

This is a repository copy of *Patterns of thought:Population variation in the associations between large-scale network organisation and self-reported experiences at rest.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/134117/>

Version: Accepted Version

Article:

Wang, Hao Ting, Bzdok, Danilo, Margulies, Daniel S et al. (4 more authors) (2018)

Patterns of thought:Population variation in the associations between large-scale network organisation and self-reported experiences at rest. *Neuroimage*. pp. 518-527. ISSN 1053-8119

<https://doi.org/10.1016/j.neuroimage.2018.04.064>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/324829283>

Patterns of thought: Population variation in the associations between large-scale network organisation and self-reported experiences at rest

Article in *NeuroImage* · May 2018

DOI: 10.1016/j.neuroimage.2018.04.064

CITATIONS

0

READS

185

7 authors, including:



Hao-Ting Wang

The University of York

9 PUBLICATIONS **60 CITATIONS**

[SEE PROFILE](#)



Danilo Bzdok

RWTH Aachen University

94 PUBLICATIONS **2,577 CITATIONS**

[SEE PROFILE](#)



Daniel S Margulies

Max Planck Institute for Human Cognitive and Brain Sciences

176 PUBLICATIONS **8,681 CITATIONS**

[SEE PROFILE](#)



Elizabeth Jefferies

The University of York

128 PUBLICATIONS **4,836 CITATIONS**

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Activation Likelihood Estimation [View project](#)



Current projects [View project](#)

Patterns of thought: population variation in the associations between large-scale network organisation and self-reported experiences at rest

Hao-Ting Wang¹, Danilo Bzdok^{2,3,4}, Daniel Margulies⁵, Cameron Craddock⁶, Michael Milham⁶, Elizabeth Jefferies¹ & Jonathan Smallwood¹ .

¹ Department of Psychology, The University of York, Heslington, York, England.

² Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Germany

³JARA-BRAIN, Jülich-Aachen Research Alliance, Germany

⁴Parietal team, INRIA, Neurospin, bat 145, CEA Saclay, 91191 Gif-sur-Yvette, France

⁵ Neuroanatomy and Connectivity Group, Max Plank Institute for Cognition and Brain Sciences, Leipzig, Germany.

⁶ Center for Biomedical Imaging and Neuromodulation, Nathan S. Kline Institute for Psychiatric Research, New York State Office of Mental Health

Address for correspondence:

Hao-Ting Wang, Department of Psychology, The University of York, Heslington, York, England, UK.

Email: haoting.wang@york.ac.uk

Abstract

Contemporary cognitive neuroscience recognises unconstrained processing varies across individuals, describing variation in meaningful attributes, such as intelligence. It may also have links to patterns of on-going experience. This study examined whether dimensions of population variation in different modes of unconstrained processing can be described by the associations between patterns of neural activity and self-reports of experience during the same period. We selected 258 individuals from a publicly available data set who had measures of resting-state functional magnetic resonance imaging, and self-reports of experience during the scan. We used machine learning to determine patterns of association between the neural and self-reported data, finding variation along four dimensions. ‘Purposeful’ experiences were associated with lower connectivity - in particular default mode and limbic networks were less correlated with attention and sensorimotor networks. ‘Emotional’ experiences were associated with higher connectivity, especially between limbic and ventral attention networks. Experiences focused on themes of ‘personal importance’ were associated with reduced functional connectivity within attention and control systems. Finally, visual experiences were associated with stronger connectivity between visual and other networks, in particular the limbic system. Some of these patterns had contrasting links with cognitive function as assessed in a separate laboratory session - purposeful thinking was linked to greater intelligence and better abstract reasoning, while a focus on personal importance had the opposite relationship. Together these findings are consistent with an emerging literature on unconstrained states and also underlines that these states are heterogeneous, with distinct modes of population variation reflecting the interplay of different large-scale networks.

24 **1 Introduction**

25 Unconstrained processing reflects important population level variation in measures of
26 cognition, affect, and demographic / lifestyle factors. Psychological studies show that almost a
27 third of on-going thought is unconstrained by events in the here-and-now (Killingsworth &
28 Gilbert, 2010) with important links to cognitive and affective processing (Mooneyham &
29 Schooler, 2013). In neuroscience, metrics defined from the brain during wakeful rest, describe
30 the organisation of neural function at both the micro and macro scale (Glasser et al., 2016;
31 Margulies et al., 2016). They also reflect individual differences in cognitive function (Finn et
32 al., 2015), psychiatric conditions (Nooner et al., 2012) and demographic / lifestyle factors
33 (Smith et al., 2015). These findings establish unconstrained neuro-cognitive processing as a
34 core element of human cognition, highlighting the need to formally understand the underlying
35 neural architecture, and the associated patterns of experience.

36 One perspective on unconstrained processing emphasises the role of memory, with
37 contributions of conceptual and episodic representations to on-going thought (Binder, Desai,
38 Graves, & Conant, 2009; Gusnard, Raichle, & Raichle, 2001). Psychological studies have
39 shown patterns of unconstrained processing have links with memory retrieval, creativity and
40 planning (Baird et al., 2012; Leszczynski et al., 2017; Medea et al., 2016; Poerio et al., 2017).
41 Such evidence raises the possibility that episodic representations anchored in the medial
42 temporal lobe (Moscovitch, Cabeza, Winocur, & Nadel, 2016) or conceptual representation
43 anchored in anterior temporal lobe (Ralph, Jefferies, Patterson, & Rogers, 2017) contribute to
44 on-going thought (Smallwood et al., 2016). It is hypothesised that these systems contribution
45 to unconstrained states may be linked to the ability for these regions to become functionally
46 decoupled from systems more directly involved in action and perception, allowing them to
47 operate in an offline manner (Smallwood, 2013). This process of decoupling may also be
48 important in neural systems closely allied to those involved in memory – the default mode
49 network (Raichle et al., 2001). These regions of transmodal cortex are relatively distant in
50 functional and structural space from systems involved in perception and action, potentially
51 facilitating their role in stimulus independent aspects of cognition (Buckner & Krienen, 2013;
52 Margulies et al., 2016; Mesulam, 1998). Together these ‘representational’ accounts of
53 unconstrained processing highlight default mode and limbic networks as important candidate
54 neural systems, especially when decoupled from systems directly involved in perception and
55 action.

56 Alternative perspectives on unconstrained thought emerge from links between types of
57 on-going experience and problems maintaining a task relevant goal in mind. This “executive-
58 failure” view (Kane & McVay, 2012; McVay & Kane, 2010) takes as a starting point evidence
59 that patterns of on-going thought, such as the experience of mind-wandering, are linked to
60 problems on tasks including sustained attention (McVay & Kane, 2009) and measures of
61 general aptitude and executive control (Mrazek et al., 2012). Task-based neuroimaging
62 investigations highlight a network of regions that increase their activity across many different
63 task situations - so called multiple demand regions (Duncan, 2010). These regions broadly
64 correspond to three well described intrinsic networks: ventral attention, dorsal attention, and
65 frontal-parietal networks. Since these systems are important for the effective performance of
66 many different tasks then dysregulation within these systems could reflect the hypothesised
67 ‘executive-failure’ contribution to aspects of on-going thought (McVay & Kane, 2010;
68 Weissman, Roberts, Visscher, & Woldorff, 2006).

69 Other aspects of unconstrained processing could reflect the importance of affective
70 processes, or different modalities of processing. On-going thought is linked to mood state:
71 Experimental inductions of mood (Smallwood, Fitzgerald, Miles, & Phillips, 2009; Smallwood
72 & O'Connor, 2011), as well as natural fluctuations (Poerio, Totterdell, & Miles, 2013; Ruby,
73 Smallwood, Engen, & Singer, 2013) impact on on-going thought. Contemporary accounts of
74 emotional processing emphasise the role of limbic regions including the amygdala (Bzdok,
75 Laird, Zilles, Fox, & Eickhoff, 2013; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012)
76 and anterior aspects the insula (Touroutoglou, Hollenbeck, Dickerson, & Barrett, 2012),
77 suggesting these regions may be important in determining affective aspects of on-going thought.
78 Psychological studies of on-going thought also suggest that another important dimension of
79 unconstrained processing may reflect the different modalities of processing (Konishi, Brown,
80 Battaglini, & Smallwood, 2017; Smallwood et al., 2016). It has been shown, for example, that
81 the visual system plays an important role in the expression of visual imagery (Ganis, Thompson,
82 & Kosslyn, 2004; Kosslyn, Ganis, & Thompson, 2001). Recent work has extended this
83 evidence to shown patterns of activity with visual regions are linked to the emergence of visual,
84 non-verbal, elements of on-going thought (Rajj & Riekkki, 2017). It is also possible that
85 sensorimotor processes may be implicated in language processing during unconstrained
86 processing, given that a role for these regions in language processing extends beyond
87 production (Bzdok et al., 2016; Pulvermuller, 2010; Pulvermuller & Fadiga, 2010).

88 **2 Current study**

89 Our study aimed to identify patterns of intrinsic connectivity associated with different
90 patterns of unconstrained states and examines their neuro-cognitive features from the
91 perspectives outlined above. We used a large publicly available dataset, containing measures
92 of resting-state functional magnetic resonance imaging (fMRI), and an accompanying self-
93 report instrument describing cognition experienced during the resting-state (Gorgolewski et al.,
94 2014; Nooner et al., 2012). We previously explored the relationships between patterns of on-
95 going thought and measures of neural activity, such as the fractional amplitude of low
96 frequency oscillations, as well as the regional homogeneity of neural activity, in a sub sample
97 of this data set (Gorgolewski et al., 2014). In this study we focused on connectivity, we applied
98 sparse canonical correlation analysis (SCCA) to obtain a conjoined decomposition of self-
99 reports of experience with matrices of whole brain connectivity data. This analysis produces
100 multivariate patterns that reflect dimensions of variation that are mutually constrained by both
101 brain and experience. In this way we capitalize on the fact that self-reports of experience during
102 scanning and descriptions of on-going neural processing provide complementary descriptions
103 of unconstrained cognition. Our analysis, therefore, helps define, at a population level, the
104 shared links between brain patterns and different types of experience. It is important to note
105 that this approach necessarily conflates state and trait related aspects of any brain-experience
106 associations that are identified in this manner, and this aspect of our design should be borne in
107 mind when interpreting our results. As our analytic approach respects the multivariate nature
108 of brain and behaviour space, it can accommodate complex many-to-many relationships
109 between patterns of connectivity and self-reports, and therefore is sensitive to the possibility of
110 complex relationships in the underlying data. We took two steps to explore the robustness of
111 the components that our study identifies. First, we use permutation testing to examine the extent
112 to which our components are different from those that would be achieved based on a null
113 distribution. Second we established whether these neuro-cognitive dimensions are associated
114 with performance on a battery of available cognitive tasks, including measures of executive
115 control and intelligence. When interpreting the results produced through our analysis it is
116 important to give greater weight to components that show evidence of robustness in both
117 comparisons.

118 We use the dimensions our analysis produces, and their links with cognitive function to
119 evaluate the perspectives on unconstrained thought outlined earlier. ‘Representational’
120 accounts emphasise links with neural systems involved in memory, such as the limbic system,

121 and regions of transmodal cortex, such as the default mode network. They highlight states with
122 lower levels of functional communication between these regions and those more directly
123 involved in external action. In contrast, ‘executive-failure’ accounts emphasise dysregulation
124 in attention and control networks as contributing to patterns of on-going thought that are linked
125 to problems in domain general task performance. Affective accounts highlight limbic regions
126 as important hubs in aspects of on-going thoughts linked to emotion. Finally, modality specific
127 influences on unconstrained thought may depend on information codes represented in regions
128 that specialise in that particular types of information, such as a role of visual cortex in
129 experiences dominated by images. Notably, some views lead to dissociable predictions with
130 respect to cognitive performance. For example, executive-failure accounts predict patterns of
131 thoughts linked to worse performance on measures of cognitive function, while
132 representational accounts makes the alternative prediction.

133 **3 Materials and Methods**

134 **3.1 Participant**

135 We analysed 258 participants (females = 162; age range 18 – 55, $M = 34.97$, $SD =$
136 12.24) obtained from the enhanced Nathan Kline Institute-Rockland sample (NKI-RS;
137 http://fcon_1000.projects.nitrc.org/indi/enhanced/). Full details of the acquisition of this
138 sample can be found in Nooner et al., 2012. We selected participants between 18 and 55
139 years old as our sample, this choice allowed us to maximise the cohesive nature of our
140 sample. All the participants have the MRI data and less than 5 missing data points among the
141 selected assessments.

142 **3.2 Cognitive measures and Questionnaires**

143 Based on prior studies examining the links between spontaneous thought and cognitive
144 performance (see Mooneyham & Schooler, 2013) we selected established neuropsychological
145 measures linked to executive control, abstract reasoning and intelligence. The measures
146 included the Delis-Kaplan Executive Function System (D-KEFS; Swanson, 2005), Wechsler
147 Abbreviated Scale of Intelligence (WASI-II; Wechsler, 1999), and Wechsler Individual
148 Achievement Test – Second Edition Abbreviated (WIAT-IIA; Wechsler, 2005). In D-KEFS
149 we selected the tower test (move accuracy ratio), colour-word interference test (errors
150 inhibition/switching), verbal fluency test (letter fluency - category fluency), design fluency

151 test (design accuracy), trail making test (sequencing errors score + set-loss errors score +
152 time-discontinue errors score), and the proverb test (a measure of abstract semantic
153 reasoning). We used the rescaled score ($M = 10$, $SD = 3$) in our analysis. Tasks measures that
154 reflected error rates (i.e. the colour-word interference test and trail making test) were
155 reversed, so that high rescaled scores indicated better task performance. All the scores were
156 transformed to z-scores.

157 **3.3 On-going cognition measure**

158 The New York Cognition Questionnaire (NYC-Q) is a self-report tool used to assess
159 the thoughts experienced at rest (Gorgolewski et al., 2014; Sanders, Wang, Schooler, &
160 Smallwood, 2017). It assesses thoughts and feelings experienced during the resting-state
161 period. The first section contains 23 questions about the content of thought. These questions
162 covers the temporal, social, emotional aspects of spontaneous thoughts that have been shown
163 to be important by prior studies (e.g. Ruby et al., 2013). Participants rated each question on a
164 scale of 1 (Completely did not describe my thoughts) to 9 (Completely did describe my
165 thoughts). The second section contains 8 questions about the forms thoughts take, capturing
166 aspects of experience such as modality and detail associated with experience that prior
167 studies suggest as important for spontaneous thoughts (Smallwood et al., 2016). Participants
168 rated each question on a scale of 1 (Completely did not characterise my experience) to 9
169 (Completely did characterise my experience). In the current study we analysed the two
170 sections together to provide single solutions that combined information on both the content
171 form of experience. The full list of questions and the corresponding labels are presented in
172 Table 1. The questionnaire was administrated once after the resting-state scan in order to
173 assess experiences during the scanning session. For the full details of the NYC-Q, please
174 refer to Gorgolewski et al., 2014. We have placed the questionnaire measure used in this
175 study along with an example self-report collection task on GitHub at the following address:
176 https://github.com/htwangtw/restingstate_thoughtreports.

177 **Table 1 The New York Cognition Questionnaire (NYC-Q)**

#	Questions	Labels
Q01	I thought about things I am currently worried about	Concerns
Q02	I thought about people I have just recently met	People
Q03	I thought of people I have known for a long time (friends)	Friend

Q04	I thought about members of my family	Family
Q05	I thought about an event that took place earlier today	Today - Past
Q06	I thought about an interaction I may possibly have in the future	Social - Future
Q07	I thought about an interaction with somebody that took place in the past	Social - Past
Q08	I thought about something that happened at a place very close to me	Near Location
Q09	I thought about something that made me feel guilty	Guilt
Q10	I thought about an event that may take place later today	Today - Plan
Q11	I thought about something that happened in the recent past (last couple of days but not today)	Recent Past
Q12	I thought about something that happened a long time ago in the past	Distant Past
Q13	I thought about something that made me angry	Anger
Q14	I thought about something that made me happy	Happiness
Q15	I thought about something that made me cheerful	Cheerfulness
Q16	I thought about something that made me calm	Calm
Q17	I thought about something that made me sad	Sadness
Q18	I thought about something that is important to me	Importance
Q19	I thought about something that could still happen today	Today - Future
Q20	I thought about something that may take place in the distant future	Distant Future
Q21	I thought about something that could take place in the near future (days or weeks but not today)	Near Future
Q22	I thought about personal worries	Worries
Q23	I thought about something that happened in a place far away from where I am now	Distant Location
Q24	In the form of images:	Image
Q25	In the form of words:	Words
Q26	Like an inner monologue or audiobook:	Monologue
Q27	Like a television program or film:	Film
Q28	Had a strong and consistent personal narrative:	Narrative
Q29	Had a clear sense of purpose:	Purpose
Q30	Vague and non-specific:	Vague
Q31	Fragmented and disjointed:	Fragment

178 **3.4 MR data processing**

179 **3.4.1 Resting-state fMRI.**

180 We used resting-state fMRI to describe the general functional organisation of the
181 brain. We selected resting-state multiband functional magnetic resonance imaging (R-
182 mfMRI; TR = 1400msec; voxel size = 2mm isotropic; duration = 10 minutes) for our
183 analysis. Functional and structural data were pre-processed using Configurable Pipeline for
184 the Analysis of Connectomes (C-PAC; <https://fcp-indi.github.io/>) to interface with FMRIB's
185 Software Library (FSL version 5.0, www.fmrib.ox.ac.uk/fsl). Individual FLAIR and T1
186 weighted structural brain images were extracted using Brain Extraction Tool (BET).
187 Structural images were linearly registered to the MNI-152 template using FMRIB's Linear
188 Image Registration Tool (FLIRT). The resting-state functional data were pre-processed and
189 analysed using the FMRI Expert Analysis Tool (FEAT). X, Y, Z displacement and the three
190 axis rotations were used to calculate the mean frame displacement (FD), characterising
191 movement of each participant during the scanning session (Power et al., 2014). Mean of the
192 absolute values for FD were later used to account for subject specific head motion. No global
193 signal regression was applied. The individual subject analysis involved: motion correction
194 using MCFLIRT; slice-timing correction using Fourier space time series phase-shifting;
195 spatial smoothing using a Gaussian kernel of FWHM 6 mm; bandpass filtering ($0.1 \text{ Hz} < f <$
196 0.009 Hz); six motion parameters (as estimated by MCFLIRT) regressed out; cerebrospinal
197 fluid and white matter signal regressed out (top five PCA components, CompCor method).

198 **3.4.2 Connectivity matrices**

199 To describe the functional architecture of the whole brain, we transformed the resting-
200 state BOLD time series into connection strength values of the different networks for each
201 participant. The whole brain parcellation was obtained from connectivity-based functional
202 parcellation created by Yeo and colleagues (Yeo et al., 2011). The 7 network parcellation was
203 used in the current study. We split the networks into two hemispheres and extracted clusters.
204 Two voxels are considered connected only if they are adjacent within the same x, y, or z
205 direction. This yielded 57 clusters from the Yeo 7 networks parcellation. The implementation
206 of spatial clusters extraction was retrieved from python library Nilearn (Abraham et al., 2014;
207 <http://nilearn.github.io/>, version 0.3.1). Next, we extracted and then averaged the time series
208 of all voxels within each cluster to create a cluster specific time series. We used these time
209 series to create region-to-region symmetrical correlation matrices representing the

210 correlations of the network signal that was computed for all the individual subjects. The off-
211 diagonal of each correlation matrix contained 1596 unique region-region connection strengths
212 (i.e., the upper or lower triangle of the network covariance matrix). This approach provided a
213 measure of connection strength of the whole brain for each participant. Finally, Fisher's r-to-
214 z transformation was applied to each network covariance matrix.

215 **3.5 Conjoint decomposition of functional connectomes and mind-** 216 **wandering measures**

217 **3.5.1 Decomposition method**

218 We performed a sparse canonical correlation analysis (SCCA; see Hastie, Tibshirani, &
219 Wainwright, 2015) on the functional connectomes and the NYC-Q reports, to yield latent
220 components that reflect multivariate patterns across neural organisation and experience (For
221 similar application, see Wang et al., 2017). SCCA maximised the linear correlation between
222 the low-rank projections of two sets of multivariate data sets with sparse model to regularise
223 the decomposition solutions a process that helps maximise the interpretability of the results.
224 The regularisation function of choice is L_1 penalty, which produces 'sparse' coefficients,
225 meaning that the canonical vectors (i.e., translating from full variables to a data matrix's low-
226 rank components of variation) will contain a number of exactly zero elements. L_1
227 regularisation conducted (i) feature selection (i.e., select only relevant components) and (ii)
228 model estimation (i.e., determine what combination of components best disentangles the
229 neuro-cognitive relationship) in an identical process. This way we handle adverse behaviours
230 of classical linear models in high-dimensional data. A reliable and robust open-source
231 implementation of the SCCA method was retrieved as R package from CRAN (PMA,
232 penalized multivariate analysis, version 1.0.9, Witten, Tibshirani, & Hastie, 2009). The
233 amount of L_1 penalty for the functional connectomes and the NYC-Q reports were chosen by
234 cross-validation. The procedure is described below.

235 **3.5.2 Model selection**

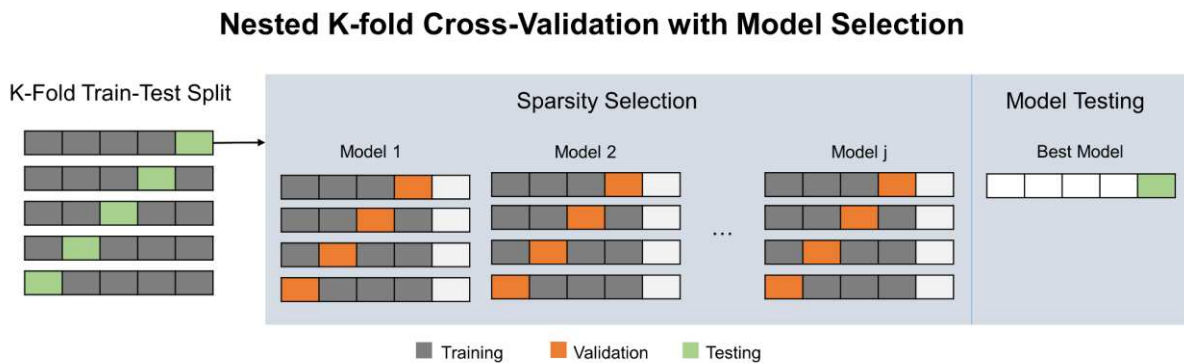
236 We employed cross-validation (CV) to select the most useful model across population
237 samples and avoid overfitting (Bzdok & Yeo, 2017). The amount of the two L_1 penalty terms
238 for the functional connectomes and the NYC-Q reports, respectively, were chosen by a nested
239 K-fold CV, where the coefficient for the penalty were chosen using a grid search to maximise
240 the quality of CV objective metric. The objective metric of choice cumulative explained

241 variances. The explained variance of each latent component was calculated using the squared
242 canonical correlation. High explained variance suggests a high pattern recovery rate between
243 the two data sets. The sparse assumption is fundamentally in conflict with the statistical goal
244 of finding components with high explained variance. Therefore we decided the number of
245 components in the model before searching for the best parameter.

246 We performed confound removal on functional connectomes and the NYC-Q reports as
247 suggested by prior studies (Smith et al., 2015). We removed the effects of nuisance variables
248 from the dataset. These confound variables were sex, age, and head motion indicated by
249 Jenkinson's mean FD (Jenkinson, Bannister, Brady, & Smith, 2002). The removal steps were
250 performed on the training set in each CV fold. We standardized the confound by calculating
251 the z-score, and also squared the three confound measures to account for potentially nonlinear
252 effects of these confounds. The 6 resulting confounds were regressed out of both data
253 matrices. The implementation of the confound removal method (Friston et al., 1994) was
254 retrieved from python library Nilearn (Abraham et al., 2014; <http://nilearn.github.io/>, version
255 0.3.1).

256 The number of latent components was determined by a preliminary analysis with no
257 sparsity and calculated the explained variances for the two datasets (i.e., brain network
258 correlations and questionnaire ratings). The explained variance increased with the number of
259 components and growth stabilised at 10 components. We selected the number of components
260 based on the point where the tangent stabilised. This led to a model of 4 components, and it
261 accounted for a total of 78% of the variance in connection strength and 29% of the variance
262 in the self-report data. Next, we determined the two coefficients for the L_1 penalty terms that
263 was associated with the best model performance with 4 latent components. We searched for
264 the best L_1 penalty values between 0.1 and 0.9 in 0.1 increments, which resulted in 81 sets of
265 parameters. For the nested K-Fold CV, we first separate the data into 5 consecutive folds after
266 shuffling the data and retained one fold as the evaluation set ($N \sim 50$); the other four folds
267 were used as the development set. The development set was further separated into 5 folds for
268 parameter selection and each fold ($N \sim 40$) was used as the validation set once. The model
269 was estimated on the training folds with all parameter sets, and after completion, we trained
270 the model with the winning parameter on the whole development set and finally tested the
271 performance on the independent, unseen evaluation set. We selected the final parameters
272 according to the best performance on the evaluation set across all folds of the outer CV loop
273 (**Figure 1**). This parameter set is used to train on the full development set and tested on the

274 evaluation set. The parameter grid search and k-fold CV was conducted by the
 275 implementation in a Python library scikit-learn (Pedregosa et al., 2011; [http://scikit-](http://scikit-learn.org/stable/)
 276 [learn.org/stable/](http://scikit-learn.org/stable/), version 0.18.2). The detailed algorithm for selecting the penalty values are
 277 presented in **Appendix: Nested K-Fold CV**.



278

279 **Figure 1. A diagram of the nested k-fold cross-validation with model selection.**

280 The model with the best test performance was selected as the final model. The final
 281 model's sparsity coefficient are 0.8 (functional connectivity) and 0.5 (self-reports), and the
 282 out-of-sample explained variance was 48%. We used the ensuing canonical vectors of the
 283 winning SCCA model to compute the latent component scores. There are two sets of
 284 canonical scores in a latent component, a weighted sum of variables forms the canonical
 285 vectors. For each latent component, we averaged the z-score of the canonical scores of the
 286 connection strength and NYC-Q as the combined scores. These scores described the
 287 summary of the experience with both the neural basis and the content reports.

288 **3.6 Test of component robustness**

289 After identifying the well performed components in compressing the brain-experience data,
 290 we examined the robustness of the four components in two different ways. The permutation
 291 test is a purely data-driven strategy that access the chance of discovering components in null
 292 samples. We also leveraged the brain-experience components to explain the cognitive
 293 functions, so that we can identify meaningful patterns by well-established cognitive
 294 measurements.

295 **3.6.1 Permutation test**

296 We used permutation testing to assess the robustness of the components identified through
 297 our analysis. We constructed a null distribution for each canonical component by holding the

298 functional connectivity data in place and randomising the row order of self-reports data. This
299 permutation scheme broke the link of individual differences in the dataset, therefore testing
300 the robustness of the components in the hypothetical population. By calculating the false-
301 discovery rate in the null distribution, we can conclude the possibility of discovering our
302 components by chance with the given penalty coefficients. Hypotheses that are accepted with
303 a 5% level of significance. In the current analyses we adopt the permutation test with the
304 FWE-corrected p-value by Smith and colleagues (2015) with data argumentation to increase
305 the size of the resampling datasets to 1000. The four components were compared to the first
306 sparse canonical correlation of the permuted sample. The low-rank components are more
307 relevant than the rest, therefore we yield more conservative p-value by comparing to the first
308 canonical correlation only. We performed 5000 permutation tests to get enough estimates for
309 4 decimal places.

310 **3.6.2 Group analysis**

311 To determine how patterns of unconstrained neuro-cognitive activity related to performance
312 on the battery of cognitive tests, we conducted an independent statistical analysis on the
313 identical subjects. A Type III multivariate multiple regression with Pillai's trace test was
314 applied to 4 individual scores for each of the latent components describing experience from
315 the SCCA were the independent variables, and the original 8 measures of cognitive
316 performance were the dependent variables that we hoped to be described by the linear
317 combination of the latent components. Pillai's trace test is considered to be the most powerful
318 and robust statistic for general use (Huberty & Olejnik, 2006). The p-values reported were
319 based on Bonferroni correction. We also performed a principal components analysis (PCA) to
320 identify the patterns of covariance among the 8 measures of cognitive performance and
321 compressed the data. The relation between the principle score and the 4 brain-experience
322 dimensions identified through SCCA was examined in a linear regression model with Pillai's
323 trace test. The analysis was conducted in R (version 3.3.1). The multivariate multiple
324 regression was conducted in R (version 3.3.1) using function 'Manova' in R package 'car'
325 (companion to applied regression, version 2.1-5).

326 **3.7 Code availability**

327 The full analysis pipeline is freely available at [https://github.com/htwangtw/patterns-of-](https://github.com/htwangtw/patterns-of-thought)
328 [thought](https://github.com/htwangtw/patterns-of-thought).

329

330 4 Results

331 4.1 Determining constituent categories of experience

332 We used Sparse Canonical Correlation Analysis (SCCA) to determine connectome-
333 wide dimensions that describe common variance shared by descriptions of brain and
334 experience. This took as input individual scores for the connections between each of the
335 regions extracted from Yeo's 7 networks parcellation and the scores of each item of the New
336 York Cognition Questionnaire (NYC-Q).

337 We applied SCCA with nested 5-fold CV as the model selection strategy. We
338 obtained a model of 4 canonical components with penalty levels of 0.8 on the functional
339 connectivity and 0.5 on the NYC-Q that indicated the best out-of-sample prediction on our
340 data (see **3.5.2 Model Selection**). The canonical correlations of the 4 latent components were
341 0.28, 0.19, 0.16, and 0.07. The latent components yielded by the best model are presented in
342 **Figure 2**. For the ease of presentation and interpretation, we summarized the components as
343 network-network connectivity instead of 57-by-57 connectivity matrices. The heat maps
344 describe the network-to-network correlations while the word clouds describe the loadings on
345 the self-report items. The components in full and the heat map for the self-report items can be
346 found in **Supplementary Materials**.

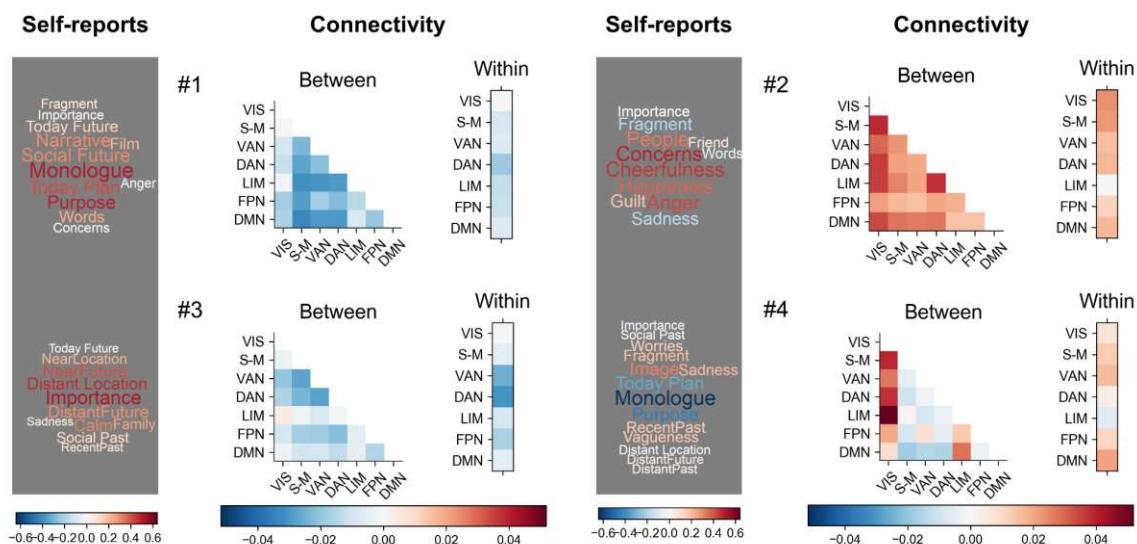
347 Component 1, describes patterns of reduced within network connectivity within all of
348 the networks studied, with this pattern most prominent in the dorsal attention network.
349 Between network connections are generally reduced, with the exception of visual to limbic.
350 Sensorimotor was decoupled from all the other systems, and, in addition, the default and
351 limbic were most decoupled from the attention networks. Experiential themes in Component
352 1 are dominated by themes related to deliberate planning with a verbal component (high
353 loadings on "words", "monologue", "today-plan", "social-future", "purpose" and
354 "deliberate"). We refer to this pattern of reports as reflecting thoughts with "purpose".

355 Component 2 is dominated by relatively higher within and between network
356 connections. Connectivity within each network was strong with the exception of the limbic
357 network. Between network connections were stronger, with this pattern most apparent in the
358 connections between limbic and ventral attention. In addition, the visual network was
359 strongly correlated with the other networks. This component is dominated by emotional
360 responses (high loadings on "anger", "guilt", "cheerfulness" and "happiness") and social

361 content (“friends” and “people”). We refer to this pattern of reports as reflecting “emotional”
 362 experience.

363 Component 3 emphasises reduced connections both between and within networks.
 364 Within network connectivity is weakest for the dorsal and ventral attention networks. Edge-
 365 to-edge connections are low, with the ventral and dorsal attention and fronto-parietal
 366 networks showing reduced correlations with each other as well as the visual and sensorimotor
 367 systems. This component was characterised by themes linked to personal “importance” with
 368 social temporal contents (“distant future”, “near future”, “social past”, “family” and “recent
 369 past”). We refer to this pattern of reports as reflecting “personal importance”.

370 Component 4 has the most heterogeneous pattern of within and between network
 371 connectivity. It is associated with stronger connections within networks with the exception of
 372 the limbic system. In addition, the visual system was strongly connected to all other
 373 networks, with this pattern most apparent for the limbic network. In contrast, lower network-
 374 to-network connectivity was observed between the default mode and sensori-motor and
 375 attention networks. This component is characterised by experiential patterns reflecting a
 376 modality difference in experience, with the highest loadings on “images” and lowest on
 377 “inner monologue”. We refer to this pattern of reports as describing “modality”.



378

379 **Figure 2. Unique neuro-cognitive dimensions of population variation revealed by sparse**
 380 **canonical correlation analysis of measures of whole brain connectivity and self-reported**

381 **descriptions of on-going experience.** The heat map describes the canonical variate of the
382 network-to-network connectivity between different Yeo networks. The connectivity matrices
383 describes the coefficients from the model, separated into within and between network
384 relationships. The word clouds reflect the coefficients on the relevant self-report items. In
385 both cases the colour bars indicate the magnitudes of the coefficients. A detailed version of
386 the canonical variates and alternative presentation of the self-report coefficients can be found
387 in **Supplementary Material Figure S1- S5.**

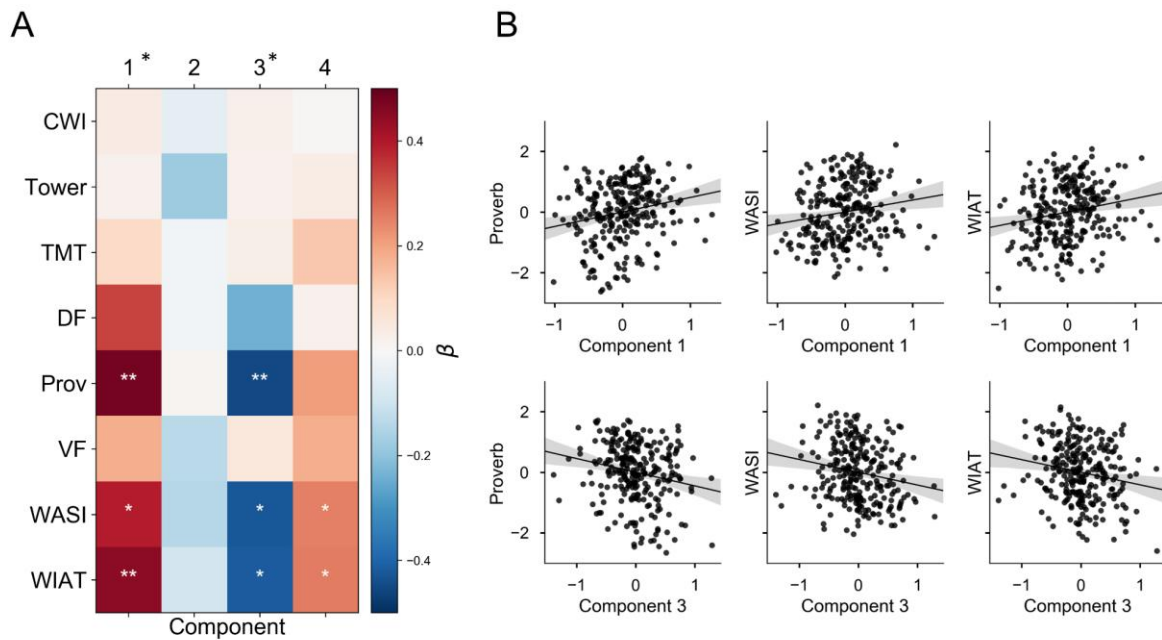
388 **4.2 The relationship between neuro-cognitive components and cognitive** 389 **function assessed in the laboratory**

390 Having documented four neuro-cognitive dimensions, we next examined the robustness of
391 the components using two complementary approaches. We first used a permutation test to
392 identify the chance of discovering components in a null samples as employed by Smith and
393 colleagues (2015). The top three components passed the permutation test and the 4th
394 component showed variance that was similar to that produced in a null sample (Component 1
395 $p = 0.0002$; Component 2 $p = 0.0010$; Component 3 $p = 0.0204$, Component 4 $p = 0.998$, $\alpha =$
396 0.05). This analysis suggests that Components 1 – 3 are unlikely to have occurred by chance.
397 Component 4 may be a Type II error and so we discuss this component in only a limited
398 manner moving forward.

399 Our next test of the robustness of our components is whether they explained unique
400 patterns of expertise in our battery of cognitive tasks. We used multiple multivariate
401 regression model in which performance on the battery of selected tasks was the dependent
402 variables and the individual scores for each of the canonical components describing
403 experience from the SCCA were the independent variables. In this analysis two of the four
404 canonical components described significant variance in our battery of tasks at multivariate
405 level: Component 1 ($F(8, 246) = 2.21, p = .027, \eta^2_p = .067$) and Component 3 ($F(8, 246) =$
406 $2.56, p = .024, \eta^2_p = .068$).

407 In the univariate results of the significant component, Component 1 was linked to
408 good performance in proverb test ($\beta = 0.48, t(251) = 3.27, p = .006, 95\% \text{ CI } [0.191 \ 0.766]$)
409 and both fluid intelligent tests WASI ($\beta = 0.39, t(251) = 2.74, p = .033, 95\% \text{ CI } [0.111$
410 $0.677]$) and WIAT ($\beta = 0.45, t(251) = 3.15, p = .009, 95\% \text{ CI } [0.167 \ 0.724]$). Component 3
411 showed a reversed pattern of the cognitive functions related to Component 1: proverb test (β
412 $= -0.45, t(251) = -0.14, p = .007, 95\% \text{ CI } [-0.176 \ -0.727]$); WASI ($\beta = -0.42, t(251) = -3.10,$

413 $p = .012$, 95% CI [-0.151 -0.693]) and WIAT ($\beta = -0.41$, $t(251) = -3.06$, $p = .012$, 95% CI [-
 414 0.148 -0.682]). The relationships between the neuro-cognitive dimensions and the pattern of
 415 relationships on the full cognitive battery and the adjusted variable scatter plots of the
 416 significant results are summarized in the form of a heat map in **Figure 3**.

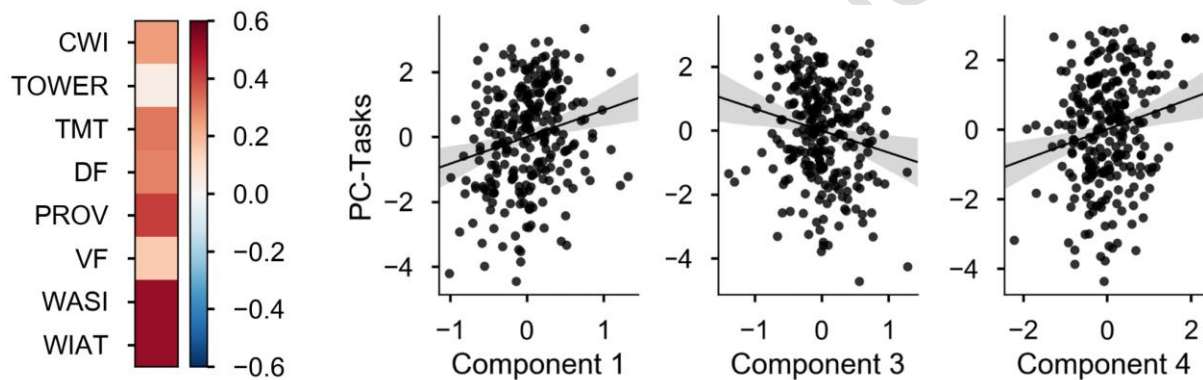


417

418 **Figure 3. The relationship between the different neural-cognitive components and the**
 419 **measures assessed in the cognitive battery.** The components 1 and 3 were significant at the
 420 multivariate level determined by multiple multivariate regression, indicated by the asterisk
 421 outside of the heat map. The cells with asterisk(s) indicates the significant results from the
 422 univariate test (bonferroni corrected) and the parameter estimates for each variable. CWI –
 423 Colour-word interference, DF – Design fluency, Pro- Proverbs, TOW – Tower of London,
 424 TMT – Trail making task, VF- Verbal Fluency, WASI – Wechsler Adult Intelligence Test,
 425 WIAT – Weschler Individual Attainment Test. P-value significant codes: 0 ****” 0.001 ****”
 426 0.01 ***”.

427 Finally, we performed a simple principle component analysis on the eight task
 428 measures to explore the associations between experience and the structure of the laboratory
 429 data. The aim of this analysis was to see if the pattern retrieved from the univariate level in
 430 the previous multiple multivariate regression was related to the internal structure of the data.
 431 Component selection was determined based on the scree plot, and we accepted one
 432 component explaining 39% of the variance. The principle component loaded on the
 433 intelligence measures and the proverb test. We fitted a linear model to this data to understand

434 the relationship to the four canonical components. The results are reported in **Figure 4**. The
 435 overall linear model was significant ($F(4, 253) = 5.43, p = .0003$). In the linear regression
 436 model, Component 1 ($\beta = 0.82, t(253) = 3.5, p = .001, 95\% \text{ CI } [0.36 \text{ } 1.29]$) showed
 437 significant contribution to explaining the task principle component. Component 3 showed a
 438 negative correlation to the task components ($\beta = -0.69, t(253) = -3.04, p = .003, 95\% \text{ CI } [-$
 439 $1.13 \text{ } -0.24]$). The relationships between tasks and the neuro-cognitive components here were
 440 similar to the ones uncovered by the multiple multivariate regression. In this analysis
 441 Component 4 ($\beta = 0.442, t(253) = 3.09, p = .002, 95\% \text{ CI } [0.16 \text{ } 0.72]$) showed a significant
 442 contribution in the regression model, but it did not pass the permutation test of robustness (p
 443 $= 0.998$). The related results should be treated cautiously. Together with our prior analysis,
 444 these results suggest that Components 1 and 3 are the most robust components identified in
 445 our study.



446
 447 **Figure 4. The principle component and its relationship to the different neural-cognitive**
 448 **components.** The heat map describes the principle component of the task battery, and the
 449 scatter plots describe the association with the components identified in our study. Component
 450 1 and 3 passed the permutation test for component robustness significantly contributed in
 451 explaining the principle component of the task. Component 4 showed a significant
 452 contribution in the regression model, but it did not pass the permutation test. The related
 453 results should be treated cautiously.

454

455 **5 Discussion**

456 We set out to describe different modes of neuro-cognitive patterns derived through the
457 simultaneous decomposition of whole brain connectivity data with self-reports of on-going
458 experience. We used a whole brain parcellation that describes cortical function in seven
459 independent networks (Yeo et al., 2011). We combined this data with self-reports of the
460 experience of our participants at rest, using a multivariate approach that allows for the
461 possibility of many-to-many mappings between neural patterns and on-going cognition. Our
462 analyses identified four stable canonical components, describing unique dimensions of neural-
463 experiential variation. Permutation testing demonstrated the statistical robustness of
464 Components 1-3. Furthermore, two components (1 and 3) described independent patterns of
465 performance in a battery of commonly used cognitive measures. This association with
466 cognitive performance that establishes a source of independent validity for these neuro-
467 cognitive components since they are related to independent measures of cognitive performance.
468 We next consider the fit between the dimensions produced by our analysis and theoretical
469 views of unconstrained neuro-cognitive processing.

470 We found evidence broadly consistent with contemporary representational accounts of
471 unconstrained processing. The neural patterns described by Component One reflect a pattern
472 of reduced correlation between regions with links to memory and representation (e.g. limbic,
473 default mode) from those with links to external behaviour (e.g. visual and sensorimotor cortex
474 and attention networks). This pattern was associated with experiences characterised by a sense
475 of purposefulness, and with verbally mediated content that was social and temporal in nature.
476 Participants high on this dimension were proficient at generating abstract semantic links and
477 performed well on measures of reasoning and intelligence. Together the features of Component
478 One support the hypothesis that the functional decoupling of systems important for memory
479 and representation are important for aspects of unconstrained cognition (Smallwood, 2013).
480 This capacity may arise from the topographical organisation of the cortex, in which neural
481 systems that can take on more transmodal properties tend to be located in regions that are more
482 distant in functional and structural terms (Buckner & Krienen, 2013; Margulies et al., 2016;
483 Mesulam, 1998). This spatial location may allow neural signals in these regions to take on
484 properties that are discrepant from the neural signal more closely tethered to inputs describing
485 the external world (Buckner & Krienen, 2013; Friston, 2013). The pattern identified by
486 Component One, therefore, may reflect a pattern of population variation describing the
487 hypothesised role of functional decoupling of memory and representational systems plays in

488 the generation of more abstract aspects of human cognition (Margulies et al., 2016; Mesulam,
489 1998). Importantly, in our prior work, limbic and default mode networks were the most distant
490 in functional connectivity terms from unimodal systems (Margulies et al., 2016).

491 Our data also highlights neural patterns that capture the hypothesised influence of
492 attention and control on on-going thought (McVay & Kane, 2010). Component 3 highlights
493 links between reduced connectivity within attention and control systems and patterns of
494 thought that emphasise personal importance. This is associated with worse performance on
495 measures of intelligence and reasoning. The combination of a focus on personally important
496 themes linked to poor performance on measures of general aptitude, captures the hallmark
497 psychological features of the current concerns X executive-failure accounts of on-going
498 thought (McVay & Kane, 2010). This view suggests that failures in attentional control lead to
499 highly personally relevant cognition to intrude into ongoing thought, leading to lapses in task
500 performance. Importantly, the neural pattern described by this component emphasises
501 dysregulated connectivity both within and between networks implicated in attention and
502 control by task-based studies (Duncan, 2010). Our prior work established that spontaneous
503 mind-wandering is linked to cortical thinning within regions linked to attention and control,
504 such as the intra-parietal sulcus (Golchert et al., 2017). Spontaneous mind-wandering has been
505 linked to worse cognitive control (Robison & Unsworth, 2018), as well as showing stronger
506 links with attention related problems, including ADHD (Seli, Smallwood, Cheyne, & Smilek,
507 2015). Together with these prior studies, our data suggests that population variation in the
508 intrinsic neural functioning within networks with an established role in external task
509 performance captures the hypothesised contribution of executive-failure to patterns of on-going
510 thought.

511 The method of decomposition used in the current study also highlighted patterns related to
512 affective processing and the modality of the experience that are similar to those seen in our
513 prior work that applied principal components analysis (PCA) to self-reported data only.
514 Component Four places experiences with visual features (“images”) in opposition to
515 experiences with verbal features (“monologue”), capturing dissociations between visual and
516 verbal thinking observed in our prior studies (Konishi et al., 2017; Medea et al., 2016;
517 Smallwood et al., 2016). The accompanying neural pattern were associated with higher
518 connectivity between the visual network with other networks, in particular the limbic system.
519 It is important to note that our permutation analysis failed to validate this component, so despite
520 its association with task performance using the PCA analysis it should be treated with relative

521 caution. Component Two loads on emotional experiences (“cheerfulness”, “anger”, “guilt” and
522 “happiness”) with the exception of those that are unhappy (“sad”). In neural terms this
523 component was characterised by high levels of connectivity, however, unlike Component Four,
524 this was highest between limbic and ventral attention networks. This pattern of coupling is
525 consistent with accounts that emphasise interactions between saliency and limbic systems in
526 affective processing (Touroutoglou et al., 2012). In the case of Component Two permutation
527 testing indicated this component was likely to be robust in statistical terms, however, we did
528 not observe associations with task performance. As with Component Four, interpretations of
529 Component Two should be made with caution in lieu of more empirical work.

530 Before closing it is worth considering several important limiting factors of our study. We
531 focused on patterns of population variance in unconstrained neuro-cognitive processing that
532 were measured once in each individual. Our study, therefore, cannot separate the influences
533 state and traits on our observed components. Treating patterns of unconstrained processing as
534 a trait is common in both the psychological (McVay & Kane, 2009; Smallwood, Ruby, &
535 Singer, 2013) and neural domains (Smith et al., 2015). Nonetheless, it remains an open
536 question how consistent these components will be across individuals over time, as well as
537 which aspects may be better described as traits. Importantly, by its very nature there are
538 dimensions of experience that our study cannot adequately address. We cannot, for example,
539 identify brain-experience associations that are highly dynamic in nature and in particular
540 those that change rapidly within an individual. Insight into this issue could be achieved by a
541 focus on dynamic rather than static connectivity (Kucyi, 2017). For example, the application
542 of techniques such as sliding window analysis (Chang & Glover, 2010) or Hidden Markov
543 models (Vidaurre, Smith, & Woolrich, 2017) to fMRI could provide information that would
544 complement our analyses. However, it may also be more important to examine these across
545 multiple sessions within the same individuals, as this would also make it most possible to
546 dissociate state from trait related influences on neural activity (Mueller et al., 2013). There
547 are also types of experience that may be difficult to assess using the measure of retrospective
548 experience sampling we have employed (Smallwood & Schooler, 2015). For important
549 features of experience, such as whether it has evolving features (Mills, Raffaelli, Irving, Stan,
550 & Christoff, 2017), or when the participant is unaware of the content of their experience
551 (Schooler, 2002), these experiential features may be best assessed using experience sampling
552 techniques that capture momentary elements of experience (Smallwood, 2013).

553

554 There are a number of methodological improvements that could enhance future studies of
555 brain-experience association. A recent benchmark study by Ciric and colleagues (Ciric et al.,
556 2017) shows that scrubbing can improve the performance of resting state analyses. Regarding
557 to the analysis pipeline, we gained hyper-parameters and best model with nested-CV an
558 approach that can help prevent overfitting (Bzdok & Yeo, 2017). There are also alternative
559 ways that could provide better tests of the robustness of the components we identified. We
560 assessed the validity of the components in three different ways; 1) with a data-driven, non-
561 parametric permutation test (Smith et al., 2015) that establishes the statistical validity of the
562 identified components and 2) by establishing the relationship between the laboratory
563 cognitive measures and 3) by consideration of their links with contemporary theoretical
564 accounts of ongoing cognition. In our study, Components 1 and 3 were statistically
565 significant in both cases and fitted well with contemporary accounts of ongoing cognition.
566 Accordingly we place encourage readers to focus on these patterns from our data. There are
567 alternative strategies that could help validate the robustness of patterns of brain-experience
568 association. One approach could be to compare the relationship between multiple sessions
569 within the same individual (Poldrack et al., 2015) and to have a larger sample that would
570 allow the reproducibility of these results through a formal split-half validation procedure. To
571 achieve this latter aim for future studies, we have placed the questionnaire measure used in
572 this study along with an example self-report collection task on GitHub at the following
573 address: https://github.com/htwangtw/restingstate_thoughtreports. We encourage interested
574 investigators to apply these measures in their resting-state investigation and to also upload the
575 resultant data onto open fMRI. These studies could be used in conjunction with the openly
576 access data used in this study to enable future investigations the opportunity to cross validate
577 experiential analyses in a more sophisticated manner than we have been able to achieve in
578 this study. The analysis pipeline of the current study can be further unified into one frame
579 work that benefits from both validation strategies. We can include the number of components
580 along with penalty coefficients in the hyper-parameters determined in the CV process, or
581 determine the best penalty terms with the first component. The permutation test will then
582 identify the reliable components occurring above chance level. After all the data-driven
583 component selection, we can examine the survived components through their relations with
584 well-documented cognitive measures and conclude the meaningful patterns. Finally, it is
585 likely that our measure of on-going thought lacks important questions regarding the content
586 of experience. It will be important, therefore, in the future to examine the relationships of the
587 type described in this study with a more exhaustive description of on-going experience. We

588 hope that by publishing our questionnaire collection task in a GitHub repository we will be
589 able to harness the power of the broader community to help generate and test plausible
590 questions for use in future studies.

591

592 **6 Acknowledgements**

593 The authors thank the Charlotte Murphy, Mladen Sormaz, Giulia Poerio, Irene De Caso
594 and the rest of the semantics and mind-wandering lab members for informative discussions;
595 Theodoros Karapanagiotidis and Deniz Vatansever for their help on fMRI data management
596 and preprocessing.

597 **7 Funding**

598 EJ was supported by ERC (SEMBIND - 283530). JS was supported by ERC
599 (WANDERINGMINDS - 646927), Volkswagen Foundation (Wandering Minds - 89440 and
600 89439) and the John Templeton Foundation, “Prospective Psychology Stage 2: A Research
601 Competition”. DB was supported by the German Research Foundation (DFG, BZ2/2-1,
602 BZ2/3-1, and BZ2/4-1), International Research Training Group (IRTG2150), Amazon AWS
603 Research Grant (2016 and 2017), the German National Academic Foundation, and the
604 START-Program of the Faculty of Medicine and an Exploratory Research Space Seed Fund,
605 RWTH Aachen University.

606 **8 Competing Interests**

607 The authors have declared that no competing interests exist.

608 **9 Appendix: Nested K-Fold CV**

- 609 1. Separate the model into 5 folds. In every iteration, 1 fold is the test set and the rest are
610 the development set.
- 611 2. For each outer fold:
- 612 a. For each parameter set to be considered:
- 613 i. Separate the development set into 5 folds. . In every iteration, 1 fold is
614 the validation set and the rest are the training set.
- 615 ii. For each inner fold:
- 616 1. Train the model on the training set
- 617 2. Calculate test error in the validation set
- 618 iii. Compute the average inner CV test error.
- 619 b. Choose the best parameter set with minimum average test error.
- 620 c. Use this parameter set to train on the development set.
- 621 d. Calculate test error in the test set
- 622 3. Determine the optimal model based on the outer fold test error
- 623 4. Train the full dataset on the optimal model
- 624

625 10 References

- 626 Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., . . . Varoquaux, G.
627 (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8.
628 Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012).
629 Inspired by distraction mind wandering facilitates creative incubation. *Psychological Science*,
630 0956797612446024.
- 631 Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A
632 critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*,
633 19(12), 2767-2796. doi:10.1093/cercor/bhp055
- 634 Buckner, R. L., & Krienen, F. M. (2013). The evolution of distributed association networks in the
635 human brain. *Trends in Cognitive Sciences*, 17(12), 648-665.
- 636 Bzdok, D., Hartwigsen, G., Reid, A., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2016). Left inferior
637 parietal lobe engagement in social cognition and language. *Neuroscience & Biobehavioral*
638 *Reviews*, 68, 319-334.
- 639 Bzdok, D., Laird, A. R., Zilles, K., Fox, P. T., & Eickhoff, S. B. (2013). An investigation of the
640 structural, connectional, and functional subspecialization in the human amygdala. *Human Brain*
641 *Mapping*, 34(12), 3247-3266.
- 642 Bzdok, D., & Yeo, B. T. (2017). Inference in the age of big data: Future perspectives on neuroscience.
643 *Neuroimage*.
- 644 Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity
645 measured with fMRI. *Neuroimage*, 50(1), 81-98. doi:10.1016/j.neuroimage.2009.12.011
- 646 Ciric, R., Wolf, D. H., Power, J. D., Roalf, D. R., Baum, G. L., Ruparel, K., . . . Davatzikos, C. (2017).
647 Benchmarking of participant-level confound regression strategies for the control of motion
648 artifact in studies of functional connectivity. *Neuroimage*, 154, 174-187.
- 649 Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for
650 intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172-179.
- 651 Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., . . . Constable, R. T.
652 (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain
653 connectivity. *Nature Neuroscience*, 18(11), 1664-1671. doi:10.1038/nn.4135
- 654 Friston, K. (2013). Life as we know it. *J R Soc Interface*, 10(86), 20130475. doi:10.1098/rsif.2013.0475
- 655 Friston, K., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994).
656 Statistical parametric maps in functional imaging: a general linear approach. *Human Brain*
657 *Mapping*, 2(4), 189-210.
- 658 Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery
659 and visual perception: an fMRI study. *Brain Research: Cognitive Brain Research*, 20(2), 226-
660 241. doi:10.1016/j.cogbrainres.2004.02.012
- 661 Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., . . . Van Essen,
662 D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), 171-
663 178. doi:10.1038/nature18933
- 664 Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntenburg, J. M., Liem, F., . . . Margulies, D. S.
665 (2017). Individual variation in intentionality in the mind-wandering state is reflected in the
666 integration of the default-mode, fronto-parietal, and limbic networks. *Neuroimage*, 146, 226-
667 235. doi:10.1016/j.neuroimage.2016.11.025
- 668 Gorgolewski, K. J., Lurie, D., Urchs, S., Kipping, J. A., Craddock, R. C., Milham, M. P., . . . Smallwood,
669 J. (2014). A correspondence between individual differences in the brain's intrinsic functional
670 architecture and the content and form of self-generated thoughts. *PloS One*, 9(5), e97176.
- 671 Gusnard, D. A., Raichle, M. E., & Raichle, M. E. (2001). Searching for a baseline: functional imaging
672 and the resting human brain. *Nature Reviews: Neuroscience*, 2(10), 685-694.
673 doi:10.1038/35094500
- 674 Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and*
675 *generalizations*: CRC press.
- 676 Huberty, C. J., & Olejnik, S. (2006). *Applied MANOVA and discriminant analysis* (Vol. 498): John
677 Wiley & Sons.

678 Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and
679 accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825-841.

680 Kane, M. J., & McVay, J. C. (2012). What Mind Wandering Reveals About Executive-Control Abilities
681 and Failures. *Current Directions in Psychological Science*, *21*(5), 348-354. doi:Doi
682 10.1177/0963721412454875

683 Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*,
684 *330*(6006), 932-932.

685 Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention wanders: Pupillometric
686 signatures of fluctuations in external attention. *Cognition*, *168*, 16-26.
687 doi:10.1016/j.cognition.2017.06.006

688 Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews:*
689 *Neuroscience*, *2*(9), 635-642. doi:10.1038/35090055

690 Kucyi, A. (2017). Just a thought: how mind-wandering is represented in dynamic brain connectivity.
691 *Neuroimage* Published online July 3, 2017.

692 Leszczynski, M., Chaieb, L., Reber, T. P., Derner, M., Axmacher, N., & Fell, J. (2017). Mind wandering
693 simultaneously prolongs reactions and promotes creative incubation. *Scientific Reports*, *7*(1),
694 10197.

695 Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of
696 emotion: a meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121-143.
697 doi:10.1017/S0140525X11000446

698 Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., . . .
699 Smallwood, J. (2016). Situating the default-mode network along a principal gradient of
700 macroscale cortical organization. *Proceedings of the National Academy of Sciences of the*
701 *United States of America*, *113*(44), 12574-12579. doi:10.1073/pnas.1608282113

702 McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: working memory capacity, goal
703 neglect, and mind wandering in an executive-control task. *Journal of experimental psychology.*
704 *Learning, memory, and cognition*, *35*(1), 196-204. doi:10.1037/a0014104

705 McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive
706 failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological*
707 *Bulletin*, *136*(2), 188-197; discussion 198-207. doi:10.1037/a0018298

708 Medea, B., Karapanagiotidis, T., Konishi, M., Ottaviani, C., Margulies, D., Bernasconi, A., . . .
709 Smallwood, J. (2016). How do we decide what to do? Resting-state connectivity patterns and
710 components of self-generated thought linked to the development of more concrete personal
711 goals. *Experimental Brain Research*. doi:10.1007/s00221-016-4729-y

712 Mesulam, M.-M. (1998). From sensation to cognition. *Brain: a journal of neurology*, *121*(6), 1013-
713 1052.

714 Mills, C., Raffaelli, Q., Irving, Z. C., Stan, D., & Christoff, K. (2017). Is an off-task mind a freely-
715 moving mind? Examining the relationship between different dimensions of thought.
716 *Consciousness and Cognition*. doi:10.1016/j.concog.2017.10.003

717 Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: a review.
718 *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie*
719 *expérimentale*, *67*(1), 11.

720 Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The
721 Hippocampus and Neocortex in Transformation. *Annual Review of Psychology*, *Vol 67*, *67*,
722 105-+. doi:10.1146/annurev-psych-113011-143733

723 Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The
724 role of mind-wandering in measurements of general aptitude. *Journal of Experimental*
725 *Psychology: General*, *141*(4), 788.

726 Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., . . . Liu, H. (2013).
727 Individual variability in functional connectivity architecture of the human brain. *Neuron*, *77*(3),
728 586-595.

729 Nooner, K. B., Colcombe, S. J., Tobe, R. H., Mennes, M., Benedict, M. M., Moreno, A. L., . . . Li, Q.
730 (2012). The NKI-Rockland sample: a model for accelerating the pace of discovery science in
731 psychiatry. *Frontiers in Neuroscience*, *6*.

732 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011).
733 Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*(Oct),
734 2825-2830.

735 Poerio, G. L., Sormaz, M., Wang, H. T., Margulies, D., Jefferies, E., & Smallwood, J. (2017). The role
736 of the default mode network in component processes underlying the wandering mind. *Social*
737 *Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsx041

738 Poerio, G. L., Totterdell, P., & Miles, E. (2013). Mind-wandering and negative mood: Does one thing
739 really lead to another? *Consciousness and Cognition*, *22*(4), 1412-1421.

740 Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., . . . Boyd, R. L.
741 (2015). Long-term neural and physiological phenotyping of a single human. *Nature*
742 *communications*, *6*, 8885.

743 Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014).
744 Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*,
745 *84*, 320-341. doi:10.1016/j.neuroimage.2013.08.048

746 Pulvermuller, F. (2010). Brain embodiment of syntax and grammar: discrete combinatorial mechanisms
747 spelt out in neuronal circuits. *Brain and Language*, *112*(3), 167-179.
748 doi:10.1016/j.bandl.2009.08.002

749 Pulvermuller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for
750 language. *Nature Reviews: Neuroscience*, *11*(5), 351-360. doi:10.1038/nrn2811

751 Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L.
752 (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of*
753 *the United States of America*, *98*(2), 676-682. doi:10.1073/pnas.98.2.676

754 Raij, T. T., & Riekkki, T. J. J. (2017). Dorsomedial prefrontal cortex supports spontaneous thinking per
755 se. *Human Brain Mapping*, *38*(6), 3277-3288. doi:10.1002/hbm.23589

756 Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational
757 bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42-55.
758 doi:10.1038/nrn.2016.150

759 Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and
760 deliberate mind-wandering. *Journal of Experimental Psychology: Learning, Memory, and*
761 *Cognition*, *44*(1), 85-98. doi:10.1037/xlm0000444

762 Ruby, F. J., Smallwood, J., Engen, H., & Singer, T. (2013). How self-generated thought shapes mood—
763 the relation between mind-wandering and mood depends on the socio-temporal content of
764 thoughts. *PloS One*, *8*(10), e77554.

765 Sanders, J. G., Wang, H.-T., Schooler, J., & Smallwood, J. (2017). Can I Get me out of my Head?
766 Exploring Strategies for Controlling the Self-Referential Aspects of the Mind-Wandering State
767 during Reading. *Quarterly Journal of Experimental Psychology*, *70*(6), 1053-1062.
768 doi:10.1080/17470218.2016.1216573

769 Schooler, J. W. (2002). Re-representing consciousness: dissociations between experience and meta-
770 consciousness. *Trends in Cognitive Sciences*, *6*(8), 339-344.

771 Seli, P., Smallwood, J., Cheyne, J. A., & Smilek, D. (2015). On the relation of mind wandering and
772 ADHD symptomatology. *Psychonomic bulletin & review*, *22*(3), 629-636.

773 Smallwood, J. (2013). Distinguishing how from why the mind wanders: a process-occurrence framework
774 for self generated thought. *Psychological Bulletin*.

775 Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds:
776 negative moods lead the mind to wander. *Emotion*, *9*(2), 271.

777 Smallwood, J., Karapanagiotidis, T., Ruby, F., Medea, B., de Caso, I., Konishi, M., . . . Jefferies, E.
778 (2016). Representing Representation: Integration between the Temporal Lobe and the Posterior
779 Cingulate Influences the Content and Form of Spontaneous Thought. *PloS One*, *11*(4),
780 e0152272. doi:10.1371/journal.pone.0152272

781 Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past: Unhappy moods lead to a
782 retrospective bias to mind wandering. *Cognition & Emotion*, 1-10.
783 doi:10.1080/02699931.2010.545263

784 Smallwood, J., Ruby, F. J., & Singer, T. (2013). Letting go of the present: mind-wandering is associated
785 with reduced delay discounting. *Consciousness and Cognition*, *22*(1), 1-7.

- 786 Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the
787 stream of consciousness. *Annual Review of Psychology*, 66, 487-518.
- 788 Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., . . . Miller,
789 K. L. (2015). A positive-negative mode of population covariation links brain connectivity,
790 demographics and behavior. *Nature Neuroscience*.
- 791 Swanson, J. (2005). The Delis-Kaplan Executive Function System: A Review. *Canadian Journal of*
792 *School Psychology*, 20(1-2), 117-128.
- 793 Touroutoglou, A., Hollenbeck, M., Dickerson, B. C., & Barrett, L. F. (2012). Dissociable large-scale
794 networks anchored in the right anterior insula subserve affective experience and attention.
795 *Neuroimage*, 60(4), 1947-1958.
- 796 Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically
797 organized in time. *Proceedings of the National Academy of Sciences of the United States of*
798 *America*, 114(48), 12827-12832. doi:10.1073/pnas.1705120114
- 799 Wang, H.-T., Poerio, G., Murphy, C., Bzdok, D., Jefferies, E., & Smallwood, J. (2017). Dimensions of
800 experience: exploring the heterogeneity of the wandering mind. *Psychological Science*,
801 0956797617728727.
- 802 Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. New York, NY: The Psychological
803 Corporation.
- 804 Wechsler, D. (2005). Wechsler Individual Achievement Test 2nd Edition (WIAT II). New York, NY:
805 The Psychological Corporation.
- 806 Weissman, D. H., Roberts, K., Visscher, K., & Woldorff, M. (2006). The neural bases of momentary
807 lapses in attention. *Nature Neuroscience*, 9(7), 971.
- 808 Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications
809 to sparse principal components and canonical correlation analysis. *Biostatistics*, 10(3), 515-534.
- 810 Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Polimeni,
811 J. R. (2011). The organization of the human cerebral cortex estimated by intrinsic functional
812 connectivity. *Journal of Neurophysiology*, 106(3), 1125-1165.

813