Further Insights on the French WISC–IV Factor Structure Through Bayesian Structural Equation Modeling

Golay, Philippe; Reverte, Isabelle; Rossier, Jérôme; Favez, Nicolas; Lecerf, Thierry

How to cite

GOLAY, Philippe et al. Further Insights on the French WISC–IV Factor Structure Through Bayesian Structural Equation Modeling. In: Psychological assessment, 2013, vol. 25, n° 2, p. 496–508. doi: 10.1037/a0030676

This publication URL: https://archive-ouverte.unige.ch/unige:28807

Publication DOI: 10.1037/a0030676

Further insight on the French WISC-IV factor structure through Bayesian structural equation modelling (BSEM)

Philippe Golay[1], Isabelle Reverte[1], Jérôme Rossier[2], Nicolas Favez[1] & Thierry Lecerf[1,2]

[1]Faculty of Psychology and Educational Sciences, University of Geneva

[2]Institute of Psychology, University of Lausanne

Author Note

Correspondence concerning this article should be addressed to Philippe Golay, FPSE - Psychology, University of Geneva, 40, Bd. du Pont d'Arve, CH-1205 Geneva , Switzerland, E-Mail: philippe.golay@unige.ch, Tel: +41 22 379 9237

Abstract

The Wechsler Intelligence Scale for children (WISC-IV) remains the most widely used test in the field of intelligence assessment. The interpretation of the WISC-IV is based on a 4-factor model which is only partially compatible with the Cattell-Horn-Carroll (CHC) model of intelligence measurement. Several confirmatory factor analytic studies (CFA) have shown that CHC-based models were more adequate than the 4-factor model on several cognitive batteries and in particular the Wechsler scales. Additionally, some degree of controversy also remains on the exact nature of constructs measured by each subtest.

In typical CFA, in order to prevent model under-identification, many loadings between latent variables and measures are fixed to zero. However, small but nonzero loadings could be equally compatible with theory. Inappropriate zero loadings can contribute to poor model fit, distorted factors and biased factor correlations. The goal of this study was to address these limitations in the comparison of CHC-based and classical models and thus get better insight on the constructs measured by each subtest of the French WISC-IV. We used Bayesian structural equation modelling (BSEM). With BSEM, zero-fixed loadings between latent variables and measures are replaced by approximate zeros based on informative, small-variance priors. BSEM is therefore a less restrictive approach that is still guided by theory. Results on a sample of 249 French-speaking Swiss children (8-12 yr) show that the CHC-based model is better than the 4-factors solution. Additionally, the BSEM-derived model is less complex than typical CFA-derived models and thus may have greater generalisability.

Keywords: WISC-IV, Bayesian structural equation modelling, confirmatory factor analysis, CHC theory

Further insight on the French WISC-IV factor structure through Bayesian structural equation modelling (BSEM)

The last decade has seen the emergence of the Cattell-Horn-Carroll theory as the mainstream approach in the field of intelligence assessment. This model has proven influential in the organisation and interpretations of several cognitive batteries (inserer exemples et ref). The last revision of the Wechsler intelligence scale for children, 4^{th} edition (WISC-IV) did nevertheless not include CHC theory in the definition of the four index scores (VCI, verbal comprehension index; POI, perceptual organisation index; WMI, working memory index & PSI, processing speed index). However confirmatory factor analytic studies (CFA) have shown that models based on the CHC framework showed closer fit to the data than the classical four factor structure. On the WISC-IV (exemples). A CHC-based structure also proved to show closer fit to the data on the French Wechsler adult intelligence scale (WAIS-III) (Golay & Lecerf, 2010). Nevertheless, it subsist some degree of controversy on the nature of constructs measured by each subtests. Arithmetic has been shown to load on the Gsm factor, on the Gf factor or on both. It has also been show that (continuer avec des examples). On the French version of the WAIS, it has been shown to load both on the Gf and Gsm factor (Golay & Lecerf, 2010) and on the French WISC-IV (ajouter ref sur article echantillon de standardisation).

The CFA approach

In typical CFA studies and in order to achieve model identification, f^2 restrictions (f = number of factors) have to be placed on the model parameters. Usually each variable only load on one factor. Cross-loadings are fixed to exact zeroes in order to attain the necessary number of restrictions for model identification. However, small but nonzero loadings could be equally compatible with theory, thus theses zero loadings can be considered as unnecessary strict restrictions to reflect the researchers' hypotheses (Muthén & Asparouhov, 2010). It has also been shown that inappropriate zero loadings can contribute to poor model fit, distorted factors and biased factor correlations (Marsh, et al., 2010). The latter issue has already been subject of great debate in the domain of personality research. Residual correlations within the big five measures (supposedly independent) have been referred as method variance artefact (McCrae, et al., 2008) because CFA may be overly restrictive: the independent cluster model requires each indicator to load on only one factor. This results in models with a huge number of exact zero loadings. Such models can show poor fit to the data because factor indicators may also measure many secondary factors. After rejection of the initial model because of inappropriate fit, the researcher can be tempted to engage in a series of modifications that may capitalize on chance. Modification indices can be used as a basis to improve the model by freeing parameters, one at a time. Nevertheless it can be argued that the use of modification indices to improve model fit is not a strictly confirmatory approach anymore. Because a serie of modification may be tested, one at a time, improvements may result of capitalizing on chance. Thus, these model refinements may not be generalised to the population. This is often refered as "overfitting" and it improves the risk that the final model has only sample-specific validity. Overly sample specific models can contribute to the lack of consensus between studies.

The BSEM approach

The goal of this study was to address CFA limitations in the comparison of CHC-based and classical models on the French WISC-IV. The second objective was to get much needed insight on the constructs measured by each subtest of the French adaptation of the battery. To compare different models, we used Bayesian structural equation modelling (BSEM). In BSEM, parameters are not viewed as fixed values but as variables. BSEM therefore replaces zero-fixed cross-loadings with approximate zeros based on informative, small-variance distributions. Cross-loadings are considered to be close but not necessary equal to zero. Major hypothesized loadings are freely estimated without prior information. The are refered as diffuse, non informative priors. Bayesian structural equation modelling is therefore less restrictive than CFA but still encompasses strong theory because researchers knowledge guides the analysis. Exploratory factor analysis (EFA) on the other hand may be seen as even less restrictive because the only way to incorporate theory is on the selection of the number of factors. BSEM then can be seen as an “in-between” CFA and EFA approach.

Because with BSEM all parameters are freed and estimated simultaneously, modifications can be implemented in a single step. Finally BSEM does not rely on large sample theory as Maximum-likelihood CFA estimation does and can accommodate heavily skewed distributions of parameters estimates. The difference between typical Maximum-Likelihood CFA and BSEM are presented in Table 1.

Method

Participants

249 french speaking children aged from 8 years to 12 years were assessed. AJOUTER 2-3 DETAILS.

The standards scores of the 15 subtests were used to conduct the analysis and there was no missing data. This dataset is part of a larger project described in Reverte, Favez, Rossier & Lecerf (submitted). For scaling reason in relation on the choice of prior variances, subtests standard scores were standardised.

Bayesian structural equation modeling

Muthen and Asparouhov (2010) give an excellent introduction to BSEM from a user perspective and the method used here closely draws on the approach described there. Yuab & MacKinnon also gives a very comprehensive introduction to the fundamental concept of Bayesian inference (Yuan & MacKinnon, 2009). The reader more interested in the technical implementation of BSEM can refers to the very detailed technical report (Asparouhov, 2010) of the software used for BSEM estimation (Mplus 6.1).

In typical Maximum-Likelihood (ML) estimation, parameters are viewed as constants. In BSEM, they are regarded as variables quantified as distributions (Yuan & MacKinnon, 2009). On the basis of previous studies, researchers can incorporate their prior knowledge into prior distributions. For example, on the basis of the literature, a researcher can assume that the score of a processing speed subtest (e.g *Coding*) is likely to only show a very small loading on a visualization factor. In ML-CFA, the usual practice would be to fix the loading to zero. In BSEM, the researcher will reflect this initial belief as a prior distribution. The parameter's distribution is believed to follow a normal distribution with a mean of zero and a small variance. The loading is essentially considered as close but not necessarly zero. After the experiment, new observed data is used to update researchers' knowledge on the parameters through Bayes theorem (for a detailed summary, see (Yuan & MacKinnon, 2009)). The posterior distribution is the updated representation of the researcher's belief after incorporating the experiment data. The greatest similarity between typical CFA and BSEM is that with CFA researchers also have to set cross-loadings to zero on the basis of their prior knowledge. These exact fixed zeros can be seen as normal distribution with zero mean and zero variance. Theses fixed loadings are also necessary because many restrictions have to be put on the parameters for the sake of model identification. In summary, the basic idea of BSEM is to acknowledge that there is little uncertainly on the value of the parameters. A prior distribution, with a defined mean of and variance, will therefore better reflect theory than a fixed parameter. Fixed zeros are considered as unnecessary strict operationalization of the researcher's hypotheses.

Informative small variance priors

The BSEM implementation follows the same logic of typical ML-CFA with two additional steps. For the WISC-IV, the researcher starts with defining which subtests are to be associated to which factors. For example, Coding, Symbols and Cancellation are set to load on the processing speed factor (PSI). In typical CFA, it would be assumed that the loadings of every other subtest on the PSI factor are zero. Within BSEM, cross-loadings are specified to follow a prior distribution of mean zero and a moderate variance. This prior is informative in the sense that it implies that the cross loading value is small or close to zero but not exactly zero. Using informative, small-variance priors for all cross-loadings brings information into the analysis which would be unidentified otherwise (Muthén & Asparouhov, 2010). The choice of the prior variance is a reflect of initial beliefs and knowledge. A very small variance may not allow some cross-loadings to sufficiently differ from zero. On the opposite, a too large variance would not contribute enough information so that the model gets closer to being non-identified (Muthén & Asparouhov, 2010). In such cases, the MCMC estimation algorithm will fail to reach convergence. The prior variance that was chosen is 0.04 which results in 95% confidence internal of ± 0.39 (figure 1) which allow small to moderate loadings. We considered it was important to allow cross-loadings to sufficiently escape of zero in order to identify potential significant cross-loadings. We also followed Muthen and Asparouhov recommendation of varying the variance priors to study the sensitivity of the results. Smaller values (0.1 & 0.3) showed similar pattern of results but at the cost of slightly worst model fit

(PPP). With larger values (0.5), the MCMC process failed to reach convergence (model unidentification).

Markov Chain Monte Carlo and convergence

The posterior distribution of Bayesian estimation was achieved through Markov Chain Monte Carlo (MCMC) algorithm with the Gibbs sampler Method (Muthen & Muthen, 2010). "*The idea behind MCMC is that the conditional distribution of one set of parameters given other sets can be used to make random draws of parameters values, ultimately resulting in an approximation of the joint distribution of all parameters*" (Muthén & Asparouhov, 2010). Three MCMC chains with 50'000 iterations were used, with different starting values and different random seeds. Theses chains were used to monitor convergence. Convergence was assessed using the Gelman-Rubin convergence diagnostic (Gelman, Carlin, Stern, & Rubin, 2004; Gelman & Rubin, 1992). The potential scale reduction factor (PSR) is computed and takes into account between- and within-chain variation. Convergence is achieved when PSR is comprised between 1 and 1.1. Such PSR values indicates that the between chain variation is small relative to the within-chain variation. Convergence did not proved to be an issue with the tested models. We checked that the PSR values were already down to convergence (<1.1) after the first half (25'000) of the iterations to insure that the chosen number of iteration was sufficiently large. The first half of the chain was discarded as a burnin phase (Muthen & Muthen, 2010) and the second part was used to estimate the posterior distribution.

Comparison of model fit

Model fit was assessed using Posterior Predictive Checking (PPC) (Gelman, Meng, & Stern, 1996). The Posterior Predictive p-value (PPP) of model fit is computed and can be used to test the structural model for misspecification. A small positive value (e.g 0.004) indicates poor fit. A PPP value around 0.5 indicates excellent fit (Muthén & Asparouhov, 2010). Contrarily to standard fit indexes such as the Root Mean Square Error of Estimation (RMSEA), there is no clear-cut PPP value that may indicate whether model fit is acceptable or not. Therefore PPP is more to be interpreted like a structural equation modeling fit index (bigger PPP indicates the better model). We additionally used the Deviance Information Criterion (DIC) (Gelman, et al., 2004; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002). . The DIC is a Bayesian generalization of the AIC which balance the largeness of the likelihood and adds a penalty for model complexity (number of parameters). The number of parameters that is used to penalize for model complexity in the DIC if the effective number of parameter, referred as pD (Asparouhov & Muthén, 2010). Models with smaller values of DIC are to be preferred.

BSEM tested models

We first tested models based on the classical 4-factor solution. Model 1 was a correlated four-factor structure (VCI, POI, WMI, PSI). *Similarities, Vocabulary, Comprehension, Information* and *Word Reasoning* were placed on the VCI factor. *Block Design, Picture Completion, Matrix Reasoning* and *Picture concept* were placed on POI. *Digit Span, Letter Number* and *Arithmetic* were put on WMI and *Coding, Symbol Search* and *Cancellation* on the PSI factor. Cross-loadings where first fixed as zero to be in line with typical CFA practice. In Model 2, cross-loadings were allowed to differ from zero and the variance of priors were set to 0.4. The objective was to assess gains in term of model fit that could be achieved when cross-loadings were freely estimated instead of being fixed to zero. We also tested higher order models that are more compatible with the existence of a general factor. Higher order models are considered more restrictive because they assume a unitary source of factor correlation. First order factors (VCI, POI, WMI, PSI) were defined the same way as in model 1 and 2. Model 3 (figure 2) did not allow cross-loadings to differ from zero and in Model 4 they were replaced by zero-mean small variance (0.4) priors.

The next step is to compare the four factor model to a CHC-based five factor model. On model 5, *Similarities, Vocabulary, Compehension, Information, Word Reasoning* were placed on the *Gc* factor. *Block design* and *Picture Completion* were put on the *Gv* factor. *Matrix Reasoning*, *Picture Concept* and *Arithmetic* were placed on the *Gf* factor. (ajouter justification du crossloading parce que Gf, RG). *Arithmetic* was also placed on the *Gsm* factor with *Digit Span* and *Letter Number*. Finally *Coding*, *Symbol Search* and *Cancellation* were placed on the *Gs* factor (figure 3).

Parameters are considered to have substantive backing when the 95% credibility interval of the parameter does not cover zero (Muthén & Asparouhov, 2010). Significant cross-loadings can then be freely estimated while keeping small variance priors for other cross-loadings in a final run. Hypothesized major loadings that failed to reach significance can also be modified at this time. They are not set to zero as it is typically done with CFA but they are rather replaced by an informative small variance cross-loadings that still allows the variable to show small non-zero loading on the factor. The goal of the estimation of this second and final model is to get slightly better estimation of the parameters distributions. Slight positive changes in model fit may occur and can be expected but are not of prime interest. Because every parameters are freed and estimated at the same time in both models, the base and the final model basically have the same meaning.

Results

The first WISC correlated four-factor structure was tested first. The first model was clearly rejected (ppp = 0.004). When cross-loading were allowed to differ from zero (model 2) the fit dramatically increased (PPP= 0.464). The value of DIC in model 2 (9617.454) was also lower than in model 1 (9645.974) indicating that freely estimated cross-loading were to be preferred. Model 3 was a higher-order variant of model 1. It was also rejected because of very low PPP value (0.005). The value of the DIC (9645.043) was also very close of model 1. Once again, by allowing cross-loadings to be freely estimated, model fit greatly increased (model 4) as indicated by a PPP of 0.456 and DIC of 9616.227. On both model 2 and 4, the 95% confidence interval of the loading of Picture concept on the POI factor failed to exclude zero (95% C.I from 0.076 to 0.421 and -0.055 to 0.417). However, and still in both model 2 and 4, the loading of Picture concept on the Gsm factor did exclude zero (95% C.I from 0.007 to 0.343 and 0.001 to 0.329), suggesting that Picture Concept may first of all measure working memory and not perceptual reasoning. Additionally, we noticed that the correlations between the 4 factors were substantially reduced when the cross-loadings were freely estimated. This indicates that the relation between a specific item and a non-target factor that would be accounted by a cross-loading was represented instead through the factor correlation between the two factors (Marsh, et al., 2010). The correlations between the 4 factor in model 1 and 3 are likely to have been inflated.

On model 2, PSI did not showed substantive correlation with the three other factors (the 95% credibility interval did cover zero). We observed the same on model 4 were the 95% confidence interval of the loading of PSI on g also failed to exclude zero. This suggests that, when adopting a 4-factor framework, PSI is indeed a poor estimate of the g factor.

Next we compared the CHC-base model (model 5) and the same model with small variance priors on the cross-loadings (model 6). Although better than both the WAIS 4-factor models, the model 5 showed poor fit (PPP = 0.028, DIC = 9632.226). On this model, Arithmetic got significant loadings on both Gf and Gsm. Free estimation of the cross-loadings (model 6) greatly increased the fit (PPP = 0.529, DIC = 9601.477). Detailed results are provided on table 3. All hypothesized major loadings except one got substantive backing because the 95% confidence interval excluded zero. Nonetheless, the loading of Arithmetic on the Gf factor was very low (0.115, 95% C.I from -0.319 to -0.519). This suggests that Arithmetic is not a measure of Gf in the French WISC-IV. No other cross-loading could be considered as substantive because the 95% C.I always included zero. Theses non-zero cross-loadings do not have much clinical significance (because their magnitude is small) but they are nevertheless statistically important for correct and non-biased estimation of the model parameters. Furthermore, saturations of first order factors were reduced in comparison of model 5 suggesting that inappropriate fixed zero increased first order factors correlations and thus loadings on the g factor. Décrire saturation des 5 facteurs sur g.

The final step was to estimate a modified and final CHC model were the free loading of Arithmetic on the Gf factor was removed. Once again, we first ran a first model in line with typical CFA practice (no cross-loadings, model 7) and a derivate with small variance priors (model 8). The loading of Arithmetic on the Gf factor was fixed to zero in model 7 and replaced by a non-zero small variance cross-loading in model 8 to still allow for small loading on the Gf factor. The first important thing to notice is that the fit of model 7 (PPP = 0.008, DIC = 9641.561) is worst than model 5 (base CHC model with arithmetic on both gsm and gf). This suggest that, with a CFA approach, the researcher would probably have kept Arithmetic on both Gsm and Gf because of better fit, resulting in a more complicated model.

The estimates of the final CHC model with cross loading (model 8) are presented in table 4. The fit slightly increased (PPP = 0.568, DIC = 9585.435), and the same pattern of substantive loadings could be seen. No major hypothesized loading were rejected and again, no cross-loadings got substantive enough backing to be freely estimated as a major loading. The goal of this final estimation was to get better parameters estimates. The meaning of the model 6 and 8 is still basically the same because Arithmetic is allowed to slightly load on the Gf factor in both the base and final model. It was nevertheless considered that freely estimating the loading of arithmetic on Gf gives too much probability that it takes an exaggerated value. The value of this loading on model 8 is somewhat smaller than in model 6. It is important to empathise that the typical CFA approach (as represented in model 5 and 7) suggests adopting a more complicated model (with Arithmetic loading on two factors) while the BSEM approach (model 6 and 8) suggest a more simple model (Arithmetic is only a measure of Gsm).

Discussion

Results show reasonably small cross-loading values. These results support a parsimonious final model. Even if the value that was chosen for the prior variance was not particularly small, no cross-loadings took exaggerated or significant values in model 6 and 8.

Comparaison avec article Reverte : 3 différences :

1. Small cross-loadings instead of exact zeroes
    a. Not necessarily clinical significance but
    b. Allow a better model estimation without factor distortion and correlation bias

2. 3 cross loadings en moins, gf sur ari (comme illustré) mais aussi gs sur les cubes et matrice sur gv. Rappeler que les liens entre construits ne passent pas que par des cross-loadings mais également à travers les correlations entre facteurs => pas necessaire d’ajouter trop de fleches (c’est toute la difference entre une matrice pattern et une matrice de structure, la première pour être correctement interprétée necessite une prise en compte des correlations entre facteurs mais retranscris de façon plus proche la conception theorique du chercheur).
3. Saturation de gf sur le facteur g revue à la baisse car les fixed zero font augmenter les correlation entre facteurs et donc les saturations sur le facteur g. Lorsqu’on en tient compte, le biais est suprimé et les saturations sont un peu réduites. Citer article puissance stat permet pas forcement de distinguer g et gf pour des raisons de puissance statistique (Matzke, Dolan, & Molenaar, 2010). Les inappropriate zero loadings sont une explication supplémentaire parce qu’elles poussent les saturations des facteurs de premier ordre sur g vers le haut.

References

Muthén, L.K. and Muthén, B.O. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén

Asparouhov, T. (2010). Bayesian analysis using Mplus. Technical appendix. Los Angeles: Muthén & Muthén.

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis of latent variable models using Mplus*: Technical report in preparation. Los Angeles: Muthén & Muthén. www.statmodel.com.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2004). Bayesian data analysis. . *Boca Rato, FL: Chapman & Hall/CRC*.

Gelman, A., Meng, X., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica, 6*, 733-759.

Gelman, A., & Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical science, 7*(4), 457-472.

Golay, P., & Lecerf, T. (2010). Orthogonal higher order structure and confirmatory factor analysis of the french wechsler adult intelligence scale (wais-iii). *Psychological assessment*.

Marsh, H., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A., Trautwein, U., et al. (2010). A New Look at the Big Five Factor Structure Through Exploratory Structural Equation Modeling. *Psychological Assessment, 22*(3), 471-491.

Matzke, D., Dolan, C. V., & Molenaar, D. (2010). The Issue of Power in the Identification of" g" with Lower-Order Factors. *Intelligence, 38*(3), 9.

McCrae, R., Yamagata, S., Jang, K., Riemann, R., Ando, J., Ono, Y., et al. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of personality and social psychology, 95*(2), 442.

Muthén, B., & Asparouhov, T. (2010). Bayesian SEM: A more flexible representation of substantive theory. *Manuscript in preparation*.

Spiegelhalter, D., Best, N., Carlin, B., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583-639.

Yuan, Y., & MacKinnon, D. (2009). Bayesian mediation analysis. *Psychological methods, 14*(4), 301-322.

Table 1. Summary of Maximum Likelihood CFA and BSEM principal differences

	CFA	BSEM
Parameters viewed as	Constants	Variables
Cross-loading	Exact zeros	Informative priors (zero mean and small variance)
Major loadings	Freely estimated	Diffuse non informative priors (zero mean and infinite variance)
Model modification	Improvement with modification indices one parameter at a time.	All parameters freed and estimated simultaneously. Modification in a single step
Parameters estimates	Assumed to be normally distributed	Based on percentiles of the posterior distribution, does not assume a normal distribution

Table 2. *Comparisons of Model fit for the French WISC-IV Structure*

Model	Number of free parameters	Posterior Predictive P-Value	Difference between observed & replicated X^2 95% C.I.		DIC	Estimated number of parameters (pD)
			Lower	Upper		
1. WISC-IV (4 factors)	51	0.004	13.776	89.296	9645.974	48.802
2. WISC-IV (4 factors) with cross-loadings (variance of priors = 0.4)	96	0.464	-40.029	42.186	9617.454	70.800
3. WISC-IV (higher order)	49	0.005	13.262	90.287	9645.043	47.335
4. WISC-IV (higher order) with cross-loadings (variance of priors = 0.4)	94	0.456	-39.380	43.376	9616.227	68.984
5. CHC-base model	51	0.028	-0.993	75.815	9632.226	49.016
6. CHC-base model with cross-loadings (variance of priors = 0.4)	110	0.529	-43.061	38.097	9601.477	58.621
7. CHC -final model	50	0.008	9.469	85.211	9641.561	48.265
8. CHC-final model with cross-loadings (variance of priors = 0.4)	110	0.568	-45.553	36.628	9585.435	44.432

Table 2. *Bayesian analysis using informative, small-variance priors for cross-loadings – CHC base model*

		Estimate	Posterior S.D	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%
Gc						
	Similarities	0.631	0.085	0.000	0.461	0.791
	Vocabulary	0.806	0.080	0.000	0.656	0.976
	Comprehension	0.722	0.084	0.000	0.569	0.901
	Information	0.743	0.084	0.000	0.586	0.917
	Word Reasoning	0.661	0.079	0.000	0.503	0.817
	Block Design	-0.069	0.098	0.238	-0.260	0.127
	Picture Completion	0.107	0.098	0.138	-0.088	0.299
	Matrix Reasoning	0.039	0.109	0.355	-0.183	0.246
	Picture Concept	0.071	0.102	0.246	-0.137	0.260
	Digit Span	-0.025	0.091	0.390	-0.209	0.151
	Letter Number Sequencing	0.023	0.097	0.405	-0.173	0.212
	Arithmetic	0.128	0.094	0.088	-0.064	0.308
	Coding	-0.078	0.088	0.184	-0.251	0.095
	Symbol Search	0.002	0.089	0.492	-0.174	0.178
	Cancellation	0.054	0.083	0.255	-0.109	0.217
Gv						
	Similarities	0.130	0.090	0.065	-0.037	0.318
	Vocabulary	-0.061	0.087	0.238	-0.234	0.112
	Comprehension	-0.117	0.088	0.086	-0.298	0.052
	Information	0.136	0.090	0.055	-0.031	0.324
	Word Reasoning	0.045	0.087	0.298	-0.126	0.216
	Block Design	0.665	0.128	0.000	0.430	0.929
	Picture Completion	0.530	0.133	0.000	0.287	0.810
	Matrix Reasoning	0.139	0.114	0.111	-0.094	0.356
	Picture Concept	-0.037	0.103	0.351	-0.256	0.153
	Digit Span	-0.083	0.099	0.192	-0.283	0.108
	Letter Number Sequencing	0.032	0.105	0.374	-0.171	0.244
	Arithmetic	-0.004	0.098	0.485	-0.198	0.191
	Coding	-0.007	0.101	0.474	-0.202	0.199
	Symbol Search	0.106	0.102	0.139	-0.089	0.313
	Cancellation	0.017	0.096	0.425	-0.163	0.213
Gf						
	Similarities	0.045	0.081	0.234	-0.071	0.248
	Vocabulary	0.010	0.073	0.426	-0.130	0.174
	Comprehension	-0.007	0.073	0.449	-0.166	0.136
	Information	-0.021	0.078	0.354	-0.208	0.109
	Word Reasoning	0.031	0.079	0.300	-0.091	0.228
	Block Design	0.004	0.082	0.474	-0.181	0.162
	Picture Completion	0.013	0.078	0.409	-0.128	0.195
	Matrix Reasoning	0.500	0.233	0.018	0.048	0.920
	Picture Concept	0.385	0.179	0.000	0.083	0.756
	Digit Span	0.022	0.080	0.351	-0.114	0.215
	Letter Number Sequencing	0.001	0.078	0.490	-0.155	0.173
	Arithmetic	0.115	0.258	0.260	-0.319	0.519
	Coding	0.001	0.076	0.491	-0.154	0.161
	Symbol Search	-0.007	0.077	0.446	-0.169	0.148
	Cancellation	0.015	0.075	0.392	-0.124	0.184
Gsm						
	Similarities	0.018	0.080	0.410	-0.135	0.178
	Vocabulary	0.059	0.080	0.222	-0.097	0.221
	Comprehension	0.049	0.081	0.263	-0.106	0.215
	Information	0.004	0.079	0.482	-0.149	0.163
	Word Reasoning	0.011	0.077	0.443	-0.143	0.163

	Block Design	0.037	0.098	0.347	-0.164	0.221
	Picture Completion	-0.098	0.091	0.124	-0.290	0.069
	Matrix Reasoning	0.023	0.099	0.404	-0.182	0.212
	Picture Concept	0.096	0.092	0.146	-0.090	0.276
	Digit Span	0.642	0.112	0.000	0.435	0.873
	Letter Number Sequencing	0.749	0.123	0.000	0.528	1.011
	Arithmetic	0.400	0.196	0.002	0.173	0.690
	Coding	0.040	0.092	0.327	-0.133	0.231
	Symbol Search	0.035	0.090	0.345	-0.141	0.215
	Cancellation	-0.131	0.085	0.053	-0.304	0.029
Gs						
	Similarities	0.000	0.070	0.497	-0.138	0.139
	Vocabulary	0.023	0.071	0.374	-0.116	0.162
	Comprehension	-0.007	0.072	0.456	-0.151	0.134
	Information	-0.003	0.071	0.485	-0.141	0.137
	Word Reasoning	-0.036	0.070	0.298	-0.178	0.099
	Block Design	0.126	0.085	0.069	-0.041	0.291
	Picture Completion	0.002	0.079	0.491	-0.157	0.153
	Matrix Reasoning	-0.028	0.080	0.356	-0.193	0.124
	Picture Concept	0.041	0.075	0.292	-0.109	0.188
	Digit Span	0.021	0.081	0.397	-0.137	0.181
	Letter Number Sequencing	-0.111	0.082	0.081	-0.278	0.045
	Arithmetic	0.106	0.074	0.074	-0.037	0.254
	Coding	0.630	0.089	0.000	0.456	0.809
	Symbol Search	0.657	0.095	0.000	0.483	0.861
	Cancellation	0.494	0.087	0.000	0.326	0.668
G						
	Gc	0.725	0.138	0.000	0.352	0.915
	Gv	0.694	0.177	0.003	0.248	0.937
	Gf	0.880	0.241	0.014	0.285	0.996
	Gsm	0.623	0.169	0.002	0.236	0.905
	Gs	0.386	0.200	0.047	-0.072	0.706

Table 3. *Bayesian analysis using informative, small-variance priors for cross-loadings – CHC final model*

		Estimate	Posterior S.D	One-Tailed P-Value	95% C.I.	
					Lower 2.5%	Upper 2.5%
Gc						
	Similarities	0.624	0.085	0.000	0.455	0.783
	Vocabulary	0.797	0.081	0.000	0.647	0.969
	Comprehension	0.719	0.086	0.000	0.564	0.903
	Information	0.750	0.088	0.000	0.589	0.934
	Word Reasoning	0.652	0.080	0.000	0.492	0.811
	Block Design	-0.066	0.098	0.246	-0.257	0.130
	Picture Completion	0.108	0.097	0.132	-0.085	0.297
	Matrix Reasoning	0.050	0.106	0.317	-0.167	0.251
	Picture Concept	0.048	0.102	0.322	-0.161	0.239
	Digit Span	-0.032	0.090	0.355	-0.212	0.141
	Letter Number Sequencing	0.018	0.096	0.423	-0.173	0.205
	Arithmetic	0.153	0.084	0.034	-0.011	0.320
	Coding	-0.078	0.087	0.181	-0.250	0.092
	Symbol Search	0.004	0.088	0.484	-0.171	0.179
	Cancellation	0.051	0.083	0.263	-0.111	0.214
Gv						
	Similarities	0.125	0.090	0.072	-0.041	0.315
	Vocabulary	-0.061	0.089	0.239	-0.236	0.116
	Comprehension	-0.114	0.089	0.092	-0.296	0.056
	Information	0.145	0.091	0.045	-0.023	0.338
	Word Reasoning	0.041	0.088	0.313	-0.132	0.215
	Block Design	0.670	0.131	0.000	0.432	0.943
	Picture Completion	0.522	0.135	0.000	0.278	0.806
	Matrix Reasoning	0.150	0.112	0.090	-0.079	0.366
	Picture Concept	-0.062	0.102	0.260	-0.281	0.125
	Digit Span	-0.090	0.099	0.171	-0.290	0.100
	Letter Number Sequencing	0.027	0.103	0.392	-0.174	0.236
	Arithmetic	0.023	0.090	0.393	-0.148	0.207
	Coding	-0.006	0.102	0.476	-0.203	0.200
	Symbol Search	0.108	0.103	0.134	-0.088	0.317
	Cancellation	0.017	0.096	0.428	-0.166	0.214
Gf						
	Similarities	0.071	0.088	0.176	-0.072	0.270
	Vocabulary	0.022	0.082	0.377	-0.134	0.199
	Comprehension	-0.011	0.083	0.436	-0.185	0.151
	Information	-0.046	0.086	0.266	-0.241	0.104
	Word Reasoning	0.054	0.088	0.236	-0.092	0.257
	Block Design	0.002	0.095	0.489	-0.209	0.177
	Picture Completion	0.027	0.089	0.364	-0.137	0.223
	Matrix Reasoning	0.469	0.204	0.004	0.103	0.895
	Picture Concept	0.476	0.166	0.000	0.176	0.810
	Digit Span	0.031	0.089	0.341	-0.131	0.230
	Letter Number Sequencing	0.001	0.089	0.496	-0.176	0.189
	Arithmetic	0.020	0.085	0.392	-0.145	0.199
	Coding	0.002	0.086	0.490	-0.172	0.180
	Symbol Search	-0.016	0.086	0.416	-0.190	0.160
	Cancellation	0.027	0.085	0.356	-0.135	0.207
Gsm						
	Similarities	0.017	0.081	0.414	-0.139	0.180
	Vocabulary	0.060	0.081	0.222	-0.098	0.225
	Comprehension	0.053	0.082	0.248	-0.103	0.221
	Information	0.012	0.081	0.438	-0.143	0.175
	Word Reasoning	0.010	0.079	0.449	-0.147	0.165

	Block Design	0.039	0.099	0.343	-0.164	0.225
	Picture Completion	-0.098	0.091	0.125	-0.292	0.069
	Matrix Reasoning	0.032	0.098	0.373	-0.174	0.216
	Picture Concept	0.080	0.094	0.197	-0.114	0.259
	Digit Span	0.646	0.111	0.000	0.444	0.877
	Letter Number Sequencing	0.748	0.123	0.000	0.530	1.016
	Arithmetic	0.440	0.100	0.000	0.251	0.642
	Coding	0.042	0.093	0.323	-0.134	0.235
	Symbol Search	0.039	0.091	0.332	-0.140	0.221
	Cancellation	-0.134	0.086	0.053	-0.309	0.030
Gs						
	Similarities	-0.001	0.071	0.492	-0.138	0.139
	Vocabulary	0.023	0.071	0.371	-0.115	0.164
	Comprehension	-0.006	0.072	0.464	-0.149	0.136
	Information	-0.001	0.071	0.496	-0.138	0.141
	Word Reasoning	-0.036	0.070	0.300	-0.178	0.099
	Block Design	0.125	0.086	0.072	-0.045	0.291
	Picture Completion	0.002	0.079	0.489	-0.156	0.153
	Matrix Reasoning	-0.024	0.079	0.376	-0.185	0.127
	Picture Concept	0.035	0.077	0.322	-0.118	0.184
	Digit Span	0.017	0.081	0.414	-0.141	0.177
	Letter Number Sequencing	-0.113	0.082	0.075	-0.280	0.042
	Arithmetic	0.114	0.074	0.058	-0.028	0.261
	Coding	0.629	0.090	0.000	0.455	0.809
	Symbol Search	0.657	0.096	0.000	0.482	0.864
	Cancellation	0.494	0.087	0.000	0.326	0.669
G						
	Gc	0.720	0.154	0.000	0.247	0.911
	Gv	0.686	0.182	0.004	0.225	0.931
	Gf	0.860	0.173	0.003	0.350	0.994
	Gsm	0.643	0.157	0.001	0.274	0.891
	Gs	0.385	0.204	0.050	-0.080	0.711

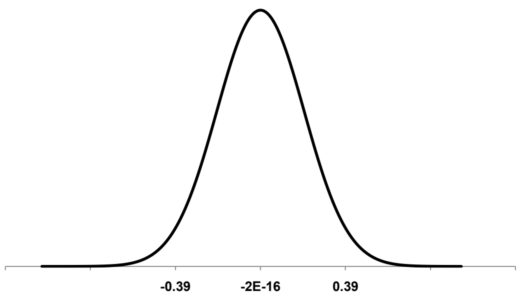
Figure captions

Figure 1. Informative normal prior for cross-loadings parameters.

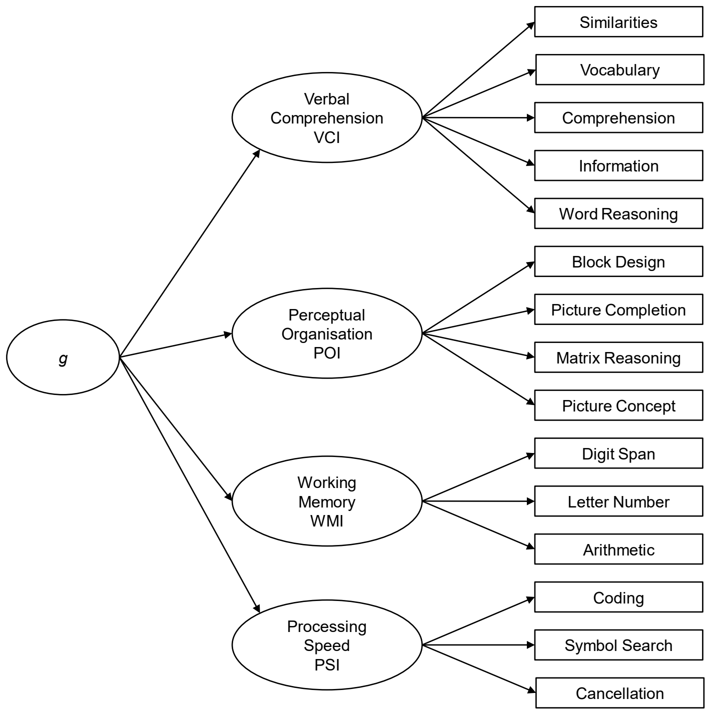
Figure 2. *Four factor higher-order model for the French WISC-IV (models 3 & 4)*

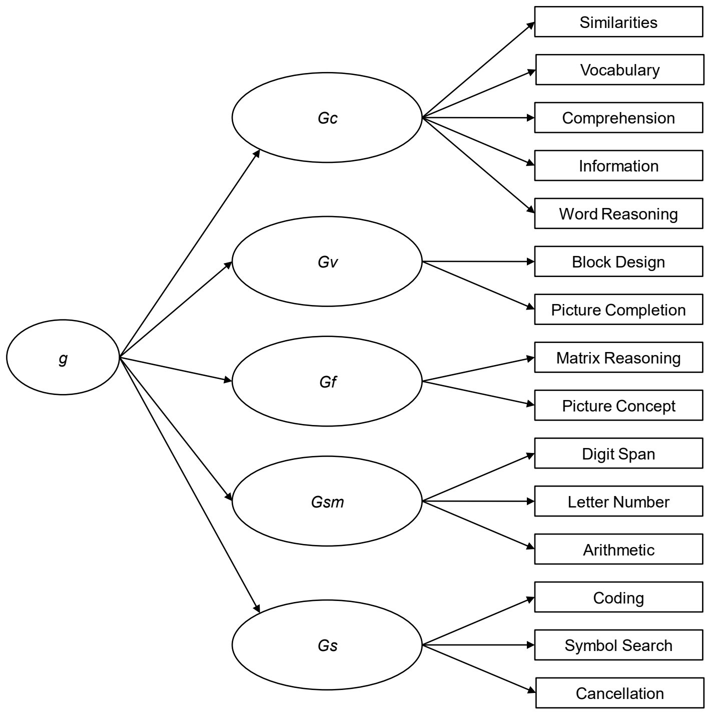
Figure 3. *CHC-based final model for the French WISC-IV (models 7 & 8)*

Prior distribution



N(0, 0.04)
95% C.I. [-0.39;+0.39]


g
Verbal
Comprehension
VCI
Perceptual
Organisation
POI
Working
Memory
WMI
Processing
Speed
PSI
Similarities
Vocabulary
Comprehension
Information
Word Reasoning
Block Design
Picture Completion
Matrix Reasoning
Picture Concept
Digit Span
Letter Number
Arithmetic
Coding
Symbol Search
Cancellation



g
Gc
Gv
Gf
Gsm
Gs
Similarities
Vocabulary
Comprehension
Information
Word Reasoning
Block Design
Picture Completion
Matrix Reasoning
Picture Concept
Digit Span
Letter Number
Arithmetic
Coding
Symbol Search
Cancellation