Supplementary Material

1. *Post-hoc* assessment of mover subgroups

X	G_1	G_2	G_3	G_4
G_1		14.33	4.62	-12.98
G_2	$< 10^{-5}$		-10.48	-31.58
G_3	$< 10^{-5}$	$< 10^{-5}$		-20.51
G_4	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	

Y	G_1	G_2	G_3	G_4
G_1		8.28	-7.92	-2.49
G_2	$< 10^{-5}$		-17.84	-11.58
G_3	$< 10^{-5}$	$< 10^{-5}$		6.01
G_4	0.013	$< 10^{-5}$	$< 10^{-5}$	

Z	G_1	G_2	G_3	G_4
G_1		6.99	-9.77	-8.16
G_2	$< 10^{-5}$		-17.87	-14.38
G_3	$< 10^{-5}$	$< 10^{-5}$		-1.35
G_4	$< 10^{-5}$	$< 10^{-5}$	0.18	

$\boldsymbol{\alpha}$	G_1	G_2	G_3	G_4
G_1		8.96	-9.67	0.19
G_2	$< 10^{-5}$		-19.67	-8.65
G_3	$< 10^{-5}$	$< 10^{-5}$		10.39
G_4	0.85	$< 10^{-5}$	$< 10^{-5}$	

$\boldsymbol{\beta}$	G_1	G_2	G_3	G_4
G_1		7.62	-4.61	-6.33
G_2	$< 10^{-5}$		-13.33	-14.87
G_3	$< 10^{-5}$	$< 10^{-5}$		-2.22
G_4	$< 10^{-5}$	$< 10^{-5}$	0.03	

$\boldsymbol{\gamma}$	G_1	G_2	G_3	G_4
G_1		32.75	13.89	13.35
G_2	$< 10^{-5}$		-14.48	-13.95
G_3	$< 10^{-5}$	$< 10^{-5}$		0.2
G_4	$< 10^{-5}$	$< 10^{-5}$	0.84	

Table 1: ***Post-hoc* analysis of spatial motion across mover groups.** For the six possible motion degrees of freedom (X, Y, Z, α, β and γ), t-statistics for each group comparison (top right diagonal elements) and associated p-values (bottom left diagonal elements).

2. Robustness to parameter changes

We verified that modifications of the parameters used in our main analyses would not substantially affect our results. We considered the followings:

- The threshold (in mm) above which a time point is censored. We compared our main choice of 0.3 mm to values of 0.2 mm (case 1), 0.5 mm (case 2) and 1 mm (case 3).
- The number of frames to excise around corrupted time points. We compared the case of solely removing the corrupted frames, as done in our main results, to the additional removal of one more frame at time $t + 1$ (case 4).
- The number of nearest neighbours used for graph construction; we contrasted our original value of 10 to alternative values of 5 (case 5) and 20 (case 6).
- The number of time bins into which to subdivide a resting-state session; in comparison with our original choice of 6 bins, we probed values ranging from 4 (case 7) to 8 (case 10).

We assessed robustness of the clustering outcomes by computing the purity measure (Yang et al., 2012), which takes a value of 1 for perfect concordance of classification, and of 0 if no data point is clustered similarly.

For each set of saliences, we computed (1) the absolute valued Spearman's correlation between the reference and output saliences, and (2) their absolute valued dot product. The absolute value enables to account for sign-flipped salience vectors across computations.

The results (computed only from session 1 data) are presented in Table 2. It can be seen that our analytical outcomes were robust to all the investigated parameter changes.

			Spearman's correlation				Dot product			
	Purity		C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4
Case 1	0.91	S_B	1.00	0.99	0.99	0.99	1.00	1.00	0.99	0.99
		S_M	0.96	0.99	0.97	0.94	0.99	0.99	0.99	0.99
Case 2	0.95	S_B	1.00	1.00	0.98	0.98	1.00	1.00	0.99	0.99
		S_M	0.97	0.98	0.95	0.95	1.00	0.99	0.99	0.98
Case 3	0.94	S_B	1.00	0.99	0.94	0.95	1.00	0.99	0.97	0.97
		S_M	0.97	0.96	0.83	0.9	0.99	0.99	0.96	0.96
Case 4	0.96	S_B	1.00	0.99	0.98	0.99	1.00	0.99	0.99	0.99
		S_M	0.99	0.97	0.95	0.94	1.00	0.99	0.99	0.99
Case 5	0.93	S_B								
		S_M								
Case 6	0.72	S_B								
		S_M								
Case 7	0.65	S_B	0.99	0.91	0.92	0.97	1.00	0.94	0.94	0.97
		S_M								
Case 8	0.69	S_B	1.00	0.99	0.99	0.99	1.00	1.00	1.00	0.99
		S_M								
Case 9	0.92	S_B	1.00	0.99	0.98	0.97	1.00	0.99	0.99	0.98
		S_M								
Case 10	0.69	S_B	0.99	0.98	0.97	0.96	1.00	0.99	0.98	0.97
		S_M								

Table 2: **Robustness of the results to parameter changes.** Cases 5 and 6 involved changes in the number of nearest neighbours, which does not alter the generation of PLS components; for this reason, robustness quantification is not provided in these cases. Similarly, spatio-temporal features generated with a different number of time bins (cases 7 to 10) do not enable to compare motion saliences. SB: behavioural saliences. SM: motion saliences.

3. Additional mathematical details

Graph theory, spectral clustering and consensus clustering

There exist many ways to cluster a dataset into a subset of distinct groups. In our case, we applied *spectral clustering* on our data matrix $\mathbf{X} \in \mathbb{R}^{S \times M}$, with S the number of data points at hand and M the number of spatio-temporal motion features. Each data point can thus be expressed as a vector $\mathbf{x_i}$ of size M, arranged as the rows of $\mathbf{X}$.

Spectral clustering necessitates (1) the definition of a graph summarising the data at hand, (2) the extraction of meaningful components summarising the data based on the graph architecture (which can be understood as a dimensionality reduction approach), and (3) classification using the extracted components. We go through these three steps in details below.

Graph definition

Let us consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is its constituting set of nodes and $\mathcal{E}$ the set of edges linking these nodes. We denote by $w_{i,j}$ the edge weight between nodes i and j; a larger value indicates a closer similarity between the nodes.

Graphs can be used to represent a wide array of systems, such as transportation networks, metabolic networks or social networks. In neuroscience, a classical approach has been to define nodes as brain regions, and edges as their structural or functional connectivity.

Here, we adopt another approach in which we define the nodes as the S sessions considered in our analyses. Edge weights were set using an N-nearest neighbour criterion, in which each node was linked to its N closest neighbours (as quantified from cosine similarity) only. Let $\mathcal{N}_i$ the N-neighbourhood of node i; edge weights were initialised as:

$$w_{i,j} = \begin{cases} e^{-\frac{d_{i,j}^2}{\sigma_i^2}} \text{ if } j \in \mathcal{N}_i \\ 0 \text{ else,} \end{cases} \tag{1}$$

where $d_{i,j}$ is the cosine distance between data points i and j, and σ_i is the average of all distances between i and its N nearest neighbours. Weights can take values between 0 (infinitely distant/non-neighbouring data points) and 1 (identical data points).

Edge weights can be efficiently summarised into the adjacency matrix $\mathbf{A}$ of the system. The adjacency matrix obtained in our main analyses (from session 1 data) is displayed in Figure 1. From this description, it is already quite clear that subdividing the data into four clusters (as explained below) provides a very good solution.

Dimensionality reduction

We first define the symmetric, positive-definite Laplacian matrix of the system as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is a diagonal matrix containing nodal degrees. The nodal degree of node i is the sum of incoming edges: $d_i = \sum_{j \in \mathcal{N}_i} w_{i,j}$. In our analyses, we considered the normalised version of the Laplacian, given by:

$$\mathbf{L_N} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}}. \tag{2}$$

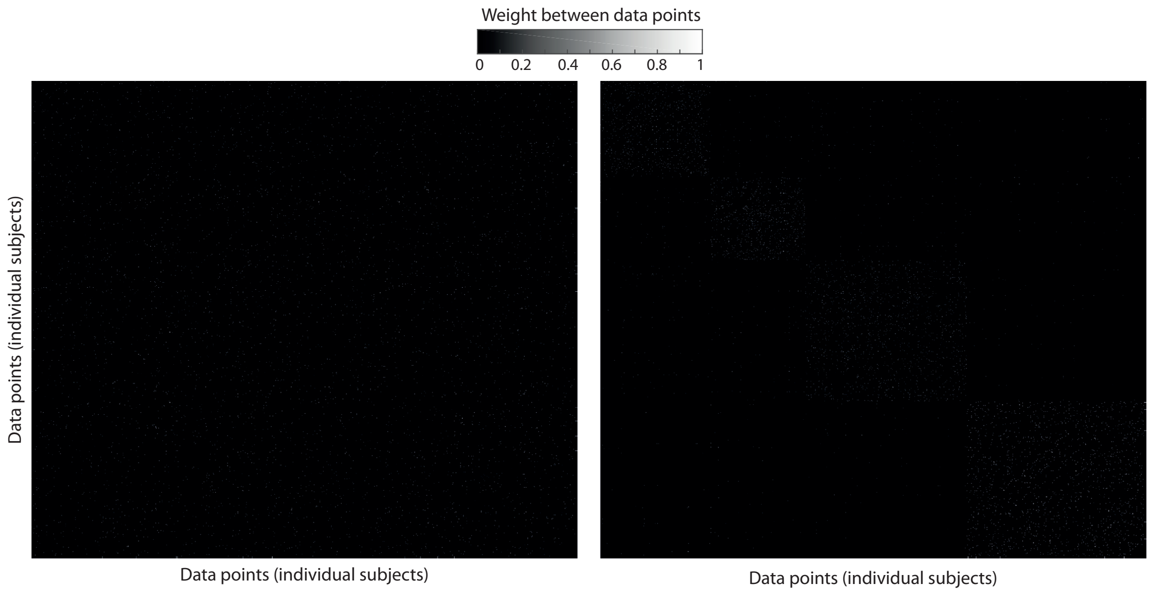

Figure 1: **Adjacency matrix of spatio-temporal movers.** Unsorted (left) and sorted (right) adjacency matrices of the data.

$\mathbf{L_N}$ can equivalently be expressed, using an eigenvalue decomposition, as $\mathbf{L_N} = \mathbf{U\Sigma U}^\top$. In this expression, $\mathbf{U}$ is the matrix of eigenvectors (arranged in columns) and $\mathbf{\Sigma}$ is a diagonal matrix containing the associated eigenvalues in its diagonal. We consider sorted eigenvalue/eigenvector pairs in decreasing eigenvalue order.

The first three eigenvectors with non-null eigenvalue happen to be an optimal basis for classification; expressing our M-dimensional data points in this three-dimensional space thus operates as a nonlinear dimensionality reduction approach. To illustrate this, we plot the representation of these three eigenvectors (for our main analyses) in Figure 2.

Clustering

To partition the data into clusters, k-means clustering is performed on the $S \times 3$ dimensionally reduced dataset. To select the optimal number of clusters, we used consensus clustering (Monti et al., 2003), a subsampling-based assessment of robustness.

In more details, the clustering process was repeatedly run (100 times) over 80% of the data points, for increasing cluster number values K. in each case, a consensus matrix summarising how frequently two data points would be clustered together was derived. Since the goal in a good clustering scheme is to either always cluster two data points together, or to never do so, the goal is to find a K for which the proportion of ambiguously clustered pairs—PAC (Senbabaoglu et al., 2014), linked to intermediate consensus values, is the lowest.

In Figure 3, we provide (for our main analyses centred on session 1 data) the consensus matrices obtained for $K = 2$ to $K = 17$. Their inspection further confirms our choice of $K = 4$.

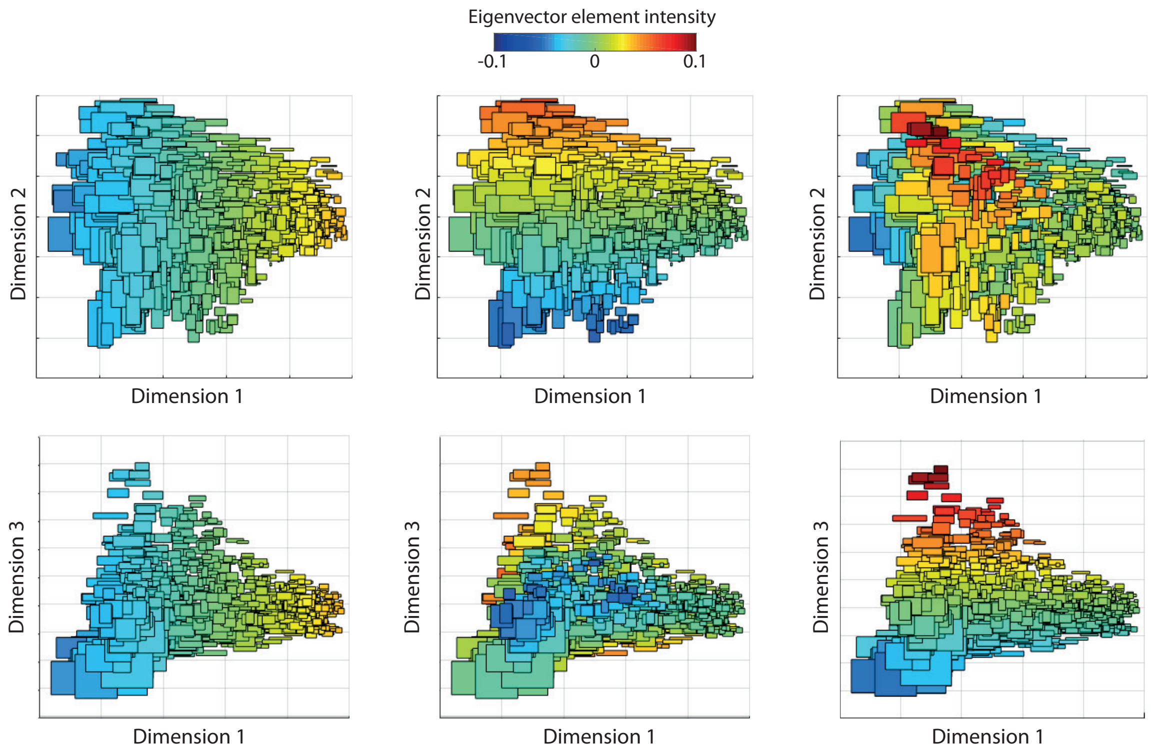

Figure 2: **First three discriminating eigenvectors.** For $\mathbf{u_1}$ (left column), $\mathbf{u_2}$ (middle column) and $\mathbf{u_3}$ (right column), representations of the eigenvectors along two selected dimensions. Each rectangle highlights one subject of the analysis, with larger width along the first, second and third dimensions respectively indicating larger average movement along the X, Y and Z axes.

Partial Least Square (PLS) analysis

PLS is a multivariate approach that enables to extract co-varying components between two types of measures. In what follows, we will be considering the matrix of spatio-temporal motion features $\mathbf{X} \in \mathbb{R}^{M \times S}$, and the matrix of behavioural domain scores $\mathbf{Y} \in \mathbb{R}^{B \times S}$, where S is the number of subjects, M the number of spatio-temporal motion features, and B the number of behavioural measures.

The goal in PLS analysis is to extract covariance components from the data. To do so, we first consider the covariance matrix between spatio-temporal motion features and behavioural domain scores:

$$\mathbf{R} = \mathbf{X}\mathbf{Y}^{\top} \in \mathbb{R}^{M \times B}. \tag{3}$$

This matrix can be equivalently expressed in the form of a singular value decomposition:

$$\mathbf{R} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top}. \tag{4}$$

In the above equation, the matrix $\mathbf{U}$ contains the left singular vectors of $\mathbf{R}$, arranged in successive columns. These vectors form an orthonormal basis; *i.e.*, $\mathbf{U}^{\top}\mathbf{U} = \mathbf{I}$. The same property applies to the right singular vectors, arranged in the columns of the matrix $\mathbf{V}$. As for $\mathbf{\Sigma}$, it is a diagonal matrix containing the singular

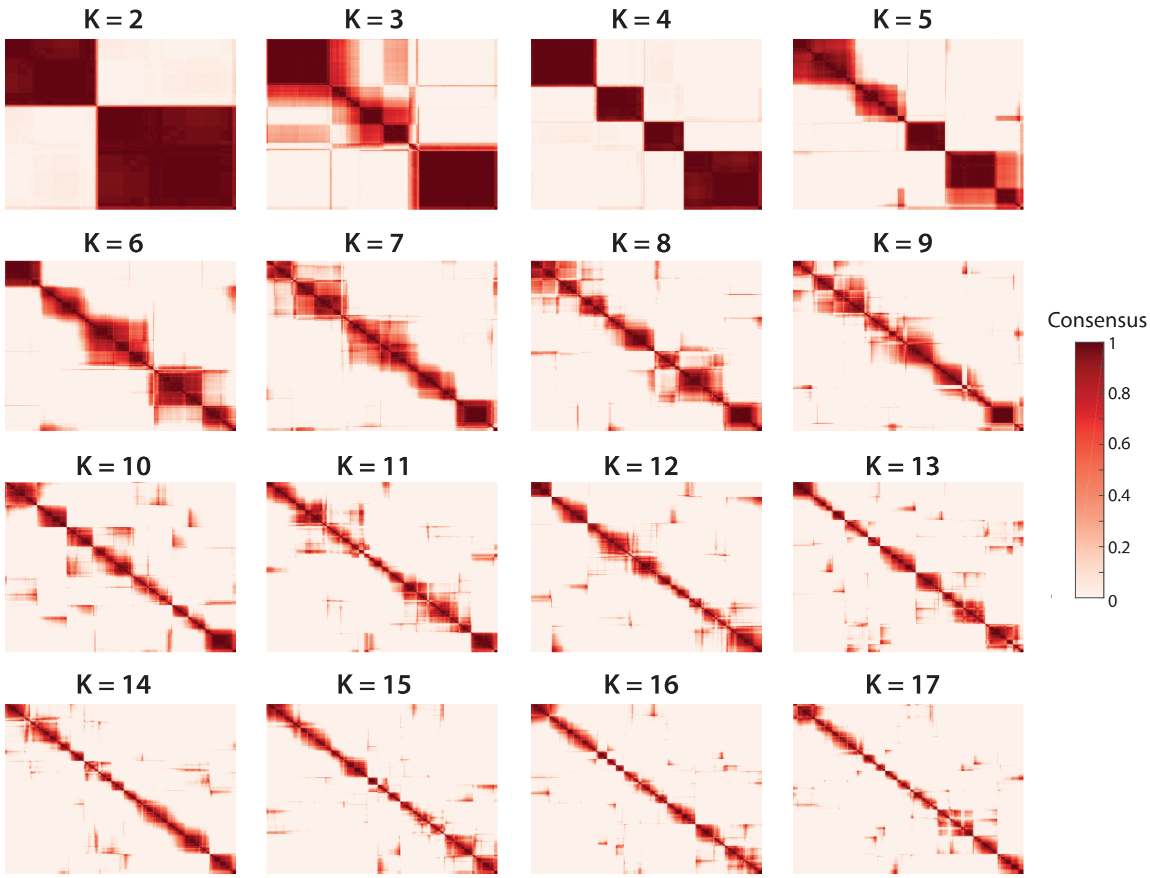

Figure 3: **Consensus matrices across cluster number values.** For K from 2 (top left) to 17 (bottom right), consensus matrices reflecting robustness of the partitioning across folds.

values $\{\sigma_i\}, i = 1, 2, ..., \min(B, M)$ as its diagonal elements. We assume here that singular vectors and singular values are sorted in decreasing singular value order.

The intuition behind this decomposition is that the full covariance between both datasets is expressed as a weighted low-rank approximation:

$$\mathbf{R} = \sum_{i=1}^{\min(B,M)} \sigma_i \mathbf{u_i} \mathbf{v_i}^\top, \tag{5}$$

where $\mathbf{u_i}$ and $\mathbf{v_i}$ are the i^{th} left and right singular vectors, respectively. Since $\mathbf{U}$ and $\mathbf{V}$ can both be seen as orthonormal bases, it follows that the strength of expression of the components in the investigated pool of subjects can be simply expressed as a projection:

$$\begin{cases} \mathbf{L_M} = \mathbf{U}^\top \mathbf{X} \\ \mathbf{L_B} = \mathbf{V}^\top \mathbf{Y}, \end{cases} \tag{6}$$

with $\mathbf{L_M}$ denoting the strength of expression of spatio-temporal motion features, or *motion latent weights*, across subjects, and $\mathbf{L_B}$ that of behavioural domain scores (called *behavioural latent weights*). The i^{th} column of $\mathbf{L_M}$ or $\mathbf{L_B}$ contains the weights associated to the i^{th} component, while the j^{th} row contains all weights associated to subject j.

4. Details on the generation of behavioural summarising measures

Selection of individual scores

We chose not to include some types of HCP scores into our analysis, because they highlighted family relationships or attributes falling beyond the scope of the present work, or were available in a too limited fraction of subjects. This included:

- Family relationships between subjects and twin status.
- Psychiatric history of the mother or father.
- Scores reflective of the menstrual cycle (only available in female subjects).
- Other irrelevant scores to the present study (for instance, ”Is the subject born in Missouri? “).

In addition, some scores were also not retained because they were considered too specific (that is, would induce overfitting if included), or overlapped with others (for example, scores reflective of task performance in a given condition were sometimes discarded if they strongly correlated with the more general performance measures available across conditions). In several cases, *age adjusted* and *unadjusted* scores are provided; both were kept and aggregated in further processing steps.

Conversion into summarising behavioural measures

Because some behavioural domains included many more scores than others, and in order to conduct a balanced analysis, we converted the scores of each domain into only one value through probabilistic principal component analysis—PPCA (Bishop, 1999), which enabled, at the same time, to fill in the few missing entries in each case.

Following PPCA, we only retained the domain scores that were sufficiently accurate, according to criteria proposed by Smith et al. (2015). Exclusion criteria were:

- More than 5 unavailable entries (in the case of domain scores that were derived from only one HCP score and did thus not undergo PPCA).
- The case $\max(\mathbf{z}) > 100\bar{z}$, with $\mathbf{z} = (\mathbf{b} - \mathbf{I}\,\mathrm{med}(\mathbf{b}))^2$, $\mathbf{b} \in \mathbb{R}^{S\times 1}$ the values for a given score across subjects, and $\mathbf{I}$ a unitary vector of appropriate size.
- More than 95% of subjects showing the same score value.

Finally, the remaining domain scores underwent rank-based inverse Gaussian transformation. The final matrix of behavioural information used for the analyses had size 951×60.

List of considered scores and derived summarising measures (HCP dataset)

Below, we provide details regarding the sets of original HCP scores that were considered together in generating a given behavioural summarising measure. We provide correlation values across subjects between individual scores and their associated output summarising measure, and also give the percentage of unavailable entries across subjects for each individual score. In the cases where only one score was available (*i.e.*, no need of a PPCA step), "n.a." is reported as a table entry.

Bodily features	*Weight*	*Height*	*Blood pressure*	$P_{NaN}[\%]$
Height	0.59	0.83		0
Weight	1.00	0.06		0
BMI	0.83	-0.44		0
BPSystolic			0.91	0.53
BPDiastolic			0.91	0.53

Table 3: **Components constructed from scores indicative of bodily features.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

Arousal	*MMSE*	*Sleep*	$P_{NaN}[\%]$
MMSE_Score	n.a.		0
PSQI_Comp1		0.73	0
PSQI_Comp2		0.52	0
PSQI_Comp3		0.56	0
PSQI_Comp4		0.58	0
PSQI_Comp5		0.53	0
PSQI_Comp6		0.28	0
PSQI_Comp7		0.46	0

Table 4: **Components constructed from scores indicative of arousal.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score. MMSE: mini mental state examination.

Cogn func	*PicSeq*	*CardSort*	*Flank*	*PMAT*	*ReadEng*	*PicVocab*	*ProcSpeed*	*DDISC*	*VSPLOT*	*VSPLOT (RT)*	*SCPT (Perf)*	*IWRD*	*ListSort*	*ER (Corr - RT)*	*ER (Sad)*	P_{NaN}[%]
PicSeq_Unadj	1.00															0
PicSeq_AgeAdj	1.00															0
CardSort_Unadj		0.99														0.21
CardSort_AgeAdj		0.99														0.21
Flanker_Unadj			1.00													0
Flanker_AgeAdj			1.00													0
PMAT24_A_CR				0.93												0
PMAT24_A_SI				-0.92												0
PMAT24_A_RTCR				0.91												0
ReadEng_Unadj					1.00											0
ReadEng_AgeAdj					1.00											0
PicVocab_Unadj						0.99										0
PicVocab_AgeAdj						0.99										0
ProcSpeed_Unadj							1.00									0
ProcSpeed_AgeAdj							1.00									0
DDISC_AUC_200								0.90								0
DDISC_AUC_40K								0.94								0
VSPLOT_TC									0.97	-0.13						0
VSPLOT_CRTE									0.27	0.95						0
VSPLOT_OFF									-0.96	0.14						0
SCPT_TP											0.24					0
SCPT_TN											1.00					0
SCPT_FP											-1.00					0
SCPT_FN											-0.24					0
SCPT_TPRT											0.59					0.32
SCPT_SEN											0.24					0
SCPT_SPEC											1.00					0
SCPT_LRNR											0.21					0
IWRD_TOT												0.83				0
IWRD_RTC												-0.75				0
ListSort_Unadj													0.99			0
ListSort_AgeAdj													1.00			0
ER40_CR														-0.74		0
ER40_CRT														0.69		0
ER40ANG															0.67	0
ER40FEAR															0.48	0
ER40HAP															0.25	0
ER40NOE															-0.05	0
ER40SAD															0.66	0

Table 5: **Components constructed from scores indicative of cognitive functions.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score. Cogn func: cognitive functions. PicSeq: picture sequence memory. CardSort: dimensional change card sort. Flank: Flanker task. PMAT: Penn progressive matrices. ReadEng: oral reading recognition. PicVocab: picture vocabulary. ProcSpeed: pattern completion processing speed. DDISC: delay discounting. VSPLOT: variable short Penn line orientation test. RT: response time. SCPT: short Penn continuous performance test. Perf: performance. IWRD: Penn word memory test. ListSort: list sorting. ER: Penn emotion recognition test. Corr: correct.

Affect	*Negative emotions*	*Positive emotions*	*Support [NEG]*	*Self efficacy - Stress*	P_{NaN}[%]
AngAffect_Unadj	0.83				0
AngHostil_Unadj	0.62				0
AngAggr_Unadj	0.32				0
FearAffect_Unadj	0.80				0
FearSomat_Unadj	0.54				0
Sadness_Unadj	0.83				0
LifeSatisf_Unadj		0.84			0
MeanPurp_Unadj		0.82			0
PosAffect_Unadj		0.80			0
Friendship_Unadj		-0.73			0
Loneliness_Unadj			0.79		0
PercHostil_Unadj			0.63		0
PercReject_Unadj			0.81		0
EmotSupp_Unadj			-0.79		0
InstruSupp_Unadj			-0.61		0
PercStress_Unadj				-0.84	0
SelfEff_Unadj				0.86	0

Table 6: **Components constructed from scores indicative of affect.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

Personality	*NEOFAC (Introverted)*	*NEOFAC (Daring)*	P_{NaN}[%]
NEOFAC_A	-0.60	0.27	0.11
NEOFAC_O	0.03	0.90	0.11
NEOFAC_C	-0.68	-0.35	0.11
NEOFAC_M	0.77	0.06	0.11
NEOFAC_E	-0.67	0.22	0.11

Table 7: **Components constructed from scores indicative of personality.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

Task performance	*Emot (RT)*	*Emot (Acc)*	*Gamb (Conservative)*	*Gamb (Daring)*	*Gamb (Bad play)*	*Langu (overall perf.)*	*Rel (RT)*	*Rel (Acc)*	*ToM*	*WM [NEG]*	P_{NaN}[%]
Emotion_Task_Acc	0.18	0.95									0
Emotion_Task_Median_RT	0.99	-0.22									0
Emotion_Task_Face_Acc	0.08	0.62									0
Emotion_Task_Face_Median_RT	0.94	-0.19									0
Emotion_Task_Shape_Acc	0.17	0.72									0
Emotion_Task_Shape_Median_RT	0.93	-0.23									0
Gambling_Task_Reward_Perc_Larger			0.44	-0.68	0.07						0
Gambling_Task_Reward_Median_RT_Larger			0.80	0.62	0.05						1.89
Gambling_Task_Reward_Perc_Smaller			-0.44	0.68	-0.07						0
Gambling_Task_Reward_Median_RT_Smaller			0.83	0.55	0.00						0.84
Gambling_Task_Punish_Perc_Larger			0.27	-0.41	0.95						0
Gambling_Task_Punish_Median_RT_Larger			0.80	0.58	0.01						0.95
Gambling_Task_Punish_Perc_Smaller			-0.27	0.41	-0.95						0
Gambling_Task_Punish_Median_RT_Smaller			0.80	0.57	0.08						0.32
Language_Task_Story_Acc						0.45					0
Language_Task_Story_Median_RT						-0.08					0
Language_Task_Story_Avg_Difficulty_Level						0.85					0
Language_Task_Math_Acc						0.82					0
Language_Task_Math_Median_RT						-0.33					0
Language_Task_Math_Avg_Difficulty_Level						0.58					0
Relational_Task_Match_Acc							-0.50	0.63			0.21
Relational_Task_Match_Median_RT							0.89	0.23			0.21
Relational_Task_Rel_Acc							-0.38	0.78			0.21
Relational_Task_Rel_Median_RT							0.72	0.54			0.21
Social_Task_Random_Perc_Random									0.62		0
Social_Task_Random_Median_RT_Random									-0.55		0.74
Social_Task_Random_Perc_TOM									-0.49		0
Social_Task_TOM_Perc_Random									-0.44		0.11
Social_Task_TOM_Perc_TOM									0.48		0.11
Social_Task_TOM_Median_RT_TOM									-0.43		0.21
WM_Task_2bk_Acc										-0.66	0.63
WM_Task_2bk_Median_RT										0.73	0.74
WM_Task_0bk_Acc										-0.65	0
WM_Task_0bk_Median_RT										0.86	0.63

Table 8: **Components constructed from scores indicative of task performance.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score. Emot: emotion task. RT: response time. Acc: accuracy. Gamb: gambling task. Langu: language task. Rel: relational task. ToM: theory of mind. WM: working memory.

Motor abilities	*Endurance*	*Gait speed*	*Dexterity*	*Strength*	$P_{NaN}[\%]$
Endurance_Unadj	1.00				0.21
Endurance_AgeAdj	1.00				0.21
GaitSpeed_Comp		n.a.			0
Dexterity_Unadj			1.00		0
Dexterity_AgeAdj			1.00		0
Strength_Unadj				1.00	0.11
Strength_AgeAdj				1.00	0.11

Table 9: **Components constructed from scores indicative of motor abilities.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

Sensory perception	*Odour*	*Pain*	*Taste*	$P_{NaN}[\%]$
Odor_Unadj	0.96			0
Odor_AgeAdj	1.00			0
PainInterf_Tscore		n.a.		0
Taste_Unadj			1.00	0.21
Taste_AgeAdj			1.00	0.21

Table 10: **Components constructed from scores indicative of sensory perception.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

Substance use	*Alcohol*	*Drugs*	$P_{NaN}[\%]$
Total_Drinks_7days	-0.82		1.05
Num_Days_Drank_7days	-0.75		0
SSAGA_Alc_D4_Dp_Sx	-0.58		0
SSAGA_Alc_D4_Ab_Sx	-0.51		0
SSAGA_Alc_12_Drinks_Per_Day	-0.67		3.79
SSAGA_Alc_12_Frq	0.83		3.68
SSAGA_Alc_12_Frq_Drk	0.77		3.68
SSAGA_Times_Used_Illicits		0.75	0
SSAGA_Times_Used_Cocaine		0.42	0
SSAGA_Times_Used_Hallucinogens		0.57	0
SSAGA_Times_Used_Opiates		0.49	0
SSAGA_Times_Used_Sedatives		0.41	0
SSAGA_Times_Used_Stimulants		0.49	0
SSAGA_Mj_Times_Used		0.93 (0.92)	0 (0)

Table 11: **Components constructed from scores indicative of substance use.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

Psychometric features	Anxiety	Withdrawal	Som prob	Thought prob	Attention prob	Aggr	Rule break	Intr	Intern	Extern	Dep	Avoid	ADHD	Inatt	Hyperresp	Antisoc	P_{NaN}[%]
ASR_Anxd_Raw	0.95																0
ASR_Anxd_Pct	0.89																0
DSM_Anxi_Raw	0.95																0
DSM_Anxi_Pct	0.92																0
ASR_Witd_Raw		1.0															0
ASR_Witd_Pct		0.96															0
ASR_Soma_Raw			0.98														0
ASR_Soma_Pct			0.95														0
DSM_Somp_Raw			0.95														0
DSM_Somp_Pct			0.95														0
ASR_Thot_Raw				1.00													0
ASR_Thot_Pct				0.96													0
ASR_Attn_Raw					1.00												0
ASR_Attn_Pct					0.98												0
ASR_Aggr_Raw						1.00											0
ASR_Aggr_Pct						0.96											0
ASR_Rule_Raw							1.00										0
ASR_Rule_Pct							0.99										0
ASR_Intr_Raw								1.00									0
ASR_Intr_Pct								0.97									0
ASR_Intn_Raw									1.00								0
ASR_Intn_T									1.00								0
ASR_Extn_Raw										1.00							0
ASR_Extn_T										1.00							0
DSM_Depr_Raw											0.89						0
DSM_Depr_Pct											0.87						0
SSAGA_Depressive_Ep											0.48						2.52
SSAGA_Depressive_Sx											0.61						2.52
DSM_Avoid_Raw												1.00					0
DSM_Avoid_Pct												0.98					0
DSM_Adh_Raw													1.00				0
DSM_Adh_Pct													0.99				0
DSM_Inat_Raw														n.a.			0
DSM_Hype_Raw															n.a.		0
DSM_Antis_Raw																1.00	0
DSM_Antis_Pct																0.96	0

Table 12: **Components constructed from scores indicative of clinical features.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score. Som prob: somatic problems. Thought prob: thought problems. Attention prob: attention problems. Aggr: aggressiveness. Intr: intrusive. Intern: internalising. Extern: externalising. Dep: depression. Avoid: avoidance. ADHD: attention deficit hyperactivity disorder. Inatt: inattention problems. Hyperresp: hyperresponsiveness. Antisoc: antisocial behaviour.

List of considered scores and derived summarising measures (UCLA dataset)

Below, we provide details regarding the sets of original UCLA dataset scores that were considered together in generating a given behavioural summarising measure. We provide correlation values across subjects between individual scores and their associated output summarising measure. There were no missing entries in the pool of subjects that we analysed.

Bodily features	*Weight*	*Height*
weight_pounds	n.a.	
height_inches		n.a.

Table 13: **Components constructed from scores indicative of bodily features.** Rows stand for individual scores, and columns for behavioural summarising measures.

Arousal	*Chronotype (MCTQ)*
chronotype	n.a.

Table 14: **Components constructed from scores indicative of arousal.** Rows stand for individual scores, and columns for behavioural summarising measures.

Cognitive functions	*Verbal learning (CVLT)*	*Intelligence (WAIS)*	*Memory (WMS)*
SDfreeRecall	0.95		
SDcuedRecall	0.95		
LDfreeRecall	0.96		
LDcuedRecall	0.96		
LDrecognition	0.72		
wais_letterNumberSequ		0.51	
wais_vocabulary		0.99	
wais_matrixReasoning		0.56	
wms_vr_immRec			0.84
wms_vr_delRec			0.95
wms_vr_recog			0.60
wms_symbolSpan			0.78
wms_digitSpan_fwd			0.27
wms_digitSpan_bwd			0.48
wms_digitSpan_seq			0.48

Table 15: **Components constructed from scores indicative of cognitive functions.** Rows stand for individual scores, and columns for behavioural summarising measures.

Task performance	*Task-switching [BAD]*	*Delay discounting (DDT)*	*Attention network (ANT)*	*Continuous performance RT (CPT)*	*Continuous performance (CPT) [BAD]*	*Response inhibition RT (SCWT)*	*Response inhibition RT (SST)*
taskswitch_acc	0.01						
taskswitch_interference	0.30						
taskswitch_switchCost	0.94						
taskswitch_residSwitchCost	0.53						
delayDisc_total		0.90					
delayDisc_mediumRewards		0.95					
delayDisc_largeRewards		0.93					
delayDisc_total		1.00					
attention_RT_conflict			n.a.				
cpt_hitRate				-0.11	-0.48		
cpt_falseAlarmRate				-0.5	0.85		
cpt_hits_medianRT				1.00	-0.03		
stroop_conflict_acc						-0.21	
stroop_conflict_RT						1.00	
stopSignal_quantileRT							n.a.

Table 16: **Components constructed from scores indicative of task performance.** Rows stand for individual scores, and columns for behavioural summarising measures.

Personality	*Personality (MPQ)*
mpq	n.a.

Table 17: **Components constructed from scores indicative of personality.** Rows stand for individual scores, and columns for behavioural summarising measures.

Substance consumption	*Tobacco*	*Alcohol*
smoking_current	n.a.	
alcohol		n.a.

Table 18: **Components constructed from scores indicative of substance consumption.** Rows stand for individual scores, and columns for behavioural summarising measures. The last column contains the percentage of unavailable data entries across subjects for each score.

References

Bishop, C. M., 1999. Bayesian pca. Curran Associates, pp. 382–388.

Monti, S., Tamayo, P., Mesirov, J., Golub, T., 2003. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Machine Learning 52 (1), 91–118.

Senbabaoglu, Y., Michailidis, G., Li, J. Z., 2014. Critical limitations of consensus clustering in class discovery. Scientific Reports 4, 6207.

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E. J., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C., Miller, K. L., 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nature Neuroscience 18 (11), 1565–1567.

Yang, Z., Hao, T., Dikmen, O., Chen, X., Oja, E., 2012. Clustering by nonnegative matrix factorization using graph random walk. In: Advances in Neural Information Processing Systems. Curran Associates, pp. 1079–1087.