COLA Registered Report: The brain basis of inconsistent language lateralisation (POWER ANALYSIS)

The COLA Consortium (Dorothy V. M. Bishop, David Carey, Margriet Groen, Jessica Hodgson, John Hudson, Emma Karlsson, Mairéad MacSweeney, Adam J. Parker, Nuala Simpson, Paul A. Thompson, Kate E. Watkins, Zoe V. J. Woodhead)

04/03/2020

# Predictions

There are three main predictions:

1. Using a modified test battery with two new tasks we will replicate our previous finding that covariance between tasks is best explained by a two-factor solution. To test this prediction, we will use an approach based on Woodhead et al (2019), but instead of having paths from all variables to both factors, we will pre-specify a two-factor model with paths from A, B and C to Factor 1, and from D, E and F to Factor 2. [add OSF link to script]
2. The two factors will be significantly correlated. Individuals who are bivariate outliers, as defined by Cook’s distance (i.e. with dissociation between scores on the two factors) will predominantly be left-handed.
3. With fMRI we predict that tasks loading on the first factor will generate frontal activity, and those loading on the second factor will show stronger posterior activation.

We document the power estimates for Prediction 1 in section 2. Prediction 2 will be contingent on the output of the analysis for prediction 1, so power cannot be accurately assessed. It should be noted that the substantial sample size proposed for the first prediction is highly likely to enable prediction 2 to be well powered. Finally, prediction 3 can use the power estimate from prediction 1 as the analysis, the only substantive change is that the measured outcomes are derived from fMRI rather than fTCD. The number of participants having measurements in fMRI is capped for pragmatic reasons at 160-200 individuals, so we will assess the power at this level.

# Power estimate

## Single group confirmatory factor analysis models (one vs two factor)

There are two considerations for power in this part: 1. Power to fit individual models with sufficient fit indices, whether one factor or two factor; 2. Power to detect the mispecified model (one factor) vs the ‘true’ model (two factor). The first power consideration can be shown to be satisfied by our proposed sample size by following Monte Carlo simulation results presented by Wolf et al. (2013). This paper presents results from a range of simple CFA models, varying numbers of factors, numbers of measured variables, and magnitudes of factor loadings.

### Power estimate for the model comparison of proposed ‘true’ model vs proposed mispecified model.

We can calculate the power to detect the model mispecification using simulation and the distribution of fit indices. We use the R package simsem (Jorgensen et al. 2018) to calculate the power for the model comparison. We compete nested models, one factor vs two factors, using data simulated from both models. We should find that the data simulated from a two factor model and then fitted to both model, the fit indices favour the more complex two factor model. We simulate multiple data sets and fit the models at each iteration, from this Monte Carlo simulation, we can determine the statistical power to detect the model that explains the data structure most accurately.

### Comparing models based on two factor simulated data (Test hypothesis)

This output shows the power as a proportion (0 = 0%; 1=100%) for various sample sizes. We can see that we have good overall power based on a likelihood ratio test. The likelihood ratio test statistic follows a Chi-squared distribution, so at each iteration of the simulation a p-value is recorded and the proportion of statistically significant results at each sample size id recorded, i.e, statistical power. We can see that we are well powered to detect this difference.

| N | Power |
| --- | --- |
| 50.000 | 1.000 |
| 89.000 | 1.000 |
| 129.000 | 1.000 |
| 168.000 | 1.000 |
| 208.000 | 1.000 |
| 247.000 | 1.000 |
| 287.000 | 1.000 |
| 326.000 | 1.000 |
| 366.000 | 1.000 |
| 405.000 | 1.000 |
| 445.000 | 1.000 |
| 484.000 | 1.000 |
| 524.000 | 1.000 |
| 563.000 | 1.000 |
| 603.000 | 1.000 |
| 642.000 | 1.000 |
| 682.000 | 1.000 |
| 721.000 | 1.000 |
| 761.000 | 1.000 |
| 800.000 | 1.000 |

### Alternative power calculation using the fit indices.

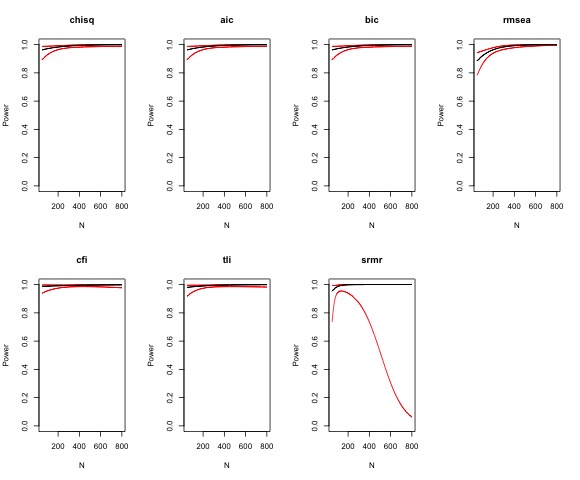
The same Monte Carlo simulation is used but the power is based on the relative fit indices. For more details, see <https://github.com/simsem/simsem/wiki/Example-25:-Nested-Model-Comparison-with-Varying-Sample-Size>.

| Fit Index | Power |
| --- | --- |
| chisq | 0.998 |
| aic | 0.998 |
| bic | 0.998 |
| rmsea | 0.999 |
| cfi | 0.998 |
| tli | 0.999 |
| srmr | 1.000 |

Two power estimates are provided, one derived by a specific sample size (Table 1) and null model (Table 2), and secondly, using a derived cutoff.

| Fit Index | Power |
| --- | --- |
| chisq | > 0.98 |
| cfi | > 0.98 |
| tli | > 0.98 |
| aic | > 0.98 |
| bic | > 0.98 |
| rmsea | > 0.98 |
| srmr | > 0.98 |

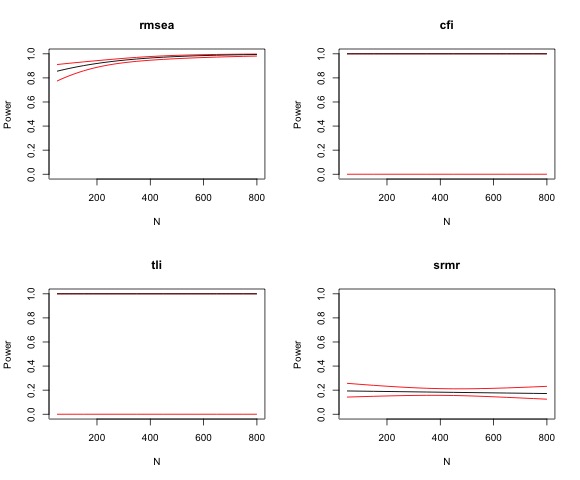
## quartz\_off\_screen   
## 2



We can also specify a user defined cutoff, for example, guidelines on minimum required fit indices’ values to ensure a satisfactory model. Here we have specified less conservative cutoffs (RMSEA = 0.05, CFI = 0.95, TLI = 0.95, SRMR = 0.06) which are commonly used in the literature (Wang and Wang 2012). It should be noted that the standardized root mean square residual (srmr) is highly sensitive to large sample sizes and will tend to smaller values, hence the odd shape power curve; therefore, we should disregard this index (Wang and Wang 2012).

| Fit Index | Power |
| --- | --- |
| cfi | > 0.98 |
| tli | > 0.98 |
| rmsea | > 0.98 |
| srmr | 0.174087643085631 |

quartz\_off\_screen   
 2



## Multigroup model (two factor multigroup CFA - left and right handers groups)

The second power calculation is based on a two factor multigroup model. We hypothesis that the left handers group will have a different size correlation between the two factor than the right handers group (0.5 and 0.8 respectively). Data is simulated based on the proposed correlation structure, and then two multigroup models are fitted to the data, unconstrained model allowing for different latent variable correlation in each group and a second model, that constrains the latent variable correlation to be the same. We then perform a likelihood ratio test on each model to determine which provides a better explanation of the data. The power is established by running the simulation and model fitting over 1000 iterations. The p value from the likelihood ratio test for each iteration is recorded and the proportion of these runs that are significant at alpha level of 5% is the statistical power.

Current power estimate for the proposed sample size and ratio of left and right handers (N left = 160, N right = 490) would be around 68%. If groups were balanced but overall sample size remained the same, power is increassed to 80%. To satisfy registered report power of 90%, we would require 450 individuals per group.

| N1 | N2 | power |
| --- | --- | --- |
| 160.000 | 490.000 | 0.680 |
| 200.000 | 450.000 | 0.735 |
| 325.000 | 325.000 | 0.803 |
| 450.000 | 450.000 | 0.905 |

# References

Jorgensen, Terrence D., Sunthud Pornprasertmanit, Patrick Miller, and Alexander Schoemann. 2018. *Simsem: Simulated Structural Equation Modeling*. <https://CRAN.R-project.org/package=simsem>.

Wang, J., and X. Wang. 2012. *Structural Equation Modeling: Applications Using Mplus*. Chichester, West Sussex, U.K: Wiley/Higher Education Press.

Wolf, E.J., K.M. Harrington, S.L. Clark, and M.W. Miller. 2013. “Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety.” *Educational and Psychological Measurement* 76 (6): 913–34. <https://doi.org/10.1177/0013164413495237>.