General Functional Connectivity: shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks

Maxwell L. Elliott[1*]
Annchen R. Knodt[1]
Megan Cooke[1]
M. Justin Kim[2]
Tracy R. Melzer[3,4]
Ross Keenan[3,5]
David Ireland[6]
Sandhya Ramrakha[6]
Richie Poulton[6]
Avshalom Caspi[1,7-9]
Terrie E. Moffitt[1,7-9]
Ahmad R. Hariri[1]

1. *Department of Psychology & Neuroscience, Duke University Box 104410, Durham, NC 27708, USA*
2. *Department of Psychology, University of Hawaii at Manoa, Honolulu, HI 96822*
3. *New Zealand Brain Research Institute, Christchurch, New Zealand*
4. *Department of Medicine, University of Otago, Christchurch, New Zealand*
5. *Christchurch Radiology Group, Christchurch, New Zealand*
6. *Dunedin Multidisciplinary Health and Development Research Unit, Department of Psychology, University of Otago, 163 Union St E, Dunedin, 9016, NZ*
7. *Social, Genetic, & Developmental Psychiatry Research Centre, Institute of Psychiatry, Psychology, & Neuroscience, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK*
8. *Department of Psychiatry & Behavioral Sciences, Duke University School of Medicine, Durham, NC 27708, USA*
9. *Center for Genomic and Computational Biology, Duke University Box 90338, Durham, NC 27708, USA*

*Correspondence: maxwell.elliott@duke.edu

Abstract: Intrinsic connectivity, measured using resting-state fMRI, has emerged as a fundamental tool in the study of the human brain. However, due to practical limitations, many studies do not collect enough resting-state data to generate reliable measures of intrinsic connectivity necessary for studying individual differences. Here we present general functional connectivity (GFC) as a method for leveraging shared features across resting-state and task fMRI and demonstrate in the Human Connectome Project and the Dunedin Study that GFC offers better test-retest reliability than intrinsic connectivity estimated from the same amount of resting-state data alone. Furthermore, at equivalent scan lengths, GFC displays higher heritability on average than resting-state functional connectivity. We also show that predictions of cognitive ability from GFC generalize across datasets, performing as well or better than resting-state or task data alone. Collectively, our work suggests that GFC can improve the reliability of intrinsic connectivity estimates in existing datasets and, subsequently, the opportunity to identify meaningful correlates of individual differences in behavior. Given that task and resting-state data are often collected together, many researchers can immediately derive more reliable measures of intrinsic connectivity through the adoption of GFC rather than solely using resting-state data. Moreover, by better capturing heritable variation in intrinsic connectivity, GFC represents a novel endophenotype with broad applications in clinical neuroscience and biomarker discovery.

Introduction

Functional magnetic resonance imaging (fMRI) has proven invaluable for identifying the neural architecture of human behavior and cognition (Betti et al., 2013; Fox et al., 2007; Huth et al., 2016). Accordingly, fMRI has been widely adopted in myriad studies seeking to deepen our understanding of the human brain in both health and disease (Cole et al., 2011; Power et al., 2013; Satterthwaite et al., 2016). Recently, fMRI studies have expanded in both scale and scope, often collecting data in thousands of individuals in an effort to adequately power the search for neural correlates of complex human traits and predictive biomarkers for mental illness (Casey et al., 2018; Elliott et al., 2018; Miller et al., 2016; Swartz et al., 2015; Thompson et al., 2014). In this context, most studies have focused on the acquisition of resting-state functional MRI to map the intrinsic connectivity of neural networks. This choice is often prompted by three considerations. First, intrinsic connectivity networks appear to be more heritable than task-elicited activity (Elliott et al., 2017; Ge et al., 2017; Winkler et al., 2010). Second, resting-state data are relatively easy to collect from informative populations of children, the elderly, and mentally ill patients (Fox, 2010; Greicius, 2008; Shehzad et al., 2009). Third, intrinsic connectivity plays a primary role in shaping task-based brain activity and associated behaviors (Cole et al., 2016, 2014; Krienen et al., 2014; Tavor et al., 2016).

While analyses of resting-state data have revealed insights about average, group-level features of the human brain (Buckner et al., 2008; Power et al., 2011; Yeo et al., 2011), progress has lagged in identifying individualized signatures, which are critical if ongoing large-scale studies are to be successful in the search for risk and disease biomarkers. This lack of progress partially reflects the generally poor test-retest reliability of intrinsic connectivity estimates in many resting-state datasets (Birn et al., 2013; Guo et al., 2012). The reliability of intrinsic connectivity measures must

be improved, as unreliability limits the ability to predict individual differences that could become clinical biomarkers (Nunnally Jr., 1970; Vul et al., 2009). In this paper we propose general functional connectivity (GFC), a simple method for combining resting-state and task fMRI data to improve the reliability of intrinsic connectivity.

Historically researchers have collected 5-10 minutes of resting-state data in response to both published recommendations (Shehzad et al., 2009; Van Dijk et al., 2010) and practical limitations (e.g., scan time, cost and subject stamina). While such data may be adequate for detecting patterns of intrinsic connectivity common across individuals, more than 25 minutes of data are needed to reliably detect individual differences in these same patterns (Anderson et al., 2011; Hacker et al., 2013; Laumann et al., 2015). This suggests that intrinsic connectivity may simply be a noisy measure of stable individual differences and that lengthier measurement is needed to identify the "true" signal (Braga and Buckner, 2017; Gordon et al., 2017; Gratton et al., 2018; Mueller et al., 2013). However, due to the high cost of scan time and the limited ability of many individuals including children, the elderly, and the mentally ill to lie still without falling asleep for tens of minutes of scanning, this level of protracted measurement is simply not feasible for a majority of studies (Power et al., 2012; Satterthwaite et al., 2013; Tagliazucchi and Laufs, 2014). To illustrate the point, two recent meta-analyses of resting-state intrinsic connectivity in depression (Kaiser et al., 2015) and schizophrenia (Dong et al., 2018) revealed average resting-state scan times of 6.53 minutes (k = 23 studies) and 6.24 minutes (k = 36 studies), respectively. It is likely that intrinsic connectivity based on these brief scans has low reliability, reducing the ability to discover individual differences and related biomarkers. If measures of intrinsic connectivity are to realize their potential as a clinical tool (Fox, 2010; Matthews et al., 2006) and contribute to personalized medicine (Collins and Varmus, 2015; Hamburg and Collins, 2010), there must first be substantial

100 improvement in their reliability, and this should be ideally achieved within the practical limitations
101 of study design and participant burden.
102
103 Although few studies collect enough resting-state data to achieve high levels of reliability, many
104 collect multiple task scans in addition to a resting-state scan. While it has been implicitly assumed
105 that task and resting-state scans are two separate measures of brain function, to be analyzed and
106 studied independently, a growing body of evidence suggests that the intrinsic connectivity
107 measured by each may have substantial overlap, sharing over 80% of the same variance (Cole et
108 al., 2014; Geerligs et al., 2015; Tavor et al., 2016). In fact, it has long been known that intrinsic
109 networks extracted from task scans are quite similar to those extracted from resting-state scans,
110 and task scans have been used for intrinsic connectivity analyses when resting-state data are absent
111 (Arfanakis et al., 2000; Fair et al., 2007; Fox et al., 2006; Smith et al., 2009). Further, mental
112 operations that occur during task states have been linked to resting-state functional networks
113 (Bzdok et al., 2016; Leech et al., 2012), and task scans have been shown to accentuate individual
114 differences in network dynamics (Finn et al., 2017; Greene et al., 2018; Satterthwaite et al., 2018).
115 Despite this evidence for a complementary relationship between resting-state and task fMRI, they
116 are typically investigated separately. Even when researchers have collected both task and resting-
117 state fMRI in the same subjects, analyses of intrinsic connectivity are almost always performed on
118 a single short resting-state scan despite the low test-retest reliability of short scan data. If it could
119 be demonstrated that the reliability of intrinsic connectivity can be improved by adding task to
120 resting-state fMRI, many researchers could immediately benefit by adopting this approach in data
121 they have already collected, and also in designing future studies. The field of individual-
122 differences neuroscience would also benefit from the resulting increase in reliability, replicability,
123 and power that would follow the widespread adoption of combining task and resting-state fMRI.

Here we hypothesized that measures of intrinsic connectivity will be more reliable when task and resting-state scans are combined into a longer dataset (i.e., GFC) compared to short resting-state scans alone. Additionally, we hypothesized that this boost in reliability with GFC will increase the amount of heritable variation detectable in intrinsic connectivity, because the detection of heritability requires reliable measurement. Finally, we hypothesized that GFC would measure a common backbone of intrinsic connectivity common across task and resting-state data allowing for better out-of-sample prediction of individual differences. We examined two existing datasets in 3 different ways to test these hypotheses. First, we investigated the test-retest reliability of different combinations of task and resting-state data in both the Human Connectome Project (HCP (Van Essen et al., 2013)), which uses state-of-the-art fMRI hardware and acquisition parameters in a highly educated, healthy sample, and in the Dunedin Study (Poulton et al., 2015), which uses more common fMRI hardware and acquisition parameters in a population-representative birth cohort. Second, we estimated the heritability of GFC and resting-state functional connectivity (RSFC) by leveraging the twin sampling design of the HCP. Third, we investigated the ability of GFC to measure generalizable individual differences in cognitive ability across samples and across task states by testing out-of-sample prediction of cognitive ability in the HCP and Dunedin Study samples.

Materials and Methods

Datasets

Human Connectome Project. This is a publicly available data set that includes 1206 participants with extensive MRI and behavioral measurement (Van Essen et al., 2013). In addition, 45 of these subjects completed the entire scan protocol a second time (referred to hereafter as "test-retest

148 sample"). All participants were free of current psychiatric or neurologic illness and were 25-35
149 years of age. In all analyses, subjects were excluded if they had truncated scans, less than 40
150 minutes of combined task data, or less than 40 minutes of resting-state data after censoring. Much
151 of the HCP dataset consists of twins and family members. We exploited this feature of the design
152 for conducting heritability analyses. For these analyses, 943 participants from 420 families met
153 inclusion criteria (144 monozygotic families, 85 dizygotic families, and 191 full sibling families).
154 To avoid bias due to family confounding in the prediction analyses, only one individual per family
155 was retained (the family member with the least motion during scanning). This resulted in 298
156 subjects left to be used in our heritability analyses. Eight subjects were removed from our test-
157 retest analyses because they had at least one fMRI scan with truncated or missing data. Four
158 subjects were removed because they did not have sufficient data (i.e., > 40 minutes of rest and >
159 40 minutes of task) after censoring. Therefore, the test-retest reliability analyses included data
160 from 33 participants.
161
162 The acquisition parameters and minimal preprocessing of these data have been described
163 extensively elsewhere (Glasser et al., 2013). In our analyses, we used the minimally preprocessed
164 data in volumetric Montreal Neurological Institute (MNI) space ("fMRIVolume" pipeline).
165 Briefly, participants underwent extensive MRI measurement that included T1 and T2 weighted
166 structural imaging, diffusion tensor imaging, and nearly two hours of resting-state and task fMRI.
167 One hour of resting-state fMRI was collected on each participant in four 15-minute scans (1200
168 time-points each) split-up into two scanning sessions over two days. In each scan session the two
169 resting-state scans were followed by task fMRI (Smith et al., 2013). Across the two sessions, each
170 participant completed seven fMRI tasks described extensively elsewhere (Barch et al., 2013).
171 Briefly, tasks were designed to identify functionally relevant "nodes" in the brain and included

172 working memory (810 timepoints, 10:02 minutes), gambling (506 timepoints, 6:24 minutes),
173 motor function (568 timepoints, 7:06 minutes), language (632 timepoints, 7:54 minutes), social
174 cognition (548 timepoints, 6:54 minutes), relational processing (464 timepoints, 6:52 minutes) and
175 emotional processing (352 timepoints, 4:32 minutes). Altogether, 4800 timepoints totaling 60
176 minutes of resting-state fMRI and 3880 timepoints totaling 48:30 minutes of task fMRI were
177 collected on each participant.
178
179 *Dunedin Study.* This is a longitudinal investigation of health and behavior in a complete birth
180 cohort of 1,037 individuals (91% of eligible births; 52% male) born between April 1972 and March
181 1973 in Dunedin, New Zealand (NZ), and eligible based on residence in the province and
182 participation in the first assessment at age three. The cohort represents the full range of
183 socioeconomic status on NZ's South Island and matches the NZ National Health and Nutrition
184 Survey on key health indicators (e.g., BMI, smoking, GP visits) (Poulton et al., 2015). The cohort
185 is primarily white; fewer than 7% self-identify as having non-Caucasian ancestry, matching the
186 South Island (Poulton et al., 2015). Assessments were carried out at birth and ages 3, 5, 7, 9, 11,
187 13, 15, 18, 21, 26, 32, and 38 years, when 95% of the 1,007 study members still alive took part.
188 Neuroimaging data collection is ongoing in participants who are now aged 45 years. We currently
189 have collected completed neuroimaging data for 756 study members. Data were excluded if the
190 study member did not complete the rest scan and all four task scans or if they did not have sufficient
191 data after censoring and bandpass filtering was applied to the resting-state and each task scan
192 separately (see below for censoring details). Data for 591 subjects survived these fMRI exclusion
193 criteria and were included in analyses. Additionally, 20 study members completed the entire scan
194 protocol a second time (mean days between scans = 79). Of these one failed to have sufficient

degrees of freedom after censoring and bandpass filtering. Therefore, the test-retest reliability analyses included data from 19 participants.

Each participant was scanned using a Siemens Skyra 3T scanner equipped with a 64-channel head/neck coil at the Pacific Radiology imaging center in Dunedin, New Zealand. High resolution structural images were obtained using a T1-weighted MP-RAGE sequence with the following parameters: TR = 2400 ms; TE = 1.98 ms; 208 sagittal slices; flip angle, 9°; FOV, 224 mm; matrix =256×256; slice thickness = 0.9 mm with no gap (voxel size 0.9×0.875×0.875 mm); and total scan time = 6:52 minutes. Functional MRI was collected during resting-state and four tasks with a series of 72 interleaved axial T2-weighted functional slices acquired using a 3-fold multi-band accelerated echo planar imaging sequence with the following parameters: TR = 2000 ms, TE = 27 ms, flip angle = 90°, field-of-view = 200 mm, voxel size = 2mm isotropic, slice thickness = 2 mm without gap.

8:16 minutes (248 TRs) of resting-state fMRI was collected immediately before the four task fMRI scans. During the resting-state scan participants were instructed to stay awake with their eyes open while looking at a grey screen. Participants completed an emotion processing task (6:40 minutes, 200 TRs), a color Stroop task (6:58 minutes, 209 TRs), a monetary incentive delay (MID) task (7:44 minutes, 232 TRs) and an episodic memory task (5:44 minutes, 172 TRs). All four tasks are described in detail in the supplement.

fMRI pre-processing

Minimal preprocessing was first applied to all data. For the HCP dataset this was done with the HCP minimal preprocessing pipeline (Glasser et al., 2013) and included correction for B0

219 distortion, realignment to correct for motion, registration to the participant's structural scan,
220 normalization to the 4D mean, brain masking, and non-linear warping to MNI space.
221
222 Minimal preprocessing steps were applied to the Dunedin Study dataset using custom processing
223 scripts. Anatomical images were skull-stripped, intensity-normalized, and nonlinearly warped to a
224 study-specific average template in the standard MNI stereotactic space using the ANTs SyN
225 registration algorithm (Avants et al., 2008; Klein et al., 2009). Time-series images were despiked,
226 slice-time-corrected, realigned to the first volume in the time-series to correct for head motion
227 using AFNI tools (Cox, 1996), coregistered to the anatomical image using FSL's Boundary Based
228 Registration (Greve and Fischl, 2009), spatially normalized into MNI space using the non-linear
229 ANTs SyN warp from the anatomical image, and resampled to 2mm isotropic voxels. Dunedin
230 Study images were additionally corrected for B0 distortions.
231
232 Time-series images for each dataset were further processed to limit the influence of motion and
233 other artifacts. Voxel-wise signal intensities were scaled to yield a time-series mean of 100 for
234 each voxel. Motion regressors were created using each participant's 6 motion correction
235 parameters (3 rotation and 3 translation) and their first derivatives (Jo et al., 2013; Satterthwaite et
236 al., 2013) yielding 12 motion regressors. Five components from white matter and cerebrospinal
237 fluid were extracted using CompCorr (Behzadi et al., 2007) and used as nuisance regressors, along
238 with the mean global signal. In the Dunedin Study dataset, images were bandpass filtered to retain
239 frequencies between .008 and .1 Hz. In the HCP dataset, images only underwent highpass filtering
240 with a cutoff of .008 Hz. High frequency signals were retained because removing high frequency
241 signals would have resulted in excessive loss of degrees of freedom due to the very low TR (.75
242 seconds) (Bright et al., 2017; Caballero-Gaudes and Reynolds, 2017). In the HCP dataset, we

243 followed the empirically derived thresholds of .39 mm frame-wise displacement or 4.9 units above
244 the median DVARS as recommended (Burgess et al., 2016). In the Dunedin Study dataset, we
245 investigated a range of framewise-displacement cutoffs using QC-RSFC plots in order to derive
246 the optimal threshold for removing motion artifact as recommended (Power et al., 2014). This
247 investigation led to thresholds of 0.35 mm framewise-displacement and 1.55 standardized
248 DVARS. In both datasets, nuisance regression, bandpass filtering, censoring, global-signal
249 regression, and smoothing of 6mm FWHM for each time-series were performed in a single
250 processing step using AFNI's 3dTproject. Identical time-series processing was applied to resting-
251 state and task time-series data with one exception.
252
253 To remove functional connectivity predominantly driven by task-evoked coactivation, signal due
254 to task structure was added as an additional nuisance covariate to all task scans and removed from
255 their time-series (Fair et al., 2007) (see supplemental for details on task modeling and regression).
256 However, we examined the impact of task structure on our estimates of intrinsic connectivity by
257 conducting parallel analyses without the regression of task structure. While task regression had a
258 small but consistent effect of lowering ICCs for intrinsic connectivity estimates (see supplemental
259 Table S1), the ability to predict cognitive ability was slightly higher when task regression was
260 applied (see supplemental Table S2). Thus, we focus on analyses with task regression in our
261 primary results.
262
263 *Effects of Global Signal Regression*
264 Global signal regression (GSR) was performed to adopt a conservative approach to motion artifact
265 reduction and used in our main analyses (Power et al., 2018). However, because use of GSR is still

266 debated (Murphy and Fox, 2017), we also preprocessed our data with parallel preprocessing
267 methods without GSR (see supplemental figure S1).
268
269 *Functional Connectivity processing*
270 We investigated whole-brain intrinsic connectivity using two parcellation schemes. In the
271 reliability and prediction analyses (described below), a 264-region parcellation was utilized. This
272 parcellation was derived in a large independent dataset (Power et al., 2011). BOLD data were
273 averaged within 5mm spheres surrounding each of the 264 coordinates in the parcellation. In the
274 heritability analyses we used a smaller 44-brain-region parcellation scheme due to computational
275 constraints imposed by the time needed to fit heritability models at each edge independently
276 (described below). This parcellation was derived from the 7-network parcellation scheme
277 described in (Yeo et al., 2011). These 7 networks were broken up into 44 spatially contiguous
278 regions across the two hemispheres. Average time series were extracted from all 44 regions. For
279 both parcellations the average time-series was extracted independently from each scan session in
280 every dataset. This allowed the time-series to be flexibly concatenated and recombined.
281 Correlation matrices were derived from these time-series using Pearson correlation, resulting in
282 34,716 edges in the Power et al. parcellation and 946 edges in the Yeo et al. parcellation.
283
284 *Test-retest reliability*
285 Intraclass correlation (ICC) was used to quantify the test-retest reliability of intrinsic connectivity
286 measures in the HCP and Dunedin Study test-retest datasets. ICC (3,1) was used in all analyses
287 (Chen et al., 2018). In the HCP dataset, the influence of the amount of data, as well as type of data
288 (resting-state or task) on ICC was investigated. In the resting-state analyses, intrinsic connectivity
289 matrices were calculated across a range of scan lengths (5, 10, 20, 30, and 40 minutes of post-

censoring data) mirrored in the test and retest samples. ICCs were then calculated for each edge of this matrix, yielding 34716 (264*263/2) unique ICC values for each scan length. A more general definition of intrinsic connectivity was derived by including task data in the intrinsic connectivity matrices. GFC was formally investigated (Figures 1, 2 and 3) by constructing datasets from a combination of task and resting-state data over a variety of scan lengths (5, 10, 20, 30 and 40 minutes of post-censoring data). In scan lengths up to 25 minutes, an exactly equal amount of data from all scan types (1/8 of total scan length from each task and resting-state) was combined. After this point, equal amounts of each task could no longer be added together because the shortest task scan no longer had sufficient timepoints remaining in all participants. Above 25 minutes, timepoints were selected at random from the pool of remaining timepoints. In the Dunedin Study, two ICC matrices were constructed. The first was built from each participant's single resting-state scan and the second from all task and resting-state scans to create the GFC matrix. Paired-sample t-tests were used to compare mean differences in reliability between ICC matrices constructed from traditional resting-state functional connectivity and GFC.

HCP fMRI data were collected in two phase-encoding directions (LR and RL). To minimize bias due to phase-encoding direction on our analyses, in each participant, for each task, timepoints were first selected from LR encoded scans until LR timepoints were exhausted. Subsequently, timepoints were selected from remaining RL-encoded scans. This ensured that all subjects had a roughly equivalent proportion of LR and RL timepoints within a given analysis. This also ensured roughly equivalent proportions of LR and RL phase encoding across test and retest correlation matrices when calculating ICCs. Selecting from LR first before RL had the additional benefit of mimicking shorter data-acquisition scans because timepoints were utilized in the order they were collected.

Heritability

For each edge of the connectivity matrix, we ran an extended twin model (Posthuma and Boomsma, 2000), which is an extension of the classic twin design in that it allows for the modeling of additional family members of twins. In the HCP study we had data from the families of monozygotic and dizygotic twins as well as siblings in families without a twin pair. In this model the variance of the phenotype (in this case correlations between brain regions) is partitioned into four sources: additive genetic factors (A), shared environmental factors (C), twin-specific factors (T), and unique environmental factors and error (E) based on the correlations between monozygotic twins, dizygotic twins, and their siblings. Due to their small numbers, we eliminated half-siblings from the model (N = 27) and capped family size at four individuals (7 families affected, N = 13). This left us with 1,078 individuals from 443 families (148 monozygotic families, 92 dizygotic Z families, 203 full-sibling families). Of these, 943 participants from 420 families (144 monozygotic families, 85 dizygotic families and 191 full-sibling families) had full scans from all tasks and rest and had at least 40 minutes of task data and 40 minutes of rest data after censoring. Heritability analyses were performed using OpenMx (Boker et al., 2011; Neale et al., 2016) in the R statistical computing environment. Each model also contained parameters to adjust for the effects of sex and age on the mean. Confidence intervals were obtained on the A, C, T, and E estimates using the mxCI command in OpenMx which obtains profile likelihood confidence intervals (Pek and Wu, 2015). For each edge, an estimate was determined to be significant if the lower bound of the confidence interval was greater than zero.

Connectome-based predictive modeling

The predictive utility of increasing reliability through our GFC measure was evaluated by testing the ability of the different intrinsic connectivity matrices to predict cognitive ability in the HCP and Dunedin Study datasets. Cognitive ability was measured in the HCP using the Raven's Progressive Matrices (PMAT24_A_CR) (Raven, 1941). This measure was adopted because it is a proxy for cognitive ability that has been shown to be predicted by intrinsic connectivity in the HCP (Dubois et al., 2018; Finn et al., 2015; Noble et al., 2017). The WAIS-IV, a well-established and validated measure of individual differences in cognitive ability (Weschler, 2008), was used in the Dunedin Study.

Cognitive ability was predicted from intrinsic connectivity data using connectome-based predictive modeling (CPM; (Shen et al., 2017)). This framework provides a general method to predict any measure from intrinsic connectivity matrices. In this approach, predictors are filtered by selecting edges that have a $p < .01$ correlation with the measure of interest. Three linear regression predictive models are then built, one from the positive features (edges positively correlated with the phenotype of interest), one from the negative features (edges negatively correlated with the phenotype of interest), and one from the combination of positive and negative features (Shen et al., 2017). Here we report results from the combined model (see supplemental tables S3, S4 and S5 for results from all models). In our CPM analyses, we used both within-sample and out-of-sample prediction. Within-sample models were used to directly compare RSFC and GFC predictions of cognitive ability using a leave-one-out cross-validation scheme. Models were trained with all participants except one and used to predict the measure in the left-out participant. This was repeated until all participants had been left out.

Out-of-sample prediction is the gold standard to test the unbiased predictive utility of models (Whelan and Garavan, 2014; Yarkoni and Westfall, 2017) and was therefore used to test the generalizability of predictive models built with RSFC, GFC, and task data. To test the influence of task state, data aggregation and scan length on predictive utility, models were trained with RSFC, GFC, and individual tasks in the Dunedin Study dataset and then used to predict cognitive ability from RSFC, GFC, and individual tasks in the HCP dataset. The Dunedin Study and HCP have three parallel tasks that tap into similar behavioral circuits. The gambling task in the HCP and the MID task in Dunedin Study both involve reward processing. The emotional processing task in the HCP and the face matching task in the Dunedin Study both involve passive processing of emotional facial expressions. Lastly, the working memory task in the HCP and the Stroop task in the Dunedin Study both tap into executive control. We restricted our analyses to comparisons of RSFC, GFC, and these 3 parallel tasks. The abilities of RSFC and GFC was contrasted with that of these parallel tasks to predict cognitive ability within tasks. In both within-sample and out-of-sample tests the Spearman correlation between predicted and true scores was adopted as an unbiased effect size measure of predictive utility. Model predictions of cognitive ability were assessed for significance using a parametric test for significance of correlations. All p-values from correlations with cognitive ability were corrected for multiple comparisons using the false discovery rate (Benjamini and Hochberg, 1995). Differences in predictive utility between models were compared using Steiger's z (Steiger, 1980). All confidence intervals for CPM prediction estimates were generated with bootstrap resampling, using AFNI's 1dCorrelate tool.

Results

What is the test-retest reliability of GFC?

We used data from the HCP to evaluate the test-retest reliability of both intrinsic connectivity derived from resting-state scans alone and of GFC derived from combinations of task and resting-state scans. Test-retest reliability, as measured by ICC, ranges from 0-1, and is commonly classified according to the following cutoffs: 0 - 0.4 = poor, 0.4 - 0.6 = moderate, 0.6 - 0.75 = good and 0.75 - 1 = excellent (Cicchetti, 1994). In the HCP, when 5 minutes (post-censoring) of resting-state data were used to measure intrinsic connectivity defined within a common atlas of 264 regions (Power et al., 2011), the reliability was generally poor (mean ICC = .28, 95% CI [.28, .28]; 71% of edges were poor , 23% moderate, 6% good, and less than 1% excellent). As more resting-state data were added, the test-retest reliability continued to increase up to the limits of the dataset at 40 minutes (mean ICC = .54, 95% CI [.54, .54]; 26% of edges poor, 32% moderate, 25% good, and 18% excellent) (Figure 1A and left panel of Figure 2).

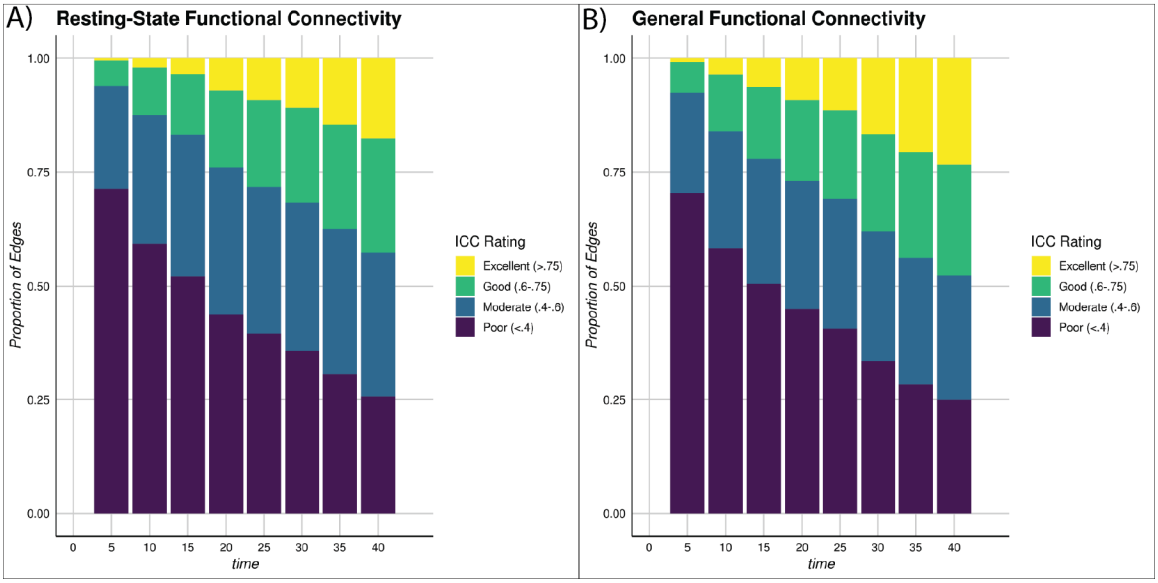
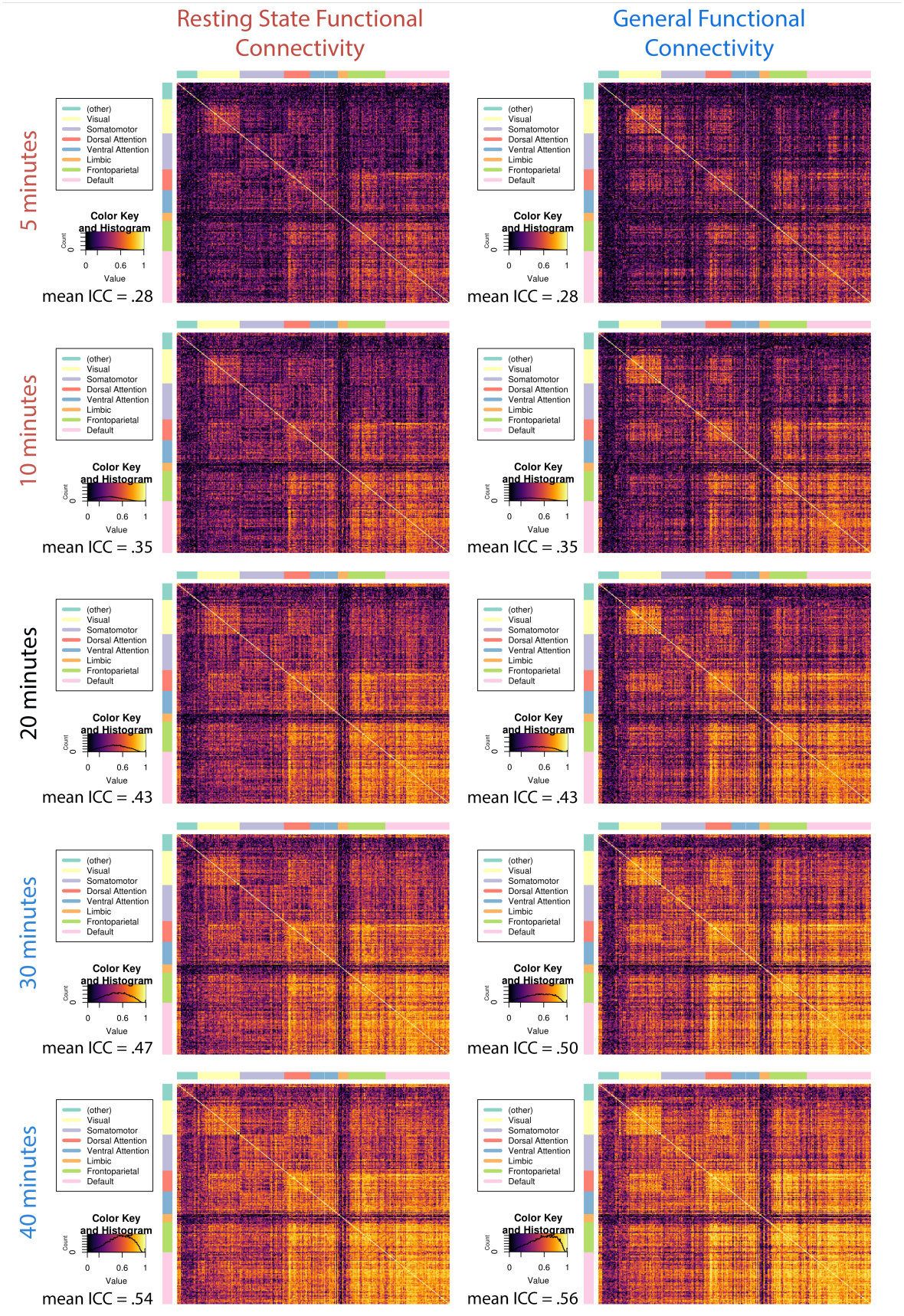


Figure 1. Test-retest reliability of individual differences in global intrinsic network connectivity increases as the amount of data used to estimate either RSFC (A) or GFC (B) increases. Stacked bar plots from the HCP dataset displaying the proportion of functional connections (i.e., edges) across neural networks as defined by (Power et al., 2011) that meet criteria for excellent, good, moderate, and poor reliability as indexed by ICCs.

GFC, defined as exactly equal parts of task and resting-state data in the HCP, followed a similar pattern of increasing reliability with increasing data: 5 minutes of data exhibited poor reliability

(mean ICC = .28, 95% CI [.28, .28]; 70% of edges were poor, 22% moderate, 7% good, and less than 1% excellent) but 40 minutes of data showed good reliability (mean ICC = .56, 95% CI [.55, .56]; 25% of edges were poor, 27% moderate, 24% good, and 23% excellent) (see Figures 1B, right panel of Figure 2). In addition, resting-state scans were not required to generate a reliable measure of GFC, as this pattern of increasing reliability held when task data alone were used to measure intrinsic connectivity (see supplemental Table S1). Furthermore, data from individual task scans alone generally resulted in poor reliability (see supplemental figures S2 and S3). A paired sample t-test comparing ICCs across all edges revealed that GFC is significantly more reliable than resting-state functional connectivity when both are estimated with 40 minutes of data ($t(34715) = 140.01$, $p < .001$). While significant, this difference is not likely meaningful as it represents only a 0.02 larger mean ICC in GFC. The low p-value is driven by the large number of degrees of freedom in the t-test comparing a mean ICC estimate from nearly 35,000 edges. Thus, we see the mean ICC of RSFC and GFC as statistically separable but practically equivalent.

This general improvement in reliability with increasing data was not unique to the HCP but was also observed in the population-representative Dunedin Study. The Dunedin Study resting-state scan of 8:16 minutes showed generally low reliability (mean ICC = .30, 95% CI [.30, .30]; 65% of edges were poor, 23% moderate, 10% good, and 2% excellent). The average reliability of the intrinsic connectivity estimates, however, improved by 49% when using GFC created in the Dunedin Study by combining the 8:16 minutes of resting-state data with 27 minutes (before censoring) of task data (mean ICC = .45, 95% CI [.45, .45]; 41% of edges were poor, 26% moderate, 19% good and 14% excellent) (Supplemental Figure S3).


Resting State Functional Connectivity
General Functional Connectivity
5 minutes
10 minutes
20 minutes
30 minutes
40 minutes
(other)
Visual
Somatomotor
Dorsal Attention
Ventral Attention
Limbic
Frontoparietal
Default
Color Key and Histogram
Count
0
0.6
1
Value
mean ICC = .28
mean ICC = .28
mean ICC = .35
mean ICC = .35
mean ICC = .43
mean ICC = .43
mean ICC = .47
mean ICC = .50
mean ICC = .54
mean ICC = .56

426

Figure 2. Test-retest reliability of the intrinsic connectivity of canonical neural networks derived from either RSFC or GFC from the HCP scales with the amount of data available for analysis. Intra-Class Correlation (ICC) matrices for RSFC (left panel) and GFC (right panel) demonstrate comparable gains in reliability with increasing amounts of data across common intrinsic networks (Power et al., 2011). 5 and 10 minutes are written in red because these are common scan lengths for resting-state scans. 30 and 40 minutes are written in blue because many researchers have collected this amount of fMRI data when resting-state and task scans are combined. To the bottom left of ICC matrices is the color key for the ICCs, with a histogram indicating the density of ICCs for the corresponding graph.

Influence of the global signal on reliability

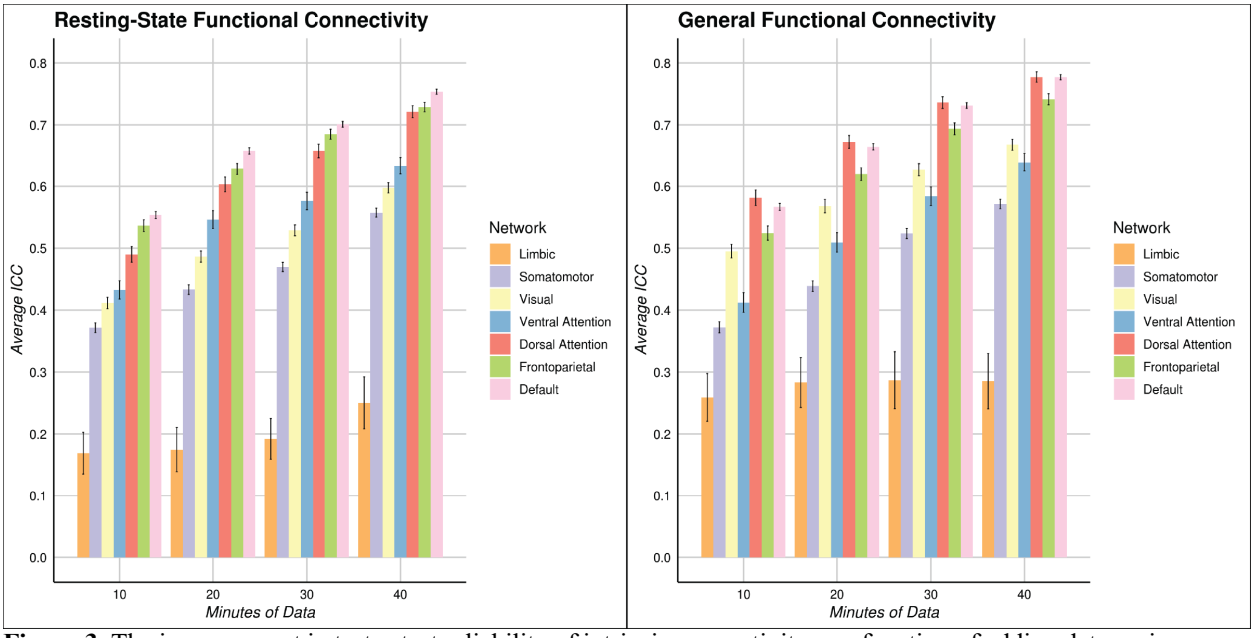
GSR is still a controversial preprocessing step. Therefore, the main reliability analyses in the HCP were rerun without GSR to test the influence of GSR on ICCs. Mean ICCs for 5 minutes of data without GSR were poor for both RSFC (mean ICC = .31, 95% CI [.31, .31]; 68% of edges were poor, 27% moderate, 4% good, and < 1% excellent) and GFC (mean ICC = .35, 95% CI [.35, .35]; 60% of edges were poor, 31% moderate, 8% good, and 1% excellent) (see supplemental figure S1). In parallel fashion to the data processed with GSR, ICCs increased as scan length increased up to 40 minutes for both RSFC (mean ICC = .61, 95% CI [.61, .61]; 8% of edges were poor, 34% moderate, 41% good, and 16% excellent) and GFC (mean ICC = .63, 95% CI [.63, .63]; 8% of edges were poor, 30% moderate, 40% good, and 22% excellent). Despite the improvement in mean ICC with additional data with and without GSR, ICCs were significantly higher when GSR was not applied: 40 minutes of GFC had an average ICC that was 0.07 higher without GSR than when GSR was applied. The same pattern was present in RSFC: 40 minutes of resting-state without GSR was more reliable on average than with GSR (rest with GSR = .61, rest without GSR = .54). While higher reliabilities without GSR could be a sign of improved detection of true individual differences in intrinsic connectivity, they may also be a sign of a stable propensity to move within the scanner, creating correlated motion artifact in test and retest data (Power et al., 2016; van Dijk et al., 2012). Because the global signal is highly susceptible to motion artifact and there is no widely accepted method to isolate and remove just the artifactual component of the global signal while retaining the neural component, we adopted a conservative approach by applying GSR signal

455 in all further analyses to ensure removal of motion related artifact present in the global signal
456 (Power et al., 2018).
457
458 *Is there network specificity to improvements in reliability*?
459
460 We next investigated network specificity in reliability by looking specifically at the reliability of
461 edges within each functional network rather than the reliability of the entire correlation matrix in
462 aggregate. As shown in Figure 3, there was clear network heterogeneity in reliability. While all
463 networks improved in reliability from 10 minutes of data to 40 minutes, some networks improved
464 more. With both RSFC and GFC, the limbic network had the lowest reliability at all scan lengths,
465 improving from 10 minutes of data (RSFC: mean ICC = .22, 95% CI [.19, .25]; GFC: mean ICC
466 = .30, 95% CI [.25, .34]) to 40 minutes of data (RSFC: mean ICC = .37, 95% CI [.33, .41]; GFC:
467 mean ICC = .39, 95% CI [.35, .44]), but failing to escape a poor mean ICC. In contrast, networks
468 that over-represent heteromodal association cortex (Margulies et al., 2016; Mesulam, 1998),
469 namely the default mode network and fronto-parietal network, on average achieved moderate
470 reliability with just 10 minutes of data for both RSFC (mean fronto-parietal network ICC = .53,
471 95% CI [.52, .53]; mean default mode network ICC = .57, 95% CI [.57, .58]) and GFC (mean
472 fronto-parietal network ICC = .55, 95% CI [.54, .56]; mean default mode network ICC = .57, 95%
473 CI [.57, .58]). Good reliability was achieved for these networks with 20 minutes of data for both
474 RSFC (mean fronto-parietal network ICC = .62, 95% CI [.61, .62]; mean default mode network
475 ICC = .66, 95% CI [.66, .67]) and GFC (mean fronto-parietal network ICC = .64, 95% CI [.63,
476 .64]; mean default mode network ICC = .65, 95% CI [.65, .66]). With 40 minutes of data, these
477 networks had mean reliabilities near the excellent range for both RSFC (mean fronto-parietal
478 network ICC = .73, 95% CI [.72, .73]; mean default mode network ICC = .75, 95% CI [.75, .75])
479 and GFC (mean fronto-parietal network ICC = .74, 95% CI [.74, .75]; mean default mode network
480 ICC = .77, 95% CI [.77, .77]). While there were some small differences in the network-specific

481 reliabilities between RSFC and GFC (e.g., see dorsal attention network in Figure 3), network

482 specific improvements with increasing scan length were similar for RSFC and GFC.

483



484
485 **Figure 3.** The improvement in test-retest reliability of intrinsic connectivity as a function of adding data varies across
486 canonical functional networks. In the left panel mean ICC for RSFC are displayed for within-network connections
487 across 7 previously defined networks (Yeo et al., 2011) for a variety of scan lengths. In the right panel the same data
488 are displayed for GFC. There are clear and consistent differences between networks. The limbic network consistently
489 has the lowest mean ICC, while the default mode network consistently has the highest mean ICC.

490

491

492 *What is the heritability of RSFC and GFC at different scan lengths?*

493 As shown in Figure 4A, the heritability of intrinsic connectivity depended on scan length. The

494 amount of variance for individual differences in RSFC attributable to the additive genetics (A)

495 component of the ACE model increased consistently as the amount of data increased from 5

496 minutes (mean A component = .09, 95% CI [.09, .1]) to 40 minutes (mean A component = .22,

497 95% CI [.21, .23]). This represents an increase in the amount of variance in RSFC attributable to

498 additive genetics of 138% as a function of increased scan length. This increase in variance

499 explained by additive genetics was also present with GFC from 5 minutes (mean A component =

500 .14, 95% CI [.13, .14]) to 40 minutes (mean A component = .28, 95% CI [.27, .29]). This

501 represented an increase in the amount of variance in GFC attributable to additive genetics of 107%

as a function of increased scan length. In addition, at equivalent scan lengths, GFC had significantly higher mean heritability across edges than RSFC (e.g., 5 minutes: $t(945) = 10.11$, $p < .001$) and 40 minutes: ($t(945) = 14.75$, $p < .001$)). Correspondingly, higher mean heritability in GFC than RSFC led to a greater percentage of heritable edges in GFC (see Figure 4B). In the ACE modeling, this pattern of increased heritability with increased scan time was associated with simultaneously decreased amount of variance explained by the E component, which comprises both non-shared environment and measurement error (Figure 4C). The E component decreased for both RSFC and GFC as data length increased from 5 minutes (RSFC mean E component = .81, 95% CI [.81, .82]; GFC mean E component = .77, 95% CI [.76, .77]) to 40 minutes (RSFC mean E component = .63, 95% CI [.63, .64]; GFC mean E component = .61, 95% CI [.60, .61]).

While the A and E components of the ACE modeling changed substantially with scan length, the shared-environment (C) (Figure 4D) and twin-specific (T) (supplemental Figure S4) components were relatively stable with scan length. The variance attributable to the C and T components was similar for 5 minutes of RSFC (mean C component = .04, 95% CI [.03, .04]; mean T component = .06, 95% CI [.05, .06]) and 5 minutes of GFC (mean C component = .04, 95% CI [.03, .04]; mean T component = .06, 95% CI [.06, .07]) as well as with 40 minutes of RSFC (mean C component = .06, 95% CI [.05, .06]; mean T component = .09, 95% CI [.08, .09]) and 40 minutes of GFC (mean C component = .05, 95% CI [.05, .06]; mean T component = .06, 95% CI [.06, .07]).

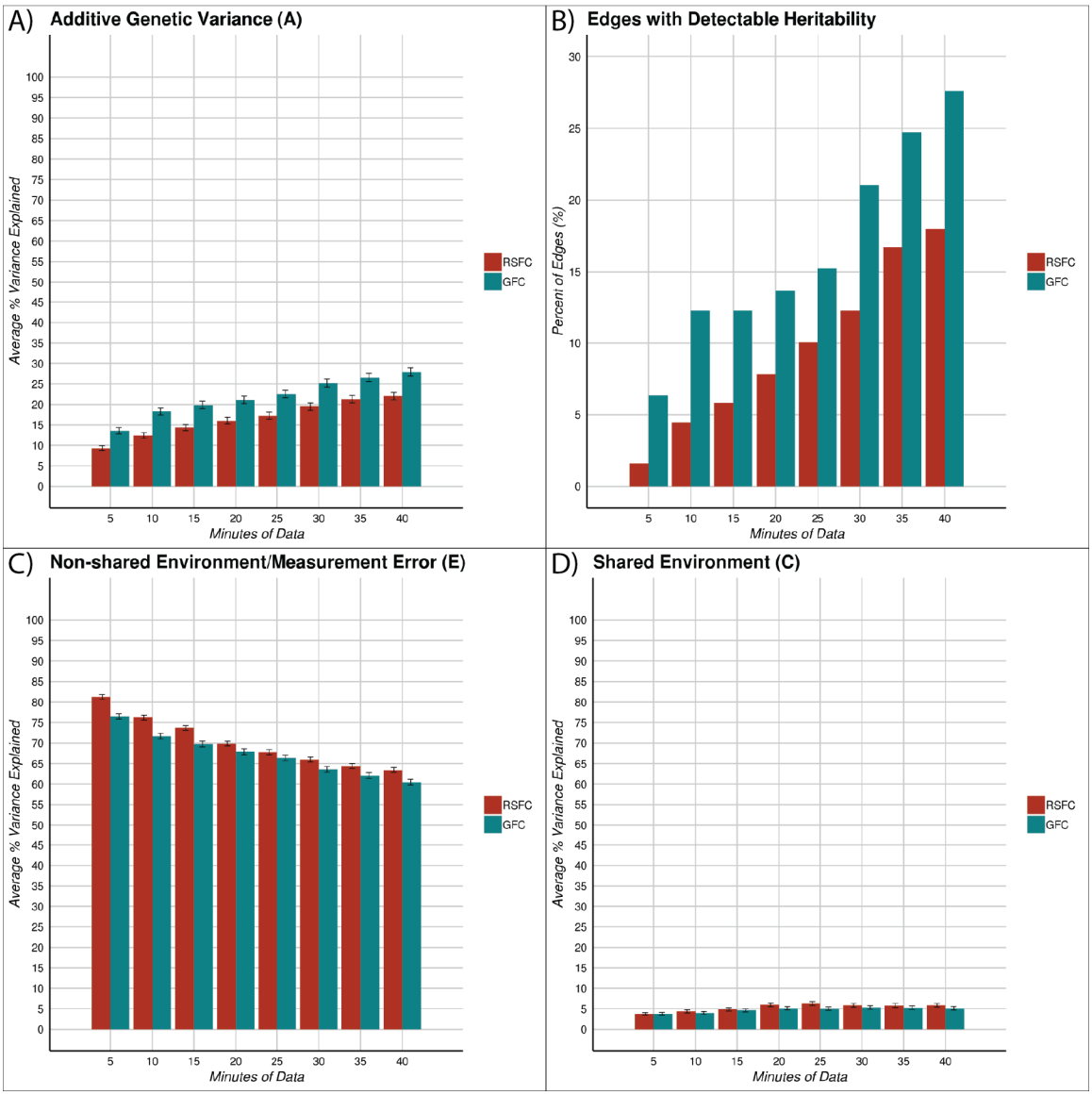


Figure 4. The heritability of RSFC and GFC vary as a function of scan length. Panels A, C and D display the mean A (additive genetics), E (non-shared environment + error) and C (shared environment) component estimates and 95% confidence intervals around those estimates, derived from ACE modeling of the twin data in the HCP. Scan length increases the A component and decreases the E component while having little effect on the C component. GFC also consistently has more variance attributable to the A component and lower variance attributable to the E component relative to RSFC. In panel B the % of heritable edges for RSFC and GFC are displayed across a variety of scan lengths. A heritable edge is defined as an edge with a lower bound of the 95% confidence interval that is larger than 0 in the ACE model.

Exploring the connection between intrinsic connectivity and cognitive ability

CPM was used to predict cognitive ability from both RSFC and GFC in both the HCP and the Dunedin Study. In the first set of analyses, within-sample predictions were made using leave-one-out cross validation in the HCP and Dunedin Study datasets separately (Figure 5, left and center

537 panels). In the HCP dataset, we found that individual differences in cognitive ability could be
538 predicted from 40 minutes of RSFC (r^2 = 4.2%, 95% CI [1.0%, 9.7%]) and from 40 minutes of
539 GFC (r^2 = 7.6%, 905% CI [2.7%, 14.2%]) (see Figure 5 and supplemental Table S3). These results
540 were replicated and extended in the Dunedin Study using scan lengths that are more representative
541 of existing datasets than the HCP. Cognitive ability could be predicted from the 8-minute RSFC
542 data (r^2 = 6.8%, 95% CI [3.5%, 11.3%]) as well as from 33 minutes of GFC (r^2 = 10.2%, 95% CI
543 [5.9%, 15.0%]; all available data; supplemental Table S4). Although in comparison with RSFC,
544 GFC had 81% greater predictive utility in the HCP dataset and 50% greater predictive utility in
545 the Dunedin Study dataset, the statistical comparison of the correlations was not statistically
546 significant in either the HCP (GFC r^2 = 7.6%, RSFC r^2 = 4.2%, Steiger's z = 1.317, p = .188) or
547 the Dunedin Study (GFC r^2 = 10.2%, RSFC r^2 = 6.8%, Steiger's z = 1.72, p = 0.085). In addition,
548 scan length had little effect on within-sample predictive utility (see supplemental figure S5).
549
550 While leave-one-out cross validation has been widely used with CPM (Finn et al., 2015; Noble et
551 al., 2017), the gold standard test of a model's generalizable predictive utility is its ability to predict
552 a phenotype in an independent sample (Bzdok and Yeo, 2017; Whelan and Garavan, 2014;
553 Yarkoni and Westfall, 2017). This is particularly important for GFC, as it purports to measure a
554 generalizable feature of intrinsic connectivity and thus correlations between individual differences
555 in GFC and phenotypes of interest should be shared across independent samples with different
556 batteries of fMRI tasks. Therefore, we used CPM to train two models (one with rest and one with
557 GFC) to predict cognitive ability in the larger Dunedin Study dataset and tested generalizability
558 by measuring the ability of these models to predict cognitive ability in the HCP dataset. The models
559 built with RSFC (r^2 = 5.6%, 95% CI [1.8%, 11.9%], p,q < .001) and GFC (r^2 = 9.5%, 95% CI
560 [4.0%, 16.3%], p,q < .001) from the Dunedin Study dataset both successfully predicted cognitive

ability in the HCP dataset (Figure 5 right panel and Figure 6 left panel). However, GFC had greater out-of-sample predictive utility than RSFC (GFC r^2 = 9.5%, RSFC r^2 = 5.6%, Steiger's z = 2.57, p = 0.010). This provides evidence that GFC is a more generalizable measure across independent samples than RSFC.

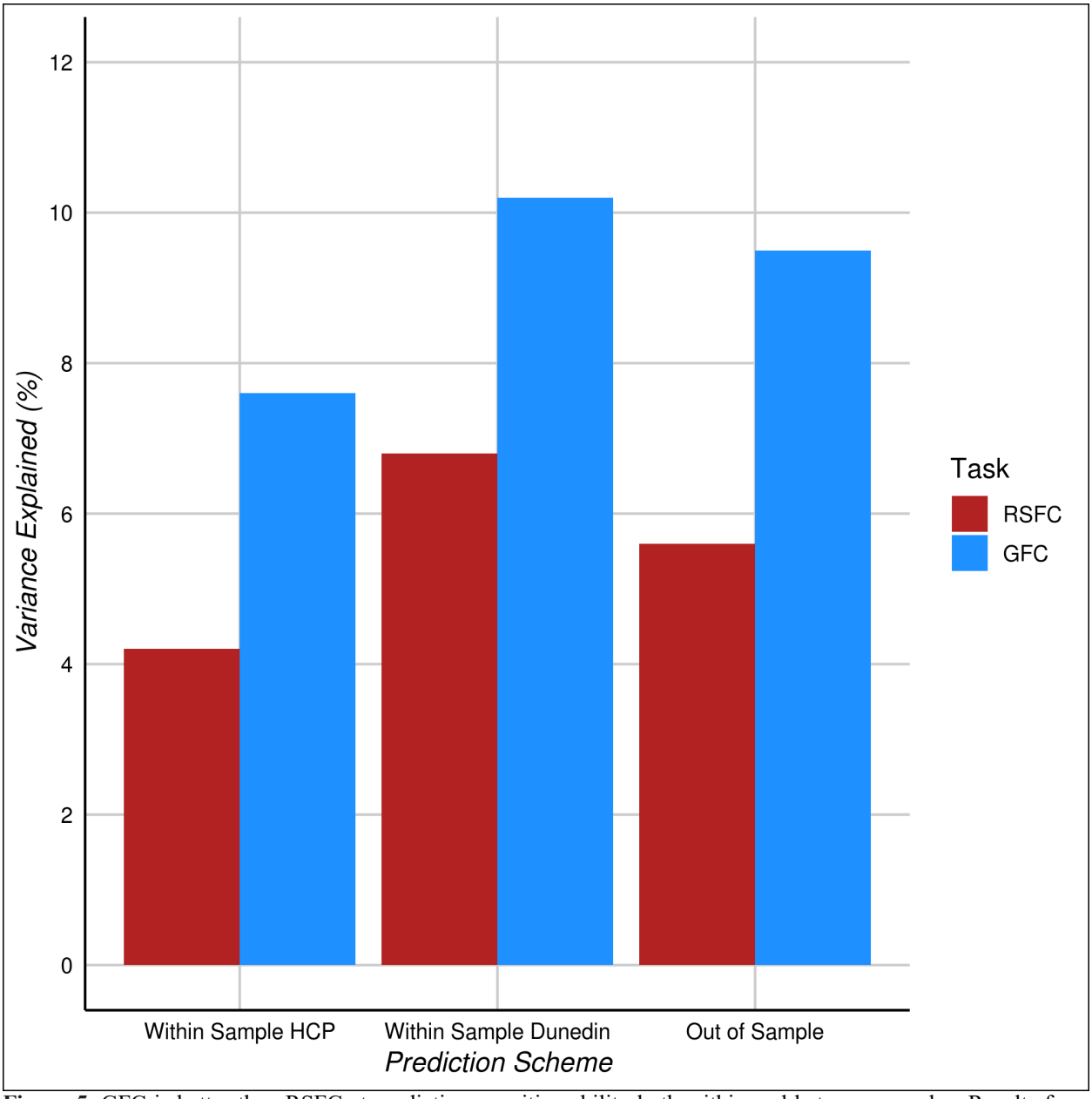


Figure 5. GFC is better than RSFC at predicting cognitive ability both within and between samples. Results from CPM models predicting cognitive ability from RSFC and GFC. The x-axis displays predictions from leave-one-out cross validation within sample and out-of-sample models trained using the Dunedin Study dataset and tested using the HCP dataset. Predictive utility is displayed as % variance explained (r^2).

572 *Does GFC predict cognitive ability better than intrinsic connectivity measured from specific tasks?*
573 Recent research has suggested that intrinsic connectivity estimates derived from task data are better
574 at predicting cognitive ability than estimates derived from resting-state data (Greene et al., 2018).
575 Our above finding that GFC outperforms RSFC at predicting cognitive ability is in line with this
576 observation. However, given that many researchers collect different tasks and resting-state scans
577 with different "task" demands (e.g., eyes-closed, fixation, etc.), we next investigated if GFC could
578 perform as well or better than specific tasks at out-of-sample prediction of cognitive ability. We
579 constructed five prediction models that were trained to predict cognitive ability using the Dunedin
580 Study dataset (the population-representative sample with the gold-standard WAIS IQ test) and
581 tested for generalizability using the HCP dataset (see supplemental figure S6 for the converse,
582 models trained in HCP and tested in the Dunedin Study). Two models were carried over from the
583 previous analyses (RSFC and GFC) and three additional task-specific models were added. These
584 consisted of models trained and tested on three parallel tasks available in both the Dunedin Study
585 and HCP (see methods for details). All models trained using the Dunedin Study dataset were able
586 to significantly predict cognitive ability in the HCP dataset (all FDR $q < .05$; Figure 6 and
587 supplemental Table S5). Prediction of cognitive ability from RSFC data ($r^2 = 5.6\%$, 95% CI [1.8%,
588 11.9%], $p,q < .001$), GFC data ($r^2 = 9.5\%$, 95% CI [4.0%, 16.3%], $p,q < .001$), emotional
589 processing data ($r^2 = 9.8\%$, 95% CI [4.3%, 17.1%], $p < .001$, $q = .001$), reward processing data (r^2
590 $= 9.1\%$, 95% CI [3.6%, 16.5%], $p < .001$, $q = .001$) and executive function data ($r^2 = 4.6\%$, 95%
591 CI [1.0%, 10.8%], $p < .001$, $q = .001$) all successfully generalized from the Dunedin Study dataset
592 to the HCP dataset.
593
594 Comparing the relative performance of models trained with each type of data revealed that GFC
595 was generalizable and performed as well or better at out-of-sample prediction of cognitive ability

compared to all tasks (Figure 6). Predictions of cognitive ability from GFC performed better than RSFC (GFC r^2 = 9.5%, RSFC r^2 = 5.6%, Steiger's z = 2.57, p = 0.01) and executive function data (GFC r^2 = 9.5%, executive function r^2 = 4.6%, Steiger's z = 2.19, p = .03), and performed as well as emotional processing data (GFC r^2 = 9.5%, emotion processing r^2 = 9.8%, Steiger's z = -.10, p = .92) and reward data (GFC r^2 = 9.5%, reward r^2 = 10.2%, Steiger's z = -.31, p = .76). Out-of-sample prediction was largely unaffected by the amount of data used in constructing the intrinsic connectivity matrices but was adversely affected when task regression was removed (see supplemental Table S2).

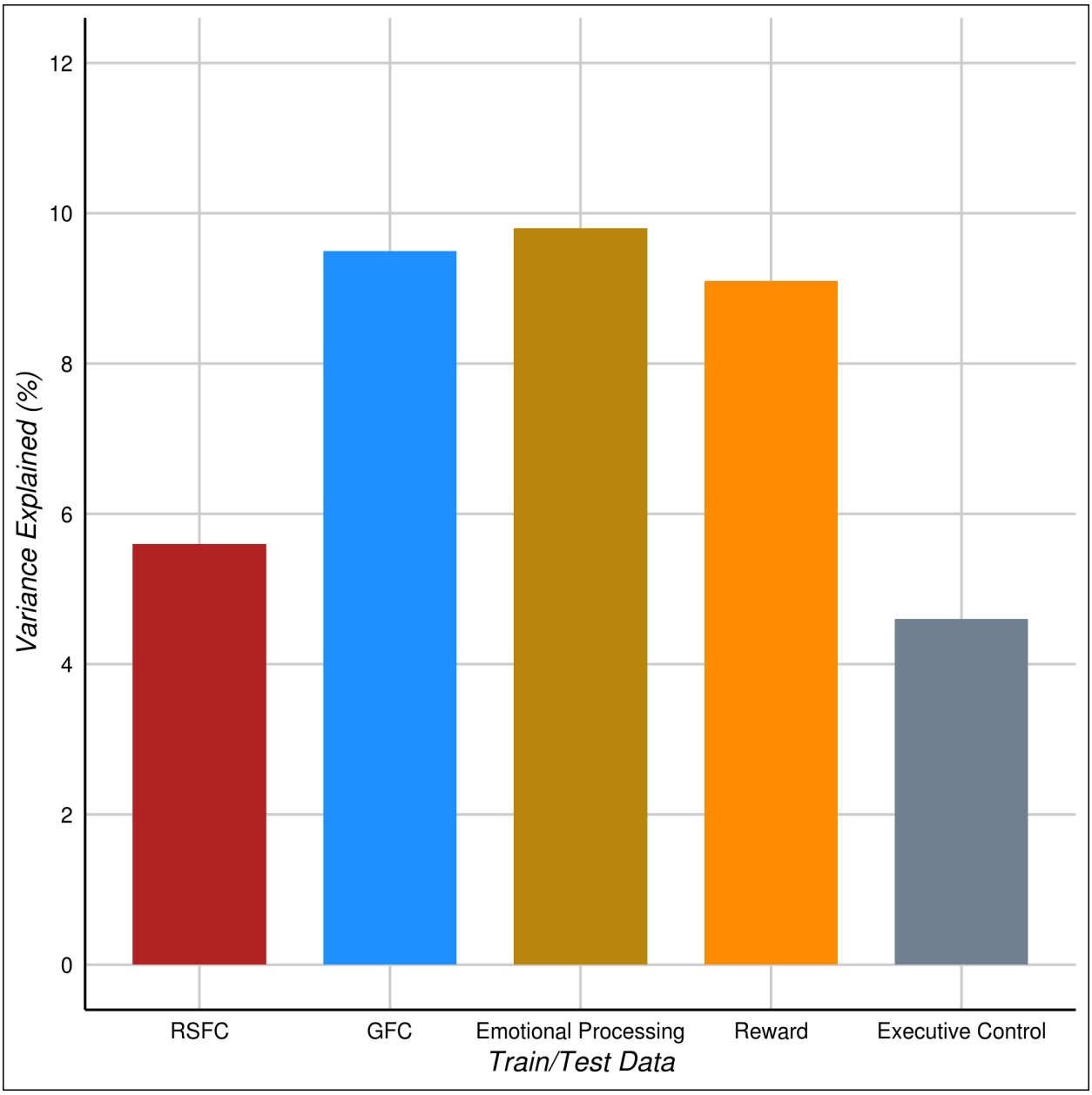


Figure 6. Out-of-sample prediction of cognitive ability for GFC is better than RSFC and as good as task-derived intrinsic connectivity. All models were trained on intrinsic connectivity data from the Dunedin Study and tested using data from the HCP. Models were trained and tested with the same type of data. With task data this meant that models were trained and tested with tasks that have a comparable (parallel) task in both the HCP and Dunedin study. Predictive utility is displayed as % variance explained (r^2).

Discussion

Here we present GFC as a method for readily improving the reliability of estimating individual differences in intrinsic connectivity by combining resting-state and task data. Across two independent datasets we found that scan length has a significant impact on the reliability of

intrinsic connectivity estimates regardless of whether the data came solely from resting-state scans (i.e., RSFC) or from a combination of resting-state and task scans (i.e., GFC). GFC, the addition of task scans to short resting-state scans, substantially increased the reliability of intrinsic connectivity over and above short resting-state scans alone. Moreover, we found that the improved reliability from aggregating data resulted in higher estimates of the heritable variation in intrinsic connectivity, and that this gain was larger for GFC than RSFC. In addition, GFC consistently performed better than RSFC and performed as well or better than intrinsic connectivity from individual task scans at predicting individual differences in a complex human trait, namely cognitive ability.

These findings have several implications for current and future neuroimaging research because they demonstrate that 1) reliable and heritable measurement of individual differences in the functional architecture of the brain is not only achievable in niche, specialty datasets with hours of resting-state scans, but also in many existing datasets if GFC is adopted; 2) mapping individual differences in behavior and cognition may be improved by adopting GFC, as demonstrated by generalizable predictive utility when studying cognitive ability; and 3) to the extent that commensurate gains in reliability and heritability may be achieved in other datasets, our results provide a reference that researchers can use to roughly estimate the gain in reliability they may achieve by adopting GFC (Figures 1, 2, S2 and S3). We next elaborate on each of these points.

Data aggregation across task and resting-state scans boosts reliability

We found that GFC, derived from the combination of resting-state and task scans, can be a reliable measure of individual differences in intrinsic connectivity. Consistent with previous studies (Birn et al., 2013; Finn et al., 2015; Laumann et al., 2015), we showed that the test-retest reliability of

640 intrinsic connectivity is highly dependent on the amount of data collected. This is true of RSFC as
641 well as GFC. With only 5-10 minutes of data, both RSFC and GFC show poor reliability in the
642 HCP and Dunedin Study datasets. Only with 30-40 minutes of data do RSFC and GFC begin to
643 broadly display good reliability. These findings follow directly from past studies that have shown
644 reliability depends on scan length (Anderson et al., 2011; Birn et al., 2013; Hacker et al., 2013),
645 that task and resting-state data share a large proportion of variance (Cole et al., 2014; Geerligs et
646 al., 2015), and that both task and resting-state data measure common individual differences in
647 intrinsic connectivity (Gratton et al., 2018). Some functional networks, however, can achieve good
648 reliability with relatively little data. For example, the default mode and fronto-parietal networks
649 achieved good reliability for within-network connections with just 20 minutes of RSFC or GFC
650 (see figure 3). In contrast, the limbic network failed to achieve even fair levels of reliability with
651 up to 40 minutes of data. These findings suggest that the amount of data needed to derive reliable
652 estimates of intrinsic connectivity for individual differences research will vary depending on the
653 network of interest.
654
655 Importantly, the current reality is that most neuroimaging studies have only 5-10 minutes of
656 resting-state data, which are further limited by sampling variability, motion-artifacts, and other
657 sources of noise (Gordon et al., 2017; Gratton et al., 2018; Power et al., 2014). Given the reliability
658 estimates reported here, it is likely that in many studies, individual differences in intrinsic
659 connectivity will be unmeasurable with resting-state data alone (Anderson et al., 2011; Hacker et
660 al., 2013). Fortunately, it is common practice for researchers to collect 10-40 minutes of task fMRI
661 in addition to resting-state. Therefore, if researchers adopt a broader definition of intrinsic
662 connectivity, such as GFC, they can combine task and resting-state data to achieve a reliable
663 measure of individual differences in intrinsic connectivity.

Data aggregation across task and resting-state scans boosts heritability

We found that individual differences in intrinsic connectivity in both RSFC and GFC were significantly attributable to additive genetic influences. With 40 minutes of data, 22% of the variance in RSFC and 28% of the variance in GFC was attributable to additive genetic effects. With equivalent amounts of data, GFC produced higher heritability estimates than RSFC, and the influence of additive genetic effects was detectable in a larger proportion of functional connections (Figure 4). In addition, we found that scan length had an impact on the heritability estimates of intrinsic connectivity in both RSFC and GFC. The average amount of variance in intrinsic connectivity that was attributable to additive genetic effects more than doubled in both RSFC (increase of 138%) and GFC (increase of 107%) as scan length increased from 5 to 40 minutes. While the heritability of intrinsic connectivity has been investigated before (Adhikari et al., 2018; Ge et al., 2017), the link between scan length and heritability has not been described. Although a reliable measure does not have to be heritable, a measure cannot have high heritability without high reliability (Ge et al., 2017). That is, measurement reliability places a ceiling on the heritability estimate of a phenotype. Therefore, our finding that the heritability of intrinsic connectivity increases with increasing reliability suggests that the low reliability of intrinsic connectivity in short fMRI scans limits estimations of the true underlying heritability of intrinsic connectivity in many datasets. More specifically, our finding suggests that researchers should consider the amount of fMRI data available for estimation of intrinsic connectivity when planning and evaluating imaging genetics research.

Our results have additional implications for genetically-informed fMRI research. Neuroimaging measures have been considered promising intermediate phenotypes that would bring researchers

688 closer to the mechanisms through which genetics lead to heritable psychological traits (Braskie et
689 al., 2011; Hariri and Weinberger, 2003; Hasler and Northoff, 2011; Meyer-Lindenberg and
690 Weinberger, 2006). To uncover these links, large datasets like ENIGMA and the UK Biobank
691 (Sudlow et al., 2015; Thompson et al., 2014), have been used to conduct genome wide association
692 studies (GWAS) of MRI-based phenotypes. While these investigations have had success in finding
693 genetic variants linked to structural MRI measures (Adams et al., 2016; Hibar et al., 2017, 2015;
694 Stein et al., 2012), they have largely failed to find significant genetic correlates of functional
695 intrinsic connectivity measures (Elliott et al., 2017). One reason for this failure may be the low
696 reliability and heritability of the short resting-state scans typically used to derive individual
697 differences measures of RSFC in these studies. Our results suggest that future GWAS may be
698 better powered to find genetic variants linked to the functional architecture of the brain by adopting
699 GFC rather than RSFC from short resting-state scans.
700
701 *How general is general functional connectivity?*
702 Previous research has shown that the predictive utility of intrinsic connectivity can be higher when
703 derived from task data than when derived from resting-state data (Greene et al., 2018). Some task
704 states in particular may reveal individual differences in intrinsic connectivity that aid in predictive
705 utility (Finn et al., 2017). While these findings provide insight into individual differences in
706 intrinsic connectivity and a new use for task data, they pose a problem for replication and
707 cumulative aggregation of findings across datasets. While most fMRI studies collect task data,
708 very few collect the same tasks. Our findings here suggest that GFC may provide a solution to this
709 problem. First, we found that like task-derived intrinsic connectivity, GFC can outperform RSFC
710 when predicting within-sample and out-of-sample variability in cognitive ability (Figures 5 and
711 6). Second, we found that GFC performed as well or better than specific tasks alone at predicting

712 cognitive ability out-of-sample. Critically, in these comparisons GFC was constructed with a
713 combination of different tasks in the training and testing datasets suggesting that GFC may
714 measure a common, generalizable backbone of intrinsic connectivity that can be applied and
715 aggregated across independent datasets with different task batteries (supplemental Tables S6 and
716 S7). While further research is needed to replicate these findings, they nevertheless suggest that
717 GFC may provide a generalizable, practical framework that researchers with resting-state and task
718 data can use to drive a cumulative, predictive neuroscience of individual differences (Button et al.,
719 2013; Szucs and Ioannidis, 2017).
720
721 *Is general functional connectivity too heterogeneous?*
722 An understandable reluctance in implementing GFC could be that it would introduce heterogeneity
723 in the estimation of intrinsic connectivity. Whereas RSFC is thought to be a common framework
724 with generalizability across datasets, the combination of different sets of tasks in different datasets
725 may introduce additional variability. Below, we provide three reasons why we think this is less
726 problematic than it first appears. We further provide specific recommendations and guidelines for
727 adopting and estimating GFC and point to features of our analyses that support the validity of GFC
728 as a robust and reliable measure of individual differences in intrinsic connectivity.
729
730 First, resting-state data are by nature heterogenous, and the resting-state is its own type of task
731 (Buckner et al., 2013). Researchers collect resting-state data under different conditions in which
732 participants close their eyes, have their eyes open, or fixate. These differences in resting-state
733 conditions influence intrinsic connectivity (Patriat et al., 2013). Despite these differences, all
734 approaches are collectively referred to as resting-state scans. In addition, factors like thought
735 content (Christoff et al., 2009; Hurlburt et al., 2015), caffeine intake (Wong et al., 2012), recent

736 cognitive tasks (Waites et al., 2005), and falling asleep (Deco et al., 2014) can introduce further
737 heterogeneity that can bias intrinsic connectivity estimates. It has even been estimated that
738 approximately 30% of participants fall asleep within the first three minutes of a resting-state scan
739 (Tagliazucchi and Laufs, 2014). Nevertheless, resting-state scans measure a common set of
740 functional networks (Yeo et al., 2011), and display trait-like features driven by stable factors such
741 as genetics (Glahn et al., 2010) and structural connectivity (Honey et al., 2009). For these reasons
742 resting-state data represent substantial promise as an individual-differences measure if enough data
743 are collected to average out sampling variability (Gratton et al., 2018). But, as already stated, few
744 researchers collect enough data to fulfill this promise. Connectivity estimates from task data are
745 also shaped by many of these stable factors (Cole et al., 2014; Krienen et al., 2014), and data from
746 different tasks predominantly measure overlapping individual differences that are only weakly
747 influenced by task demands (Gratton et al., 2018). Admittedly, GFC will vary, to some extent,
748 across samples because each study collects a unique combination of resting-state and task data.
749 However, RSFC will vary too. Moreover, we found that disparate task and resting-state scans
750 collectively measure reliable, heritable individual differences with generalizable out-of-sample
751 predictive utility in the form of GFC.
752
753 Second, the benefits of GFC converged across two different samples, with different scanners,
754 scanning parameters, and demographics. The HCP represents a healthy sample of highly-educated
755 individuals (Van Essen et al., 2013), while the Dunedin Study represents a population-
756 representative birth cohort with a broad range of mental and physical health conditions,
757 socioeconomic status, and full representation of variability in many complex traits (Poulton et al.,
758 2015). Across samples, the data came from 11 different tasks (HCP = 7, Dunedin Study = 4), many
759 with unrelated cognitive demands, different scan lengths, stimuli, behavioral requirements, and

instructions. The data were processed with different preprocessing schemes and software ("minimally preprocessed" in the HCP and custom scripts in the Dunedin Study). While the amount of scan time was tightly controlled in the HCP analyses (i.e., equal time after censoring), in the Dunedin Study motion censoring led to unequal scan lengths across participants as is the case in many analyses of intrinsic connectivity. Despite these many differences between datasets, the results demonstrate that GFC can achieve good test-retest reliability, inter-rater reliability (different preprocessing schemes), out-of-sample reliability (convergence across datasets and out-of-sample prediction), and parallel forms reliability (different tasks in each sample) (Dubois and Adolphs, 2016). While we present evidence in favor of the generalizability of GFC, future research should further investigate the heterogeneity in GFC to find combinations of tasks that most efficiently estimate individual differences in intrinsic connectivity (Finn et al., 2017; Shah et al., 2016).

Third, reliability fundamentally limits the ability to detect associations between any two measures (Nunnally Jr., 1970; Vul et al., 2009). Therefore, any investigation mapping intrinsic connectivity to individual differences in behavior, cognition, or disease is fundamentally limited by the reliability of the intrinsic connectivity measures. Relatedly, statistical power to detect true effects depends on the anticipated effect size, which is in turn dependent on reliability of each measurement (Kanyongo et al., 2007; Williams and Zimmerman, 1989). Given that many studies only have 5-10 minutes of resting-state data and consequently poor reliability of intrinsic connectivity measures, they have drastically reduced statistical power (Button et al., 2013; Ioannidis, 2008, 2005), and likely tenuous correlates of individual differences unless they adopt a framework like GFC. Previous research has found that collectively a large number of unreliable edges can achieve multivariate reliability that results in predictive utility (Noble et al., 2017).

While multivariate predictive utility may be achievable in studies with poor univariate edge-wise reliability (Figure S5), the interpretability and clinical utility of these studies will be diminished by low univariate reliability. This is because interpretability and clinical utility depend on isolating specific brain areas or functional connections that can be said to be "most" important to the phenotype of interest. This cannot be done without decent univariate reliability because accurate estimation of true parameters in a prediction model depends on reliability (Cremers et al., 2017; Vul et al., 2009; Yarkoni and Westfall, 2017). For example, if researchers want to characterize biomarkers of mental illness that could be targeted with interventions like transcranial magnetic stimulation (Fox et al., 2014) or real-time neurofeedback (Caria et al., 2012), they will need to isolate specific circuits that with confidence can be said to be "altered" and relevant to the disease process in mental illness. Actionable identification of these biomarkers will only be possible with adequate univariate reliabilities. GFC provides a practical solution to this problem, while also addressing power and reliability issues that have been identified more generally in neuroscience (Button et al., 2013; Szucs and Ioannidis, 2017) and beyond (Aarts et al., 2015; Errington et al., 2014; Ioannidis, 2005; Ioannidis et al., 2001; Munafò et al., 2017). GFC provides a tool for researchers to repurpose existing and future datasets to better contribute to a cumulative science of individual differences (Dubois and Adolphs, 2016) with clinical relevance (Fox, 2010). For these reasons, we think that in many studies the gains in reliability offered by GFC will outweigh potential task-induced "heterogeneity."

Recommendations

Based on the above summary, we generally recommend that researchers seeking to maximize the reliability and generalizable predictive utility of GFC should aggregate as much data as possible, across all rest and task data, that they have at their disposal. Our results further suggest that with a

808 minimum of 20 minutes of GFC data, researchers can achieve good reliability in some networks
809 of broad interest (e.g., default, frontoparietal and dorsal attention networks). Furthermore, a strict
810 balance of the amount of data from each task and resting-state scan is not necessary to achieve the
811 benefits of GFC. This is demonstrated by the increased reliability of GFC despite unequal task
812 lengths in the Dunedin Study dataset and in the HCP dataset by continued improvement in
813 reliability despite timepoints being selected at random from the remaining data after reaching a
814 minimum of 25 minutes of data. However, we urge caution in applying GFC to datasets that are
815 heavily skewed (> 50%) towards resting-state or one specific task. Although we expect that
816 deriving GFC from this type of dataset will boost reliability, the generalizability of these findings
817 to GFC derived from other datasets may be limited because of the overrepresentation of a specific
818 type of data. Lastly, our results suggest that researchers developing new protocols may benefit
819 from collecting multiple short task scans rather than a single longer task scan (supplemental tables
820 S1 and S5).
821
822 *Conclusion*
823 Here we propose GFC as a method for improving the reliability of estimating individual
824 differences in the intrinsic architecture of functional networks. We demonstrate that, when the
825 amount of data available for analysis is held constant, the combination of resting-state and task
826 data is at least as reliable as resting-state alone. Additionally, when the same amount of data is
827 available, GFC produces higher heritability estimates than RSFC. Many researchers who have less
828 than 25 minutes of resting-state data but have additionally collected task data on the same
829 participants can immediately boost reliability, heritability, and power by adopting GFC as a
830 measure of intrinsic connectivity when studying individual differences. Our findings also suggest
831 that future data collection should consider naturalistic fMRI (Hasson et al., 2010, 2004; Huth et

832 al., 2016; Lahnakoski et al., 2012; Vanderwal et al., 2017) and engaging tasks in addition to
833 resting-state scans when generating data for estimating intrinsic connectivity. This may be
834 especially true when studying children, elderly, or mentally ill participants, as these groups often
835 cannot lie still for 25 minutes or more of resting-state scanning (Power et al., 2012; Satterthwaite
836 et al., 2013; Yuan et al., 2009). Our current findings also suggest that researchers need not think
837 of collecting resting-state and task data as a zero-sum competition for scan time when designing
838 studies. Instead, our findings demonstrate that task and resting-state data provide complementary
839 measures that together can lead to more reliable, heritable measurement of individual differences
840 in intrinsic connectivity. We think that collecting adequate amounts (i.e., > 25 minutes) of high-
841 quality fMRI data should be emphasized over the traditional task-rest dichotomy, when individual
842 differences are of interest. If fMRI is going to be a part of a cumulative neuroscience of individual
843 differences with clinical relevance, psychometric measurement quality, especially reliability, must
844 be at the forefront of study design.

845 **Acknowledgement.** The Dunedin Multidisciplinary Health and Development Study is supported
846 by the NZ HRC and NZ MBIE. This research was supported by National Institute on Aging
847 grant R01AG032282, R01AG049789, and UK Medical Research Council grant MR/P005918/1.
848 Additional support was provided by the Jacobs Foundation. MLE is supported by the National
849 Science Foundation Graduate Research Fellowship under Grant No. NSF DGE-1644868. Thank
850 you to members of the Advisory Board for the Dunedin Neuroimaging Study. We thank the
851 Dunedin Study members, unit research staff, and Study founder Phil Silva. The study protocol
852 was approved by the institutional ethical review boards of the participating universities. Study
853 members gave informed consent before participating. The authors declare no competing financial
854 interests.
855
856 **Author Contributions.** Conceptualization, M.L.E.; Methodology, Software and Formal
857 Analysis, M.L.E., A.R.K and M.C.; Data Curation, M.L.E., A.R.K., A.C., and T.E.M.;
858 Resources, T.R.M., R.K., S.R., R.P., A.C., T.E.M. and A.R.H., Writing, M.L.E., A.R.K., M.J.K.,
859 M.C., T.R.M., R.K., D.I., S.R., R.P., A.C., T.E.M. and A.R.H.; Supervision, A.C., T.E.M. and
860 A.R.H.; Funding Acquisition, R.P., A.C., T.E.M. and A.R.H.
861

References

Aarts, A.A., Anderson, J.E., Anderson, C.J., Attridge, P.R., Attwood, A., Axt, J., Babel, M., Bahník, Š., Baranski, E., Barnett-Cowan, M., Bartmess, E., Beer, J., Bell, R., Bentley, H., Beyan, L., Binion, G., Borsboom, D., Bosch, A., Bosco, F.A., Bowman, S.D., Brandt, M.J., Braswell, E., Brohmer, H., Brown, B.T., Brown, K., Brüning, J., Calhoun-Sauls, A., Callahan, S.P., Chagnon, E., Chandler, J., Chartier, C.R., Cheung, F., Christopherson, C.D., Cillessen, L., Clay, R., Cleary, H., Cloud, M.D., Conn, M., Cohoon, J., Columbus, S., Cordes, A., Costantini, G., Alvarez, L.D.C., Cremata, E., Crusius, J., DeCoster, J., DeGaetano, M.A., Penna, N.D., Den Bezemer, B., Deserno, M.K., Devitt, O., Dewitte, L., Dobolyi, D.G., Dodson, G.T., Donnellan, M.B., Donohue, R., Dore, R.A., Dorrough, A., Dreber, A., Dugas, M., Dunn, E.W., Easey, K., Eboigbe, S., Eggleston, C., Embley, J., Epskamp, S., Errington, T.M., Estel, V., Farach, F.J., Feather, J., Fedor, A., Fernández-Castilla, B., Fiedler, S., Field, J.G., Fitneva, S.A., Flagan, T., Forest, A.L., Forsell, E., Foster, J.D., Frank, M.C., Frazier, R.S., Fuchs, H., Gable, P., Galak, J., Galliani, E.M., Gampa, A., Garcia, S., Gazarian, D., Gilbert, E., Giner-Sorolla, R., Glöckner, A., Goellner, L., Goh, J.X., Goldberg, R., Goodbourn, P.T., Gordon-McKeon, S., Gorges, B., Gorges, J., Goss, J., Graham, J., Grange, J.A., Gray, J., Hartgerink, C., Hartshorne, J., Hasselman, F., Hayes, T., Heikensten, E., Henninger, F., Hodsoll, J., Holubar, T., Hoogendoorn, G., Humphries, D.J., Hung, C.O.Y., Immelman, N., Irsik, V.C., Jahn, G., Jäkel, F., Jekel, M., Johannesson, M., Johnson, L.G., Johnson, D.J., Johnson, K.M., Johnston, W.J., Jonas, K., Joy-Gaba, J.A., Kappes, H.B., Kelso, K., Kidwell, M.C., Kim, S.K., Kirkhart, M., Kleinberg, B., Knežević, G., Kolorz, F.M., Kossakowski, J.J., Krause, R.W., Krijnen, J., Kuhlmann, T., Kunkels, Y.K., Kyc, M.M., Lai, C.K., Laique, A., Lakens, D., Lane, K.A., Lassetter, B., Lazarević, L.B., Le Bel, E.P., Lee, K.J., Lee, M., Lemm, K., Levitan, C.A., Lewis, M., Lin, L., Lin, S., Lippold, M., Loureiro, D., Luteijn, I., MacKinnon, S., Mainard, H.N., Marigold, D.C., Martin, D.P., Martinez, T., Masicampo, E.J., Matacotta, J., Mathur, M., May, M., Mechin, N., Mehta, P., Meixner, J., Melinger, A., Miller, J.K., Miller, M., Moore, K., Möschl, M., Motyl, M., Müller, S.M., Munafo, M., Neijenhuijs, K.I., Nervi, T., Nicolas, G., Nilsonne, G., Nosek, B.A., Nuijten, M.B., Olsson, C., Osborne, C., Ostkamp, L., Pavel, M., Penton-Voak, I.S., Perna, O., Pernet, C., Perugini, M., Pipitone, R.N., Pitts, M., Plessow, F., Prenoveau, J.M., Rahal, R.M., Ratliff, K.A., Reinhard, D., Renkewitz, F., Ricker, A.A., Rigney, A., Rivers, A.M., Roebke, M., Rutchick, A.M., Ryan, R.S., Sahin, O., Saide, A., Sandstrom, G.M., Santos, D., Saxe, R., Schlegelmilch, R., Schmidt, K., Scholz, S., Seibel, L., Selterman, D.F., Shaki, S., Simpson, W.B., Sinclair, H.C., Skorinko, J.L.M., Slowik, A., Snyder, J.S., Soderberg, C., Sonnleitner, C., Spencer, N., Spies, J.R., Steegen, S., Stieger, S., Strohminger, N., Sullivan, G.B., Talhelm, T., Tapia, M., Te Dorsthorst, A., Thomae, M., Thomas, S.L., Tio, P., Traets, F., Tsang, S., Tuerlinckx, F., Turchan, P., Valášek, M., Van't Veer, A.E., Van Aert, R., Van Assen, M., Van Bork, R., Van De Ven, M., Van Den Bergh, D., Van Der Hulst, M., Van Dooren, R., Van Doorn, J., Van Renswoude, D.R., Van Rijn, H., Vanpaemel, W., Echeverría, A.V., Vazquez, M., Velez, N., Vermue, M., Verschoor, M., Vianello, M., Voracek, M., Vuu, G., Wagenmakers, E.J., Weerdmeester, J., Welsh, A., Westgate, E.C., Wissink, J., Wood, M., Woods, A., Wright, E., Wu, S., Zeelenberg, M., Zuni, K., 2015. Estimating the reproducibility of psychological science. Science (80-.). 349. https://doi.org/10.1126/science.aac4716

Adams, H.H.H., Hibar, D.P., Chouraki, V., Stein, J.L., Nyquist, P.A., Rentería, M.E., Trompet, S., Arias-Vasquez, A., Seshadri, S., Desrivières, S., Beecham, A.H., Jahanshad, N., Wittfeld, K., Van Der Lee, S.J., Abramovic, L., Alhusaini, S., Amin, N., Andersson, M.,

Arfanakis, K., Aribisala, B.S., Armstrong, N.J., Athanasiu, L., Axelsson, T., Beiser, A., Bernard, M., Bis, J.C., Blanken, L.M.E., Blanton, S.H., Bohlken, M.M., Boks, M.P., Bralten, J., Brickman, A.M., Carmichael, O., Chakravarty, M.M., Chauhan, G., Chen, Q., Ching, C.R.K., Cuellar-Partida, G., Braber, A. Den, Doan, N.T., Ehrlich, S., Filippi, I., Ge, T., Giddaluru, S., Goldman, A.L., Gottesman, R.F., Greven, C.U., Grimm, O., Griswold, M.E., Guadalupe, T., Hass, J., Haukvik, U.K., Hilal, S., Hofer, E., Hoehn, D., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Kasperaviciute, D., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Liao, J., Liewald, D.C.M., Lopez, L.M., Luciano, M., Macare, C., Marquand, A., Matarin, M., Mather, K.A., Mattheisen, M., Mazoyer, B., McKay, D.R., McWhirter, R., Milaneschi, Y., Mirza-Schreiber, N., Muetzel, R.L., Maniega, S.M., Nho, K., Nugent, A.C., Loohuis, L.M.O., Oosterlaan, J., Papmeyer, M., Pappa, I., Pirpamer, L., Pudas, S., Pütz, B., Rajan, K.B., Ramasamy, A., Richards, J.S., Risacher, S.L., Roiz-Santiañez, R., Rommelse, N., Rose, E.J., Royle, N.A., Rundek, T., Sämann, P.G., Satizabal, C.L., Schmaal, L., Schork, A.J., Shen, L., Shin, J., Shumskaya, E., Smith, A. V., Sprooten, E., Strike, L.T., Teumer, A., Thomson, R., Tordesillas-Gutierrez, D., Toro, R., Trabzuni, D., Vaidya, D., Van Der Grond, J., Van Der Meer, D., Van Donkelaar, M.M.J., Van Eijk, K.R., Van Erp, T.G.M., Van Rooij, D., Walton, E., Westlye, L.T., Whelan, C.D., Windham, B.G., Winkler, A.M., Woldehawariat, G., Wolf, C., Wolfers, T., Xu, B., Yanek, L.R., Yang, J., Zijdenbos, A., Zwiers, M.P., Agartz, I., Aggarwal, N.T., Almasy, L., Ames, D., Amouyel, P., Andreassen, O.A., Arepalli, S., Assareh, A.A., Barral, S., Bastin, M.E., Becker, D.M., Becker, J.T., Bennett, D.A., Blangero, J., Van Bokhoven, H., Boomsma, D.I., Brodaty, H., Brouwer, R.M., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Bulayeva, K.B., Cahn, W., Calhoun, V.D., Cannon, D.M., Cavalleri, G.L., Chen, C., Cheng, C.Y., Cichon, S., Cookson, M.R., Corvin, A., Crespo-Facorro, B., Curran, J.E., Czisch, M., Dale, A.M., Davies, G.E., De Geus, E.J.C., De Jager, P.L., De Zubicaray, G.I., Delanty, N., Depondt, C., Destefano, A.L., Dillman, A., Djurovic, S., Donohoe, G., Drevets, W.C., Duggirala, R., Dyer, T.D., Erk, S., Espeseth, T., Evans, D.A., Fedko, I.O., Fernández, G., Ferrucci, L., Fisher, S.E., Fleischman, D.A., Ford, I., Foroud, T.M., Fox, P.T., Francks, C., Fukunaga, M., Gibbs, J.R., Glahn, D.C., Gollub, R.L., Göring, H.H.H., Grabe, H.J., Green, R.C., Gruber, O., Gudnason, V., Guelfi, S., Hansell, N.K., Hardy, J., Hartman, C.A., Hashimoto, R., Hegenscheid, K., Heinz, A., Le Hellard, S., Hernandez, D.G., Heslenfeld, D.J., Ho, B.C., Hoekstra, P.J., Hoffmann, W., Hofman, A., Holsboer, F., Homuth, G., Hosten, N., Hottenga, J.J., Pol, H.E.H., Ikeda, M., Ikram, M.K., Jack, C.R., Jenkinson, M., Johnson, R., Jönsson, E.G., Jukema, J.W., Kahn, R.S., Kanai, R., Kloszewska, I., Knopman, D.S., Kochunov, P., Kwok, J.B., Lawrie, S.M., Lemaître, H., Liu, X., Longo, D.L., Longstreth, W.T., Lopez, O.L., Lovestone, S., Martinez, O., Martinot, J.L., Mattay, V.S., McDonald, C., McIntosh, A.M., McMahon, K.L., McMahon, F.J., Mecocci, P., Melle, I., Meyer-Lindenberg, A., Mohnke, S., Montgomery, G.W., Morris, D.W., Mosley, T.H., Mühleisen, T.W., Müller-Myhsok, B., Nalls, M.A., Nauck, M., Nichols, T.E., Niessen, W.J., Nöthen, M.M., Nyberg, L., Ohi, K., Olvera, R.L., Ophoff, R.A., Pandolfo, M., Paus, T., Pausova, Z., Penninx, B.W.J.H., Pike, G.B., Potkin, S.G., Psaty, B.M., Reppermund, S., Rietschel, M., Roffman, J.L., Romanczuk-Seiferth, N., Rotter, J.I., Ryten, M., Sacco, R.L., Sachdev, P.S., Saykin, A.J., Schmidt, R., Schofield, P.R., Sigurdsson, S., Simmons, A., Singleton, A., Sisodiya, S.M., Smith, C., Smoller, J.W., Soininen, H., Srikanth, V., Steen, V.M., Stott, D.J., Sussmann, J.E., Thalamuthu, A., Tiemeier, H., Toga, A.W., Traynor, B.J., Troncoso, J., Turner, J.A., Tzourio, C., Uitterlinden, A.G., Hernández, M.C.V., Van Der Brug, M., Van Der Lugt, A., Van Der Wee, N.J.A., Van Duijn, C.M., Van Haren, N.E.M., Van't Ent, D., Van Tol, M.J., Vardarajan, B.N., Veltman, D.J., Vernooij, M.W., Völzke, H., Walter,

H., Wardlaw, J.M., Wassink, T.H., Weale, M.E., Weinberger, D.R., Weiner, M.W., Wen, W., Westman, E., White, T., Wong, T.Y., Wright, C.B., Zielke, H.R., Zonderman, A.B., Deary, I.J., Decarli, C., Schmidt, H., Martin, N.G., De Craen, A.J.M., Wright, M.J., Launer, L.J., Schumann, G., Fornage, M., Franke, B., Debette, S., Medland, S.E., Ikram, M.A., Thompson, P.M., 2016. Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nat. Neurosci. 19, 1569–1582. https://doi.org/10.1038/nn.4398

Adhikari, B.M., Jahanshad, N., Shukla, D., Glahn, D.C., Blangero, J., Fox, P.T., Reynolds, R.C., Cox, R.W., Fieremans, E., Veraart, J., Novikov, D.S., Nichols, T.E., Hong, L.E., Thompson, P.M., Kochunov, P., 2018. Comparison of heritability estimates on resting state fMRI connectivity phenotypes using the ENIGMA analysis pipeline. Hum. Brain Mapp. https://doi.org/10.1002/hbm.24331

Anderson, J.S., Ferguson, M.A., Lopez-Larson, M., Yurgelun-Todd, D., 2011. Reproducibility of single-subject functional connectivity measurements. Am. J. Neuroradiol. 32, 548–555. https://doi.org/10.3174/ajnr.A2330

Arfanakis, K., Cordes, D., Haughton, V.M., Moritz, C.H., Quigley, M.A., Meyerand, M.E., 2000. Combining independent component analysis and correlation analysis to probe interregional connectivity in fMRI task activation datasets. Magn. Reson. Imaging 18, 921–930. https://doi.org/10.1016/S0730-725X(00)00190-9

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12, 26–41. https://doi.org/10.1016/j.media.2007.06.004

Barch, D.M., Burgess, G.C., Harms, M.P., Petersen, S.E., Schlaggar, B.L., Corbetta, M., Glasser, M.F., Curtiss, S., Dixit, S., Feldt, C., Nolan, D., Bryant, E., Hartley, T., Footer, O., Bjork, J.M., Poldrack, R., Smith, S., Johansen-Berg, H., Snyder, A.Z., Van Essen, D.C., 2013. Function in the human connectome: Task-fMRI and individual differences in behavior. Neuroimage 80, 169–189. https://doi.org/10.1016/j.neuroimage.2013.05.033

Behzadi, Y., Restom, K., Liau, J., Liu, T.T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37, 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. https://doi.org/10.2307/2346101

Betti, V., DellaPenna, S., de Pasquale, F., Mantini, D., Marzetti, L., Romani, G.L., Corbetta, M., 2013. Natural scenes viewing alters the dynamics of functional connectivity in the human brain. Neuron 79, 782–797. https://doi.org/10.1016/j.neuron.2013.06.022

Birn, R.M., Molloy, E.K., Patriat, R., Parker, T., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., 2013. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. Neuroimage 83, 550–558. https://doi.org/10.1016/j.neuroimage.2013.05.099

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T., Mehta, P., Fox, J., 2011. OpenMx: An Open Source Extended Structural Equation Modeling Framework. Psychometrika 76, 306–317. https://doi.org/10.1007/s11336-010-9200-6

Braga, R.M., Buckner, R.L., 2017. Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. Neuron 95, 457–471.e5. https://doi.org/10.1016/j.neuron.2017.06.038

Braskie, M.N., Ringman, J.M., Thompson, P.M., 2011. Neuroimaging measures as

endophenotypes in Alzheimer's disease. Int. J. Alzheimers. Dis. https://doi.org/10.4061/2011/490140

Bright, M.G., Tench, C.R., Murphy, K., 2017. Potential pitfalls when denoising resting state fMRI data using nuisance regression. Neuroimage 154, 159–168. https://doi.org/10.1016/j.neuroimage.2016.12.027

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The Brain's Default Network. Ann. N. Y. Acad. Sci. 1124, 1–38. https://doi.org/10.1196/annals.1440.011

Buckner, R.L., Krienen, F.M., Yeo, B.T.T., 2013. Opportunities and limitations of intrinsic functional connectivity MRI. Nat. Neurosci. https://doi.org/10.1038/nn.3423

Burgess, G.C., Kandala, S., Nolan, D., Laumann, T.O., Power, J.D., Adeyemo, B., Harms, M.P., Petersen, S.E., Barch, D.M., 2016. Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data from the Human Connectome Project. Brain Connect. 6, 669–680. https://doi.org/10.1089/brain.2016.0435

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013. Power failure: Why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365–376. https://doi.org/10.1038/nrn3475

Bzdok, D., Varoquaux, G., Grisel, O., Eickenberg, M., Poupon, C., Thirion, B., 2016. Formal Models of the Network Co-occurrence Underlying Mental Operations. PLoS Comput. Biol. 12. https://doi.org/10.1371/journal.pcbi.1004994

Bzdok, D., Yeo, B.T.T., 2017. Inference in the age of big data: Future perspectives on neuroscience. Neuroimage 155, 549–564. https://doi.org/10.1016/j.neuroimage.2017.04.061

Caballero-Gaudes, C., Reynolds, R.C., 2017. Methods for cleaning the BOLD fMRI signal. Neuroimage 154, 128–149. https://doi.org/10.1016/j.neuroimage.2016.12.018

Caria, A., Sitaram, R., Birbaumer, N., 2012. Real-time fMRI: A tool for local brain regulation. Neuroscientist. https://doi.org/10.1177/1073858411407205

Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C.A., Wager, T.D., Banich, M.T., Speer, N.K., Sutherland, M.T., Riedel, M.C., Dick, A.S., Bjork, J.M., Thomas, K.M., Chaarani, B., Mejia, M.H., Hagler, D.J., Daniela Cornejo, M., Sicat, C.S., Harms, M.P., Dosenbach, N.U.F., Rosenberg, M., Earl, E., Bartsch, H., Watts, R., Polimeni, J.R., Kuperman, J.M., Fair, D.A., Dale, A.M., 2018. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev. Cogn. Neurosci. https://doi.org/10.1016/j.dcn.2018.03.001

Chen, G., Taylor, P.A., Haller, S.P., Kircanski, K., Stoddard, J., Pine, D.S., Leibenluft, E., Brotman, M.A., Cox, R.W., 2018. Intraclass correlation: Improved modeling approaches and applications for neuroimaging. Hum. Brain Mapp. 39, 1187–1206. https://doi.org/10.1002/hbm.23909

Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., Schooler, J.W., 2009. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proc. Natl. Acad. Sci. 106, 8719–8724. https://doi.org/10.1073/pnas.0900234106

Cicchetti, D. V., 1994. Guidlines, Criteria, and Rules of Thumb for Evalauting Normed and Standardized Assessment Instruments in Psychology. Psychol. Assess. 6, 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cole, M.W., Anticevic, A., Repovs, G., Barch, D., 2011. Variable global dysconnectivity and individual differences in schizophrenia. Biol. Psychiatry 70, 43–50. https://doi.org/10.1016/j.biopsych.2011.02.010

1054 Cole, M.W., Bassett, D.S., Power, J.D., Braver, T.S., Petersen, S.E., 2014. Intrinsic and task-
1055 evoked network architectures of the human brain. Neuron 83, 238–251.
1056 https://doi.org/10.1016/j.neuron.2014.05.014
1057 Cole, M.W., Ito, T., Bassett, D.S., Schultz, D.H., 2016. Activity flow over resting-state networks
1058 shapes cognitive task activations. Nat. Neurosci. 19, 1718–1726.
1059 https://doi.org/10.1038/nn.4406
1060 Collins, F.S., Varmus, H., 2015. A New Initiative on Precision Medicine. N. Engl. J. Med. 372,
1061 793–795. https://doi.org/10.1056/NEJMp1500523
1062 Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic
1063 resonance neuroimages. Comput. Biomed. Res. 29, 162–73.
1064 https://doi.org/10.1006/cbmr.1996.0014
1065 Cremers, H.R., Wager, T.D., Yarkoni, T., 2017. The relation between statistical power and
1066 inference in fMRI 1–20. https://doi.org/10.1371/journal.pone.0184923
1067 Deco, G., Hagmann, P., Hudetz, A.G., Tononi, G., 2014. Modeling resting-state functional
1068 networks when the cortex falls asleep: Local and global changes. Cereb. Cortex 24, 3180–
1069 3194. https://doi.org/10.1093/cercor/bht176
1070 Dong, D., Wang, Y., Chang, X., Luo, C., Yao, D., 2018. Dysfunction of Large-Scale Brain
1071 Networks in Schizophrenia: A Meta-analysis of Resting-State Functional Connectivity.
1072 Schizophr. Bull. 44, 168–181. https://doi.org/10.1093/schbul/sbx034
1073 Dubois, J., Adolphs, R., 2016. Building a Science of Individual Differences from fMRI. Trends
1074 Cogn. Sci. 20, 425–443. https://doi.org/10.1016/j.tics.2016.03.014
1075 Dubois, J., Galdi, P., Han, Y., Paul, L.K., Adolphs, R., Angeles, L., Systems, I., Dubois, J., 2018.
1076 Resting-state functional brain connectivity best predicts the personality dimension of
1077 openness to experience. bioRxiv 215129. https://doi.org/10.1101/215129
1078 Elliott, L., Sharp, K., Alfaro-Almagro, F., Douaud, G., Miller, K., Marchini, J., Smith, S., 2017.
1079 The genetic basis of human brain structure and function: 1,262 genome-wide associations
1080 found from 3,144 GWAS of multimodal brain imaging phenotypes from 9,707 UK Biobank
1081 participants. bioRxiv.
1082 Elliott, M.L., Romer, A., Knodt, A.R., Hariri, A.R., 2018. A Connectome-wide Functional
1083 Signature of Transdiagnostic Risk for Mental Illness. Biol. Psychiatry 84, 452–459.
1084 https://doi.org/10.1016/j.biopsych.2018.03.012
1085 Errington, T.M., Iorns, E., Gunn, W., Tan, F.E. lisabet., Lomax, J., Nosek, B.A., 2014. An open
1086 investigation of the reproducibility of cancer biology research. Elife 3.
1087 https://doi.org/10.7554/eLife.04333
1088 Fair, D.A., Schlaggar, B.L., Cohen, A.L., Miezin, F.M., Dosenbach, N.U.F., Wenger, K.K., Fox,
1089 M.D., Snyder, A.Z., Raichle, M.E., Petersen, S.E., 2007. A method for using blocked and
1090 event-related fMRI data to study "resting state" functional connectivity. Neuroimage 35,
1091 396–405. https://doi.org/10.1016/j.neuroimage.2006.11.051
1092 Finn, E.S., Scheinost, D., Finn, D.M., Shen, X., Papademetris, X., Constable, R.T., 2017. Can
1093 brain state be manipulated to emphasize individual differences in functional connectivity?
1094 Neuroimage 160, 140–151. https://doi.org/10.1016/j.neuroimage.2017.03.064
1095 Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X.,
1096 Constable, R.T., 2015. Functional connectome fingerprinting: Identifying individuals using
1097 patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671.
1098 https://doi.org/10.1038/nn.4135
1099 Fox, M.D., 2010. Clinical applications of resting state functional connectivity. Front. Syst.
1100 Neurosci. 4. https://doi.org/10.3389/fnsys.2010.00019
1101 Fox, M.D., Buckner, R.L., Liu, H., Chakravarty, M.M., Lozano, A.M., Pascual-Leone, A., 2014.

Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.1405003111

Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., 2007. Intrinsic Fluctuations within Cortical Systems Account for Intertrial Variability in Human Behavior. Neuron 56, 171–184. https://doi.org/10.1016/j.neuron.2007.08.023

Fox, M.D., Snyder, A.Z., Zacks, J.M., Raichle, M.E., 2006. Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. Nat. Neurosci. 9, 23–25. https://doi.org/10.1038/nn1616

Ge, T., Holmes, A.J., Buckner, R.L., Smoller, J.W., Sabuncu, M.R., 2017. Heritability analysis with repeat measurements and its application to resting-state functional connectivity. Proc. Natl. Acad. Sci. 114, 5521–5526. https://doi.org/10.1073/pnas.1700765114

Geerligs, L., Rubinov, M., Cam-CAN, Henson, R.N., 2015. State and Trait Components of Functional Connectivity: Individual Differences Vary with Mental State. J. Neurosci. 35, 13949–13961. https://doi.org/10.1523/JNEUROSCI.1324-15.2015

Glahn, D.C., Winkler, A.M., Kochunov, P., Almasy, L., Duggirala, R., Carless, M.A., Curran, J.C., Olvera, R.L., Laird, A.R., Smith, S.M., Beckmann, C.F., Fox, P.T., Blangero, J., 2010. Genetic control over the resting brain. Proc Natl Acad Sci U S A 107, 1223–1228. https://doi.org/10.1073/pnas.0909969107

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J.M., Coalson, R.S., Nguyen, A.L., McDermott, K.B., Shimony, J.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Nelson, S.M., Dosenbach, N.U.F., 2017. Precision Functional Mapping of Individual Human Brains. Neuron 95, 791–807.e7. https://doi.org/10.1016/j.neuron.2017.07.011

Gratton, C., Laumann, T.O., Nielsen, A.N., Greene, D.J., Gordon, E.M., Gilmore, A.W., Nelson, S.M., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., Dosenbach, N.U.F., Petersen, S.E., 2018. Functional Brain Networks Are Dominated by Stable Group and Individual Factors, Not Cognitive or Daily Variation. Neuron 439–452. https://doi.org/10.1016/j.neuron.2018.03.035

Greene, A.S., Gao, S., Scheinost, D., Constable, R.T., 2018. Task-induced brain state manipulation improves prediction of individual traits. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-04920-3

Greicius, M., 2008. Resting-state functional connectivity in neuropsychiatric disorders. Curr. Opin. Neurol. 21, 424–30. https://doi.org/10.1097/WCO.0b013e328306f2c5

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48, 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Guo, C.C., Kurth, F., Zhou, J., Mayer, E.A., Eickhoff, S.B., Kramer, J.H., Seeley, W.W., 2012. One-year test-retest reliability of intrinsic connectivity network fMRI in older adults. Neuroimage 61, 1471–1483. https://doi.org/10.1016/j.neuroimage.2012.03.027

Hacker, C.D., Laumann, T.O., Szrama, N.P., Baldassarre, A., Snyder, A.Z., Leuthardt, E.C., Corbetta, M., 2013. Resting state network estimation in individual subjects. Neuroimage 82, 616–633. https://doi.org/10.1016/j.neuroimage.2013.05.108

Hamburg, M.A., Collins, F.S., 2010. The Path to Personalized Medicine - Perspective. N. Engl. J. Med. 363, 301–304. https://doi.org/10.1056/NEJMp1006304

Hariri, A.R., Weinberger, D.R., 2003. Imaging genomics. Br. Med. Bull. https://doi.org/10.1093/bmb/65.1.259

Hasler, G., Northoff, G., 2011. Discovering imaging endophenotypes for major depression. Mol. Psychiatry. https://doi.org/10.1038/mp.2011.23

Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. Trends Cogn. Sci. https://doi.org/10.1016/j.tics.2009.10.011

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject Synchronization of Cortical Activity during Natural Vision. Science (80-.). 303, 1634–1640. https://doi.org/10.1126/science.1089506

Hibar, D.P., Adams, H.H.H., Jahanshad, N., Chauhan, G., Stein, J.L., Hofer, E., Renteria, M.E., Bis, J.C., Arias-Vasquez, A., Ikram, M.K., Desrivières, S., Vernooij, M.W., Abramovic, L., Alhusaini, S., Amin, N., Andersson, M., Arfanakis, K., Aribisala, B.S., Armstrong, N.J., Athanasiu, L., Axelsson, T., Beecham, A.H., Beiser, A., Bernard, M., Blanton, S.H., Bohlken, M.M., Boks, M.P., Bralten, J., Brickman, A.M., Carmichael, O., Chakravarty, M.M., Chen, Q., Ching, C.R.K., Chouraki, V., Cuellar-Partida, G., Crivello, F., Den Braber, A., Doan, N.T., Ehrlich, S., Giddaluru, S., Goldman, A.L., Gottesman, R.F., Grimm, O., Griswold, M.E., Guadalupe, T., Gutman, B.A., Hass, J., Haukvik, U.K., Hoehn, D., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Jørgensen, K.N., Karbalai, N., Kasperaviciute, D., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Liewald, D.C.M., Lopez, L.M., Luciano, M., MacAre, C., Marquand, A.F., Matarin, M., Mather, K.A., Mattheisen, M., McKay, D.R., Milaneschi, Y., Muñoz Maniega, S., Nho, K., Nugent, A.C., Nyquist, P., Loohuis, L.M.O., Oosterlaan, J., Papmeyer, M., Pirpamer, L., Pütz, B., Ramasamy, A., Richards, J.S., Risacher, S.L., Roiz-Santiañez, R., Rommelse, N., Ropele, S., Rose, E.J., Royle, N.A., Rundek, T., Sämann, P.G., Saremi, A., Satizabal, C.L., Schmaal, L., Schork, A.J., Shen, L., Shin, J., Shumskaya, E., Smith, A. V., Sprooten, E., Strike, L.T., Teumer, A., Tordesillas-Gutierrez, D., Toro, R., Trabzuni, D., Trompet, S., Vaidya, D., Van Der Grond, J., Van Der Lee, S.J., Van Der Meer, D., Van Donkelaar, M.M.J., Van Eijk, K.R., Van Erp, T.G.M., Van Rooij, D., Walton, E., Westlye, L.T., Whelan, C.D., Windham, B.G., Winkler, A.M., Wittfeld, K., Woldehawariat, G., Wolf, C., Wolfers, T., Yanek, L.R., Yang, J., Zijdenbos, A., Zwiers, M.P., Agartz, I., Almasy, L., Ames, D., Amouyel, P., Andreassen, O.A., Arepalli, S., Assareh, A.A., Barral, S., Bastin, M.E., Becker, D.M., Becker, J.T., Bennett, D.A., Blangero, J., Van Bokhoven, H., Boomsma, D.I., Brodaty, H., Brouwer, R.M., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Bulayeva, K.B., Cahn, W., Calhoun, V.D., Cannon, D.M., Cavalleri, G.L., Cheng, C.Y., Cichon, S., Cookson, M.R., Corvin, A., Crespo-Facorro, B., Curran, J.E., Czisch, M., Dale, A.M., Davies, G.E., De Craen, A.J.M., De Geus, E.J.C., De Jager, P.L., De Zubicaray, G.I., Deary, I.J., Debette, S., Decarli, C., Delanty, N., Depondt, C., Destefano, A., Dillman, A., Djurovic, S., Donohoe, G., Drevets, W.C., Duggirala, R., Dyer, T.D., Enzinger, C., Erk, S., Espeseth, T., Fedko, I.O., Fernández, G., Ferrucci, L., Fisher, S.E., Fleischman, D.A., Ford, I., Fornage, M., Foroud, T.M., Fox, P.T., Francks, C., Fukunaga, M., Gibbs, J.R., Glahn, D.C., Gollub, R.L., Göring, H.H.H., Green, R.C., Gruber, O., Gudnason, V., Guelfi, S., Håberg, A.K., Hansell, N.K., Hardy, J., Hartman, C.A., Hashimoto, R., Hegenscheid, K., Heinz, A., Le Hellard, S., Hernandez, D.G., Heslenfeld, D.J., Ho, B.C., Hoekstra, P.J., Hoffmann, W., Hofman, A., Holsboer, F., Homuth, G., Hosten, N., Hottenga, J.J., Huentelman, M., Pol, H.E.H., Ikeda, M., Jack, C.R., Jenkinson, M., Johnson, R., Jönsson, E.G., Jukema, J.W., Kahn, R.S., Kanai, R., Kloszewska, I., Knopman, D.S., Kochunov, P., Kwok, J.B., Lawrie, S.M., Lemaître, H., Liu, X., Longo, D.L., Lopez, O.L., Lovestone, S., Martinez, O., Martinot, J.L., Mattay, V.S., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Mecocci, P., Melle,

I., Meyer-Lindenberg, A., Mohnke, S., Montgomery, G.W., Morris, D.W., Mosley, T.H., Mühleisen, T.W., Müller-Myhsok, B., Nalls, M.A., Nauck, M., Nichols, T.E., Niessen, W.J., Nöthen, M.M., Nyberg, L., Ohi, K., Olvera, R.L., Ophoff, R.A., Pandolfo, M., Paus, T., Pausova, Z., Penninx, B.W.J.H., Pike, G.B., Potkin, S.G., Psaty, B.M., Reppermund, S., Rietschel, M., Roffman, J.L., Romanczuk-Seiferth, N., Rotter, J.I., Ryten, M., Sacco, R.L., Sachdev, P.S., Saykin, A.J., Schmidt, R., Schmidt, H., Schofield, P.R., Sigursson, S., Simmons, A., Singleton, A., Sisodiya, S.M., Smith, C., Smoller, J.W., Soininen, H., Steen, V.M., Stott, D.J., Sussmann, J.E., Thalamuthu, A., Toga, A.W., Traynor, B.J., Troncoso, J., Tsolaki, M., Tzourio, C., Uitterlinden, A.G., Hernández, M.C.V., Van Der Brug, M., Van Der Lugt, A., Van Der Wee, N.J.A., Van Haren, N.E.M., Van 'T Ent, D., Van Tol, M.J., Vardarajan, B.N., Vellas, B., Veltman, D.J., Völzke, H., Walter, H., Wardlaw, J.M., Wassink, T.H., Weale, M.E., Weinberger, D.R., Weiner, M.W., Wen, W., Westman, E., White, T., Wong, T.Y., Wright, C.B., Zielke, R.H., Zonderman, A.B., Martin, N.G., Van Duijn, C.M., Wright, M.J., Longstreth, W.T., Schumann, G., Grabe, H.J., Franke, B., Launer, L.J., Medland, S.E., Seshadri, S., Thompson, P.M., Ikram, M.A., 2017. Novel genetic loci associated with hippocampal volume. Nat. Commun. 8. https://doi.org/10.1038/ncomms13624

Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivières, S., Jahanshad, N., Toro, R., Wittfeld, K., Abramovic, L., Andersson, M., Aribisala, B.S., Armstrong, N.J., Bernard, M., Bohlken, M.M., Boks, M.P., Bralten, J., Brown, A.A., Mallar Chakravarty, M., Chen, Q., Ching, C.R.K., Cuellar-Partida, G., Den Braber, A., Giddaluru, S., Goldman, A.L., Grimm, O., Guadalupe, T., Hass, J., Woldehawariat, G., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Olde Loohuis, L.M., Luciano, M., MacAre, C., Mather, K.A., Mattheisen, M., Milaneschi, Y., Nho, K., Papmeyer, M., Ramasamy, A., Risacher, S.L., Roiz-Santiañez, R., Rose, E.J., Salami, A., Sämann, P.G., Schmaal, L., Schork, A.J., Shin, J., Strike, L.T., Teumer, A., Van Donkelaar, M.M.J., Van Eijk, K.R., Walters, R.K., Westlye, L.T., Whelan, C.D., Winkler, A.M., Zwiers, M.P., Alhusaini, S., Athanasiu, L., Ehrlich, S., Hakobjan, M.M.H., Hartberg, C.B., Haukvik, U.K., Heister, A.J.G.A.M., Hoehn, D., Kasperaviciute, D., Liewald, D.C.M., Lopez, L.M., Makkinje, R.R.R., Matarin, M., Naber, M.A.M., Reese McKay, D., Needham, M., Nugent, A.C., Pütz, B., Royle, N.A., Shen, L., Sprooten, E., Trabzuni, D., Van Der Marel, S.S.L., Van Hulzen, K.J.E., Walton, E., Wolf, C., Almasy, L., Ames, D., Arepalli, S., Assareh, A.A., Bastin, M.E., Brodaty, H., Bulayeva, K.B., Carless, M.A., Cichon, S., Corvin, A., Curran, J.E., Czisch, M., De Zubicaray, G.I., Dillman, A., Duggirala, R., Dyer, T.D., Erk, S., Fedko, I.O., Ferrucci, L., Foroud, T.M., Fox, P.T., Fukunaga, M., Raphael Gibbs, J., Göring, H.H.H., Green, R.C., Guelfi, S., Hansell, N.K., Hartman, C.A., Hegenscheid, K., Heinz, A., Hernandez, D.G., Heslenfeld, D.J., Hoekstra, P.J., Holsboer, F., Homuth, G., Hottenga, J.J., Ikeda, M., Jack, C.R., Jenkinson, M., Johnson, R., Kanai, R., Keil, M., Kent, J.W., Kochunov, P., Kwok, J.B., Lawrie, S.M., Liu, X., Longo, D.L., McMahon, K.L., Meisenzahl, E., Melle, I., Mohnke, S., Montgomery, G.W., Mostert, J.C., Mühleisen, T.W., Nalls, M.A., Nichols, T.E., Nilsson, L.G., Nöthen, M.M., Ohi, K., Olvera, R.L., Perez-Iglesias, R., Bruce Pike, G., Potkin, S.G., Reinvang, I., Reppermund, S., Rietschel, M., Romanczuk-Seiferth, N., Rosen, G.D., Rujescu, D., Schnell, K., Schofield, P.R., Smith, C., Steen, V.M., Sussmann, J.E., Thalamuthu, A., Toga, A.W., Traynor, B.J., Troncoso, J., Turner, J.A., Valdés Hernández, M.C., Van 'T Ent, D., Van Der Brug, M., Van Der Wee, N.J.A., Van Tol, M.J., Veltman, D.J., Wassink, T.H., Westman, E., Zielke, R.H., Zonderman, A.B., Ashbrook, D.G., Hager, R., Lu, L., McMahon, F.J., Morris, D.W., Williams, R.W., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Cahn, W., Calhoun, V.D.,

Cavalleri, G.L., Crespo-Facorro, B., Dale, A.M., Davies, G.E., Delanty, N., Depondt, C., Djurovic, S., Drevets, W.C., Espeseth, T., Gollub, R.L., Ho, B.C., Hoffmann, W., Hosten, N., Kahn, R.S., Le Hellard, S., Meyer-Lindenberg, A., Müller-Myhsok, B., Nauck, M., Nyberg, L., Pandolfo, M., Penninx, B.W.J.H., Roffman, J.L., Sisodiya, S.M., Smoller, J.W., Van Bokhoven, H., Van Haren, N.E.M., Völzke, H., Walter, H., Weiner, M.W., Wen, W., White, T., Agartz, I., Andreassen, O.A., Blangero, J., Boomsma, D.I., Brouwer, R.M., Cannon, D.M., Cookson, M.R., De Geus, E.J.C., Deary, I.J., Donohoe, G., Fernández, G., Fisher, S.E., Francks, C., Glahn, D.C., Grabe, H.J., Gruber, O., Hardy, J., Hashimoto, R., Hulshoff Pol, H.E., Jönsson, E.G., Kloszewska, I., Lovestone, S., Mattay, V.S., Mecocci, P., McDonald, C., McIntosh, A.M., Ophoff, R.A., Paus, T., Pausova, Z., Ryten, M., Sachdev, P.S., Saykin, A.J., Simmons, A., Singleton, A., Soininen, H., Wardlaw, J.M., Weale, M.E., Weinberger, D.R., Adams, H.H.H., Launer, L.J., Seiler, S., Schmidt, R., Chauhan, G., Satizabal, C.L., Becker, J.T., Yanek, L., Van Der Lee, S.J., Ebling, M., Fischl, B., Longstreth, W.T., Greve, D., Schmidt, H., Nyquist, P., Vinke, L.N., Van Duijn, C.M., Xue, L., Mazoyer, B., Bis, J.C., Gudnason, V., Seshadri, S., Ikram, M.A., Martin, N.G., Wright, M.J., Schumann, G., Franke, B., Thompson, P.M., Medland, S.E., 2015. Common genetic variants influence human subcortical brain structures. Nature 520, 224–229. https://doi.org/10.1038/nature14101

Honey, C.J., Honey, C.J., Sporns, O., Sporns, O., Cammoun, L., Cammoun, L., Gigandet, X., Gigandet, X., Thiran, J.P., Thiran, J.P., Meuli, R., Meuli, R., Hagmann, P., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc. Natl. Acad. Sci. U. S. A. 106, 2035–40. https://doi.org/10.1073/pnas.0811168106

Hurlburt, R.T., Alderson-Day, B., Fernyhough, C., Kï¿½hn, S., 2015. What goes on in the resting-state? A qualitative glimpse into resting-state experience in the scanner. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.01535

Huth, A.G., De Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532, 453–458. https://doi.org/10.1038/nature17637

Ioannidis, J.P.A., 2008. Why Most Discovered True Associations Are Inflated. Epidemiology 19, 640–648. https://doi.org/10.1097/EDE.0b013e31818131e7

Ioannidis, J.P.A., 2005. Why most published research findings are false. PLoS Med. https://doi.org/10.1371/journal.pmed.0020124

Ioannidis, J.P.A., Ntzani, E.E., Trikalinos, T.A., Contopoulos-Ioannidis, D.G., 2001. Replication validity of genetic association studies. Nat. Genet. 29, 306–309. https://doi.org/10.1038/ng749

Jo, H.J., Gotts, S.J., Reynolds, R.C., Bandettini, P.A., Martin, A., Cox, R.W., Saad, Z.S., 2013. Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state FMRI. J. Appl. Math. 2013. https://doi.org/10.1155/2013/935154

Kaiser, R.H., Andrews-Hanna, J.R., Wager, T.D., Pizzagalli, D.A., 2015. Large-Scale Network Dysfunction in Major Depressive Disorder A Meta-analysis of Resting-State Functional Connectivity. JAMA psychiatry 72, 603–611. https://doi.org/10.1001/jamapsychiatry.2015.0071

Kanyongo, G.Y., Brook, G.P., Kyei-Blankson, L., Gocmen, G., 2007. Reliability and Statistical Power: How Measurement Fallibility Affects Power and Required Sample Sizes for Several Parametric and Nonparametric Statistics. J. Mod. Appl. Stat. Methods 6, 81–90. https://doi.org/10.22237/jmasm/1177992480

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D.,

Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R. V., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 46, 786–802. https://doi.org/10.1016/j.neuroimage.2008.12.037

Krienen, F.M., Yeo, B.T.T., Buckner, R.L., 2014. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 369, 20130526-. https://doi.org/10.1098/rstb.2013.0526

Lahnakoski, J.M., Glerean, E., Salmi, J., Jääskeläinen, I.P., Sams, M., Hari, R., Nummenmaa, L., 2012. Naturalistic fMRI Mapping Reveals Superior Temporal Sulcus as the Hub for the Distributed Brain Network for Social Perception. Front. Hum. Neurosci. 6. https://doi.org/10.3389/fnhum.2012.00233

Laumann, T.O., Gordon, E.M., Adeyemo, B., Snyder, A.Z., Joo, S.J., Chen, M.Y., Gilmore, A.W., McDermott, K.B., Nelson, S.M., Dosenbach, N.U.F., Schlaggar, B.L., Mumford, J.A., Poldrack, R.A., Petersen, S.E., 2015. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. Neuron 87, 658–671. https://doi.org/10.1016/j.neuron.2015.06.037

Leech, R., Braga, R., Sharp, D.J., 2012. Echoes of the Brain within the Posterior Cingulate Cortex. J. Neurosci. 32, 215–222. https://doi.org/10.1523/JNEUROSCI.3689-11.2012

Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., Bezgin, G., Eickhoff, S.B., Castellanos, F.X., Petrides, M., Jefferies, E., Smallwood, J., 2016. Situating the default-mode network along a principal gradient of macroscale cortical organization. Proc. Natl. Acad. Sci. U. S. A. 113, 12574–12579. https://doi.org/10.1073/pnas.1608282113

Matthews, P.M., Honey, G.D., Bullmore, E.T., 2006. Applications of fMRI in translational medicine and clinical practice. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn1929

Mesulam, M.M., 1998. From sensation to cognition. Brain. https://doi.org/10.1093/brain/121.6.1013

Meyer-Lindenberg, A., Weinberger, D.R., 2006. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn1993

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523. https://doi.org/10.1038/nn.4393

Mueller, S., Wang, D., Fox, M.D., Yeo, B.T.T., Sepulcre, J., Sabuncu, M.R., Shafee, R., Lu, J., Liu, H., 2013. Individual Variability in Functional Connectivity Architecture of the Human Brain. Neuron 77, 586–595. https://doi.org/10.1016/j.neuron.2012.12.028

Munafò, M.R., Nosek, B.A., Bishop, D.V.M., Button, K.S., Chambers, C.D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E.J., Ware, J.J., Ioannidis, J.P.A., 2017. A manifesto for reproducible science. Nat. Hum. Behav. 1, 1–9. https://doi.org/10.1038/s41562-016-0021

Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. Neuroimage 154, 169–173. https://doi.org/10.1016/j.neuroimage.2016.11.052

Neale, M.C., Hunter, M.D., Pritkin, J., Zahery, M., Brick, T.R., Kirkpatrick, R.M., Estabrook, R., Bates, T.C., Maes, H.H., Boker, S.M., 2016. OpenMx 2.0: Extended Structural Equation and Statistical Modeling. Psychometrika 81, 535–549. https://doi.org/10.1007/s11336-014-9435-8

Noble, S., Spann, M.N., Tokoglu, F., Shen, X., Constable, R.T., Scheinost, D., 2017. Influences on the Test–Retest Reliability of Functional Connectivity MRI and its Relationship with

Behavioral Utility. Cereb. Cortex 1–15. https://doi.org/10.1093/cercor/bhx230
Nunnally Jr., J.C., 1970. Introduction to psychological measurement., Introduction to psychological measurement. McGraw-Hill, New York, NY, US.
Patriat, R., Molloy, E.K., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., Birn, R.M., 2013. The effect of resting condition on resting-state fMRI reliability and consistency: A comparison between resting with eyes open, closed, and fixated. Neuroimage 78, 463–473. https://doi.org/10.1016/j.neuroimage.2013.04.013
Pek, J., Wu, H., 2015. Profile Likelihood-Based Confidence Intervals and Regions for Structural Equation Models. Psychometrika 80, 1123–1145. https://doi.org/10.1007/s11336-015-9461-1
Poldrack, R.A., Baker, C.I., Durnez, J., Gorgolewski, K.J., Matthews, P.M., Munafò, M.R., Nichols, T.E., Poline, J.B., Vul, E., Yarkoni, T., 2017. Scanning the horizon: Towards transparent and reproducible neuroimaging research. Nat. Rev. Neurosci. 18, 115–126. https://doi.org/10.1038/nrn.2016.167
Posthuma, D., Boomsma, D.I., 2000. A note on the statistical power in extended twin designs. Behav. Genet. 30, 147–158.
Poulton, R., Moffitt, T.E., Silva, P.A., 2015. The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. Soc. Psychiatry Psychiatr. Epidemiol. https://doi.org/10.1007/s00127-015-1048-8
Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018
Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional Network Organization of the Human Brain. Neuron 72, 665–678. https://doi.org/10.1016/j.neuron.2011.09.006
Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048
Power, J.D., Plitt, M., Gotts, S.J., Kundu, P., Voon, V., Bandettini, P.A., Martin, A., 2018. Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. Proc. Natl. Acad. Sci. 115, 201720985. https://doi.org/10.1073/pnas.1720985115
Power, J.D., Plitt, M., Laumann, T.O., Martin, A., 2016. Sources and implications of whole-brain fMRI signals in humans. Neuroimage 146, 0–1. https://doi.org/10.1016/j.neuroimage.2016.09.038
Power, J.D., Schlaggar, B.L., Lessov-Schlaggar, C.N., Petersen, S.E., 2013. Evidence for hubs in human functional brain networks. Neuron 79, 798–813. https://doi.org/10.1016/j.neuron.2013.07.035
Raven, J.C., 1941. Standardization of progressive matrices. Br. J. Med. Psychol. 19, 137–150. https://doi.org/10.1111/j.2044-8341.1941.tb00316.x
Satterthwaite, T.D., Cook, P.A., Bruce, S.E., Conway, C., Mikkelsen, E., Satchell, E., Vandekar, S.N., Durbin, T., Shinohara, R.T., Sheline, Y.I., 2016. Dimensional depression severity in women with major depression and post-traumatic stress disorder correlates with fronto-amygdalar hypoconnectivty. Mol. Psychiatry 21, 894–902. https://doi.org/10.1038/mp.2015.149
Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., Wolf, D.H., 2013. An improved

framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64, 240–256. https://doi.org/10.1016/j.neuroimage.2012.08.052

Satterthwaite, T.D., Xia, C.H., Bassett, D.S., 2018. Personalized Neuroscience: Common and Individual-Specific Features in Functional Brain Networks. Neuron 98, 243–245. https://doi.org/10.1016/j.neuron.2018.04.007

Shah, L.M., Cramer, J.A., Ferguson, M.A., Birn, R.M., Anderson, J.S., 2016. Reliability and reproducibility of individual differences in functional connectivity acquired during task and resting state. Brain Behav. 6, 1–15. https://doi.org/10.1002/brb3.456

Shehzad, Z., Kelly, A.M.C., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X., Milham, M.P., 2009. The resting brain: Unconstrained yet reliable. Cereb. Cortex 19, 2209–2229. https://doi.org/10.1093/cercor/bhn256

Shen, X., Finn, E.S., Scheinost, D., Rosenberg, M.D., Chun, M.M., Papademetris, X., Constable, R.T., 2017. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat. Protoc. 12, 506–518. https://doi.org/10.1038/nprot.2016.178

Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., Kelly, M., Laumann, T., Miller, K.L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A.Z., Vu, A.T., Woolrich, M.W., Xu, J., Yacoub, E., Uğurbil, K., Van Essen, D.C., Glasser, M.F., 2013. Resting-state fMRI in the Human Connectome Project. Neuroimage 80, 144–168. https://doi.org/10.1016/j.neuroimage.2013.05.039

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.0905267106

Steiger, J.H., 1980. Tests for comparing elements of a correlation matrix. Psychol. Bull. 87, 245–251. https://doi.org/10.1037/0033-2909.87.2.245

Stein, J.L., Medland, S.E., Vasquez, A.A., Hibar, D.P., Senstad, R.E., Winkler, A.M., Toro, R., Appel, K., Bartecek, R., Bergmann, Ø., Bernard, M., Brown, A.A., Cannon, D.M., Chakravarty, M.M., Christoforou, A., Domin, M., Grimm, O., Hollinshead, M., Holmes, A.J., Homuth, G., Hottenga, J.J., Langan, C., Lopez, L.M., Hansell, N.K., Hwang, K.S., Kim, S., Laje, G., Lee, P.H., Liu, X., Loth, E., Lourdusamy, A., Mattingsdal, M., Mohnke, S., Maniega, S.M., Nho, K., Nugent, A.C., O'brien, C., Papmeyer, M., Pütz, B., Ramasamy, A., Rasmussen, J., Rijpkema, M., Risacher, S.L., Roddey, J.C., Rose, E.J., Ryten, M., Shen, L., Sprooten, E., Strengman, E., Teumer, A., Trabzuni, D., Turner, J., Van Eijk, K., Van Erp, T.G.M., Van Tol, M.J., Wittfeld, K., Wolf, C., Woudstra, S., Aleman, A., Alhusaini, S., Almasy, L., Binder, E.B., Brohawn, D.G., Cantor, R.M., Carless, M.A., Corvin, A., Czisch, M., Curran, J.E., Davies, G., De Almeida, M.A.A., Delanty, N., Depondt, C., Duggirala, R., Dyer, T.D., Erk, S., Fagerness, J., Fox, P.T., Freimer, N.B., Gill, M., Göring, H.H.H., Hagler, D.J., Hoehn, D., Holsboer, F., Hoogman, M., Hosten, N., Jahanshad, N., Johnson, M.P., Kasperaviciute, D., Kent, J.W., Kochunov, P., Lancaster, J.L., Lawrie, S.M., Liewald, D.C., Mandl, R., Matarin, M., Mattheisen, M., Meisenzahl, E., Melle, I., Moses, E.K., Mühleisen, T.W., Nauck, M., Nöthen, M.M., Olvera, R.L., Pandolfo, M., Pike, G.B., Puls, R., Reinvang, I., Rentería, M.E., Rietschel, M., Roffman, J.L., Royle, N.A., Rujescu, D., Savitz, J., Schnack, H.G., Schnell, K., Seiferth, N., Smith, C., Steen, V.M., Hernández, M.C.V., Van Den Heuvel, M., Van Der Wee, N.J., Van Haren, N.E.M., Veltman, J.A., Völzke, H., Walker, R., Westlye, L.T., Whelan, C.D., Agartz, I., Boomsma, D.I., Cavalleri,

G.L., Dale, A.M., Djurovic, S., Drevets, W.C., Hagoort, P., Hall, J., Heinz, A., Jack, C.R., Foroud, T.M., Le Hellard, S., Macciardi, F., Montgomery, G.W., Poline, J.B., Porteous, D.J., Sisodiya, S.M., Starr, J.M., Sussmann, J., Toga, A.W., Veltman, D.J., Walter, H., Weiner, M.W., Bis, J.C., Ikram, M.A., Smith, A. V., Gudnason, V., Tzourio, C., Vernooij, M.W., Launer, L.J., Decarli, C., Seshadri, S., Andreassen, O.A., Apostolova, L.G., Bastin, M.E., Blangero, J., Brunner, H.G., Buckner, R.L., Cichon, S., Coppola, G., De Zubicaray, G.I., Deary, I.J., Donohoe, G., De Geus, E.J.C., Espeseth, T., Ferréndez, G., Glahn, D.C., Grabe, H.J., Hardy, J., Hulshoff Pol, H.E., Jenkinson, M., Kahn, R.S., Mcdonald, C., Mcintosh, A.M., Mcmahon, F.J., Mcmahon, K.L., Meyer-Lindenberg, A., Morris, D.W., Müller-Myhsok, B., Nichols, T.E., Ophoff, R.A., Paus, T., Pausova, Z., Penninx, B.W., Potkin, S.G., Sämann, P.G., Saykin, A.J., Schumann, G., Smoller, J.W., Wardlaw, J.M., Weale, M.E., Martin, N.G., Franke, B., Wright, M.J., Thompson, P.M., 2012. Identification of common variants associated with human hippocampal and intracranial volumes. Nat. Genet. 44, 552–561. https://doi.org/10.1038/ng.2250

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R., 2015. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Med. 12. https://doi.org/10.1371/journal.pmed.1001779

Swartz, J.R., Knodt, A.R., Radtke, S.R., Hariri, A.R., 2015. A neural biomarker of psychological vulnerability to future life stress. Neuron 85, 505–511. https://doi.org/10.1016/j.neuron.2014.12.055

Szucs, D., Ioannidis, J.P.A., 2017. Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS Biol. 15. https://doi.org/10.1371/journal.pbio.2000797

Tagliazucchi, E., Laufs, H., 2014. Decoding Wakefulness Levels from Typical fMRI Resting-State Data Reveals Reliable Drifts between Wakefulness and Sleep. Neuron 82, 695–708. https://doi.org/10.1016/j.neuron.2014.03.020

Tavor, I., Parker Jones, O., Mars, R.B., Smith, S.M., Behrens, T.E., Jbabdi, S., 2016. Task-free MRI predicts individual differences in brain activity during task performance. Science 352, 216–20. https://doi.org/10.1126/science.aad8127

Thompson, P.M., Stein, J.L., Medland, S.E., Hibar, D.P., Vasquez, A.A., Renteria, M.E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., Wright, M.J., Martin, N.G., Agartz, I., Alda, M., Alhusaini, S., Almasy, L., Almeida, J., Alpert, K., Andreasen, N.C., Andreassen, O.A., Apostolova, L.G., Appel, K., Armstrong, N.J., Aribisala, B., Bastin, M.E., Bauer, M., Bearden, C.E., Bergmann, Ø., Binder, E.B., Blangero, J., Bockholt, H.J., Bøen, E., Bois, C., Boomsma, D.I., Booth, T., Bowman, I.J., Bralten, J., Brouwer, R.M., Brunner, H.G., Brohawn, D.G., Buckner, R.L., Buitelaar, J., Bulayeva, K., Bustillo, J.R., Calhoun, V.D., Cannon, D.M., Cantor, R.M., Carless, M.A., Caseras, X., Cavalleri, G.L., Chakravarty, M.M., Chang, K.D., Ching, C.R.K., Christoforou, A., Cichon, S., Clark, V.P., Conrod, P., Coppola, G., Crespo-Facorro, B., Curran, J.E., Czisch, M., Deary, I.J., de Geus, E.J.C., den Braber, A., Delvecchio, G., Depondt, C., de Haan, L., de Zubicaray, G.I., Dima, D., Dimitrova, R., Djurovic, S., Dong, H., Donohoe, G., Duggirala, R., Dyer, T.D., Ehrlich, S., Ekman, C.J., Elvsåshagen, T., Emsell, L., Erk, S., Espeseth, T., Fagerness, J., Fears, S., Fedko, I., Fernández, G., Fisher, S.E., Foroud, T., Fox, P.T., Francks, C., Frangou, S., Frey, E.M., Frodl, T., Frouin, V., Garavan, H., Giddaluru, S., Glahn, D.C., Godlewska, B., Goldstein, R.Z., Gollub, R.L., Grabe, H.J., Grimm, O., Gruber, O., Guadalupe, T., Gur, R.E., Gur, R.C., Göring, H.H.H., Hagenaars, S., Hajek, T., Hall, G.B., Hall, J., Hardy, J.,

Hartman, C.A., Hass, J., Hatton, S.N., Haukvik, U.K., Hegenscheid, K., Heinz, A., Hickie, I.B., Ho, B.C., Hoehn, D., Hoekstra, P.J., Hollinshead, M., Holmes, A.J., Homuth, G., Hoogman, M., Hong, L.E., Hosten, N., Hottenga, J.J., Hulshoff Pol, H.E., Hwang, K.S., Jack, C.R., Jenkinson, M., Johnston, C., Jönsson, E.G., Kahn, R.S., Kasperaviciute, D., Kelly, S., Kim, S., Kochunov, P., Koenders, L., Krämer, B., Kwok, J.B.J., Lagopoulos, J., Laje, G., Landen, M., Landman, B.A., Lauriello, J., Lawrie, S.M., Lee, P.H., Le Hellard, S., Lemaître, H., Leonardo, C.D., Li, C. shan, Liberg, B., Liewald, D.C., Liu, X., Lopez, L.M., Loth, E., Lourdusamy, A., Luciano, M., Macciardi, F., Machielsen, M.W.J., MacQueen, G.M., Malt, U.F., Mandl, R., Manoach, D.S., Martinot, J.L., Matarin, M., Mather, K.A., Mattheisen, M., Mattingsdal, M., Meyer-Lindenberg, A., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Meisenzahl, E., Melle, I., Milaneschi, Y., Mohnke, S., Montgomery, G.W., Morris, D.W., Moses, E.K., Mueller, B.A., Muñoz Maniega, S., Mühleisen, T.W., Müller-Myhsok, B., Mwangi, B., Nauck, M., Nho, K., Nichols, T.E., Nilsson, L.G., Nugent, A.C., Nyberg, L., Olvera, R.L., Oosterlaan, J., Ophoff, R.A., Pandolfo, M., Papalampropoulou-Tsiridou, M., Papmeyer, M., Paus, T., Pausova, Z., Pearlson, G.D., Penninx, B.W., Peterson, C.P., Pfennig, A., Phillips, M., Pike, G.B., Poline, J.B., Potkin, S.G., Pütz, B., Ramasamy, A., Rasmussen, J., Rietschel, M., Rijpkema, M., Risacher, S.L., Roffman, J.L., Roiz-Santiañez, R., Romanczuk-Seiferth, N., Rose, E.J., Royle, N.A., Rujescu, D., Ryten, M., Sachdev, P.S., Salami, A., Satterthwaite, T.D., Savitz, J., Saykin, A.J., Scanlon, C., Schmaal, L., Schnack, H.G., Schork, A.J., Schulz, S.C., Schür, R., Seidman, L., Shen, L., Shoemaker, J.M., Simmons, A., Sisodiya, S.M., Smith, C., Smoller, J.W., Soares, J.C., Sponheim, S.R., Sprooten, E., Starr, J.M., Steen, V.M., Strakowski, S., Strike, L., Sussmann, J., Sämann, P.G., Teumer, A., Toga, A.W., Tordesillas-Gutierrez, D., Trabzuni, D., Trost, S., Turner, J., Van den Heuvel, M., van der Wee, N.J., van Eijk, K., van Erp, T.G.M., van Haren, N.E.M., van ’t Ent, D., van Tol, M.J., Valdés Hernández, M.C., Veltman, D.J., Versace, A., Völzke, H., Walker, R., Walter, H., Wang, L., Wardlaw, J.M., Weale, M.E., Weiner, M.W., Wen, W., Westlye, L.T., Whalley, H.C., Whelan, C.D., White, T., Winkler, A.M., Wittfeld, K., Woldehawariat, G., Wolf, C., Zilles, D., Zwiers, M.P., Thalamuthu, A., Schofield, P.R., Freimer, N.B., Lawrence, N.S., Drevets, W., 2014. The ENIGMA Consortium: Large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav. 8, 153–182. https://doi.org/10.1007/s11682-013-9269-5

Van Dijk, K.R.A., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2010. Intrinsic Functional Connectivity As a Tool For Human Connectomics: Theory, Properties, and Optimization. J. Neurophysiol. 103, 297–321. https://doi.org/10.1152/jn.00783.2009

van Dijk, K.R.A., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59, 431–438. https://doi.org/10.1016/j.neuroimage.2011.07.044

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn Human Connectome Project: An overview. Neuroimage 80, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

Vanderwal, T., Eilbott, J., Finn, E.S., Craddock, R.C., Turnbull, A., Castellanos, F.X., 2017. Individual differences in functional connectivity during naturalistic viewing conditions. Neuroimage 157, 521–530. https://doi.org/10.1016/j.neuroimage.2017.06.027

Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition1. Perspect. Psychol. Sci. 4, 274–290. https://doi.org/10.1111/j.1745-6924.2009.01125.x

Waites, A.B., Stanislavsky, A., Abbott, D.F., Jackson, G.D., 2005. Effect of prior cognitive state on resting state networks measured with functional connectivity. Hum. Brain Mapp. 24, 59–68. https://doi.org/10.1002/hbm.20069

Weschler, D., 2008. Wechsler Adult Intelligence Scale - Fourth Edition. Stat. Solut. 1–3.

Whelan, R., Garavan, H., 2014. When optimism hurts: Inflated predictions in psychiatric neuroimaging. Biol. Psychiatry 75, 746–748. https://doi.org/10.1016/j.biopsych.2013.05.014

Williams, R.H., Zimmerman, D.W., 1989. Statistical power analysis and reliability of measurement. J. Gen. Psychol. 116, 359–369. https://doi.org/10.1080/00221309.1989.9921123

Winkler, A.M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P.T., Duggirala, R., Glahn, D.C., 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. Neuroimage 53, 1135–1146. https://doi.org/10.1016/j.neuroimage.2009.12.028

Wong, C.W., Olafsson, V., Tal, O., Liu, T.T., 2012. Anti-correlated networks, global signal regression, and the effects of caffeine in resting-state functional MRI. Neuroimage 63, 356–364. https://doi.org/10.1016/j.neuroimage.2012.06.035

Yarkoni, T., Westfall, J., 2017. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. Perspect. Psychol. Sci. 12, 1100–1122. https://doi.org/10.1177/1745691617693393

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106, 1125–1165. https://doi.org/10.1152/jn.00338.2011.

Yuan, W., Altaye, M., Ret, J., Schmithorst, V., Byars, A.W., Plante, E., Holland, S.K., 2009. Quantification of head motion in children during various fMRI language tasks. Hum. Brain Mapp. 30, 1481–1489. https://doi.org/10.1002/hbm.20616