately, if the bifactor model tends to fit better than the higher-order model, but both models yield very comparable solutions (once the higher-order model is Schmid–Leiman decomposed), then the question addressed in this investigation would be justly described as relatively trivial.

Based on the results of this investigation, arguably substantial differences in the group-level factor loadings were observed at moderate levels of proportionality constraint violation (CoV = .37 to .49). Based on previous empirical research with field data, it may be suggested that important differences can be observed between the higher-order model and bifactor model solutions, as well. For example, based on the WAIS-IV normative sample (Wechsler, 2008), Gignac and Watkins (2013) found that the higher-order model suggested consistently (across age groups) that Arithmetic loaded (S–L decomposed) onto a WMC factor in a manner very comparable to other WMC subtests (≈.35). By contrast, the corresponding bifactor model solutions suggested consistently that the Arithmetic subtest loaded onto a WMC factor very weakly or not at all. The Figure Weights and Matrix Reasoning subtests, two very good indicators of g_f, also evidenced differential results between the two models such that both subtests were observed to load onto the *g* factor model more substantially as estimated within a bifactor model. Consequently, Gignac and Watkins (2013) suggested that the WAIS-IV bifactor model results were more theoretically congruent than the corresponding higher-order results. It is also possible for a higher-order model to suggest the presence of a particular group-level specific factor which is not observed in the corresponding bifactor model (e.g., Gignac, 2007, 2008). Thus, very substantial differences can be observed between the two solutions. As a general statement, it may be suggested that the proportionality constraint imposed by the higher-order model has a tendency to cause indicators to be related to factors in a homogenous manner, arguably unnaturally so (Brunner, 2008).

All of the higher-order models tested in this investigation were observed to be associated with very good levels of model fit according to conventional standards (Hu & Bentler, 1999), even when the proportionality constraint was violated. For example, the higher-order model tested on the Violated 5 correlation matrix ($N = 1000$) was observed to be associated with a TLI value of .995. However, arguably, substantial differences were nonetheless observed in the higher-order model and bifactor model solutions when tested on the Violated 5 correlation matrix. Correspondingly, the higher-order model of the WAIS-IV tested in Gignac and Watkins (2013) was observed to fit reasonably well across the age groups (TLI ≈ .940). However, the bifactor model was observed to fit better (ΔTLI ≈ .025). Such results suggest that even when a higher-order model fits reasonably well, according to conventional standards (Hu & Bentler, 1999), there may be benefits to conducting further testing via the bifactor model, in order to evaluate the impact of the proportionality constraint. In practice, it is likely that the violations of the proportionality constraint imposed by the higher-order model will be distributed across several of the group-level factors, rather than concentrated on a single group-level factor. Consequently, any single violation might not be expected to be as substantial as that observed in this investigation for the Violated 4 and 5 correlation matrix conditions.

Although the results of this simulation suggest that the proportionality constraint imposed by the higher-order model may sometimes make it difficult to endorse from a model-fit perspective, it should be acknowledged that the higher-order model does fit better than the bifactor model in some cases with field data (e.g., MacCann, Joseph, Newman, & Roberts, 2014). Furthermore, the higher-order model does enjoy some unique attractive qualities. In particular, the higher-order model is the only model that allows for the estimation of the association between all group-level factors and a general factor of intelligence, simultaneously. Thus, in some cases, the higher-order model may be clearly the preferred model with respect to addressing a particular hypothesis (e.g., Gignac, 2014b; Gignac & Watkins, 2015). Additionally, theoretically, the higher-order model is the only model that implies that the group-level factors of intelligence are inter-correlated because of *g*: a notion that arguably holds a lot of currency in the broader intelligence literature currently (Reynolds & Keith, 2013). Based on current SEM methodologies, whether the *g* factor emerges due to correlations between broad abilities (higher-order model) or correlations between subtests (bifactor model) cannot be confirmed or disconfirmed, unfortunately. Finally, based on a small amount of unpublished simulation research I have conducted, it would appear that the bifactor model may require somewhat greater sample sizes (≈ 10%) than that required by the higher-order model, in order to identify corresponding factor loadings as statistically significant (i.e., bifactor loadings in comparison to Schmid–Leiman decomposed loadings). Correspondingly, as the bifactor model has more estimated parameters than a corresponding higher-order model, it may be associated with greater risks of overparameterisation (negative variance estimates; Maydeu-Olivares & Coffman, 2006).

6.1. Limitations

This investigation examined only one parameter associated with the higher-order model: the proportionality constraint. Although the manipulation of this single parameter evidenced an essentially perfect association with the degree of difference in model fit between the higher-order model and the bifactor model, it remains a possibility that one or more other parameters could be manipulated in such a way as to compliment, or possibly even mediate, the effect observed in this investigation. Given the current analytical on the distinction between the higher-order model and the bifactor model (Schmiedek & Li, 2004; Yung et al., 1999), it is difficult to generate speculations on what those parameters might be exactly.

Although the number of simulated correlation matrices examined in this investigation was greater than that of most previous investigations in the area (Chen et al., 2006; Mulaik & Quartetti, 1997; Murray & Johnson, 2013), it was not an entirely comprehensive simulation. Models with different numbers of group-level factors, different numbers of indicators per group-level factor, as well as additional methods to compare model fit, may potentially yield different results to those reported in this investigation. It will be noted that the model used in this investigation was relatively small, as it included only three group-level factors defined by only three indicators each. Such a model was selected to help isolate precisely the effect of violating the proportionality constraint, as it becomes increasingly difficult to disturb the proportionality constraint in larger correlation matrices, without also causing the bifactor model to become associated with a certain level of model misfit (i.e., cause both models to be associated with misfit for reasons other than the proportionality constraint). Consequently, the results of this simulation are limited to the degree that the reported effect generalises beyond models characterised by latent variables that are just-identified. In practice, when group-level factors are defined by four or more indicators per group-level factor, it would be expected that there would be greater chances to violate the proportionality constraint imposed by the higher-order model. Consequently, the bifactor model would be expected to be more likely to be observed to fit better than the higher-order model in such larger models. Conversely, when models are defined by several group-level factors defined by only two indicators, the higher-order model may be more likely to be associated with the better fit than the bifactor model, as there are simply less opportunities to violate the proportionality constraint imposed by the higher-order model (e.g., MacCann et al., 2014).

7. Conclusion

It is hoped that the results of this investigation may be viewed as a clear description for when the bifactor model may be expected to fit better than the higher-order model. Divergent views in the area of intelligence with respect to preferences for the higher-order model, the bifactor model, and even the correlated factors model may be expected to continue, of course, until additional types of evidence is provided (i.e., not simply based on fit statistics/indices). Interestingly, it will be noted that the bifactor model appears to have become embraced in the area of psychometrics and personality in recent years (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; Gibbons et al., 2007; Gruhl, Erosheva, & Crane, 2013; Morin, Arens, & Marsh, in press; Reise, 2012; Rodriguez, Reise, & Haviland, in press). Although one can only speculate, it may be suggested that psychometricians and personality researchers view the bifactor model favourably because they ascribe less theoretical distinction between a superordinate general factor and a breadth general factor. Given that the only statistically testable distinction between the higher-order model and the bifactor model appears to be an arguably implausible proportionality constraint, it may be time for intelligence researchers to consider adopting a similar perspective. More work which attempts to integrate the bifactor model into existing intelligence theory is encouraged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.intell.2016.01.006.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csáki (Eds.), *Second International Symposium on Information Theory* (pp. 267–281). Budapest, Hungary: Akadémiai Kiadó.

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, *40*, 37–47.

Arbuckle, J. L. (2012). *Amos (Version 21.0)*. Chicago: IBM SPSS.

Beaujean, A. A., Parkin, J., & Parker, S. (2014). Comparing Cattell–Horn–Carroll factor models: Differences between bifactor and higher order factor models in predicting language achievement. *Psychological Assessment*, *26*, 789–805.

Bechtel, R. B., Verdugo, V. C., & de Queiroz Pinheiro, J. (1999). Environmental belief systems United States, Brazil, and Mexico. *Journal of Cross-Cultural Psychology*, *30*(1), 122–128.

Browne, M. W., MacCallum, R. C., Kim, C. T., Andersen, B. L., & Glaser, R. (2002). When fit indices and residuals are incompatible. *Psychological Methods*, *7*(4), 403–421.

Brunner, M. (2008). No g in education? *Learning and Individual Differences*, *18*(2), 152–165.

Brunner, M., & Süß, H. -M. (2005). Analyzing the reliability of multidimensional measures: An example from intelligence research. *Educational and Psychological Measurement*, *65*, 227–240.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304.

Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer, & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements*. Gottingen, Germany: Hogrefe Publishers (in press).

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

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J. P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219–251.

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189–225.

Crawford, J. R., Deary, I. J., Allan, K. M., & Gustafsson, J. E. (1998). Evaluating competing models of the relationship between inspection time and psychometric intelligence. *Intelligence*, *26*(1), 27–42.

Gibbons, R. D., Bock, R. D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K., ... Stover, A. (2007). Full-information item bifactor analysis of graded response data. *Applied Psychological Measurement*, *31*(1), 4–19.

Gignac, G. E. (2005). Revisiting the factor structure of the WAIS-R insights through nested factor modeling. *Assessment*, *12*(3), 320–329.

Gignac, G. E. (2006a). A confirmatory examination of the factor structure of the Multidimensional Aptitude Battery (MAB): Contrasting oblique, higher-order, and nested factor models. *Educational and Psychological Measurement*, *66*(1), 136–145.

Gignac, G. E. (2006b). The WAIS-III as a nested factors model: A useful alternative to the more conventional oblique and higher-order models. *Journal of Individual Differences*, *27*, 73–86.

Gignac, G. E. (2007). Multi-factor modeling in individual differences research: Some recommendations and suggestions. *Personality and Individual Differences*, *42*(1), 37–48.

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: g as superordinate or breadth factor? *Psychology Science*, *50*(1), 21–43.

Gignac, G. E. (2014b). Fluid intelligence shares closer to 60% of its variance with working memory capacity and is a better indicator of general intelligence. *Intelligence*, *47*, 122–133.

Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, *48*(5), 639–662.

Gignac, G. E., & Watkins, M. (2015). There may be nothing special about the association between working memory capacity and fluid intelligence. *Intelligence*, *52*, 18–23.

Gignac, G. E., Palmer, B. R., Manocha, R., & Stough, C. (2005). An examination of the factor structure of the Schutte self-report emotional intelligence (SSREI) scale via confirmatory factor analysis. *Personality and Individual Differences*, *39*(6), 1029–1042.

Gignac, G. E. (2014a). Dynamic mutualism versus *g* factor theory: An empirical test. *Intelligence*, *42*, 89–97.

Golay, P., & Lecerf, T. (2011). Orthogonal higher order structure and confirmatory factor analysis of the French Wechsler Adult Intelligence Scale (WAIS-III). *Psychological Assessment*, *23*, 143–152.

Gruhl, J., Erosheva, E. A., & Crane, P. K. (2013). A semiparametric approach to mixed outcome latent variable models: Estimating the association between cognition and regional brain volumes. *Annals of Applied Statistics*, *7*(4), 2361–2383.

Gustafsson, J., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, *28*, 407–434.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*(3), 179–203.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*, 42–54.

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

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

Jensen, A. R., & Weng, L. J. (1994). What is a good *g*? *Intelligence*, *18*(3), 231–258.

Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 581–614) (2nd ed.). New York: Guilford Press.

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*(2), 358–374.

Marsh, H. W., & Balla, J. (1994). Goodness of fit in confirmatory factor analysis: The effects of sample size and model parsimony. *Quality and Quantity*, *28*(2), 185–217.

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares, & J. J. McArdle (Eds.), *Contemporary psychometrics: A Festschrift for Roderick P. McDonald* (pp. 225–340). Mahwah, NJ: Erlbaum.

Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods*, *11*(4), 344–362.

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

Mohlman, J., & Zinbarg, R. E. (2000). The structure and correlates of anxiety sensitivity in older adults. *Psychological Assessment*, *12*(4), 440.

Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are Fit indices biased in favor of bi-factor models in cognitive ability research?: a comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. *Journal of Intelligence*, *3*(1), 2–20.

Morin, A. J., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling* (in press).

Mulaik, S. A., & Quartetti, D. A. (1997). First order or higher order general factor? *Structural Equation Modeling*, *4*, 193–211.

Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, *41*(5), 407–422.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Reeve, C. L., & Blacksmith, N. (2009). Identifying *g*: A review of current factor analytic practices in the science of mental abilities. *Intelligence*, *37*(5), 487–494.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696.

Reynolds, M. R., & Keith, T. Z. (2013). Measurement and statistical issues in child psychological assessment. In D. H. Saklofske, V. L. Schwean, & C. R. Reynolds (Eds.), *Oxford handbook of child and adolescent assessment* (pp. 48–83). New York: Oxford University Press.

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, *23*(1), 51–67.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment* (in press).

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*, 53–61.

Schmiedek, F., & Li, S. C. (2004). Toward an alternative representation for disentangling age-associated differences in general and specific cognitive abilities. *Psychology and Aging*, *19*(1), 40–56.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.
Smith, R. H., Parrott, W. G., Diener, E. F., Hoyle, R. H., & Kim, S. H. (1999). Dispositional envy. *Personality and Social Psychology Bulletin*, *25*(8), 1007–1020.
Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10.
Van Der Maas, H. L., Dolan, C. V., Grasman, R. 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*(4), 842–861.
Watkins, M. W., & Kush, J. C. (2002). Confirmatory factor analysis of the WISC-III for students with learning disabilities. *Journal of Psychoeducational Assessment*, *20*(1), 4–19.
Wechsler, D. (2008). *Wechsler adult intelligence scale—Fourth edition: Technical and interpretive manual*. San Antonio, TX: Pearson Assessment.
Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013). Technical and practical issues in the structure and clinical invariance of the Wechsler Scales: A rejoinder to commentaries. *Journal of Psychoeducational Assessment*, *31*(2), 235–243.
Yung, Y. F., McLeod, L. D., & Thissen, D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128.
Zinbarg, R. E., Revelle, W., & Yovel, . (2007). Estimating Ω_h for structures containing two group factors: Perils and prospects. *Applied Psychological Measurement*, *31*(2), 135–157.