Neuron, Volume *76*

Supplemental Information

Fractionating Human Intelligence

Adam Hampshire, Roger R. Highfield, Beth L. Parkin, and Adrian M. Owen

Table S1. PCA of activation levels in 2275 MD voxels during the performance of 12 cognitive tasks

		Initial eigenvalues		Variance explained after rotation	
Component	Total	% of Variance	Cumulative %	% of Variance	Cumulative %
1	**9.567**	**79.721**	**79.721**	**45.161**	**45.161**
2	**1.211**	**10.095**	**89.816**	**44.655**	**89.816**
3	0.517	4.311	94.127		
4	0.32	2.664	96.791		
5	0.132	1.103	97.894		
6	0.065	0.539	98.434		
7	0.056	0.467	98.901		
8	0.046	0.382	99.283		
9	0.029	0.239	99.522		
10	0.022	0.18	99.701		
11	0.02	0.163	99.865		
12	0.016	0.135	100		

Extraction Method: Principal Component Analysis.
Rotation method: Varimax

Table S2. Test-retest reliabilities

Task	N	Retest reliability (Pearson's correlation)	Learning effects % improvement
Spatial Span	647	0.62	0.46
Visuospatial Working Memory	804	0.57	1.62
Self Ordered Search	1113	0.66	4.99
Paired Associates	1131	0.45	-0.38
Spatial Planning	1150	0.87	3.75
Spatial Rotation	1122	0.7	5.43
Feature Match	1132	0.57	4.09
Interlocking Polygons	905	0.6	7.91
Deductive Reasoning	1138	0.73	1.55
Digit Span	1022	0.64	1.33
Verbal Reasoning	1148	0.89	2.24
Color-Word Remapping	1151	0.92	4.9
average		*0.69*	*3.16*

The reliabilities of the tasks were calculated from a separate population sample, collected on the CambridgeBrainSciences.com web site prior to the study reported here. Data were standardised so that for a given task, there was unit deviation and zero mean. Correlations were then calculated between the first and second instances in which participants chose to undertake a task.

Table S3. PCA of behavioral performance in 44,600 Internet users

	Initial Eigenvalues			Variance explained after rotation	
Component	Total	% of Variance	Cumulative %	% of Variance	Cumulative %
1	**3.277**	**27.31**	**27.31**	**17.072**	**17.072**
2	**1.119**	**9.326**	**36.636**	**15.819**	**32.891**
3	**1.008**	**8.397**	**45.033**	**12.142**	**45.033**
4	0.876	7.303	52.336		
5	0.828	6.9	59.236		
6	0.769	6.41	65.645		
7	0.759	6.323	71.968		
8	0.732	6.101	78.07		
9	0.706	5.881	83.951		
10	0.685	5.704	89.656		
11	0.658	5.485	95.14		
12	0.583	4.86	100		

Extraction Method: Principal Component Analysis.
Rotation method: Varimax

Table S4. Task-component loadings from the PCA of data simulated from the task-functional brain activation levels

	F1	F2	F3
Spatial Span	0.51	0.16	-0.48
Visuospatial Working Memory	0.57	0.11	0.05
Self Ordered Search	0.59	0.10	0.14
Paired Associates	0.37	0.00	0.56
Spatial Planning	0.48	0.45	0.24
Spatial Rotation	0.25	0.47	0.06
Feature Match	0.34	0.47	0.12
Interlocking Polygons	0.23	0.50	0.08
Deductive Reasoning	-0.14	0.76	0.00
Digit Span	0.68	0.21	0.12
Verbal Reasoning	0.08	0.32	0.58
Color-Word Remapping	0.21	0.47	0.47

Table S5. IFS and IFO activations relative to rest

	IFO	IFS
Spatial Span	0.59	0.08
Digit Span	0.62	0.24
Visuospatial Working Memory	0.71	0.35
Self Ordered Search	0.68	0.41
Paired Associates	0.52	0.53
Feature Match	0.15	0.38
Interlocking Polygons	0.14	0.34
Spatial Rotation	0.50	0.73
Spatial Planning	0.51	0.73
Verbal Reasoning	0.39	0.80
Color-Word Remapping	0.33	0.80
Deductive Reasoning	0.08	1.09

Supplemental Experimental Procedures

1) Detailed statistical methods

FMRI data were collected using the standard 2 second protocol on the 3T Siemens Tim Trio scanner at the Robarts Imaging Centre in the University of Western Ontario. Tasks were displayed on a projector screen, visible from the bore of the MRI scanner via a mirror. Responses were taken with a custom MRI compatible tracker ball mouse. FMRI data were pre-processed using the SPM 2005 software package (Wellcome Department of Cognitive Neurology). Time-courses were motion and slice-time corrected, Normalized to the standard MNI template, smoothed with an 8 mm Gaussian kernel, and high-passed filtered prior to analysis (cut-off 180s). Individual participant's data were modelled in SPM5 using the general linear model. Models consisted of 12 sets of predictor functions, each of which included timings for the task vs. baseline blocks convolved with the standard hemodynamic response function, six movement parameters, and the resting baseline constants. (Note- analysis with 36 predictor functions, one for each block of each task, revealed no significant cross-block differences in task-related activation within MD cortex at the group level). Region of interest (ROI) analyses were carried out using the MarsBaR toolbox (Brett et al., 2002), which averages data across all voxels within a given ROI. MD network ROIs and the MD ROI mask were provided by Duncan J from a previously published meta-analysis of neuroimaging studies in which difficulty was parametrically modulated (Duncan and Owen, 2000). PCAs of task related activation levels across MD voxels were carried out in the SPSS software package and ICA was carried out in Matlab, a high-level mathematical programming environment.

Individual tests for the imaging and Internet studies were programmed in the Adobe Flex development environment. In the imaging study, the tests ran as stand alone software within the AIR runtime environment. In the behavioural study, the tests were embedded in a custom-built web site programmed in

Microsoft ASP.net. Prior to analysis, behavioural scores for each test were standardized by subtraction of the population mean to centre scores around zero and division by the standard deviation to ensure unit deviation. Participants who had extreme outlier scores over 4SDs from the mean on any single test or who made nonsensical responses to any demographic questions were excluded case-wise. Participants above the age of 70 and below the age of 12 were excluded due to low numbers (N<100). Regression of the MDwm and MDr vectors onto the behavioural data, permutation modelling, and simulated data based on the brain imaging data were generated using custom software programmed in Matlab. Overall mean scores were calculated for each individual by first averaging the standardized scores across all 12 tasks and then standardizing the resultant values across all participants to ensure unit deviation. Individuals STM, Reasoning and Verbal factor scores were calculated using regression in SPSS and were also in standardized units. This approach ensured that the mean scores and factor scores were not biased towards any one test and that results from all GLMs were in standard deviation units. Generalized linear modelling (GLM) and principal components analysis (PCA) of the behavioural and simulated data sets were carried out using the SPSS software package. Unless stated otherwise, PCAs used the default Kaiser convention of including only those factors with eigenvalues >1 and an orthogonal Varimax rotation. 2^{nd} order factor analyses were carried out in SPSS using Promax oblique rotation. GLMs used the default Maximum Likelihood Estimate algorithm and the Wald test of statistical significance.

2) Brief descriptions of the 12 tests

The 12 cognitive tasks are based on classical paradigms from the cognitive psychology literature.

(All tests are available for evaluation at www.CambridgeBrainSciences.com)

Spatial Span is based on the Corsi Block Tapping Task (Corsi, 1972), a classical tool for measuring spatial short-term memory capacity. 16 squares are displayed in a 4 * 4 grid. A sub-set of the squares flash in a random sequence at a rate of 1 flash every 900 ms. Subsequently, the mouse cursor is displayed and a tone cues the participant to repeat the sequence by clicking on the squares in the same order in which they flashed. The test starts with four flashes and difficulty on subsequent trials is dynamically varied. If the participant responds correctly, the length of the next sequence increases by one flash, otherwise the length of the next sequence is one flash shorter. The test finishes after 3 errors. Maximum level = 16, minimum level = 2. Outcome measure = maximum level achieved. Population mean = 6.15, standard deviation (SD) = 1.07.

Visuospatial Working Memory is based on a task from the non-human primate literature (Inoue and Matsuzawa, 2007). Sets of numbered squares are displayed all at the same time at random locations within an invisible 5*5 grid. After a variable interval (number of squares * 900 ms), the numbers are removed leaving just the blank squares visible and a tone cues the participant to respond by clicking on the squares in ascending numerical sequence. The level of difficulty is dynamically varied, starting with just 2 numbered squares on the first trial and then increasing or decreasing by 1 depending on whether or not the response is correct. The test finishes after three errors. Maximum level = 25 and minimum level = 2. Outcome measure = maximum level achieved. Population mean = 7.85, SD = 1.154.

Self-Ordered Search is based on a test that is widely used to measure strategy during search behaviour (Collins et al., 1998). Sets of boxes are displayed on the screen in random locations within an invisible 5*5 grid. The participant must find a hidden ‘token’ by clicking on the boxes one at a time to reveal their contents. When the token is found, it is hidden within another box. On any given trial, the token will not appear within the same box twice, thus, the participant must search the boxes until the token has been found once within each box. If they search the same empty box twice whilst looking for the token, or search a box in which the token has previously been found, this is an error and the trial ends, subsequent to which a new trial begins with one less box to search. If the token is found once in each box without any errors being made, a new trial begins with one extra box to search. After three errors the test ends. The test starts with just four boxes. Maximum level = 25 and minimum level = 2. Outcome measure = maximum level achieved. Population mean = 8.23, SD = 2.10.

Paired Associates is based on a paradigm that is commonly used to assess memory impairments in aging clinical populations (Gould et al., 2005). Boxes are displayed at random locations on an invisible 5*5 grid. The boxes open one after another to reveal an enclosed object. Subsequently, the objects are displayed in random order in the centre of the grid and the participant must click on the boxes that contained them. If the participant remembers all of the object-location paired associates correctly then the difficulty level on the subsequent trial increases by 1 object-box pair, otherwise it decreases by 1. After three errors the test ends. The test starts with just two boxes. Maximum level = 24 minimum level = 2. Outcome measure = maximum level achieved. Population mean = 5.28, SD = 1.13.

Spatial Planning is based on the Tower of London Task (Shallice, 1982), which is widely used to measure executive function. Numbered beads are positioned on a tree shaped frame. The participant repositions the beads so that they are configured in ascending numerical order running from left to right and top to

bottom of the tree. To gain maximum points, the participant must solve as many problems as possible in as few moves as possible within 3 minutes. Problems become progressively harder with the total number of moves required and the planning complexity increasing in steps. Trials are aborted if the participant makes more than twice the number of moves required to solve the problem. After each trial, the total score is incremented by adding the minimum number of moves required * 2 – the number of moves actually made, thereby rewarding efficient planning. The first problem can be solved in just 3 moves. Outcome measures = total score. Population mean = 64 SD = 10.185.

Spatial Rotation tasks are often used for measuring the ability to manipulate objects spatially in mind (Silverman et al., 2000). In this variant, two grids of colored squares are displayed to either side of the screen with one of the grids rotated by a multiple of 90 degrees. When rotated, the grids are either identical or differ by the position of just one square. In order to gain maximum points, the participant must indicate whether the grids are identical, solving as many problems as possible within 90 seconds. If the response is correct, the total score increases by the number of squares in the grid and subsequent trials have more squares. If the response is incorrect the total score decreases by the number of squares in the grid and subsequent trials have fewer squares. The first grids contain 4 colored squares each. Outcome measure = total score. Population mean = 88.72, SD = 36.32.

Feature Match is based on the classic feature search tasks that have historically been used to measure attentional processing (Treisman and Gelade, 1980). Two grids are displayed on the screen, each containing a set of abstract shapes. In half of the trials the grids differ by just one shape. In order to gain maximum points, the participant must indicate whether or not the grid's contents are identical, solving as many problems as possible within 90 seconds. If the response is correct, the total score increases by the number of shapes in the grid and the number of shapes in subsequent trials increases. If the response is

incorrect the total score decreases by the number of shapes in the grid and subsequent trials have fewer shapes. The first grids contain two abstract shapes each. Outcome measure = total score. Population mean = 131.35, SD = 32.79.

Interlocking Polygons is based on the Interlocking Pentagons task, which is often used in the assessment of age related disorders (Folstein et al., 1975). A pair of overlapping polygons is displayed on one side of the screen. In order to gain maximum points, the participant must indicate whether a polygon displayed on the other side of the screen is identical to one of the interlocking polygons, solving as many problems as possible within 90 seconds. If responses are correct, the total score increases by the difficulty level and the differences between the polygons becomes increasingly subtle. If the responses are incorrect the total score decreases by the difficulty level and the difference between the polygons become more pronounced. Main outcome measure = total score. Population mean = 51.41, SD = 24.86.

Deductive Reasoning is based on a sub-set of problems from the Cattell Culture Fair Intelligence Test (Cattell, 1949). A 3* 3 grid of cells is displayed on the screen. Each cell contains a variable number of copies of a colored shape. The features that make up the objects in each cell (color, shape, number of copies) are related to each other according to a set of rules. The participant must deduce the rules that relate the object features and select the one cell whose contents do not correspond to those rules. To gain maximum points, the participant must solve as many problems as possible within 90 seconds. If the response is correct, the total score increases by one point and the next problem is more complex. If the response is incorrect, the total score decreases by 1 point. Outcome measure = total correct. Population mean = 10.43, SD = 3.31.

Digit Span is a computerised variant on the verbal working memory component of the WAIS-R intelligence test (Weschler, 1981). Participants view a sequence of digits that appear on the screen one after another. Subsequently, they repeat

the sequence of numbers by entering them on the keyboard. Difficulty is dynamically varied with the number of digits to remember increasing or decreasing by 1 depending on whether the participant got the previous trial correct. The test ends after 3 errors. Maximum level = 25 and minimum level = 2. Outcome measure = maximum level achieved. Population mean = 7.22, SD = 1.52.

Verbal Reasoning is based on Alan Baddeley's 3 minute grammatical reasoning test (Baddeley, 1968). Problems of the form "The square is not encapsulated by the circle" are displayed on the screen and the participant must indicate whether the statement correctly describes a pair of objects displayed in the centre of the screen. In order to achieve maximum points, the participant must solve as many problems as possible within 90 seconds. Total score increases or decreases by 1 after each trial depending on whether responses are correct. The first trial has four numbers. The maximum level is 25 and the minimum level is 2. Outcome measure = total score. Population mean = 17.38, SD = 5.01.

Color-Word Remapping is a more challenging variant on the Stroop test (Stroop, 1935). A colored word is displayed at the top of the screen, for example the word RED drawn in blue ink. The participants must indicate which of two colored words at the bottom of the screen describes the color that the word at the top of the screen is drawn in. The color word mappings may be congruent, incongruent, or doubly incongruent, depending on whether or not the colors that a given words describes matches the color that it is drawn in. To gain maximum points, the participant must solve as many problems as possible within 90 seconds. The total score increases or decreases by 1 after each trial depending on whether they responded correctly. Outcome measure = total score. Population mean = 30.92, SD = 13.01.

Supplemental References

Baddeley, A.D. (1968). A three-minute reasoning test based on grammatical transformation. Psychometric science *10*, 341-342.

Cattell, R.B. (1949). Culture Free Intelligence Test, Scale 1, Handbook (Champaign, Illinois: Institute of Personality and Abilit).

Collins, P., Roberts, A.C., Dias, R., Everitt, B.J., and Robbins, T.W. (1998). Perseveration and strategy in a novel spatial self-ordered sequencing task for nonhuman primates: effects of excitotoxic lesions and dopamine depletions of the prefrontal cortex. Journal of cognitive neuroscience *10*, 332-354.

Corsi, P.M. (1972). Human memory and the medial region of the brain. . In PhD thesis (Montreal, McGill).

Folstein, M.F., Folstein, S.E., and McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res *12*, 189-198.

Gould, R.L., Brown, R.G., Owen, A.M., Bullmore, E.T., Williams, S.C., and Howard, R.J. (2005). Functional neuroanatomy of successful paired associate learning in Alzheimer's disease. Am J Psychiatry *162*, 2049-2060.

Inoue, S., and Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. Curr Biol *17*, R1004-1005.

Shallice, T. (1982). Specific impairments of planning. Philos Trans R Soc Lond B Biol Sci *298*, 199-209.

Silverman, I.I., Choi, J., Mackewn, A., Fisher, M., Moro, J., and Olshansky, E. (2000). Evolved mechanisms underlying wayfinding. further studies on the hunter-gatherer theory of spatial sex differences. Evol Hum Behav *21*, 201-213.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology *18*, 643-642.

Treisman, A.M., and Gelade, G. (1980). A feature-integration theory of attention. Cogn Psychol *12*, 97-136.

Weschler, D. (1981). Wechsler Adult Intelligence Scale--Revised (The Psychological Corporation).