“A distributed brain network predicts general intelligence from resting-state human neuroimaging data.” by Julien Dubois, Paola Galdi, Lynn K. Paul, and Ralph Adolphs

Supplementary material

Test Name	Variable(s)	Test Activity / Outcome
NIH Toolbox Dimensional Change Card Sort Test	**CardSort_Unadj** [Unadjusted Scale Score]	Two target pictures are presented that vary along two dimensions (e.g., shape and color). Participants are asked to match a series of bivalent test pictures (e.g., yellow balls and blue trucks) to the target pictures, first according to one dimension (e.g., color) and then, after a number of trials, according to the other dimension (e.g., shape). “Switch” trials are also employed, in which the participant must change the dimension being matched. For example, after 4 straight trials matching on shape, the participant may be asked to match on color on the next trial and then go back to shape, thus requiring the cognitive flexibility to quickly choose the correct stimulus. Scoring is based on a combination of accuracy and reaction time, and the test takes approximately 4 minutes to administer.
NIH Toolbox Flanker Inhibitory Control and Attention Test	**Flanker_Unadj** [Unadjusted Scale Score]	The test requires the participant to focus on a given stimulus while inhibiting attention to arrows flanking it. Sometimes the middle stimulus is pointing in the same direction as the “flankers” (congruent) and sometimes in the opposite direction (incongruent). Scoring is based on a combination of accuracy and reaction time, and the test takes approximately 3 minutes to administer.
NIH Toolbox List Sorting Working Memory Test	**ListSort_Unadj** [Unadjusted Scale Score]	This task assesses working memory and requires the participant to sequence different visually- and orally-presented stimuli. Pictures of different foods and animals are displayed with both a sound clip and written text that name the item. The task has two different conditions: 1-List and 2-List. In the 1-List condition, participants are required to order a series of objects (either food or animals) in size order from smallest to largest. In the 2-List condition, participants are presented both food and animals and are asked to report the food in size order, followed by the animals in size order. Participants have two practice items, in which the images briefly “flash” sequentially on the screen in each condition.
NIH Toolbox Picture Sequence Memory Test	**PicSeq_Unadj** [Unadjusted Scale Score]	The Picture Sequence Memory Test is a measure developed for the assessment of episodic memory. It involves recalling increasingly lengthy series of illustrated objects and activities that are presented in a particular order on the computer screen. The participants are asked to recall the sequence of pictures that is demonstrated over two learning trials; sequence length varies from 6-18 pictures, depending on age. Participants are given credit for each adjacent pair of pictures (i.e., if pictures in locations 7 and 8 and placed in that order and adjacent to each other anywhere – such as slots 1 and 2 – one point is awarded) they correctly place, up to the maximum value for the sequence, which is one less than the sequence length (if there are 18 pictures in the sequence, the maximum score is 17, because that is the number of adjacent pairs of pictures that exist). The test takes approximately 7 minutes to administer.
NIH Toolbox Picture Vocabulary Test	**PicVocab_Unadj** [Unadjusted Scale Score]	This measure of receptive vocabulary is administered in a computerized adaptive format. The respondent is presented with an audio recording of a word and four photographic images on the computer screen and is asked to select the picture that most closely matches the meaning of the word. This test takes approximately 4 minutes to administer.
NIH Toolbox Pattern Comparison Processing Speed Test	**ProcSpeed_Unadj** [Unadjusted Scale Score]	This test measures speed of processing by asking participants to discern whether two side-by-side pictures are the same or not. Participants' raw score is the number of items correct in a 90-second period. The items are designed to be simple to most purely measure processing speed. The test overall takes approximately 3 minutes to administer.
NIH Toolbox Oral Reading Recognition Test	**ReadEng_Unadj** [Unadjusted Scale Score]	The participant is asked to read and pronounce letters and words as accurately as possible. The test administrator scores them as right or wrong. The test is given via a computerized adaptive format and requires approximately 3 minutes.
Penn CNB Penn Progressive Matrices	**PMAT24_A_CR** [# of Correct Resp]	Fluid intelligence is measured using Raven’s Progressive Matrices . We use Form A of an abbreviated version of the Raven’s developed by Gur and colleagues. Participants are presented with patterns made up of 2x2, 3x3 or 1x5 arrangements of squares, with one of the squares missing. The participant must pick one of five response choices that best fits the missing square on the pattern. The task has 24 items and 3 bonus items, arranged in order of increasing difficulty. However, the task discontinues if the participant makes 5 incorrect responses in a row.
Penn CNB Penn Word Memory Test	**IWRD_TOT** [Total # of Correct Resp]	Verbal episodic memory is measured using Form A of the Penn Word Memory Test. Participants are shown 20 words and asked to remember them for a subsequent memory test. They are then shown 40 words (the 20 previously presented words and 20 new words matched on memory related characteristics). They decide whether they have seen the word previously by choosing among “definitely yes,” “probably yes,” “probably no,” and “definitely no.”
Penn CNB Variable Short Penn Line Orientation	**VSPLOT_TC** [total # correct]	Spatial orientation processing is measured using the Variable Short Penn Line Orientation Test. Participants are shown two lines with different orientations. They have to rotate one of the lines (a moveable blue one) so that is parallel to the other line (a fixed red line). The rotation of the blue line is accomplished by clicking buttons on the keyboard that rotate the lines either clockwise or counterclockwise. Across trials, the lines vary in their relative location on the screen, though the distance between the centers of the two lines is always the same. The length of the red line is always the same, but the length of the blue line can be either short or long. There are a total of 24 trials

Supplementary Table 1. List of cognitive measures included in our analyses.

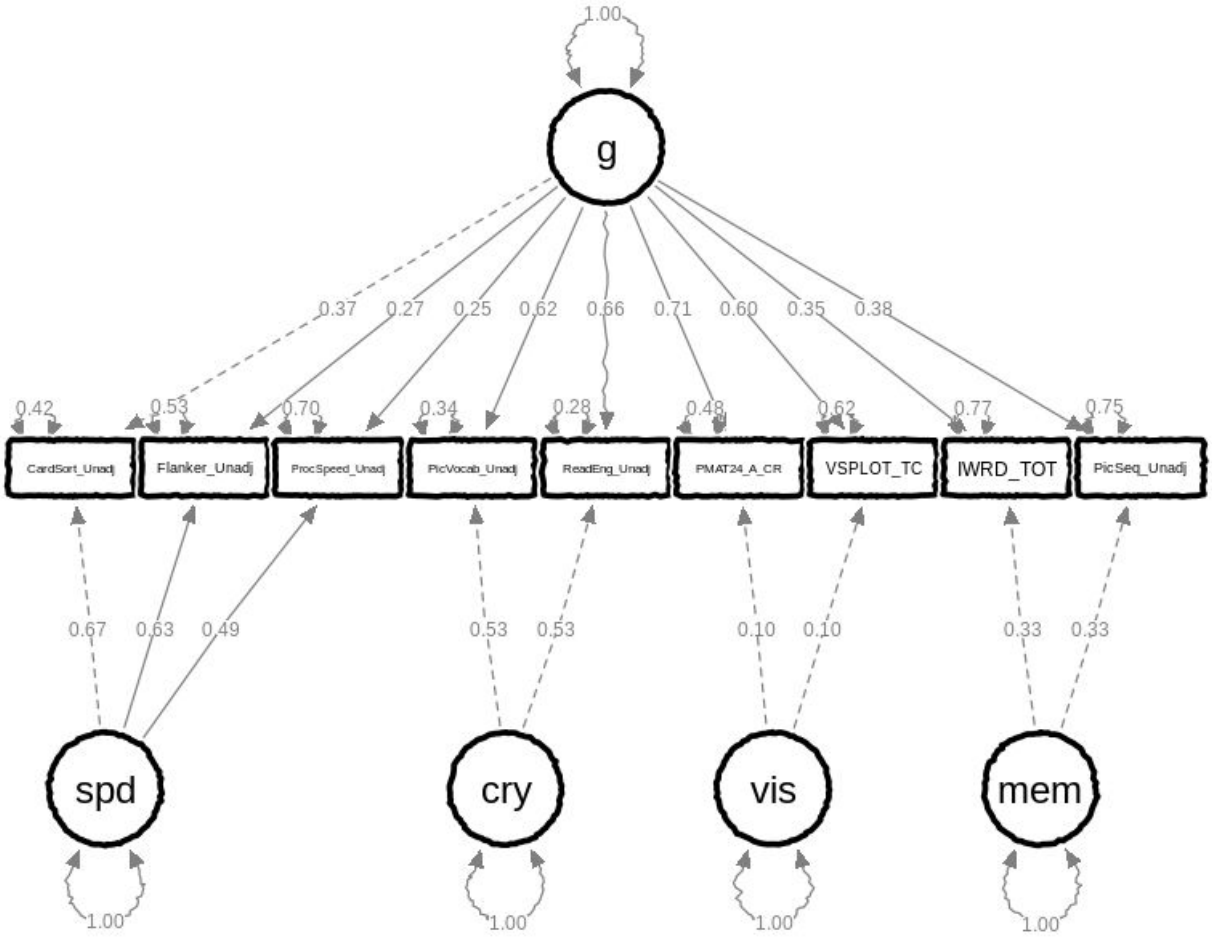
Confirmatory factor analysis

Though deriving factor scores from an EFA is often done by empirical researchers, it is theoretically preferable to use a confirmatory factor analysis (CFA) framework. A major difference is that a CFA model typically restricts cross-loadings (an observed variable loading on several latent factors), while the EFA allows them; this can reduce the size of the *g* factor in the EFA.

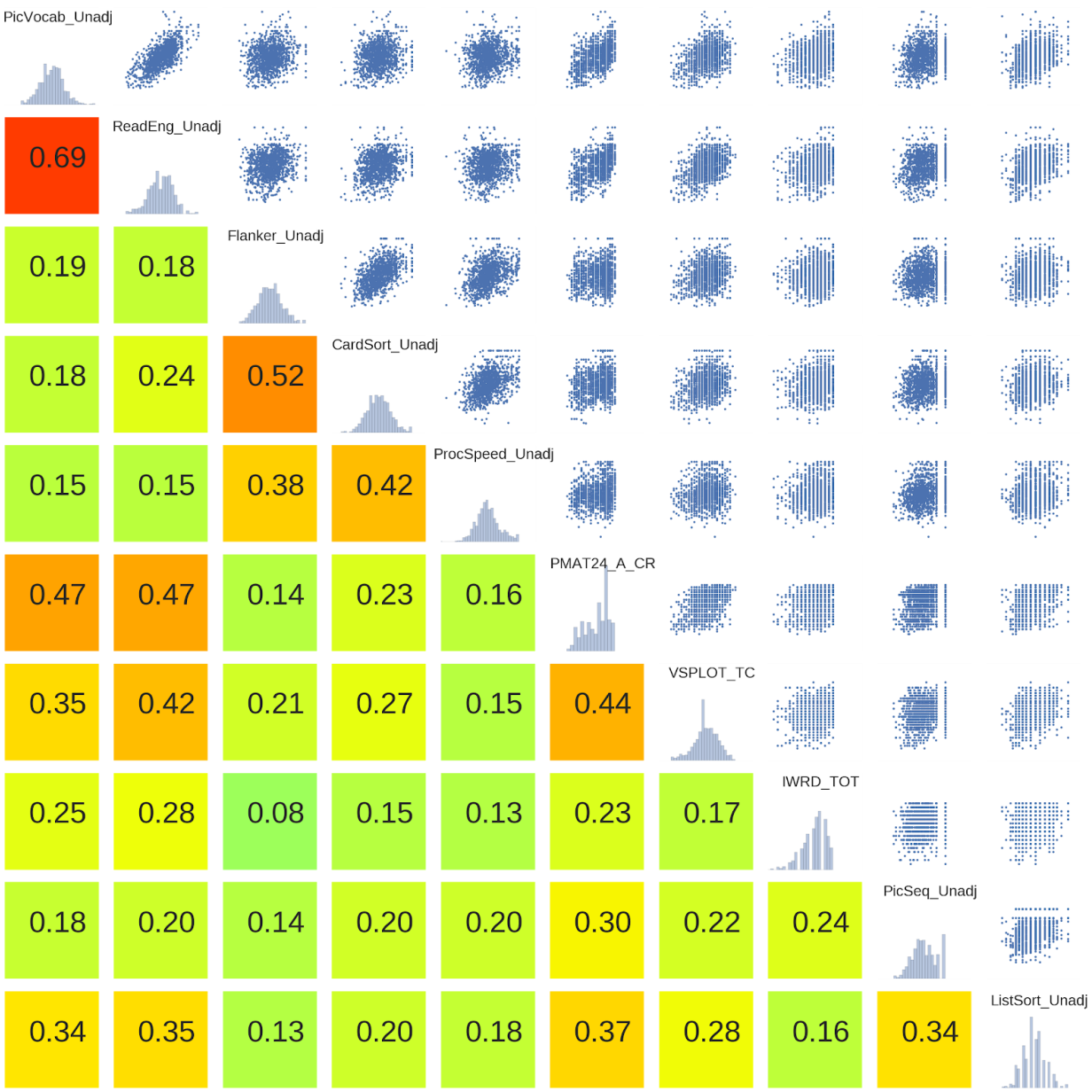
Using the EFA solution, we specified a bi-factor model, including a general factor (loading on all tasks) and four group factors (loading on subsets of tasks), in a confirmatory factor analysis. The model does not allow for any cross-loadings of a task on several factors, and group factors are orthogonal to one another and to the general factor. As some of the group factors are only defined by two indicators, it was necessary to impose a constraint for the purposes of identification. We fixed the unstandardized loadings for both tasks to 1.0 in this case. We initially found that the lavaan model did not converge, and identified that the issue lied with the ListSort_Unadj task score. We ran the CFA without ListSort_Unadj, with the following lavaan syntax:

```
#g-factor
g       =~ CardSort_Unadj + Flanker_Unadj + ProcSpeed_Unadj + PicVocab_Unadj +
    ReadEng_Unadj + PMAT24_A_CR + VSPLOT_TC + IWRD_TOT + PicSeq_Unadj
#Domain factors
spd    =~ CardSort_Unadj + Flanker_Unadj + ProcSpeed_Unadj
cry    =~ 1*PicVocab_Unadj + 1*ReadEng_Unadj
vis    =~ 1*PMAT24_A_CR    + 1*VSPLOT_TC
mem =~ 1*IWRD_TOT      + 1*PicSeq_Unadj
#Domain factors are not correlated with g
g ~~ 0*spd
g ~~ 0*cry
g ~~ 0*vis
g ~~ 0*mem
#Domain factors are not correlated with one another
spd ~~ 0*cry
spd ~~ 0*vis
spd ~~ 0*mem
cry ~~ 0*vis
cry ~~ 0*mem
vis ~~ 0*mem
```

This CFA model converged after 49 iterations, and the fit was very good with CFI=0.974, RMSEA=0.052, SRMR=0.032, BIC = 27820.2. The standardized solution is depicted in **Supplementary Figure 1**. The general factor was found to account for 64.0% of the variance (coefficient omega_hierarchical ω_h), while group factors accounted for 17.2% of the variance. We derived the factor scores for *g* using the regression method; we found that the scores derived from the CFA were almost perfectly correlated with the scores derived from the EFA (**Figure 1**), r=0.99.

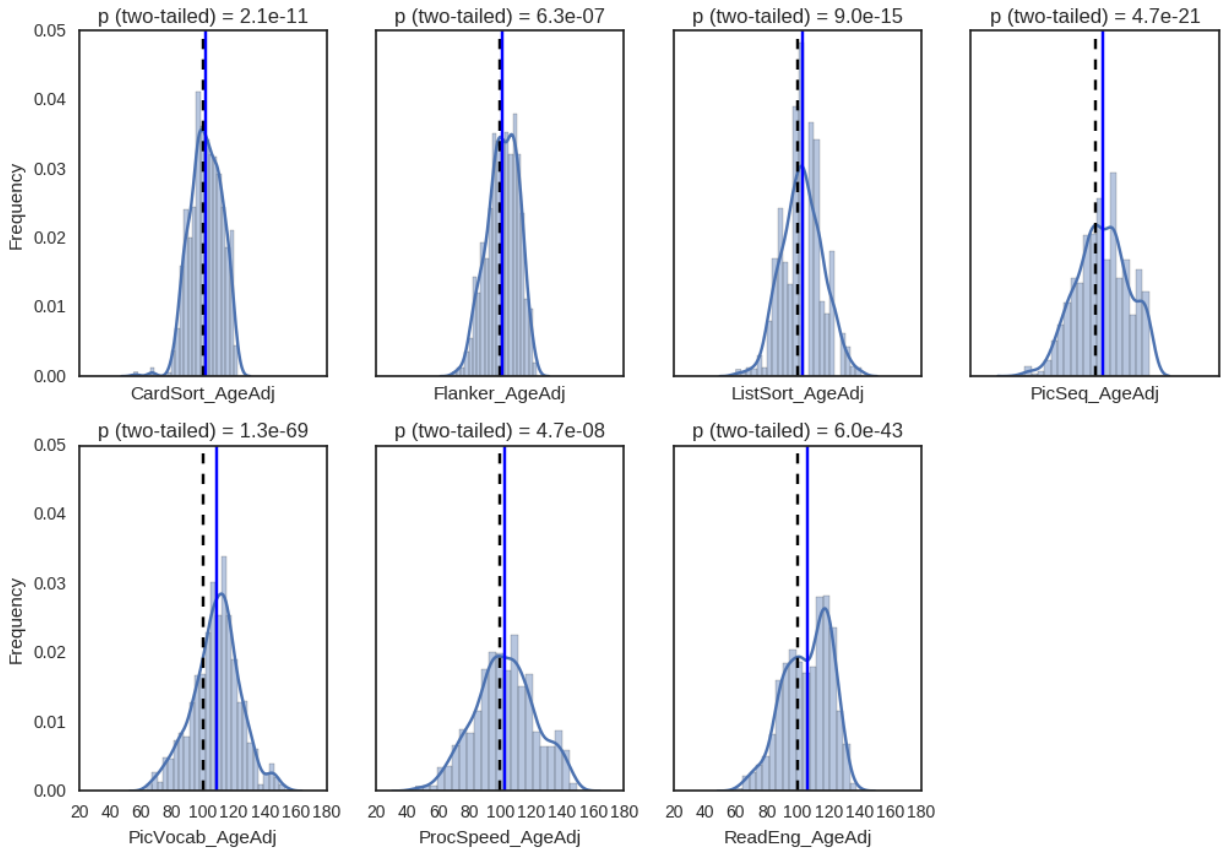


Supplementary Figure 1. Standardized solution for our confirmatory factor analysis of the HCP cognitive task scores. The CFA omits *ListSort_Unadj* which prevented the model from converging.



Supplementary Figure 2. Correlations between (z-scored) HCP cognitive task scores. On the diagonal, the distribution of each of the 10 cognitive variables in shown. Below the diagonal, the Pearson correlation is displayed, together with a color visualizing the strength of the relationship. Above the diagonal, a scatter plot is displayed for each pair of variables, with x- and y- axes between -4 and 4 (standard deviation of all variables is 1 due to z-scoring).

Supplementary Figure 3. Correlation of the general intelligence factor scores with PMAT24_A_CR scores, and with potential confounds, in the sample of subjects used for prediction analyses (N=884). All of these variables except for *PMAT24_A_CR* scores were regressed out of the training set data to obtain an unconfounded measure of g.



Supplementary Figure 4. Distributions of the age-normed scores of HCP subjects on NIH-toolbox cognitive tasks. The blue line shows the mean score in our subject sample, which is greater than 100 for all tests, while the black dashed line shows the mean in the normative population. For all tests, a 1-sample Student's t-test indicates that the mean is significantly higher than 100.