NeuroImage 102 (2014) 424–434

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

# Heritability of head motion during resting state functional MRI in 462 healthy twins

Baptiste Couvy-Duchesne [a,b,c,*], Gabriëlla A.M. Blokland [a,b], Ian B. Hickie [d], Paul M. Thompson [e], Nicholas G. Martin [a], Greig I. de Zubicaray [b], Katie L. McMahon [c], Margaret J. Wright [a]

[a] QIMR Berghofer Medical Research Institute, Brisbane, Australia
[b] School of Psychology, University of Queensland, Brisbane, Australia
[c] Centre for Advanced Imaging, University of Queensland, Brisbane, Australia
[d] Brain & Mind Research Institute, University of Sydney, Australia
[e] Imaging Genetics Center, Laboratory of Neuro-Imaging, Dept. of Neurology & Psychiatry, UCLA School of Medicine, Los Angeles, CA, USA

ARTICLE INFO

*Article history:*
Accepted 5 August 2014
Available online 13 August 2014

*Keywords:*
Head motion
Resting state fMRI
Twin study
Heritability
Broca's area

ABSTRACT

Head motion (HM) is a critical confounding factor in functional MRI. Here we investigate whether HM during resting state functional MRI (RS-fMRI) is influenced by genetic factors in a sample of 462 twins (65% female; 101 MZ (monozygotic) and 130 DZ (dizygotic) twin pairs; mean age: 21 (SD = 3.16), range 16–29). Heritability estimates for three HM components—mean translation (MT), maximum translation (MAXT) and mean rotation (MR)—ranged from 37 to 51%. We detected a significant common genetic influence on HM variability, with about two-thirds (genetic correlations range 0.76–1.00) of the variance shared between MR, MT and MAXT. A composite metric (HM-PC1), which aggregated these three, was also moderately heritable ($h^2$ = 42%). Using a subsample (N = 35) of the twins we confirmed that mean and maximum translational and rotational motions were consistent "traits" over repeated scans (r = 0.53–0.59); reliability was even higher for the composite metric (r = 0.66). In addition, phenotypic and cross-trait cross-twin correlations between HM and resting state functional connectivities (RS-FCs) with Brodmann areas (BA) 44 and 45, in which RS-FCs were found to be moderately heritable (BA44: $h^2$ = 0.23 (sd = 0.041), BA45: $h^2$ = 0.26 (sd = 0.061)), indicated that HM might not represent a major bias in genetic studies using FCs. Even so, the HM effect on FC was not completely eliminated after regression. HM may be a valuable endophenotype whose relationship with brain disorders remains to be elucidated.



## Introduction

Head motion (HM), defined as small head movements (from μm to a few mm), can be detected in every individual, despite the restraint of the head during magnetic resonance imaging (MRI). HM is routinely computed and analysed in resting state functional MRI (RS-fMRI) studies, where it is known to be a confounding factor in the measurement of brain connectivity. Participants with excessive HM are often excluded, and spurious sources of variance caused by HM are removed using linear regression. However, it has been shown that these steps do not remove all the bias introduced by HM in functional connectivity (FC) analyses (Bright and Murphy, 2013; Mowinckel et al., 2012; Power et al., 2012, 2014; Satterthwaite et al., 2013; Satterthwaite et al., 2012; Van Dijk et al., 2012; Yan et al., 2013). For example, in young adults with greater head motion, FC measures in long-range networks were reduced and local FC between nearby voxels increased. This suggests that HM may weaken the long-range signal and create false positive local correlations, at least in those with relatively high HM (Van Dijk et al., 2012; Yan et al., 2013). Similar results have been reported in adolescents using various FC analysis methods (seed based, amplitude of low-frequency fluctuation (ALFF & fALFF) or Independent Component Analysis) (Satterthwaite et al., 2012). In addition, motion has been associated with long term (up to 10 s) BOLD signal changes in grey matter, white matter and cerebrospinal fluid, drawing a more complex picture of the spurious effect of motion (Power et al., 2014). This translates into RS-FC, with a manifest motion bias up to 10 s after motion (Power et al., 2014), longer than previously reported (Satterthwaite et al., 2013) and a lowering of the reliability of the FC estimates (Yan et al., 2013).

HM is fairly consistent across RS-fMRI sessions, with reliability estimates in the moderate range ($r_{MT}$ = 0.66) (Van Dijk et al., 2012). This

*Abbreviations:* HM, head motion; MT, mean translation; MAXT, maximum translation; MR, mean rotation; NUMO, number of movements greater that 0.1 mm; FC, functional connectivity; RS, resting state; BA, Brodmann area; MZ, monozygotic; DZ, dizygotic; ICC, intra class correlation; IFC, inferior frontal cortex; MFC, middle frontal cortex; SMA, supplementary motor area; IPC, inferior parietal cortex; SPC, superior parietal cortex; ITC, inferior temporal cortex; MTC, middle temporal cortex; OC, occipital cortex.

* Corresponding author at: 300 Herston Road, 4006 Herston, Queensland, Australia.
*E-mail address:* baptiste.couvyduchesne@qimrberghofer.edu.au (B. Couvy-Duchesne).

http://dx.doi.org/10.1016/j.neuroimage.2014.08.010
1053-8119/

suggests that HM is a trait that is consistent enough to study. In addition, an association of HM with age has been reported. HM is higher in younger individuals (8–23 years old) (Satterthwaite et al., 2012) and increases with age in older individuals aged 61 years and over (Seto et al., 2001). Furthermore, males tend to exhibit more head movement than females (Van Dijk et al., 2012).

So far, only a few studies have investigated HM in neurological or psychiatric diseases associated with motor control difficulties. Patients recovering from stroke with a hemiparesis exhibit greater task-related HM, perhaps because they may need to recruit more proximal muscles to respond to a stimulus (Seto et al., 2001). In a more general way, task-related fMRI is likely to provoke more motion when individuals (both cases and controls) know they have made a mistake ("whoops" phenomenon), complicating analyses of error-monitoring and related processes (Epstein et al., 2007). This limitation can be even more important in schizophrenia or ADHD, in which impulsiveness on Go-NoGo tasks is increased (Greene et al., 2008). In addition, the proportion of ADHD and autism subjects excluded because of excessive (gross) head movements is significantly higher than for controls (Durston et al., 2003; Epstein et al., 2007; Jones et al., 2010; Yu-Feng et al., 2007). In schizophrenia, head rotation in particular appears to be significantly higher among patients than healthy controls (Bullmore et al., 1999).

Here we investigated the "trait" aspect of HM during RS-fMRI, by analysing a sample of 462 healthy adolescent and young adult twins (mean age: 21 years). In the twin design, we included both monozygotic (MZ) and dizygotic (DZ) twins. This enables the familial similarities in a trait to be decomposed into genetic (heritability) and environmental sources (Neale and Cardon, 1992; Verweij et al., 2012). We estimated the heritability of three translational and rotational HM measures (Van Dijk et al., 2012). Phenotypic correlations between HM parameters were decomposed into common (shared) and specific (non-shared) sources of genetic and environmental variance. In a post-hoc analysis we then examined how strongly genetic influences on HM affect FC, by focusing on the language production networks, organised around Broca's area (Broca, 1861), which comprises Brodmann areas (BA) 44 and 45. These resting state (RS) networks have been characterised previously using seed-based approach (Kelly et al., 2010; Tomasi and Volkow, 2012) and parallel ICA (Jamadar et al., 2013). In addition, we examined the reliability of the HM measures in a sub-sample of the twins who were scanned twice.

## Materials and methods

### Participants

Data were collected as part of the Queensland Twin Imaging Study (QTIMS), which has acquired MRI scans on more than a thousand individuals, including twins and their non-twin siblings (de Zubicaray et al., 2008). For the present study, we included all twin pairs aged between 16 and 30 (years) for which RS-fMRI scans were available. We excluded individuals (n = 30) with very large overall HM (i.e., maximal rigid body parameter (Collignon et al., 1995) greater than 3 mm or 2°, the centre of rotation being the centre of the FOV) prior to any analyses. This censoring was performed on absolute motion (measured from a volume of reference) as available in Statistical Parametric Mapping SPM8 (Friston et al., 2006).[1] Our final sample comprised 101 (73 female and 28 male) monozygotic (MZ) and 130 (49 female, 26 male and 55 opposite-sex) dizygotic (DZ) twin pairs (65% females, mean age = 21, SD = 3.16, range 16–29). Sixteen percent of the twins were aged 16 (38 pairs, 15 MZ, 23 DZ). In addition, a sub-sample (35 twin individuals (7 MZ and 3 DZ pairs) and 15 siblings; mean age = 22.5, SD = 2.5, 58% females) of the participants were scanned twice for RS-fMRI in order to assess the reliability of the measurements. The median interval between the two scans was ~3 months (median = 96 days, range 35–203).

Zygosity was established objectively by typing nine independent microsatellite polymorphisms (PIC > 0.7) in the ProfilerPlus™ set using standard methods, and was later confirmed for >80% of the sample genotyped on the 610 K Illumina SNP chip (Medland et al., 2009). Participants were screened (by self-report) for significant medical, psychiatric or neurological conditions, including head injuries, a current or past diagnosis of substance abuse, and for current use of medication that was likely to affect cognition. Written informed consent was obtained from all participants. The study was approved by the ethics review boards of the Queensland Institute of Medical Research, the University of Queensland, and Uniting Health Care, Wesley Hospital, Brisbane. Participants received an honorarium, in appreciation of their time.

### Image collection

The images were acquired on a 4 T Bruker Medspec whole-body scanner in Brisbane, Australia. RS-fMRI was performed with a repetition time TR = 2100 ms, echo time TE = 30 ms, flip angle = 90°, field of view FOV = 230 mm, and total acquisition length of 5 min: 19 s. Thirty-six 3 mm-thick transverse slices, with 0.6 mm gap, were acquired per volume, yielding a voxel size of 3.6 × 3.6 × 3.0 mm. In total, 150 volumes were collected, with the first 5 volumes discarded from the analysis to allow time for steady state to occur. During the scan participants were asked to close their eyes, empty their minds, and to try not to fall asleep. Participants who reported having fallen asleep were excluded, to ensure a consistent experimental procedure. The RS scan was part of a larger protocol lasting approximately 60 min, including a 3D T1-weighted scan to which the functional scans were coregistered. Structural scans were acquired with TR = 1500 ms, TE = 3.35 ms, TI = 700 ms, flip angle = 8°, 256 or 240 (coronal or sagittal) slices, FOV = 240 mm, 256 × 256 × 256 (or 256 × 256 × 240) matrix, slice thickness = 0.9 mm and voxel size 0.9 $mm^3$.

### Head motion measurement

Six head motion parameters were obtained from rigid body transformation (Collignon et al., 1995) of fMRI volumes with Statistical Parametric Mapping SPM8 (Friston et al., 2006). Three translation (sagittal or left-right, coronal or front-back, and axial or up-down directions) and three rotation parameters (pitch, roll and yaw) were extracted by comparing each of the remaining 145 volumes of the time series to the first. For each participant, 4 aggregated measures of the HM time series were then computed, adapted from the description by (Van Dijk et al., 2012): mean translational head motion (MT), maximum translational motion (MAXT), number of movements greater than 0.1 mm (NUMO) and mean rotation (MR).

We calculated the motion from frame to frame (i.e. relative transformation parameters), and not the absolute parameters that measure the displacement from a reference frame. This approach accounts for most of the movement that is responsible for the bias in RS-FC. While absolute motion parameters that capture the position of the head are responsible for shifts in the intensity of the BOLD signal, these have little effect on RS-FC (Power et al., 2014). Calculating the motion metrics from the differenced time series of rigid body parameters also has the added advantage of stabilising the mean and variance over time (1st order

[1] Using absolute (versus frame to frame) rigid body parameters for "gross motion" exclusion is not the most pertinent approach in fMRI as this may lead to more individuals being excluded and does not guarantee that the motion occurring during a repetition time is smaller to the voxel size. In our case, using frame to frame parameters, we would have excluded only 9 individuals, and these were all excluded using absolute motion parameters. We therefore believe that our use of absolute motion should have a limited impact on the results and the statistical power.

integration often results in enhancing processes stationarity). MT and MAXT were calculated for each participant using the following formulas:

$$MT = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}$$

$$MAXT = \max_i\left(\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}\right).$$

MR was defined similarly with $\Delta\boldsymbol{\theta}$, $\Delta\boldsymbol{\varphi}$ ***and*** $\Delta\phi$ the differenced rotation angles ($\Delta\boldsymbol{\theta_i} = \boldsymbol{\theta_i} - \boldsymbol{\theta_{i-1}}$) and N = 144 the number of frames.

$$MR = \frac{1}{N}\sum_{i=1}^{N}\left[\cos^{-1}\left(\frac{\cos\Delta\theta_i\cos\Delta\varphi_i + \cos\Delta\theta_i\cos\Delta\phi_i + \cos\Delta\phi_i\cos\Delta\varphi_i + \sin\Delta\theta_i\sin\Delta\varphi_i\sin\Delta\phi_i - 1}{2}\right)\right]$$

Variables were log-transformed to ensure normality of the distribution. In the following MR, MT, MAXT and NUMO refer to the log-transformed variables, if not stated otherwise. NUMO did not reach normality but was strongly correlated with MT ($r = 0.92$) and therefore was not analysed further. Out of the 462 individuals, five observations (1 for MT, 4 for MAXT) were classified as outliers (>3 SD from the mean) and Winsorised (by imputing their values to ±3SD). None of the outliers were from the same participant or the same twin pair. Using the same criterion, we checked the data for twin-pair outliers by considering the mean value of the pair. We identified 2 outlying twin pairs, one for MT and one for MAXT. For MT, the pair of DZ opposite sex twins, aged 20 years, additionally showed high rotational movements (top 10% of the distribution). For MAXT, the outlying MZ female pair, aged 16 years, had normal MT and MR (just above average). In addition, both of these twin pairs were identified as outliers in their zygosity group. Thus, for these pairs, MT and MAXT respectively were set to missing, due to their possible strong impact on variance and mean estimation.Using a principal component analysis (PCA) we then showed the first principal component (PC) of the PCA (HM-PC1) represented a "size effect", keeping most of the information carried by MR, MT and MAXT, and defined a composite metric of motion (accounting for both rotational and translational motion). The second PC (HM-PC2) differentiated rotation from translation, where individuals exhibited asymmetric motion (rotation without translation and conversely) with the third PC capturing the residual variance and extreme motion (MAXT vs MR and MT) (see Inline Supplementary Figure S1).

Inline Supplementary Fig. S1 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

Several composite HM metrics have been proposed to summarise the information from the rigid body parameters (Jenkinson et al., 2002; Power et al., 2012; Van Dijk et al., 2012; Yan et al., 2013). A recent comparison (Yan et al., 2013) showed very good performance of Jenkinson et al. 's composite metric against Yan et al.'s metric of reference. Mathematically, the 2 approaches are very similar: they consist of averaging, over all the voxels of the brain, the displacement occurring between two frames. Their only difference is that Yan et al. average the displacement of each voxel while Jenkinson et al. assume that the brain is a sphere of radius 80 mm. Thus, the very strong agreement between metrics (Yan et al., 2013) suggests that the simplification of the brain into a sphere or radius 80 mm is valid, at least for young adults. Here, we use a sample very close in age to Yan et al. (2013), and can therefore reasonably assume the equivalence of the Yan et al. and Jenkinson et al. metrics. We chose to use the Jenkinson et al. metric, which is the faster of the two to compute, as the motion measurement of reference, to expand on prior work that compared the metrics (Yan et al., 2013). We showed that the Power et al. metric (aka. FD: "framewise displacement"), HM-PC1 as well as MT were highly correlated with the metric of reference ($r_{Pearson} > 0.90$, $\tau_{Kendall} > 0.80$) and that all metrics showed similar reliability and heritability. We suggest that residual differences between MT, HM-PC1, Power et al. and Jenkinson et al. arise from the weight given to rotation over translation, since all the metrics are built upon equivalent norms (2-norm (Euclidian) for Jenkinson et al., Yan et al., MT, MR and HM-PC1; 1-norm ("taxicab") for Power et al. metric). As previously found, we also show that there is poor agreement between Yan et al.'s metric and their definition of mean translation, and we attribute this to their choice of parameterisation (Inline Supplementary Table S1). Thus, using the Jenkinson et al. metric as a reference, we found only very small differences in the correlations between metrics. We attribute this to the different weight given to rotation over translation, which depends on the brain radius for Power et al. and on the PCA decomposition for HM-PC1. Ultimately, and despite these small differences, we show very similar estimates, reliability and heritability for the metrics (Inline Supplementary Table S1). Further, even if validity and comparability still remain to be assessed in adolescents and younger children, the robustness shown here for young adults suggests that all composite metrics (and MT) might also be a good proxy for HM in younger subjects. (Detailed results are presented in Inline Supplementary Table S1 and Inline Supplementary Figure S2).

Inline Supplementary Table S1 and Fig. S2 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

*Resting-state functional image processing*

Resting state fMRI images were processed using the DPARSF-A (Chao-Gan and Yu-Feng, 2010) toolbox for SPM8 (Wellcome Department of Imaging Neuroscience, 2009). The first 5 images were removed and slice timing correction was applied. Functional volumes were realigned and coregistered with the structural scans using DARTEL (Ashburner, 2007). Then, using General Linear Model (GLM), we regressed out sources of physiological (white matter, cerebrospinal fluid) and non-physiological noise (trend, HM rigid body parameters, global mean signal) from the BOLD signal and band pass filtered ([0.01–0.08 Hz]) after linear regression (Weissenbacher et al., 2009).

Functional connectivity maps were computed in a seed-based analysis using BA44 and BA45, as the seed regions. The seeds were defined using cytoarchitectonic maximum probability maps (MPMs) as masks (from SPM8's Anatomy toolbox (Zilles and Amunts, 2010)) merged with the grey matter mask created from structural scans (inclusive masking). The average time course in each seed region was calculated and the correlations were estimated with all the voxels of the brain (r-map). Fisher's z transformation ($z = \text{arctanh}(r)$) was applied to the correlation map to centre and normalise the distribution (z-map).

We performed a one-sample t-test voxel-wise among the participants' z-maps to discard the voxels showing non-significant z-correlations, correcting for multiple testing using a topological estimate of the Family Wise Error (FWE) rate: the Euler characteristic (Worsley et al., 1992), which takes into account the volume tested and the smoothness (i.e., local correlation) of the image (Nichols and Hayasaka, 2003). Outlier detection and deletion was performed voxel-wise. For each observation a chi-squared test (chi-squared distance, 1 degree of freedom) was performed with a risk alpha set at 0.001 and outlier values set to missing.

We further masked the FC results according to their reliability using the sub-sample of participants scanned twice, adopting an Intra-Class Correlation (ICC), threshold of >0.4, with random effect on subjects. Analyses were performed using the R package "irr" (Gamer et al., 2010). This step aimed to limit the number of false positive in our analysis by ensuring that the observed FC was corresponding to a robust co-activation with the seed. In doing so, we assume that RS-FC must be fairly consistent across a short period of time (3 months in median between test and retest MRI), even in presence of brain plasticity or brain maturation. This strong assumption may partially explain why the overall reliability of the RS-FC was never strong (<0.8 across the brain) and why it can be justified to set a "low" reliability threshold (ICC > 0.4 is, at

best, described as "fairly reliable" in the literature). It is true that this approach may exclude from the analysis the regions of BA44 and BA45's networks that evolve the most (or the faster) with age, but we think it is a fair price to pay considering the gain in power that the pruning of voxels represents.

For BA45, 31,671 voxels (67% of the 47,279 grey matter voxels) showed a significant correlation with the seed, 2902 of which reliably (ICC > 0.4) defined BA45's network. The 10 largest clusters (size greater than 50 voxels) comprised 64% of BA45's reliable network (1855 voxels out of 2902) and allowed a good description of the network. From these 10 clusters we identified 8 brain regions using SMP8's Anatomy toolbox (Eickhoff et al., 2007): Inferior Frontal cortex (IFC; BA44, BA45 and pars Orbitalis), Middle Frontal cortex (MFC), Supplementary Motor Areas (SMA; BA6); Inferior Parietal (IPC: PGa, PF, PFm, hIP1-3), Superior Parietal (SPC: BA7a, BA7pc), Inferior Temporal (ITC) and Middle Temporal cortex (MTC); Occipital/Visual cortex (OC; BA17, 18, hOC5).

For BA44, 32,775 voxels passed the significance threshold, of which only 4422 reliably defined BA44's network. Again, the 10 largest clusters (size > 50) comprised most (73%) of the network that covered 9 brain regions: Inferior Frontal cortex (IFC; BA44, BA45 and right pars Orbitalis), Middle Frontal cortex (MFC), Supplementary Motor Area (SMA; BA6); Inferior Parietal (IPC: PGa, PF, PFm, PFcm, PFt), Superior Parietal (SPC: BA7M, 5M, 5Ci, 5L); Inferior Temporal (ITC) and Middle Temporal cortex (MTC); Occipital/Visual cortex (OC; BA17, 18, hOC5) and Cerebellum (Lobule VIIa, crus 1-2).

Similar areas were activated in the left and right hemispheres, except for the Cerebellum Lobule where only the right side of the cerebellum was correlated with BA44. However, the extent of the activation was greater in the left hemisphere. Furthermore, all the clusters, except for the occipital cortex and superior parietal areas, were positively correlated with the seeds.

These findings replicate those of resting-state FC studies on Broca's language area, with similar positive correlation of Broca's area with the IFC (pars opercularis, triangularis and orbitalis) (Jamadar et al., 2013; Kelly et al., 2010; Tomasi and Volkow, 2012), the MFC (Jamadar et al., 2013; Tomasi and Volkow, 2012), the SMA (pre-SMA and BA6 (Kelly et al., 2010)); the IPC (supramarginal and angular gyrus) (Jamadar et al., 2013; Kelly et al., 2010; Tomasi and Volkow, 2012), the ITC (Jamadar et al., 2013; Tomasi and Volkow, 2012) and the MTC (Kelly et al., 2010). Significant FC between Broca's area and the cerebellum (crus) has been reported (Tomasi and Volkow, 2012), however the authors used a 3.4 $cm^3$ seed centred in BA45 while we only observed an association with BA44. We did not replicate findings of positive association with the superior frontal cortex (medial frontal gyrus with BA8, 9, 32 (Kelly et al., 2010), BA8 (Tomasi and Volkow, 2012)). Kelly et al. reported positive correlation with the caudal part of the superior temporal gyrus and sulcus while Tomasi and Volkow reported non-significant correlation, coherent with our observations.

Anti-correlations replicated previous results showing anti-activation in the SPC (BA7 (Jamadar et al., 2013), BA7 and BA5 (Tomasi and Volkow, 2012)) and with the OC (Tomasi and Volkow, 2012).

### *Genetic analyses of head motion*

Saturated models were fitted to compare HM means, variances and covariances for the five zygosity groups (MZ males, MZ females, DZ males, DZ females and DZ opposite-sex), with age, age$^2$ and sex included as covariates. Model fit was assessed based on the difference in the −2 log likelihood between the full model (ACE or ADE) and any nested model (AE, CE or DE and E), which follows a chi-squared distribution under the null hypothesis (of no difference of fit). At the univariate level, using two zygosity groups (i.e., MZ and DZ, after ensuring the homogeneity of each of these groups in the saturated model), we decomposed the variance of each of the HM measures (MR, MT, MAXT and HM-PC1) into additive genetic (A), common environmental (C), and unique environmental (E) sources of variance. We also estimated a dominant genetic (D) effect (instead of the shared environment) when the MZ and DZ correlation indicated a possible non-additive effect ($r_{DZ} < 0.5 \times r_{MZ}$). The dominant effect models the alleles' interaction at one locus (dominance) or at different loci (epistasis).

We then examined the covariation among MT, MAXT, and MR using a Cholesky decomposition followed by independent and common pathway modelling (Neale and Cardon, 1992; Verweij et al., 2012). Cholesky decomposition is the standard general approach to decompose variance into genetic and environmental sources, so we used this model to test the significance of shared A and C influences and to estimate the genetic correlations. In our Cholesky decomposition, the three dimensional genetic and environmental variance/covariance matrices are decomposed into the product of a lower triangular matrix and its transpose. This decomposition involves a first factor that influences all variables, a second factor (independent of the first) that influences the second and third variables, and a third factor (independent of the two first) that influences only the third variable.

Common and Independent pathway models represent a different approach from the Cholesky decomposition in that they distinguish, for each variable, the shared sources of variance from the specific. However, the Independent pathway model contains the same number of parameters than the Cholesky and is equivalent in term of fit. The common pathway model is different from the Independent pathway model, in that the co-variation between the HM measures is determined by a latent variable: a global HM factor. This latent HM factor has genetic and environmental sources of variance, which account for a proportion of the MR, MT and MAXT variance (Arseneault et al., 2003). The Bayesian Information Criterion (BIC) was used to compare model fit between these two models, taking into account the number of parameters to estimate.[2] All genetic modelling was performed in OpenMx (Boker et al., 2011) using a full information maximum likelihood (FIML) estimator, under the R 3.1.0 distribution (R Development Core Team, 2012).

Reliability of the four HM metrics was assessed using the ICC, which corresponds to a mixed effect model (random effect of subjects, fixed effect of experiment). Since ICC coefficients are highly sensitive to outliers, extreme values in the series were identified, based on visually inspecting the test-retest scatter plots. All the identified outliers were excluded from the reliability analysis if their value was greater than ±3SD from the mean. Two unrelated individuals were excluded for MR and HM-PC1, none for MT and MAXT.

### *Genetic multivariate modelling of head motion and heritable functional connectivity*

Using FC measures extracted from the same dataset as HM we ran a voxel-wise heritability analysis in OpenMx (Boker et al., 2011), fitting univariate ACE models with HM, age and sex as covariates, for each voxel of BA45 and BA44's networks.

Homogeneity of sampling across groups (female–male, MZ–DZ), covariate effects and significance of the heritability estimates were tested voxel-wise (likelihood ratio test) and corrected for multiple testing, controlling the FWER. We used the Monte-Carlo simulation protocol of 3DClustSim (AFNI) (Cox, 1996), to estimate the volume threshold such that the chance of larger clusters of contiguously significant voxels (individual level $\alpha = 5\%$; i.e., "discovery clusters") occurring at random is smaller than 5%. The FWHM, as an estimate of spatial correlation, was calculated from the square root of the network residual map from the GLM regression using 3dFWHMx (AFNI) (Cox, 1996) and used as a 3DClustSim input. The significance volume threshold was 47 voxels for BA45's network, and 58 voxels for BA44.

In each network, to explore common sources of variance between HM and all heritable FC measures, we estimated cross-trait (phenotypic) and cross-trait cross-twin (ct–ct) correlations. At first, we used HM-

[2] The BIC often leads to select more parsimonious models (than the AIC) as it penalises more heavily the model complexity (number of parameters).

PC1 as a composite measure of HM. Then, we made the distinction between rotational and translational confound by calculating the correlations with MR and MT separately. Thus, for each FC and HM measurement (HM-PC1, MR and MT), we tested, using a likelihood ratio test whether: a) the MZ and DZ phenotypic correlations are equal and null, and b) the MZ and DZ ct–ct correlations are equal. The ct–ct correlations can shed light on the genetic and environmental contributions to the phenotypic correlations and represent a simple way of assessing the presence of genetic correlation. Indeed, where phenotypic correlation is observed, a MZ ct–ct correlation greater than DZ indicates a genetic effect, a DZ ct–ct correlation greater than half the MZ suggests a significant effect of the shared environment. Finally, if the traits are driven by independent individual environmental factors the ct–ct correlations should be null. Phenotypic and ct–ct correlations, as well as the p-values, were estimated using OpenMx (Boker et al., 2011), which takes into account the relatedness in the sample.

We used the same previous multiple testing approach to identify regions of significant correlation. The significance threshold and smoothness were re-estimated on the sub-maps of significant heritability, with a resulting threshold of 29 voxels for BA45's connectivity map and 24 voxels for BA44.

## Results

### *Head motion characteristics of the sample*

Table 1 shows the motion characteristics for the sample. The HM means and variance were similar to those reported in a previous study (Van Dijk et al., 2012) and exhibited similar magnitude with a mean MT = 0.064 (SD = 0.025). Assumption testing supported homogeneity of means and variances, across zygosity and sex. For MR there was a subtle significant negative effect of age (−0.013 [−0.014,−0.012]; p-values = 0.02) and age$^2$ (−0.00027 [−0.00028,−0.00026]; p-values = 0.03) but no difference between males and females. Age, age$^2$ and sex were not significant for either MAXT, MT and HM-PC1. HM measures were moderately to highly correlated (Table 1).

Test–retest ICCs for all four HM measures indicated moderately good reliability, improved by the exclusion of outliers on MR ($ICC_{MR}$ = 0.46, $ICC_{HM\text{-}PC1}$ = 0.65, in the overall sample, $ICC_{MR}$ = 0.53 and $ICC_{HM\text{-}PC1}$ = 0.66 when excluding outliers) (Table 1).

### *Genetic modelling of head motion*

Univariate modelling indicated neither a significant common environmental (C), nor a dominant genetic effect (D) on the HM measures (ACE vs AE model for MT: p = 1, for MAXT: p = 0.71; ADE vs AE model for MR: p = 0.82, for HM-PC1: p = 0.46) (Table 2). Similarly, using Cholesky decomposition of MR, MT and MAXT, common environmental effects were not significant. Therefore, all subsequent modelling included A and E components only, knowing that we may slightly overestimate the additive genetic effect, as it would include the non-significant C or D variance.

Additive genetic estimates in the AE Cholesky model were highly significant (p < 0.001). High genetic correlations ($r_g$) ($r_{gMT-MAXT}$ = 0.99 [0.65,1.00]; $r_{gMR-MAXT}$ = 0.77 [−0.32, 1.00] and $r_{gMT-MR}$ = 0.82 [0.18,1.00]), indicated that the three HM measures shared, to a large extent, common additive genetic factor(s). The environmental correlations ($r_e$) were lower ($r_{eMT-MAXT}$ = 0.64 [0.55, 0.74], $r_{eMR-MAXT}$ = 0.56 [0.45, 0.66] and $r_{eMT-MR}$ = 0.61 [0.50, 0.70]). Thus, the environmental factors influencing MT head motion were mostly common to the 3 measures.

A common pathway model provided a similar overall goodness-of-fit to the independent pathway decomposition ($AIC_{Independent}$ = 2891.8 and $AIC_{Common}$ = 2891.9). However, with only 14 parameters to estimate (versus 15) the common pathway model was the simplest for a similar fit, and minimised the BIC criterion ($BIC_{Independent}$ = −1730.7 versus −1724.0). Therefore, we used the more parsimonious AE common pathway model to examine the covariation between the HM measures, and to test the significance of common (Ac) and specific additive genetic sources of variance (As). Both specific and shared A and E estimates were significant in the model (p < 0.001). Estimates are presented in Fig. 1.

The latent HM factor explained a high proportion of the observed variance: 70% for MT, 72% for MAXT and 50% for MR. This result confirms that a substantial amount of the variance in MT, MAXT and MR is due to a common source. Forty-six percent of the variance in this latent HM factor was due to an additive genetic factor (A) and 54% due to non-shared environmental (E) variance, some of which is correlated measurement error. The variance explained by specific factors was as low as 30% (19% As and 11% Es) for MT and 28% (2% As and 26% Es) for MAXT. Specific variance was slightly higher for MR (50% [14% As and 36% Es]).

**Table 1**
Motion characteristics of the twin sample.

| | MT (mm) | MAXT (mm) | MR (degrees) | HM-PC1 |
|---|---|---|---|---|
| Descriptive statistics | | | | |
| Raw Mean (SD) | 0.06 (0.02) | 0.27 (0.22) | 0.04 (0.01) | |
| Range | [0.03,0.28] | [0.06,1.35] | [0.02,0.10] | |
| Log-transformed Mean (SD) | −2.83 (0.35) | −1.54 (0.65) | −7.32 (0.32) | 0.00 (1.50) |
| Range | [−3.63,−1.89] | [−2.75,0.30] | [−8.24,−6.39] | [−3.63,4.46] |
| Covariate effects | | | | |
| p-Value$_{Age}$ | 0.61 | 0.54 | 0.02* | 0.18 |
| p-Value$^2_{Age}$ | 0.67 | 0.57 | 0.03* | 0.22 |
| p-Value$_{Sex}$ | 1 | 1 | 1 | 1 |
| Reliability | | | | |
| ICC [95% CI] | 0.59 [0.32,0.77] | 0.53 [0.24,0.73] | 0.53 [0.24, 0.74] | 0.66 [0.42,0.82] |
| Phenotypic correlations [95% CI] | | | | |
| MT | – | 0.71 [0.66,0.75] | 0.58 [0.52,0.64] | 0.88 [0.86,0.90] |
| MAXT | | – | 0.60 [0.54,0.66] | 0.89 [0.87,0.91] |
| MR | | | – | 0.83 [0.80,0.86] |
| Twin correlations | | | | |
| $r_{MZ}$ [95% CI] | 0.50 [0.39,0.60] | 0.33 [0.20,0.45] | 0.29 [0.16,0.41] | 0.38 [0.20,0.54] |
| $r_{DZ}$ [95% CI] | 0.25 [0.13,0.36] | 0.19 [0.07,0.30] | 0.14 [0.02,0.26] | 0.18 [0.01, 0.34] |

Means, standard deviations, and ranges of the HM measurements (raw and log-transformed), as well as sex, age$^2$ and age effects, repeatability coefficients (ICC), Phenotypic correlations (Pearson) and ML twin correlations are presented. Exclusion of 2 outliers improved the reliability of MR and HM-PC1 ($ICC_{MR}$ = 0.46 [0.15, 0.68] and $ICC_{HM\text{-}PC1}$ = 0.65 [0.40, 0.80] in the overall sample). No outlier was identified in test–retest MT and MAXT distributions. Our sample exhibited similar amount of motion to that reported previously (Van Dijk et al., 2012).
* p-Value < 0.05.

**Table 2**
Model fit and parameter estimates of the univariate genetic models for HM.

| Variable | Model | Parameter estimates | | | Model fit | | | |
|---|---|---|---|---|---|---|---|---|
| | | A | C/D | E | df | −2LL | AIC | p-Values |
| **MT** | ACE | 0.54 [0.21,0.65] | 0.00 [0.00,0.24] | 0.46 [0.35,0.61] | 4 | 302.06 | −609.9 | |
| | **AE** | **0.54 [0.39,0.65]** | | **0.46 [0.35,0.61]** | **3** | **302.06** | **−611.9** | **p = 1** |
| | CE | | 0.35 [0.23,0.46] | 0.65 [0.54,0.77] | 3 | 310.51 | −603.5 | p = 0.003 |
| | E | | | | 2 | 340.25 | −575.7 | p < 0.001 |
| **MAXT** | ACE | 0.26 [0.00,0.48] | 0.07 [0.00,0.35] | 0.67 [0.51,0.85] | 4 | 886.35 | −25.7 | |
| | **AE** | **0.34 [0.19,0.48]** | | **0.66 [0.52,0.81]** | **3** | **886.48** | **−27.5** | **p = 0.71** |
| | CE | | 0.26 [0.13,0.37] | 0.74 [0.63,0.87] | 3 | 887.54 | −26.5 | p = 0.27 |
| | E | | | | 2 | 903.26 | −12.7 | p < 0.001 |
| **MR** | ADE | 0.19 [0.00,0.42] | 0.09 [0.00,0.44] | 0.72 [0.56,0.89] | 4 | 236.14 | −679.9 | |
| | **AE** | **0.27 [0.10,0.42]** | | **0.73 [0.58,0.90]** | **3** | **236.19** | **−681.8** | **p = 0.82** |
| | E | | | | 2 | 246.22 | −673.8 | p = 0.006 |
| | ACE | 0.28 [0.00,0.42] | 0.00 [0.00,028] | 0.72 [0.58,0.90] | 4 | 234.46 | −677.5 | |
| **HM-PC1** | ADE | 0.32 [0.01,0.55] | 0.08 [0.00,0.29] | 0.60 [0.45,0.78] | 4 | 1653.13 | 745.1 | |
| | **AE** | **0.42 [0.25,0.56]** | | **0.58 [0.44,0.75]** | **3** | **1653.67** | **743.7** | **p = 0.46** |
| | E | | | | 2 | 1673.05 | 761.4 | p < 0.001 |
| | ACE | 0.39 [0.00,0.53] | 0.00 [0.00,0.00] | 0.61 [0.47,0.78] | 4 | 1653.32 | 745.3 | |

The parameter estimates are followed by the 95% CI between brackets. All the estimates are standardised (percentage of variance). df gives the degree of freedom i.e. the number of estimates in the model. −2LL is −2 times the log-likelihood of the model, AIC the Akaike criterion for each model. The p-value corresponds to the likelihood ratio test of the full model versus any nested model. For each HM variable, the best model (p-value > 0.05 and AIC minimal) appears in bold.

Following this evidence for large common underlying genetics (and environmental) factors in rotational and translational movement we investigated the heritability difference between MT ($h^2 = 0.51$) and MR ($h^2 = 0.37$) by coming back to the distributions of frame-to-frame rigid body parameters. It became clear that the extent (amplitude) of movement during an MRI session is not the same in all directions: translation along the X axis (left–right translation) and rotation around the Y ("maybe" rotation) and Z axis ("no" rotation) exhibited reduced variance (See Inline Supplementary Figure S3). These differences in variance are likely due, at least in part, to the head coil restraining rotation more than translation.

In addition, heritability of MAXT ($h^2 = 0.35$) was also lower than for MT, likely arising from MAXT being (by construct) more sensitive than the mean to the presence of rare and extreme movement, which results in greater within-pair variability.

Inline Supplementary Fig. S3 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

### *Genetic modelling of Broca's functional connectivity*

Saturated univariate models estimated on each voxel indicated no mean or variance differences for either sex or zygosity groups. Common or shared environment (C) had no significant influence across the brain and was dropped from the models. There was a significant effect of age and sex in 3 regions of BA45's network (see Inline Supplementary Table S2, Inline Supplementary Figure S4). The principal components of HM were not significant across the brain (the biggest cluster of significant voxels for HM-PC1 effect was of size k = 32 voxels, the discovery clusters were even smaller for HM-PC2: $k \leq 28$ and HM-PC3: $k \leq 24$, none of them reaching the significance threshold). Fitting AE models we identified 3 regions of BA45's network with significant heritability: left IPC (k = 254 voxels, $h^2 = 0.27$ (sd = 0.069)), left SMA (k = 133, $h^2 = 0.23$ (0.048)), right IPC (k = 82, $h^2 = 0.25$ (0.043)) and right IFC (k = 53, $h^2 = 0.22$ (0.034)) (see Fig. 2 and Inline Supplementary Table S2).

Inline Supplementary Table S2 and Fig. S4 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

In BA44's network, age, sex and the first PC of HM were kept as covariates. 5 regions showed a significant effect of age or sex on FC (see Inline Supplementary Table S3, Inline Supplementary Figure S5). In one region there was a significant effect of the first principal component of HM (HM-PC1) that corresponds to HM effect size and therefore can be interpreted as the global amount of motion. The region was located in left/right SPC (k = 98, $\overline{\beta} = 0.025$ (0.0069)). We identified three regions of the BA44 network that showed significant heritability, and contained 4 significant clusters: left SMA (k = 123, $h^2 = 0.23$ (sd = 0.043)), IPC (k = 90, $h^2 = 0.23$ (0.036); k = 69, $h^2 = 0.23$ (0.042)) and left/right OC (k = 62, $h^2 = 0.21$ (0.045)) (see Fig. 3 and Inline Supplementary Table S3).

Inline Supplementary Table S3 and Fig. S5 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

Histograms of p-values for BA44 and BA45's ACE model parameters were coherent with the previous results: the p-values exhibit a uniform distribution when no significant effect. Conversely, a significant effect is associated with a peak in the histogram for low p-values (see Inline Supplementary Figures S4 and S5). In BA45's network, even if HM-PC1 p-values histogram may have suggested a significant effect, no cluster of voxels passed multiple testing correction.[3]

Inline Supplementary Figs. S6 and S7 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

### *Sources of covariation between head motion and functional connectivity*

No clusters of phenotypic correlation between FC and HM-PC1 passed the significance threshold for BA44 (24 voxels) and BA45 (29 voxels). In BA44's heritable network, the significant correlations were positioned randomly in the network ($k \leq 2$ voxels for the discovery clusters of phenotypic correlation, $k \leq 7$ for the ct–ct correlation, $k \leq 2$ for the genetic correlations). The absence of a significant correlation was confirmed by the histogram of p-values (See Inline Supplementary Figure S8). We also examined the correlation between FC with both MR and MT, and again, no significant correlations were observed.

Inline Supplementary Fig. S8 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

In BA45's heritable network, the size of the discovery clusters for phenotypic correlation reached k = 24 voxels, $k \leq 17$ for the ct–ct correlations but only isolated genetic correlations were observed ($k \leq 2$). Therefore, we could not reject the null hypothesis of no correlation between FC and HM across the brain. The histograms of p-values may have suggested a significant phenotypic correlation (See Inline

[3] In our case, the histogram of p-values can be misleading and a non uniform distribution can correspond to non-significant effect (when the significant voxels do not cluster). This is because we use topological properties (clusters) to correct for multiple testing instead of correcting voxel-wise.

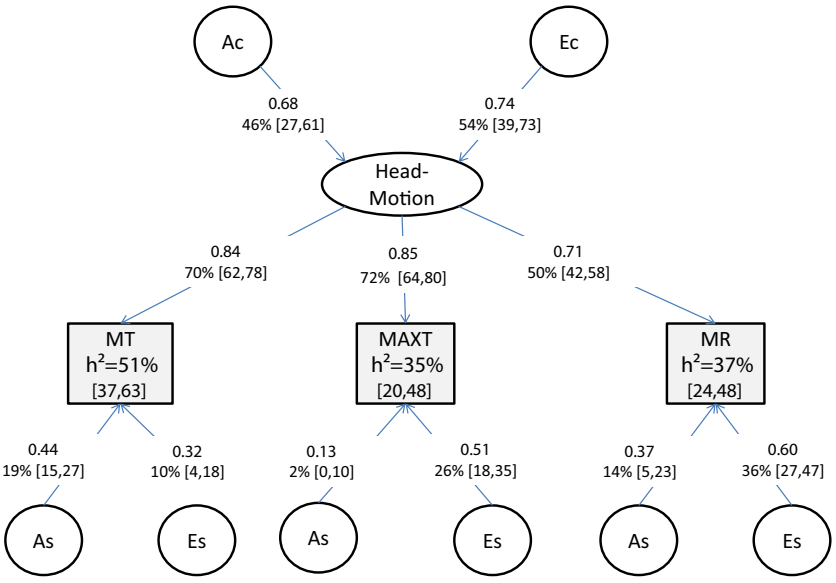


**Fig. 1.** The Common pathway model showing parameter estimates and covariation between MT, MAXT and MR. In the path diagram, the square boxes represent the observed variables (phenotypes), and the circles the latent variables (A, C, E). $A_c$ and $E_c$ are the Additive genetic and Environmental effects common to the 3 HM measures through the latent HM factor (oval). $A_s$ and $E_s$ are, for each HM variable the specific additive genetic and environment effects. Path coefficients (standardised) are presented in bold. Below each is the percentage of variance for the HM measurement and the latent variable, followed by its 95% confidence interval. Heritability estimates for each HM measure are shown below each HM variable name.

Supplementary Figure S9); however no area passed the significance threshold (that accounts for multiple testing).

We broke down the HM into rotational and translational components by considering the correlations of FC with MT and MR. MT showed no significant correlation with the FC in heritable voxels. However, we identified a area, in the left IPC, showing significant differences in MZ and DZ phenotypic correlation between FC and MR ($k = 42$ voxels, $\overline{rMZ} = 0.090$ ($sd = 0.061$), $\overline{rDZ} = -0.16(sd = 0.039)$). The sign difference between MZ and DZ correlations suggested a false positive, which was confirmed by the low ct–ct and non-significant genetic correlations (on $k = 42$ voxels, $\overline{ctctMZ} = 0.056(sd = 0.050)$, $\overline{ctctDZ} = -0.033(sd = 0.042)$ and only 3 voxels were significantly correlated with MR).

Inline Supplementary Fig. S9 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.08.010.

## Discussion

Here, we show for the first time that individual differences in small movements of the head during RS-fMRI are influenced by genetic factors. Using a healthy population sample, comprising 462 largely young adult twins, we demonstrated significant low to moderate heritability of three HM measures: mean translation, mean rotation, and maximum translation. The strong covariation across MR, MR and MAXT was captured by a latent head motion factor for which half of the variance was found to be due to genes. Heritability of HM-PC1 (0.42) was very similar to the latent Head Motion factor (0.46) from the Common Pathway model, suggesting HM-PC1 captured well the genetic dimension of head motion. We confirmed that HM measurements are reliable, and showed that there was little effect of age or sex on small head movements.

Our finding that each of the three HM measures is significantly heritable (MT = 54%, MAXT = 35%, MR = 37%) is consistent with the proposal that small movements of the head are stable traits (Van Dijk et al., 2012). Interestingly, we found that a large proportion of translational and rotational HM is under the control of the same genetic factors (genetic correlations ranging 0.76 to 1.00). Common variability, represented by a latent HM factor, was equally explained by the genes and the environment ($h^2 = 46\%$ [0.27, 0.61]). Thus, variation in HM in the population can be partially explained by an additive effect of several genetic loci common to the 3 HM measures. In contrast, specific genetic influences explained only a small proportion of the variance in HM (19% for MT; 2% for MAXT and 14% for MR). In addition to a common genetic influence, we also showed that both translation and rotation head movements were significantly influenced by the same environmental factors (environmental correlations about 0.6) some of which may reflect correlated measurement error. Specific environmental factors explained a minor proportion of the variance in MT (11%) but a greater amount of the extreme and rotational displacements (26% for MAXT, 36% for MR). This was likely to be the case for MAXT since it may capture some voluntary movement. One must keep in mind that all these results were obtained for HM constrained by the head coil; a study of unconstrained HM might reveal a slightly different pattern of genetic and environmental factors.

Our heritability estimates for HM are consistent with those for another non-voluntary movement such as eye blink startle reflex suppression, which has a heritability of 50% (Anokhin et al., 2003). However, our estimates for HM are somewhat less than those reported for resting tremor (i.e., 93–99%) (Lorenz et al., 2004), which is caused by contraction of opposing muscle groups (Walker, 1990), or eye blink startle magnitude (i.e., 70%) (Anokhin et al., 2003). Eye blink startle reflex suppression and resting tremor are frequently associated with Parkinson's disease (Anokhin et al., 2003; Lorenz et al., 2004). Furthermore, eye blink startle reflex suppression is implicated in the biological bases of schizophrenia and has been proposed as a possible endophenotype for genetic studies (Anokhin et al., 2003). Both types of non-voluntary movement are associated with impaired cerebellum connectivity (Manto et al., 2012; Walker, 1990).

Moderate test–retest reliability for all HM metrics (range 0.53–0.66) confirmed the results of a prior study (Van Dijk et al., 2012). Higher test–retest reliability (ICC = 0.73) has been reported for the Yan et al. metric of HM (Yan et al., 2013). However, the reliability of the concurrent metrics on the same sample was not reported and thus could also be higher (e.g. due to better controlled experimental parameters across sessions or longer RS scans). We also found similar magnitude of HM in the twins to those previously reported (Van Dijk et al., 2012), despite having possibly made different choices for the centre of rotation and the motion used (absolute or frame-wise). However, contrary to their study, we found little indication of any sex difference in HM. Similarly, and only for rotational motion did we find a subtle decrease with age. This may reflect the fact that the twins in our sample were largely

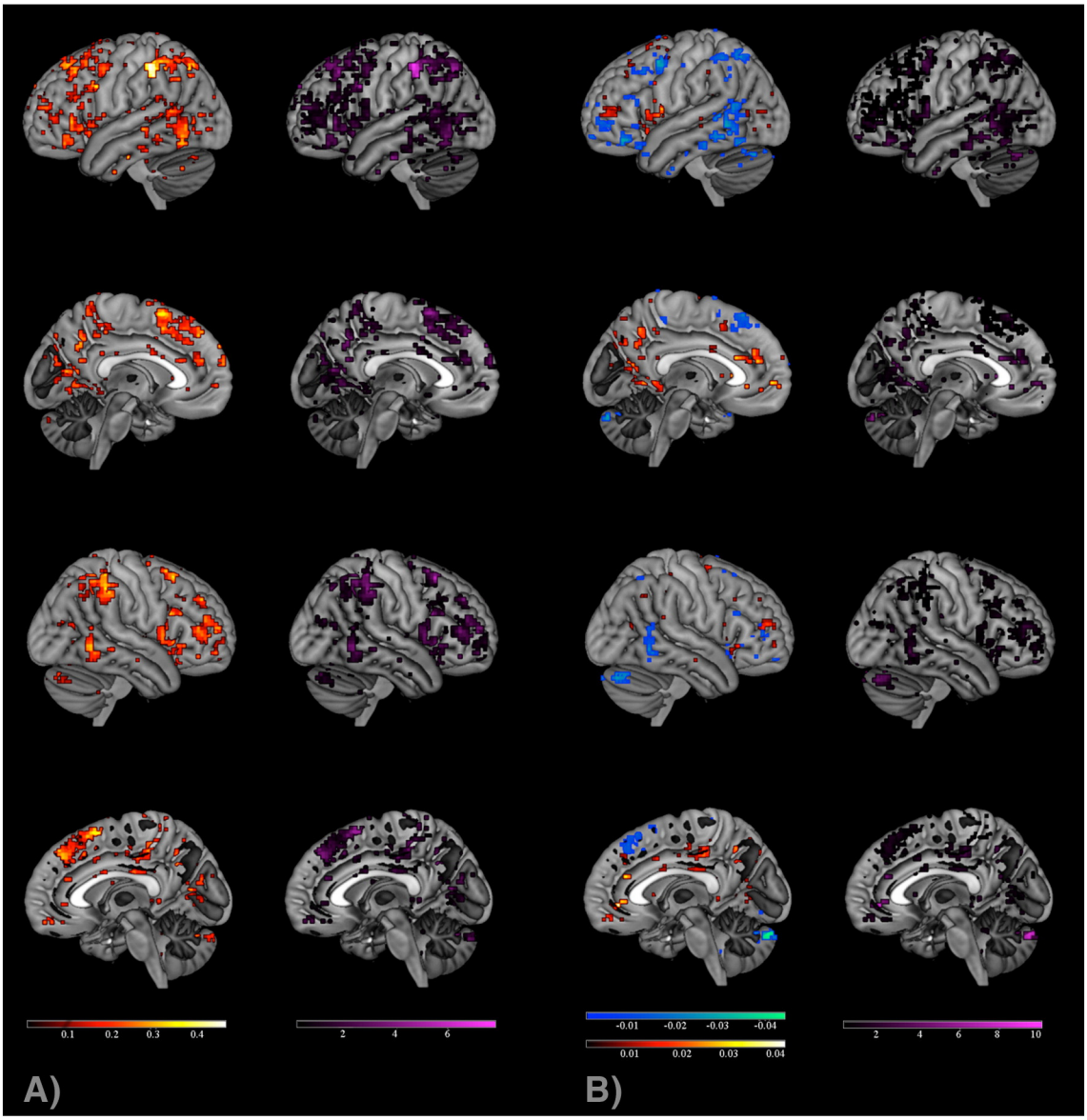


**Fig. 2.** Heritability map of BA45's network and effect of head motion on the network's FC. (A) Heritability map of the BA45's network (heat-map) and corresponding −log(p-values) map (purple). (B) HM-PC1 effect on BA45's FC with beta maps and corresponding −log(p-values). The negative betas are shown in cold colours (with strongest effect corresponding to green), positive betas are plotted with warm colours (the strongest effect being white). Therefore, all the colorbars rank effect/significance from low to high (i.e. left to right). Effect size and significance are plotted voxel-wise. Heritability and HM effects are projected onto brain maps after averaging the signal (using 12 voxels depth resolution) using MRIcron (Rorden and Brett, 2000). We plotted the top 80% of voxels for heritability and motion effect to facilitate the reading of the image. No thresholding was applied on the p-values maps and significance of all the voxels considered is presented.

young adults, with the youngest participants aged 16 (i.e. mean age = 21, SD = 3.16, range 16–29). A prior study that reported significant linear and quadratic effects of age (Satterthwaite et al., 2012) had a much younger sample (8–23 years old). Further studies are needed to investigate the effects of age and sex on HM, but our findings suggest that at least in healthy young adults, variability in HM due to age and sex are likely to be small.

HM is a known confound for RS-fMRI. Linear regression of HM is insufficient in removing all bias (Bright and Murphy, 2013; Mowinckel et al., 2012; Power et al., 2014; Satterthwaite et al., 2013; Van Dijk et al., 2012; Yan et al., 2013) so the real effect of motion on FC is still unknown and is likely to be expressed differently according to the FC metric used (Van Dijk et al., 2012; Yan et al., 2013). Here, in a post-hoc analysis we investigated the extent that heritable HM could confound seed-based FC heritability studies by creating a bias related to sample zygosity. Overall, we found that the FC and HM correlations indicated little common genetic sources of variance between those traits, indicating no major motion induced bias in the genetic estimates of FC. However, some sparse residual associations were found in BA44's network (HM-PC1 was a significant covariate in SPC). While these results require confirmation on a larger sample (greater power) and at a larger scale (we focussed on seed-based resting-state FC and did not investigate more complex [non-linear] relationships), our findings should encourage controlling for and minimising HM influences on FC estimates. Thus, including HM measures as covariates (in pre-processing and further analyses) can reduce the confound with fMRI measurements (Beall and Lowe, 2014; Bright and Murphy, 2013; Mowinckel et al., 2012; Power et al., 2012; Power et al., 2014; Satterthwaite et al., 2013; Yan et al., 2013), and even more when slice-wise motion is used (Beall and Lowe, 2014), but knowing the complex and wide spread effect of HM on brain signal (Power et al., 2014) and its association with other regressors (Hallquist et al., 2013; Power et al., 2014), a complete removal of the spurious effect of motion through regression only is illusory. To this effect, a new approach, based on post-hoc deletion

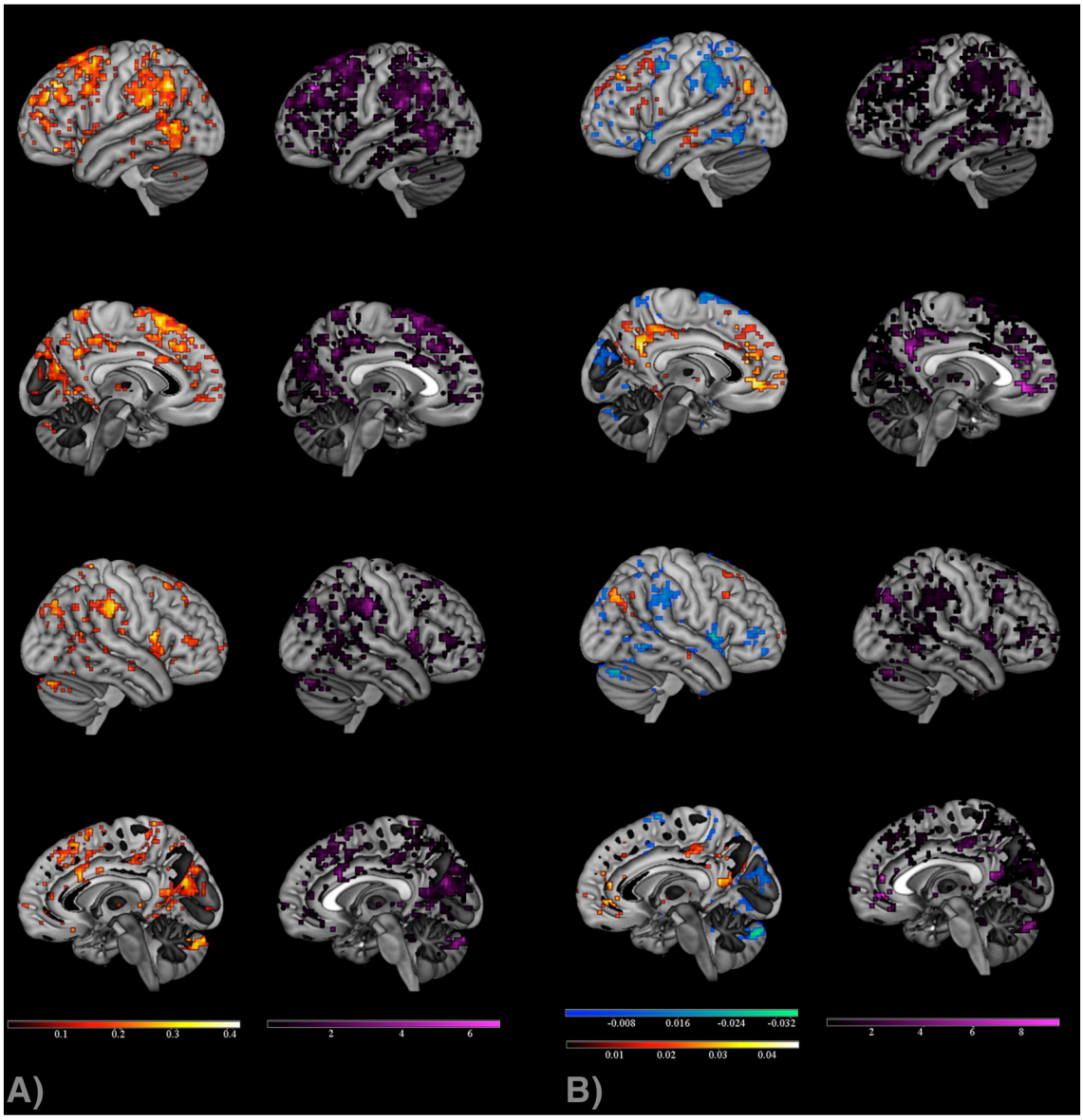


**Fig. 3.** Heritability map of BA44's network and effect of head motion on the network's FC. (A) Heritability map of the BA44's network (heat-map) and corresponding −log(p-values) map (purple). (B) HM-PC1 effect on BA44's FC with beta maps and corresponding −log(p-values). The negative betas are shown in cold colours (with strongest effect corresponding to green), positive betas are plotted with warm colours (the strongest effect being white). Therefore, all the colorbars rank effect/significance from low to high (i.e. left to right). Effect size and significance are plotted voxel-wise. Heritability and HM effects are projected onto brain maps after averaging the signal (using 12 voxels depth resolution) using MRIcron (Rorden and Brett, 2000). We plotted the top 80% of voxels for heritability and motion effect to facilitate the reading of the image. No thresholding was applied on the p-values maps and significance of all the voxels considered is presented.

(censoring) of highly noisy volumes, has been developed to reduce the residual effect of HM on FC and shows significant promise (Power et al., 2012; Power et al., 2014).

As a consequence, we can interpret our findings of a significant genetic influence on FC for both BA44 and BA45's network with greater confidence. In previous MRI studies, a similar level of heritability has been reported on RS-FC based graph theory metrics (van den Heuvel et al., 2013) and group-ICA FC (Glahn et al., 2010; Korgaonkar et al., 2014). This is the first genetic study of seed based RS-FC and we provided heritability maps of the Broca's networks at rest. They showed that several regions across Broca's resting state networks were heritable with the genetic influence on FC for BA45 and BA44 distributed differently. It would be of interest for other studies to replicate these findings as would heritability studies of additional RS brain network. In addition, it could be worthwhile investigating the extent of any genetic covariation between RS-FC and cognitive ability scores, such as vocabulary or language skills (Hart et al., 2010; Haworth et al., 2009). This could shed light on the network's organisation and performance.

Besides, the moderate heritability of HM and FC has to be considered in light of the (moderate) reliability of these traits. Reliability (or repeatability) represents, in most cases, an upper bound for heritability estimates (Dohm, 2002). Here, for both type of measures the similarity of MZ correlation and reliability estimates suggests that most of the reliable signal is heritable and it would be of interest to see how the heritability estimates evolve when the reliability of the measurement is increased (by acquiring longer RS fMRI for instance, or by using a more reliable metric if available).

Finally, together with the work showing an association between gross HM level and case status for diseases such as schizophrenia (Bullmore et al., 1999), autism (Jones et al., 2010) and ADHD (Durston et al., 2003; Epstein et al., 2007; Yu-Feng et al., 2007), we suggest that HM may be considered a possible endophenotype for brain disorders,

especially if the diseases are known or suspected to impact motor functions.

## Conclusion

We estimated, for the first time, the heritability of HM measured during RS-fMRI. HM was found to be a trait rather than a state, with good reliability and a significant genetic influence. Most of the variance in the three HM measures (mean translation, rotation and maximum translation) was due to common genetic and environmental influences, through a latent HM factor. Genetic and environmental influences specific to each HM measure were significant, but could be biased upwards in our experimental protocol. No clear sex or age effects on movement amplitude were observed in our largely young adult sample. Further, while post-hoc analyses showed that FC in multiple areas of Broca's network is heritable, there was little evidence of shared genetic influences between HM and FC of this network.

HM could help to describe and characterise heritable brain disorders, and therefore may be a valuable endophenotype in future fMRI studies.

## Acknowledgments

This study was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, USA (Grant RO1HD050735), and the National Health and Medical Research Council (NHMRC), Australia (Project Grant 496682). Zygosity typing was funded by the Australian Research Council (ARC) (Grants A7960034, A79906588, A79801419, and DP0212016). Greig de Zubicaray is supported by an ARC Future Fellowship (FT0991634). Baptiste Couvy-Duchesne was supported by the Region Bretagne (Ulysses grants for internship), the University of Queensland (UQI PhD scholarship) and the Queensland Institute of Medical Research. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, The National Institutes of Health, NHMRC, or ARC.

We are very grateful to the twins for their generosity of time and willingness to participate in our studies. We thank research nurses Marlene Grace and Ann Eldridge for twin recruitment, research assistants Lachlan Strike, Kori Johnson, Aaron Quiggle, and Natalie Garden, and radiographers Matthew Meredith, Peter Hobden, Kate Borg, Aiman Al Najjar, and Anita Burns for data acquisition, David Butler and Daniel Park for IT support. We also thank the *three* anonymous reviewers whose comments/suggestions helped improve and strengthen the results presented in this manuscript.

## References

Anokhin, A.P., Heath, A.C., Myers, E., Ralano, A., Wood, S., 2003. Genetic influences on prepulse inhibition of startle reflex in humans. Neurosci. Lett. 353, 45–48.

Arseneault, L., Moffitt, T.E., Caspi, A., Taylor, A., Rijsdijk, F.V., Jaffee, S.R., Ablow, J.C., Measelle, J.R., 2003. Strong genetic effects on cross-situational antisocial behaviour among 5-year-old children according to mothers, teachers, examiner-observers, and twins' self-reports. J. Child Psychol. Psychiatry 44, 832–848.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38, 95–113.

Beall, E.B., Lowe, M.J., 2014. SimPACE: generating simulated motion corrupted BOLD data with synthetic-navigated acquisition for the development and evaluation of SLOMOCO: a new, highly effective slicewise motion correction. NeuroImage 101C, 21–34.

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.

Bright, M.G., Murphy, K., 2013. Removing motion and physiological artifacts from intrinsic BOLD fluctuations using short echo data. NeuroImage 64, 526–537.

Broca, P., 1861. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). Bull. Soc. Anat. 6, 330–357.

Bullmore, E., Brammer, M., Rabe-Hesketh, S., Curtis, V., Morris, R., Williams, S., Sharma, T., McGuire, P., 1999. Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI. Hum. Brain Mapp. 7, 38–48.

Chao-Gan, Y., Yu-Feng, Z., 2010. DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. Front. Syst. Neurosci. 4, 13.

Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G., 1995. Automated multi-modality image registration based on information theory. Information Processing in Medical Imaging,pp. 264–274.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

de Zubicaray, G.I., Chiang, M.C., McMahon, K.L., Shattuck, D.W., Toga, A.W., Martin, N.G., Wright, M.J., Thompson, P.M., 2008. Meeting the challenges of neuroimaging genetics. Brain Imaging Behav. 2, 258–263.

Development Core Team, R., 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Dohm, M.R., 2002. Repeatability estimates do not always set an upper limit to heritability. Funct. Ecol. 16, 273–280.

Durston, S., Tottenham, N.T., Thomas, K.M., Davidson, M.C., Eigsti, I.M., Yang, Y., Ulug, A.M., Casey, B., 2003. Differential patterns of striatal activation in young children with and without ADHD. Biol. Psychiatry 53, 871–878.

Eickhoff, S.B., Paus, T., Caspers, S., Grosbras, M.-H., Evans, A.C., Zilles, K., Amunts, K., 2007. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. NeuroImage 36, 511–521.

Epstein, J.N., Casey, B., Tonev, S.T., Davidson, M., Reiss, A.L., Garrett, A., Hinshaw, S.P., Greenhill, L.L., Vitolo, A., Kotler, L.A., 2007. Assessment and prevention of head motion during imaging of patients with attention deficit hyperactivity disorder. Psychiatry Res. Neuroimaging 155, 75–82.

Friston, K.J., Penny, W.D., Ashburner, J., Kiebel, S.J., Nichols, T.E., 2006. Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press, London.

Gamer, M., Lemon, J., Fellows, I., Sing, P., 2010. irr: Various Coefficients Of Interrater Reliability and Agreement. Retrieved 17/08, 2013, from http://CRAN.R-project.org/package=irr.

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.

Greene, C.M., Braet, W., Johnson, K.A., Bellgrove, M.A., 2008. Imaging the genetics of executive function. Biol. Psychol. 79, 30–42.

Hallquist, M.N., Hwang, K., Luna, B., 2013. The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. NeuroImage 82, 208–225.

Hart, S.A., Petrill, S.A., Dush, C.M., 2010. Genetic influences on language, reading, and mathematics skills in a national sample: an analysis using the National Longitudinal Survey of Youth. Lang. Speech Hear. Serv. Sch. 41, 118–128.

Haworth, C.M., Kovas, Y., Harlaar, N., Hayiou-Thomas, M.E., Petrill, S.A., Dale, P.S., Plomin, R., 2009. Generalist genes and learning disabilities: a multivariate genetic analysis of low performance in reading, mathematics, language and general cognitive ability in a sample of 8000 12-year-old twins. J. Child Psychol. Psychiatry 50, 1318–1325.

Jamadar, S., Powers, N.R., Meda, S.A., Calhoun, V.D., Gelernter, J., Gruen, J.R., Pearlson, G.D., 2013. Genetic influences of resting state fMRI activity in language-related brain regions in healthy controls and schizophrenia patients: a pilot study. Brain Imaging Behav. 7, 15–27.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17, 825–841.

Jones, T.B., Bandettini, P.A., Kenworthy, L., Case, L.K., Milleville, S.C., Martin, A., Birn, R.M., 2010. Sources of group differences in functional connectivity: an investigation applied to autism spectrum disorder. NeuroImage 49, 401–414.

Kelly, C., Uddin, L.Q., Shehzad, Z., Margulies, D.S., Castellanos, F.X., Milham, M.P., Petrides, M., 2010. Broca's region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. Eur. J. Neurosci. 32, 383–398.

Kelly, A.M., Di Martino, A., Uddin, L.Q., Shehzad, Z., Gee, D.G., Reiss, P.T., Margulies, D.S., Castellanos, F.X., Milham, M.P., 2009. Development of anterior cingulate functional connectivity from late childhood to early adulthood. Cereb. Cortex 19, 640–657.

Korgaonkar, M.S., Ram, K., Williams, L.M., Gatt, J.M., Grieve, S.M., 2014. Establishing the resting state default mode network derived from functional magnetic resonance imaging tasks as an endophenotype: a twins study. Hum. Brain Mapp. 35, 3893–3902.

Lorenz, D., Frederiksen, H., Moises, H., Kopper, F., Deuschl, G., Christensen, K., 2004. High concordance for essential tremor in monozygotic twins of old age. Neurology 62, 208–211.

Manto, M., Bower, J.M., Conforto, A.B., Delgado-Garcia, J.M., da Guarda, S.N., Gerwig, M., Habas, C., Hagura, N., Ivry, R.B., Marien, P., Molinari, M., Naito, E., Nowak, D.A., Oulad Ben Taib, N., Pelisson, D., Tesche, C.D., Tilikete, C., Timmann, D., 2012. Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. Cerebellum 11, 457–487.

Medland, S.E., Duffy, D.L., Wright, M.J., Geffen, G.M., Hay, D.A., Levy, F., Van-Beijsterveldt, C.E. M., Willemsen, G., Townsend, G.C., White, V., 2009. Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. Neuropsychologia 47, 330–337.

Mowinckel, A.M., Espeseth, T., Westlye, L.T., 2012. Network-specific effects of age and in-scanner subject motion: a resting-state fMRI study of 238 healthy adults. NeuroImage 63, 1364–1373.

Neale, M.C., Cardon, L.R., 1992. Methodology for Genetic Studies of Twins and Families. Kluwer Academic Pub.

Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat. Methods Med. Res. 12, 419–446.

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.

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.

Satterthwaite, T.D., Wolf, D.H., Loughead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., Gur, R.E., 2012. Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. NeuroImage 60, 623–632.

Rorden, C., Brett, M., 2000. Stereotaxic display of brain lesions. Behav. Neurol. 12, 191–200.

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.

Seto, E., Sela, G., McIlroy, W.E., Black, S.E., Staines, W.R., Bronskill, M.J., McIntosh, A.R., Graham, S.J., 2001. Quantifying head motion associated with motor tasks used in fMRI. NeuroImage 14, 284–297.

Tomasi, D., Volkow, N.D., 2012. Resting functional connectivity of language networks: characterization and reproducibility. Mol. Psychiatry 17, 841–854.

van den Heuvel, M.P., van Soelen, I.L., Stam, C.J., Kahn, R.S., Boomsma, D.I., Hulshoff Pol, H. E., 2013. Genetic control of functional brain network efficiency in children. Eur. Neuropsychopharmacol. 23, 19–23.

Uddin, L.Q., Supekar, K., Menon, V., 2010. Typical and atypical development of functional human brain networks: insights from resting-state FMRI. Front Syst. Neurosci. 4, 21.

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.

Verweij, K.J., Mosing, M.A., Zietsch, B.P., Medland, S.E., 2012. Estimating heritability from twin studies. Methods Mol. Biol. 850, 151–170.

Walker, H.K., 1990. The Origins of the History and Physical Examination.

Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., Windischberger, C., 2009. Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. NeuroImage 47, 1408–1416.

Wellcome Department of Imaging Neuroscience, 2009. SPM8. Queen Square, London, UK.

Worsley, K.J., Evans, A.C., Marrett, S., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. J. Cereb. Blood Flow Metab. 12, 900–931.

Yan, C.G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R.C., Di Martino, A., Li, Q., Zuo, X.N., Castellanos, F.X., Milham, M.P., 2013. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. NeuroImage 76, 183–201.

Yu-Feng, Z., Yong, H., Chao-Zhe, Z., Qing-Jiu, C., Man-Qiu, S., Meng, L., Li-Xia, T., Tian-Zi, J., Yu-Feng, W., 2007. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev. 29, 83–91.

Zilles, K., Amunts, K., 2010. Centenary of Brodmann's map—conception and fate. Nat. Rev. Neurosci. 11, 139–145.