Supplementary File B

1 Factor analysis formulation

One of the goals of factor analysis methods is to quantify the association of each variable with more fundamental entities that explain the correlations among observations [1]. Let the vector $\mathbf{y_i} \in \mathbb{R}^{m\times 1}$ be the i-th observation of m variables, and $\mathbf{z_i} \in \mathbb{R}^{p\times 1}$ the unobserved measures of p factors. The *common factor model* states:

$$\mathbf{y_i} = \mathbf{F}\mathbf{z_i} + \mathbf{e_i} \tag{1}$$

Or component-wise:

$$y_{ij} = \sum_{k=1}^{m} f_{jk} z_{ik} + e_{ij} \tag{2}$$

Matrix $\mathbf{F} \in \mathbb{R}^{m\times p}$ contains the *factor loadings*, also known as the factor structure. Residuals $e_{\cdot j} \in \mathbb{R}^{1\times m}$ contain the portion of the j-th variable that is not defined by the factors and matrix of residuals correlations $\mathbf{C_e} \in \mathbb{R}^{m\times m}$ is assumed diagonal. Therefore, components of the correlation matrix between observed variables $\mathbf{C_y} \in \mathbb{R}^{m\times m}$ are given by:

$$\begin{aligned} [\mathbf{C_y}]_{jl} &= \mathrm{corr}(y_{\cdot j}, y_{\cdot l}) \\ &= \sum_{s=1}^{m}\sum_{t=1}^{m} f_{js} f_{lt} [\mathbf{C_z}]_{st} \quad \text{for } j \neq l \end{aligned} \tag{3}$$

Exploratory Factor Analysis (EFA) consists in estimating the matrices $\mathbf{F}$, $\mathbf{C_z}$, and $\mathbf{C_e}$ to find out an underlying factor structure, while Confirmatory Factor Analysis (CFA) estimates the coefficients for an hypothesized model from Equation 3. The resulting factor structure matrix $\mathbf{F} \in \mathbb{R}^{m\times p}$ quantifies the influence of the p unobserved factors over the m observed variables.

1.1 Factor scores

Given the factor model, $\mathbf{F}$, $\mathbf{C_z}$, and $\mathbf{C_e}$, the unobserved factor scores $\mathbf{z}$ can be estimated with various methods [2] that are based on the linear relations between factors $\mathbf{z_i}$ and standardized observations $\mathbf{y_i}$. That is, the estimated vector of factors, $\widehat{\mathbf{z_i}}$, for a particular case i is estimated as:

$$\widehat{\mathbf{z_i}} = \mathbf{W}\mathbf{y_i} \tag{4}$$

The *weight matrix* $\mathbf{W} \in \mathbb{R}^{p\times m}$ that minimizes the sum of squares of the uniqueness [3], *i.e.*, the portion of the observations variance that is not explained by the factors, is given by:

$$\mathbf{W} = \left(\mathbf{F}^T \mathbf{C_e}^{-1} \mathbf{F}\right)^{-1} \mathbf{F}^T \mathbf{C_e}^{-1} \tag{5}$$

where $\mathbf{C_e}$ is the residuals covariance matrix.

2 Factor analysis in this work

Inputs:

- $\mathbf{y_i} \in \mathbb{R}^{35\times 1}$: vector of observed sub-scores transformed into standardized regression based (SRB) z-scores for each subject i.
- $\mathbf{C_y} \in \mathbb{R}^{35\times 35}$: matrix of correlations between SRB z-scores.

Outputs:

- $\mathbf{F} \in \mathbb{R}^{35\times 6}$: factor structure matrix quantifying the influence of each domain in the observed SRB z-scores.
- $\mathbf{W} \in \mathbb{R}^{6\times 35}$: matrix of weights for domain score calculation.
- $\mathbf{z_i} \in \mathbb{R}^{6\times 1}$: vector of un-observed domain scores for each subject i.

Table 1: Weight of each sub-score in the calculation of the 6 domain scores (Matrix $\mathbf{W}^T$). Complete description of each sub-score code is presented in Supplementary File A.

	Domain					
Sub-score	Memory	Language	Executive	Visuospatial	Orientation	Attention
Q1SCORE	0.126	0	0	0	0	0
Q4SCORE	0.125	0	0	0	0	0
MOCADLREC	0.061	0	0	0	0	0
RAVLT.IMMED	0.15	0	0	0	0	0
AVTOT6	0.128	0	0	0	0	0
AVTOTB	0.052	0	0	0	0	0
AVDEL30MIN	0.068	0	0	0	0	0
AVDELTOT	0.052	0	0	0	0	0
LIMMTOTAL	0.085	0	0	0	0	0
LDELTOTAL	0.113	0	0	0	0	0
MMRECALL	0.039	0	0	0	0	0
Q5SCORE	0	0.097	0	0	0	0
MOCANAM	0	0.093	0	0	0	0
BMNOCUE	0	0.291	0	0	0	0
CATANIMSC	0	0.306	0	0	0	0
Q13SCORE	0	0	0.026	0	0	0
TRAASCOR	0	0	0.065	0	0	0
TRABSCOR	0	0	0.224	0	0	0
MOCASERIAL	0	0	0.036	0	0	0
TRAILS	0	0	0.025	0	0	0
CLOCKSCOR	0	0	0	0.395	0	0
COPYSCOR	0	0	0	0.229	0	0
MOCACLOCK	0	0	0	0.448	0	0
Q3SCORE	0	0	0	0.157	0	0
CUBE	0	0	0	0.164	0	0
MMDRAW	0	0	0	0.048	0	0
Q7SCORE	0	0	0	0	0.331	0
MMORITIME	0	0	0	0	0.313	0
MMORISPACE	0	0	0	0	0.114	0
MOCAORI	0	0	0	0	0.418	0
Q9SCORE	0	0	0	0	0	0.074
Q10SCORE	0	0	0	0	0	0.092
Q11SCORE	0	0	0	0	0	0.204
Q12SCORE	0	0	0	0	0	0.148
Q2SCORE	0	0	0	0	0	0.085

References

[1] R. Cudeck, "Exploratory factor analysis," in *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (H. E. Tinsley and S. D. Brown, eds.), pp. 265 – 296, San Diego: Academic Press, 2000.

[2] J. W. Grice, "Computing and evaluating factor scores.," *Psychological Methods*, vol. 6, 2001.

[3] M. S. Bartlett, "The statistical conception of mental factors," *British Journal of Psychology. General Section*, vol. 28, pp. 97–104, jul 1937.