



UvA-DARE (Digital Academic Repository)

How to Compare Psychometric Factor and Network Models

Kan, K.J.; de Jonge, H.; van der Maas, H.L.J.; Levine, S.Z.; Epskamp, S.

DOI

[10.3390/jintelligence8040035](https://doi.org/10.3390/jintelligence8040035)

Publication date

2020

Document Version

Final published version

Published in

Journal of Intelligence

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Kan, K. J., de Jonge, H., van der Maas, H. L. J., Levine, S. Z., & Epskamp, S. (2020). How to Compare Psychometric Factor and Network Models. *Journal of Intelligence*, 8(4), [35]. <https://doi.org/10.3390/jintelligence8040035>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Commentary

How to Compare Psychometric Factor and Network Models

Kees-Jan Kan ^{1,*} , Hannelies de Jonge ¹, Han L. J. van der Maas ², Stephen Z. Levine ³ 
and Sacha Epskamp ²

¹ Research Institute of Child Development and Education, University of Amsterdam, 1018 WS Amsterdam, The Netherlands; h.dejonge@uva.nl

² Department of Psychology, University of Amsterdam, 1018 WS Amsterdam, The Netherlands; h.l.j.vandermaas@uva.nl (H.L.J.v.d.M.); s.epskamp@uva.nl (S.E.)

³ Department of Community Mental Health, University of Haifa, Haifa 3498838, Israel; slevine@univ.haifa.ac.il

* Correspondence: k.j.kan@uva.nl

Received: 7 June 2020; Accepted: 24 September 2020; Published: 2 October 2020



Abstract: In memory of Dr. Dennis John McFarland, who passed away recently, our objective is to continue his efforts to compare psychometric networks and latent variable models statistically. We do so by providing a commentary on his latest work, which he encouraged us to write, shortly before his death. We first discuss the statistical procedure McFarland used, which involved structural equation modeling (SEM) in standard SEM software. Next, we evaluate the penta-factor model of intelligence. We conclude that (1) standard SEM software is not suitable for the comparison of psychometric networks with latent variable models, and (2) the penta-factor model of intelligence is only of limited value, as it is nonidentified. We conclude with a reanalysis of the Wechsler Adult Intelligence Scale data McFarland discussed and illustrate how network and latent variable models can be compared using the recently developed R package Psychometrics. Of substantive theoretical interest, the results support a network interpretation of general intelligence. A novel empirical finding is that networks of intelligence replicate over standardization samples.

Keywords: psychometric network analysis; factor analysis; latent variable modeling; intelligence; model comparison; replicating networks

1. Introduction

On 29 April 2020, the prolific researcher Dr. Dennis John McFarland sadly passed away. The *Journal of Intelligence* published one of his last contributions: “The Effects of Using Partial or Uncorrected Correlation Matrices When Comparing Network and Latent Variable Models” (McFarland 2020). We here provide a commentary on this paper, which McFarland encouraged us to write after personal communication, shortly before his death. The aim of our commentary is threefold. The first is to discuss the statistical procedures McFarland (2020) used to compare network and latent variable models; the second to evaluate the “penta-factor model” of intelligence; and the third to provide a reanalysis of two data sets McFarland (2020) discussed. With this reanalysis, we illustrate how we advance the comparison of network and latent variable models using the recently developed R package Psychometrics (Epskamp 2020b). Before we do so, we first express our gratitude to Dr. McFarland for sharing his data, codes, output files, and thoughts shortly before he passed away, as these were insightful in clarifying what had been done precisely, and so helped us write this commentary.

As McFarland (2020) pointed out, psychometric network modeling (of Gaussian data, e.g., (Epskamp et al. 2017)) and latent variable modeling (Kline 2015) are alternatives to each other, in the sense that both can be applied to describe or explain the variance–covariance structure of observed

variables of interest. As he continued, psychometric network models are typically depicted graphically, where so-called “nodes” represent the variables in the model and—depending on the choice of the researcher—“edges” represent either zero-order correlations or partial correlations among these variables. It is indeed recommended (Epskamp and Fried 2018) to opt for partial correlations and also to have removed those edges that are considered not different from zero or are deemed spurious (false positives), as McFarland (2020) mentioned. This procedure of removing edges is commonly referred to as “pruning” (Epskamp and Fried 2018). We would have no difficulty with the position that probably most of the psychometric network models of Gaussian data can be interpreted as “reduced” (i.e., pruned) partial correlation matrices. This also holds for the network models of intelligence that McFarland (2020) discussed (Kan et al. 2019; Schmank et al. 2019).

Through personal communication, and as his R script revealed, it became clear that, unfortunately, McFarland (2020) mistook how network models of Gaussian data are produced with the R function *glasso* (Friedman et al. 2019) and, as a consequence, how fit statistics are usually obtained. Although we are unable to verify any longer, we believe that this explains McFarland’s (2020, p. 2) concern that in latent variable and psychometric network analyses “disparate correlation matrices” are being compared, i.e., zero-order correlation matrices versus partial correlation matrices. To clarify to the audience, what happens in psychometric network analysis of Gaussian data (Epskamp et al. 2017) is that two variance–covariance matrices are being compared, (1) the variance–covariance matrix implied by the network (Σ) versus (2) the variance–covariance matrix that was observed in the sample (S).¹ This procedure parallels the procedure in structural equation modeling (Kline 2015) in which the variance–covariance matrix implied by a latent variable model is compared to the observed sample variance–covariance matrix. In terms of *standardized* matrices, a comparison is thus always between the *zero-order correlation matrix implied by the model* and the *observed zero-order correlation matrix*, no matter whether a latent variable model is fitted in SEM software or a psychometric network model in software that allows for psychometric network analysis. This also applies when the partial correlation matrix is pruned (and Σ is modeled as $\Sigma = \Delta(I - \Omega)^{-1}\Delta$, where Ω contains [significant] partial correlations on the off diagonals, Δ is a weight matrix, and I is an identity matrix; see (Epskamp et al. 2017)). We concur with McFarland’s (2020, p. 2) statement that “in order to have a meaningful comparison of models, it is necessary that they be evaluated using identical data (i.e., correlation matrices)”, but thus emphasize that this is standard procedure.

The procedure McFarland (2020) used to compare psychometric network models and latent variable models deviated from this standard procedure in several aspects. Conceptually, his aim was as follows:

- Consider a Gaussian graphical model (GGM) and one or more factor models as being competing models of intelligence.
- Fit both type of models (in standard SEM software) on the zero-order correlation matrix.
 - Obtain the fit statistics of the models (as reported by this software).
 - Compare these statistics.
- Fit both types of models (in standard SEM software) on the (pruned) partial correlation matrix derived from an exploratory network analysis.
 - Obtain the fit statistics of the models (as reported by this software).
 - Compare these statistics.

¹ This was also the case in the studies of Kan et al. (2019) and Schmank et al. (2019) that McFarland (2020) discussed. Although McFarland (2020, p. 2) stated that “from the R code provided by Schmank et al. (2019), it is clear that the fit of network models to partial correlations were compared to the fit of latent variable models to uncorrected correlation matrices”, the R code of Schmank et al. (2019, see <https://osf.io/j3cvz/>), which is an adaption of ours (Kan et al. 2019, see <https://github.com/KJKan/nwsem>) shows that this is incorrect. Correct is that the variance–covariance matrix implied by the network, $\Sigma = \Delta(I - \Omega)^{-1}\Delta$, was compared to the observed sample variance–covariance matrix, S.

Applying this procedure on Wechsler Adult Scale of Intelligence (Wechsler 2008) data, McFarland (2020) concluded that although the GGM fitted the (pruned) partial correlation matrix better, the factor models fitted the zero-order correlation matrix better.

We strongly advise against using this procedure as it will result in uninterpretable results and invalid fit statistics for a number of models. For instance, when modeling the zero-order correlation matrix, the result is invalid in the case of a GGM (i.e., a pruned partial correlation matrix is fitted): To specify the relations between observed variables in the software as covariances (standardized: zero-order correlations), while these relations in the GGM represent partial correlations, is a misrepresentation of that network model. Since the fit obtained from a misspecified model must be considered invalid, the fit statistics of the psychometric network model in McFarland's (2020) Table 1 need to be discarded.

In the case McFarland (2020) modeled partial correlations, more technical detail is required. This part of the procedure was as follows:

- Transform the zero-order correlation to a partial correlation matrix.
- Use this partial correlation matrix as the input of the R *glasso* function.
- Take the R *glasso* output matrix as the input matrix in the standard SEM software (SAS CALIS in (McFarland 2020; SAS 2010)).
- Within this software, specify the relations between the variables according to the network and factor models, respectively.
- Obtain the fit statistics of the models as reported by this software.

This will also produce invalid statistics and for multiple reasons. Firstly, the R *glasso* function expects a variance–covariance matrix or zero-order correlation as the input, after which it will return a (sparse) inverse covariance matrix. When a partial correlation matrix is used as the input, the result is an inverse matrix of the partial correlation matrix. Such a matrix is not a partial correlation matrix (nor is it a zero-order correlation matrix), hence it is not the matrix that was intended to be analyzed. In addition, even if this first step would be corrected, meaning a zero-order correlation matrix would be used as the input of R *glasso*, such that the output would be an inverse covariance matrix (standardized: a (pruned) partial correlation matrix), the further analysis would still go awry. For example, standard SEM software programs, including SAS CALIS, expect as the input either a raw data matrix or a sample variance–covariance matrix (standardized: zero-correlation matrix). Such software does not possess the means to compare partial correlation matrices or inverse covariance matrices, and is, therefore, unable to produce the likelihoods of these matrices. In short, using the output of R *glasso* as the input of SAS CALIS leads to uninterpretable results because the likelihood computed by the software is incomparable to the likelihood of the data; all fit statistics in McFarland's (2020) Table 2 are also invalid.

Only the SEM results of the factor models in McFarland's (2020) Table 1 are valid, except for the penta-factor model, as addressed below. Before we turn to this model, and speaking about model fit, we deem it appropriate to first address the role of fit in choosing between competing statistical models and so between competing theories. This role is not to *lead* but to *assist* the researcher (Kline 2015). If the analytic aim is to explain the data (rather than summarize them), any model under consideration needs to be in line with the theories being put to the test. This formed a reason as to why—in our effort to compare the mutualism approach to general intelligence with g theory—we (Kan et al. 2019) did not consider fitting a bifactor model. Bifactor models may look attractive but are problematic for both statistical and conceptual reasons (see Eid et al. 2018; Hood 2008). Decisive for our consideration not to fit a bifactor model was that such a model cannot be considered in line with g theory (Hood 2008; Jensen 1998), among others, because essential measurement assumptions are violated (Hood 2008).

Note that this decision was not intended to convey that the bifactor model cannot be useful at all, theoretically or statistically. Consider the following example. Suppose a network model is the true underlying data generating mechanism and that the edges of this network are primarily positive, then one of the implications would be the presence of a positive manifold of correlations. The presence

of this manifold makes it, in principle, possible to decompose the variance in any of the network’s variables into the following *variance components*: (1) a general component, (2) a unique component, and (3) components that are neither general nor unique (denoting variance that is shared with some but not all variables). A bifactor model may then provide a satisfactory *statistical summary* of these data. Yet, if so, it would not represent the data-generating mechanism. The model would *describe* but not *explain* the variance–covariance structure of the data.

The “penta-factor model” [McFarland \(2020\)](#) considered, which is depicted in Figure 1, is an extension of the bifactor model (see also [McFarland 2017](#)). Rather than one general factor, it includes multiple (i.e., four) general factors. As a conceptual model, we would have the same reservations as with the bifactor model.² As a statistical model, we argue against it, even when intended as a means to carry out a variance decomposition.

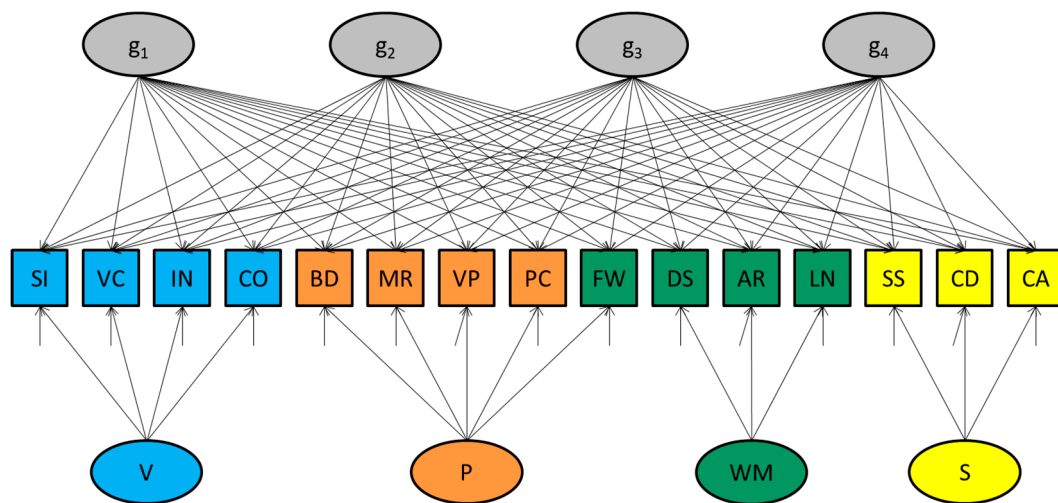


Figure 1. Graphical representation of the (nonidentified) penta-factor model. Note: abbreviations of subtests: SI—Similarities; VC—Vocabulary; IN—Information; CO—Comprehension; BD—Block Design; MR—Matrix Reasoning; VP—Visual Puzzles; FW—Figure Weights; PC—Picture Completion; DS—Digit Span; AR—Arithmetic; LN—Letter–Number Sequencing; SS—Symbol Search; CD—Coding; CA—Cancellation. Abbreviations of latent variables: V—Verbal factor; P—Perceptual factor; WM—Working Memory factor; S—Speed factor; g_1 to g_4 —general factors.

Firstly, a confirmatory factor model containing more than one general factor is not identified and produces invalid fit statistics (if at all the software reports fit statistics, rather than a warning or error). To solve the nonidentification, one would need to apply additional constraints, as is the case in exploratory factor modeling. Although SEM software other than SAS CALIS ([Rosseel 2012](#); [Jöreskog and Sörbom 2019](#)) was able to detect the nonidentification of the penta-factor model, the SAS CALIS software did not give an explicit warning or error when rerunning McFarland’s scripts. The software did reveal problems that relate to it, however. For instance, the standard errors of (specific) parameters could not be computed. This is symptomatic of a nonidentified model.

Secondly, even if we would have accepted the standard errors and fit measures that were reported as correct, the pattern of parameter values themselves indicate that the model is not what it seems

² In comparing *g* theory ([Jensen 1998](#)) with process overlap theory ([Kovacs and Conway 2016](#)), [McFarland \(2017\)](#) argued that (1) a superior fit of the penta-factor model over a hierarchical *g* model provides evidence against the interpretation of the general factor as a single, unitary variable, as in *g* theory, while (2) the penta-factor model is (conceptually speaking) in line with process overlap theory. We note here that process overlap theory is a variation of sampling theory ([Bartholomew et al. 2009](#)), in which the general factor of intelligence is a composite variable and arises, essentially, as a result of a measurement problem (see also [Kan et al. 2016](#)): cognitive tests are not unidimensional, but always sample from multiple cognitive domains. Important measurement assumptions are thus indeed being violated.

(see Table 1, which summarizes McFarland 2020's results concerning the WAIS-IV data of 20–54 years old). Inspection of the factor loadings reveals that only one of the four factors that were specified as general—namely factor g_4 —turns out to be actually general (the loadings on this factor are generally (all but one) significant). Factor g_3 turns out to be a Speed factor (only the loadings of the Speed tests are substantial and statistically significant). As a result, the theoretical Speed factor that was also included in the model became redundant (none of the loadings on this factor are substantial or significant). Other factors are also redundant, including factor g_1 (none of its loadings are substantial or significant). Furthermore, only one of the subtests has a significant loading on g_2 , which makes this factor specific rather than general.

Despite our conclusion that part of McFarland's (2020) results are invalid, we regard his efforts as seminal and inspiring, as his paper belongs to the first studies that aim to provide a solution to the issue of how to compare network and latent variable models statistically. The question becomes: Is there a valid alternative to compare psychometric network and latent variable models? In order to answer this question, we return to the fact that standard SEM programs do not possess the means to fit GGMs. This shortcoming was the motivation as to why in previous work, we (Kan et al. 2019; Epskamp 2019) turned to programming in the R software package OpenMx (Neale et al. 2016). This package allows for user-specified matrix algebraic expressions of the variance–covariance structure and is so able to produce adequate fit indices for network models. The fit statistics of these models can be directly compared to those of competing factor models. Alternatively, one could also use the recently developed and more user-friendly R package Psychometrics (Epskamp 2020b). Psychometrics has been explicitly designed to allow for both latent variable modeling and psychometric network modeling (and combinations thereof) and for comparing fit statistics of both types of models (Epskamp 2020b).

We here illustrate how a valid comparison between latent variable and network models can be conducted using this software. Codes, data, and an accompanying tutorial are available at GitHub (<https://github.com/kjkan/mcfarland>).

Table 1. McFarland’s (2020) results of the penta-factor model fitted on the WAIS-IV US standardization data of 20–54 years old.

<i>Factor</i>		<i>Verbal</i>			<i>Perceptual</i>			<i>Working</i>			<i>Processing</i>			<i>Error</i>		
		<i>Comprehension</i>			<i>Organization</i>			<i>Memory</i>			<i>Speed</i>			<i>Variance</i>		
Subtest (Abbreviation)		Est.	SE	sig.	Est.	SE	sig.	Est.	SE	sig.	Est.	SE	sig.	Est.	SE	sig.
1	Similarities (SI)	−0.3413	0.0998	***										0.2928	0.0198	***
2	Vocabulary (VC)	−0.4869	0.1317	***										0.1847	0.0243	***
3	Information (IN)	−0.7114	0.2994	*										0.0001	1	
4	Comprehension (CO)	−0.3553	0.1066	***										0.2635	0.0203	***
5	Block Design (BD)				0.5063	0.0578	***							0.3003	0.0480	***
6	Matrix Reasoning (MR)				0.0842	0.0575								0.4388	0.0312	***
7	Visual Puzzles (VP)				0.4316	0.0584	***							0.3493	0.0346	***
8	Figure Weights (FW)				0.1238	0.0507	*							0.3858	0.0255	***
9	Picture Completion (PC)				0.2638	0.0532	***							0.6145	0.0331	***
10	Digit Span (DS)							−0.3849	0.2596					0.3256	0.2446	
11	Arithmetic (AR)							0.1040	0.2586					0.0001	1	
12	Letter–Number Sequencing (LN)							−0.5303	0.4294					0.2344	0.4409	
13	Symbol Search (SS)										−0.6189	0.8171		0.0001	1	
14	Coding (CD)										−0.2186	0.1188		0.5036	0.1391	***
15	Cancelation (CA)										0.1866	0.5088		0.4025	0.3731	
<i>Factor</i>		<i>General Factor</i>			<i>General Factor</i>			<i>General Factor</i>			<i>General Factor</i>					
		<i>g1</i>			<i>g2</i>			<i>g3</i>			<i>g4</i>					
Subtest (Abbreviation)		Est.	SE	sig.	Est.	SE	sig.	Est.	SE	sig.	Est.	SE	sig.	Est.	SE	sig.
1	Similarities (SI)	0.1584	1.9584		0.0596	2.7983		−0.2318	0.5437		0.7130	0.1692	***			
2	Vocabulary (VC)	0.0951	3.0534		0.1458	1.7943		−0.2079	0.1715		0.7104	0.1032	***			
3	Information (IN)	−0.1982	4.0028		−0.2636	2.9837		−0.0322	0.4997		0.6197	0.1529	***			
4	Comprehension (CO)	0.1793	2.2841		0.0725	3.1929		−0.2698	0.6062		0.7072	0.1914	***			
5	Block Design (BD)	0.0198	1.8986		−0.1006	0.2779		0.0472	0.3089		0.6562	0.0574	***			
6	Matrix Reasoning (MR)	0.0652	3.0642		−0.2035	1.0725		−0.0564	0.6956		0.7108	0.1338	*			
7	Visual Puzzles (VP)	0.0474	2.5845		−0.1580	0.6918		0.0039	0.5324		0.6612	0.0999	***			
8	Figure Weights (FW)	−0.0264	2.4213		−0.1714	0.3925		−0.1174	0.2968		0.7450	0.0591	***			
9	Picture Completion (PC)	0.1317	2.1678		−0.0885	1.5528		0.1853	0.7582		0.5064	0.1370	***			
10	Digit Span (DS)	−0.0857	1.8393		0.1220	1.3714		0.0011	0.6273		0.7099	0.1222	***			
11	Arithmetic (AR)	−0.5003	3.5156		0.2038	7.3568		−0.1221	2.6049		0.8260	0.5251				
12	Letter–Number Sequencing (LN)	−0.1229	1.0698		0.0517	1.8421		−0.0398	0.6581		0.6819	0.1310	***			
13	Symbol Search (SS)	0.1119	0.9418		0.1328	0.5802		0.5088	0.1748	**	0.5726	0.0128	***			
14	Coding (CD)	0.0763	1.2949		0.1579	0.5028		0.2938	0.0066	***	0.5758	0.0129	***			
15	Cancelation (CA)	0.0828	1.1700		0.1277	0.0029	***	0.5838	0.0131	***	0.4458	0.0100	***			

Note: ¹ not computed; * significant at $\alpha = 0.05$; ** significant at $\alpha = 0.01$; *** significant at $\alpha = 0.001$.

2. Method

2.1. Samples

The data included the WAIS-IV US and Hungarian standardization sample correlation matrices (Wechsler 2008) that were network-analyzed by Kan et al. (2019) and Schmank et al. (2019), respectively. The sample sizes were 1800 (US) and 1112 (Hungary).

2.2. Analysis

In R (R Core Team 2020, version 4.0.2), using Psychonetrics' (version 0.7.2) default settings, we extracted from the US standardization sample a psychometric network: That is, first the full partial correlation matrix was calculated, after which this matrix was pruned (at $\alpha = 0.01$). We refer to the skeleton (adjacency matrix) of this pruned matrix as "the network model". As it turned out, the zero-order correlation matrix implied by the network model fitted the observed zero-order correlation matrix in the (US) correlation matrix well ($\chi^2(52) = 118.59, p < 0.001$; CFI = 1; TLI = 0.99; RMSEA = 0.026, CI₉₀ = 0.20–0.33), which was to be expected since the network was derived from that matrix. One may compare this to the extraction of an exploratory factor model and refitting this model back as a confirmatory model on the same sample data. Other than providing a check if a network is specified correctly or to obtain fit statistics of this model, we do not advance such a procedure (see also Kan et al. 2019).

Next, we fitted the network model (truly confirmatory) on the observed Hungarian sample. Such a procedure is comparable to the cross-validation of a factor model (or to a test for configural measurement invariance). Apart from the network model, we also fitted on the Hungarian sample matrix the factor models McFarland (2020) considered, i.e., the WAIS-IV measurement model (Figure 5.2 in the US manual (Wechsler 2008)), a second-order g factor model, which explains the variance–covariance structure among the latent variables being measured, and the bifactor model. Due to its nonidentification (see above), the penta-factor model was not taken into consideration.

The network and factor models were evaluated according to their absolute and relative fit and following standard evaluation criteria (Schermelleh-Engel et al. 2003): RMSEA values ≤ 0.05 (and TLI and CFI values > 0.97) were considered to indicate a good absolute fit, RMSEA values between 0.05 and 0.08 as an adequate fit, and RMSEA values between 0.08 and 0.10 as mediocre. RMSEA values > 0.10 were considered unacceptable. TLI and CFI values between 0.95 and 0.97 were considered acceptable. The model with the lowest AIC and BIC values was considered to provide the best relative fit, hence as providing the best summary of the data.

The comparison between the second-order g factor and the psychometric network models was considered to deliver an answer to the following question. Which among these two models provides the best theoretical explanation of the variance–covariance structure among the observed variables? Is it g theory (Jensen 1998) or a network approach towards intelligence (consistent, for example, with the mutualism theory (Van Der Maas et al. 2006))?

3. Results

The absolute and relative fit measures of the models are shown in Table 2. From this table, one can observe the following. The absolute fit of the network model was good, meaning that the network model extracted from the US sample replicated in the Hungarian sample. In addition, both the AIC and BIC values indicated a better relative fit for the psychometric network than the g -factor model. The same conclusion may be derived from the RMSEA values and their nonoverlapping 90% confidence intervals. According to the criteria mentioned above, these results would thus favor the network approach over g theory.

Table 2. Results of the model comparisons based on analyses in software package Psychonetrics.

Model	χ^2 (df)	p-Value	CFI	TLI	RMSEA [CI ₉₀]	AIC	BIC
Network	129.96 (52)	<0.001	0.99	0.99	0.037 [0.029–0.045]	36,848.56	37,189.51
Bifactor	263.41 (75)	<0.001	0.98	0.97	0.048 [0.041–0.054]	36,966.01	37,266.85
Measurement	369.47 (82)	<0.001	0.96	0.97	0.056 [0.050–0.062]	37,058.07	37,323.81
Hierarchical g	376.56 (84)	<0.001	0.97	0.96	0.056 [0.050–0.062]	37,061.16	37,316.87

Note: preferred model in bold.

The AIC and BIC values also suggested that among all models considered, the network model provided the best statistical summary, as it also outperformed the measurement model and the bifactor model. A graphical representation of the network model is given in Figure 2.

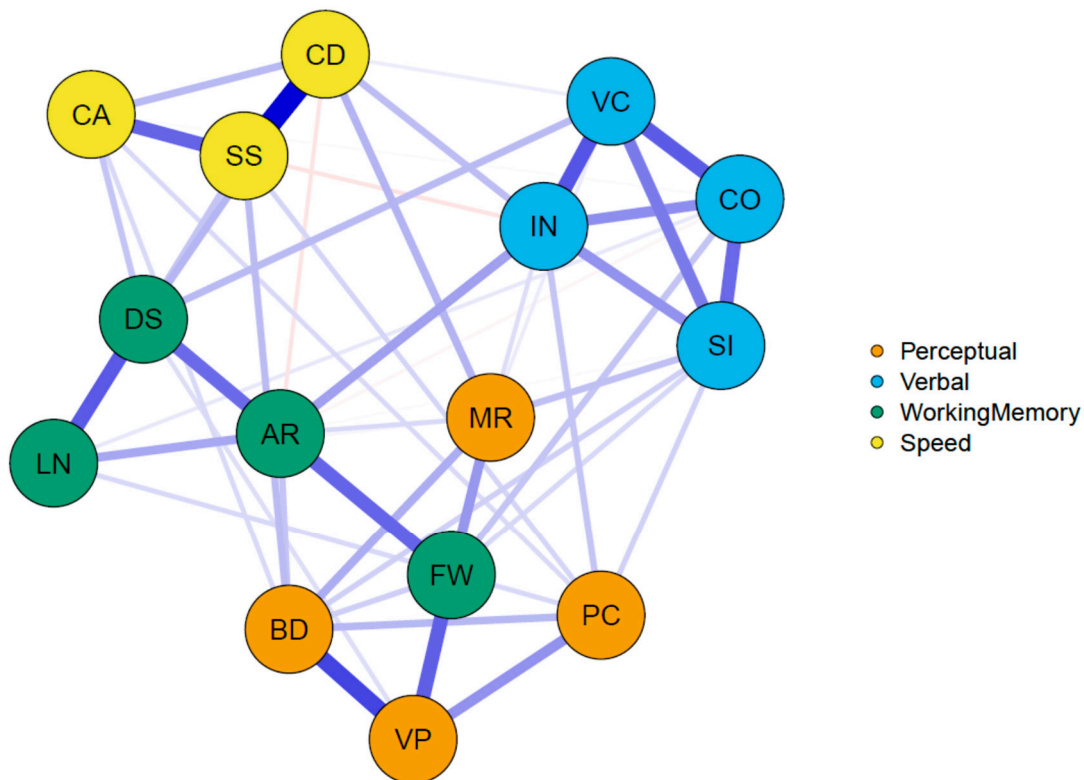


Figure 2. Graphical representation of the confirmatory WAIS network model. Note: abbreviations of Verbal subtests: SI—Similarities; VC—Vocabulary; IN—Information; CO—Comprehension; Perceptual subtests: BD—Block Design; MR—Matrix Reasoning; VP—Visual Puzzles; PC—Picture Completion; Working Memory subtests: DS—Digit Span; AR—Arithmetic; LN—Letter–Number Sequencing; FW—Figure Weights; and Speed subtests: SS—Symbol Search; CD—Coding; CA—Cancellation.

4. Conclusions

We contribute two points regarding the comparison of latent variable and psychometric network models. First, we pointed out that the penta-factor model is nonidentified and is therefore of only limited value as a model of general intelligence. Second, we proposed an alternative (and correct) method for comparing psychometric networks with latent variable models. If researchers want to compare network and latent variable models, we advance the use of the R package Psychonetrics (Epskamp 2020a, 2020b). Potential alternative ways we have not evaluated in this commentary may exist or are under development. These alternatives may include the recently developed partial correlation likelihood test (van Bork et al. 2019), for example.

The results of our (re)analysis are of interest in their own right, especially from a substantive theoretical perspective. Although one may still question to what extent the psychometric network model fitted better than factor models that were not considered here, the excellent fit for the WAIS network is definitely in line with a network interpretation of intelligence (Van Der Maas et al. 2006). The novel finding that psychometric network models of intelligence replicated over samples is also of importance, especially in view of the question of whether or not psychometric networks replicate in general. This question is a hotly debated topic in other fields, e.g., psychopathology (see Forbes et al. 2017; Borsboom et al. 2017). As we illustrated here and in previous work (Kan et al. 2019), non-standard SEM software such as OpenMx (Neale et al. 2016) and Psychonetrics (Epskamp 2020a, 2020b) provides a means to investigate this topic.

Author Contributions: Conceptualization, K.-J.K., H.L.J.v.d.M., S.Z.L. and S.E.; methodology, K.-J.K. and S.E.; formal analysis, K.-J.K. and S.E.; writing—original draft preparation, K.-J.K.; writing—review and editing, H.d.J., H.L.J.v.d.M., S.Z.L. and S.E.; visualization, K.-J.K., H.d.J. and S.E. All authors have read and agreed to the published version of the manuscript.

Funding: Sacha Epskamp is funded by the NWO Veni grant number 016-195-261.

Acknowledgments: As members of the scientific community studying intelligence, we thank McFarland for his contributions to our field. We thank Eiko Fried for bringing McFarland (2020) to our attention. We thank Denny Borsboom, Eiko Fried, Mijke Rhemtulla, and Riet van Bork for sharing their thoughts on the subject. We thank our reviewers for their valuable comments and so for the improvement of our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Bartholomew, David J., Ian J. Deary, and Martin Lawn. 2009. A new lease of life for Thomson's bonds model of intelligence. *Psychological Review* 116: 567. [CrossRef] [PubMed]
- Borsboom, Denny, Eiko I. Fried, Sacha Epskamp, Lourens J. Waldorp, Claudia D. van Borkulo, Han LJ van der Maas, and Angélique OJ Cramer. 2017. False alarm? A comprehensive reanalysis of "Evidence that psychopathology symptom networks have limited replicability" by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology* 126: 989–99. [CrossRef]
- Eid, Michael, Stefan Krumm, Tobias Koch, and Julian Schulze. 2018. Bifactor models for predicting criteria by general and specific factors: Problems of nonidentifiability and alternative solutions. *Journal of Intelligence* 6: 42. [CrossRef]
- Epskamp, Sacha. 2019. Lvnnet: Latent Variable Network Modeling. R Package Version 0.3.5. Available online: <https://cran.r-project.org/web/packages/lvnnet/index.html> (accessed on 28 September 2020).
- Epskamp, S. 2020a. Psychometric network models from time-series and panel data. *Psychometrika* 85: 1–26. [CrossRef] [PubMed]
- Epskamp, Sacha. 2020b. Psychonetrics: Structural Equation Modeling and Confirmatory Network Analysis. R Package Version 0.7.1. Available online: <https://cran.r-project.org/web/packages/psychonetrics/index.html> (accessed on 28 September 2020).
- Epskamp, Sacha, and Eiko I. Fried. 2018. A tutorial on regularized partial correlation networks. *Psychological Methods* 23: 617. [CrossRef] [PubMed]
- Epskamp, Sacha, Mijke Rhemtulla, and Denny Borsboom. 2017. Generalized network psychometrics: Combining network and latent variable models. *Psychometrika* 82: 904–27. [CrossRef] [PubMed]
- Forbes, Miriam K., Aidan GC Wright, Kristian E. Markon, and Robert F. Krueger. 2017. Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology* 126: 969–88. [CrossRef] [PubMed]
- Friedman, J., T. Hastie, and R. Tibshirani. 2019. Glasso: Graphical Lasso-Estimation of Gaussian Graphical Models. R Package Version 1. Available online: <https://cran.r-project.org/web/packages/glasso/index.html> (accessed on 28 September 2020).
- Hood, Steven Brian. 2008. *Latent Variable Realism in Psychometrics*. Bloomington: Indiana University.
- Jensen, Arthur Robert. 1998. *The g Factor: The Science of Mental Ability*. Westport: Praeger.

- Jöreskog, Karl G., and Dag Sörbom. 2019. *LISREL 10: User's Reference Guide*. Chicago: Scientific Software International.
- Kan, Kees-Jan, Han LJ van der Maas, and Rogier A. Kievit. 2016. Process overlap theory: Strengths, limitations, and challenges. *Psychological Inquiry* 27: 220–28. [[CrossRef](#)]
- Kan, Kees-Jan, Han LJ van der Maas, and Stephen Z. Levine. 2019. Extending psychometric network analysis: Empirical evidence against g in favor of mutualism? *Intelligence* 73: 52–62. [[CrossRef](#)]
- Kline, Rex B. 2015. *Principles and Practice of Structural Equation Modeling*. New York: Guilford Publications.
- Kovacs, Kristof, and Andrew R.A. Conway. 2016. Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry* 27: 151–77. [[CrossRef](#)]
- McFarland, Dennis J. 2017. Evaluation of multidimensional models of WAIS-IV subtest performance. *The Clinical Neuropsychologist* 31: 1127–40. [[CrossRef](#)] [[PubMed](#)]
- McFarland, Dennis J. 2020. The Effects of Using Partial or Uncorrected Correlation Matrices When Comparing Network and Latent Variable Models. *Journal of Intelligence* 8: 7. [[CrossRef](#)] [[PubMed](#)]
- Neale, Michael C., Michael D. Hunter, Joshua N. Pritikin, Mahsa Zahery, Timothy R. Brick, Robert M. Kirkpatrick, Ryne Estabrook, Timothy C. Bates, Hermine H. Maes, and Steven M. Boker. 2016. OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika* 81: 535–49. [[CrossRef](#)] [[PubMed](#)]
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.
- Rosseel, Yves. 2012. Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* 48: 1–36. [[CrossRef](#)]
- SAS. 2010. *SAS/STAT 9.22 User's Guide*. Cary: SAS Institute.
- Schermelleh-Engel, Karin, Helfried Moosbrugger, and Hans Müller. 2003. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research* 8: 23–74.
- Schmank, Christopher J., Sara Anne Goring, Kristof Kovacs, and Andrew RA Conway. 2019. Psychometric Network Analysis of the Hungarian WAIS. *Journal of Intelligence* 7: 21. [[CrossRef](#)] [[PubMed](#)]
- van Bork, R., M. Rhemtulla, L. J. Waldorp, J. Kruis, S. Rezvanifar, and D. Borsboom. 2019. Latent variable models and networks: Statistical equivalence and testability. *Multivariate Behavioral Research*, 1–24. [[CrossRef](#)]
- Van Der Maas, Han LJ, Conor V. Dolan, Raoul PPP Grasman, Jelte M. Wicherts, Hilde M. Huizenga, and Maartje EJ Raijmakers. 2006. A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review* 113: 842–61. [[CrossRef](#)] [[PubMed](#)]
- Wechsler, David. 2008. *Wechsler Adult Intelligence Scale*, 4th ed. San Antonio: NCS Pearson.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).