Supplementary Material

Increased Integration Between Default Mode and Task-Relevant Networks in Children with ADHD is Associated with Impaired Response Control

Kelly A. Duffy[1,*,†], Keri S. Rosch[2,3,4,*], Mary Beth Nebel[2,5], Karen E. Seymour[2,4,6,††], Martin A. Lindquist[7], James J. Pekar[8,9], Stewart H. Mostofsky[2,4,5], Jessica R. Cohen[1,10,11]

[*] *Co-first authors*
[1] *Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.*
[2] *Center for Neurodevelopmental and Imaging Research, Kennedy Krieger Institute, Baltimore, MD, USA.*
[3] *Department of Neuropsychology, Kennedy Krieger Institute, Baltimore, MD, USA.*
[4] *Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, USA.*
[5] *Department of Neurology, Johns Hopkins University, Baltimore, MD, USA.*
[6] *Department of Mental Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.*
[7] *Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.*
[8] *F.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD, USA.*
[9] *Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD, USA.*
[10] *Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.*
[11] *Carolina Institute for Developmental Disabilities, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.*
[†] *Kelly Duffy is now with the Department of Psychology, University of Minnesota, Minneapolis, MN, USA.*
[††] *Karen Seymour is now with the Center for Scientific Review, National Institutes of Health, Bethesda, MD, USA. This work was prepared while Karen Seymour was employed at Johns Hopkins University and Kennedy Krieger Institute. The opinions expressed in this article are the author's own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government.*

1. FMRIPREP Processing Steps

Functional data were first minimally preprocessed using FMRIPREP version 1.0.7 (Esteban et al., 2019), a Nipype (Gorgolewski et al., 2011, 2017) based tool. Each T1w (T1-weighted) volume was corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction v2.1.0 (using the OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.0 (Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM), and gray-matter (GM) was performed on the brain-extracted T1w using fast (Zhang et al., 2001; FSL v5.0.9).

Functional data was slice time corrected using 3dTshift from AFNI v16.2.07 (Cox, 1996) and motion corrected using mcflirt (Jenkinson et al., 2002; FSL v5.0.9). This was followed by co-registration to the corresponding T1w using boundary-based registration (Greve and Fischl, 2009) with 9 degrees of freedom, using bbregister (FreeSurfer v6.0.0). Motion correcting transformations, BOLD-to-T1w transformation, and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation.

Physiological noise regressors were extracted from mean signal within CSF and WM. Framewise displacement (FD; Power et al., 2014) was calculated for each functional run using the implementation of Nipype.

Many internal operations of FMRIPREP use Nilearn (Abraham et al., 2014), principally within the BOLD-processing workflow. For more details of the pipeline see http://fmriprep.readthedocs.io/en/latest/workflows.html.

2. Relationship between In-Scanner Motion and ADHD Symptoms and Task Performance

As stated in the *Main Text*, we chose to not match participants in the attention-deficit/hyperactivity disorder (ADHD) and typically developing (TD) groups on mean raw FD, quantified before motion correction, given evidence that head motion is correlated with ADHD symptomatology due to shared genetic factors (Couvy-Duchesne et al., 2016). To determine whether this relationship was the case in our sample, we correlated mean raw FD with ADHD symptomatology (hyperactivity/impulsivity and inattention) and metrics of go/no-go (GNG) performance often impaired in ADHD (commission errors [ComErr] and the log-transformed tau component of the response time distribution) across all participants. We used the false discovery rate (FDR) correction for 4 comparisons. We found that mean raw FD was significantly correlated with symptoms of hyperactivity/impulsivity ($r = 0.29$, corrected $p = .016$) and inattention ($r = 0.30$, corrected $p = .016$), with a nonsignificant trend toward a correlation with GNG ComErr ($r = 0.21$, corrected $p = .070$). Mean FD was not correlated with tau ($r = 0.12$, corrected $p = .28$). The finding that raw mean FD was related to behavioral metrics of interest in our sample underscores the importance of allowing group differences in raw mean FD in our participants.

3. Motion-Related Quality Assurance

Due to the fact that we did not match groups on in-scanner motion, we took extra steps to ensure that group differences in motion did not drive our findings. Importantly, while children with ADHD had significantly higher mean FD than TD children before any data processing (see Table 1 in the *Main Text*), after processing FD significantly decreased for both ADHD and TD children for both the scrubbed and the unscrubbed data (ADHD scrubbed: $t(42) = 15.46$, $p <$ 1E-18, mean FD post-processing = 0.018; ADHD unscrubbed: $t(42) = 15.64$, $p <$ 1E-18, mean FD post-processing = 0.018; TD scrubbed: $t(42) = 18.46$, $p <$ 1E-21, mean FD post-processing = 0.015; TD unscrubbed: $t(42) = 19.03$, $p <$ 1E-21, mean FD post-processing = 0.016). The groups no longer had significantly different FD after processing for the unscrubbed data ($t(84) = 1.70$, $p = .092$), although they did for the scrubbed data ($t(84) = 2.00$, $p = .048$). Further, to ensure that motion did not differentially impact the quality of the post-processed data across groups, we calculated recommended post-processing quality control metrics (Ciric et al., 2017, 2018) using the same 264-node functional atlas (Power et al., 2011) used to develop these metrics. We calculated the metrics separately for each group and for the scrubbed and the unscrubbed data. We compared the groups on the relationship between residual (post-processing) motion and functional connectivity (FC) strength (scrubbed: $r = 0.144$ for ADHD; $r = 0.138$ for TD; unscrubbed: $r = 0.114$ for ADHD; $r = 0.127$ for TD), the percentage of FC edges significantly related to motion (scrubbed: 0.006% for ADHD; 0% for TD; unscrubbed: 0% for ADHD; 0% for TD), and the relationship between edge length and motion (scrubbed: r = -0.132 for ADHD; r = -0.068 for TD; unscrubbed: r = -0.024 for ADHD; r = 0.004 for TD). Our values for percentage of FC edges related to motion and the relationship between edge length and motion are either comparable or lower numerically than those reported in the literature (values reported for the recommended processing pipelines: % edges 0.28% – 10.29%; edge length-motion relationship -0.243 – -0.116), and our values for the relationship between motion and FC strength are slightly higher (our range 0.127 – 0.144 versus a range of 0.042 – 0.059 for the recommended processing pipelines). We note that despite our relationship between motion and FC strength being higher, it is still quite low and comparable across groups. See Figure S1 for depiction of quality control metrics. Critically, no brain-related outcomes of interest were correlated with mean FD after processing (r-values ranged from -0.14 – 0.06; all p-values FDR-corrected for 6 brain-related outcomes of interest $> .84$).

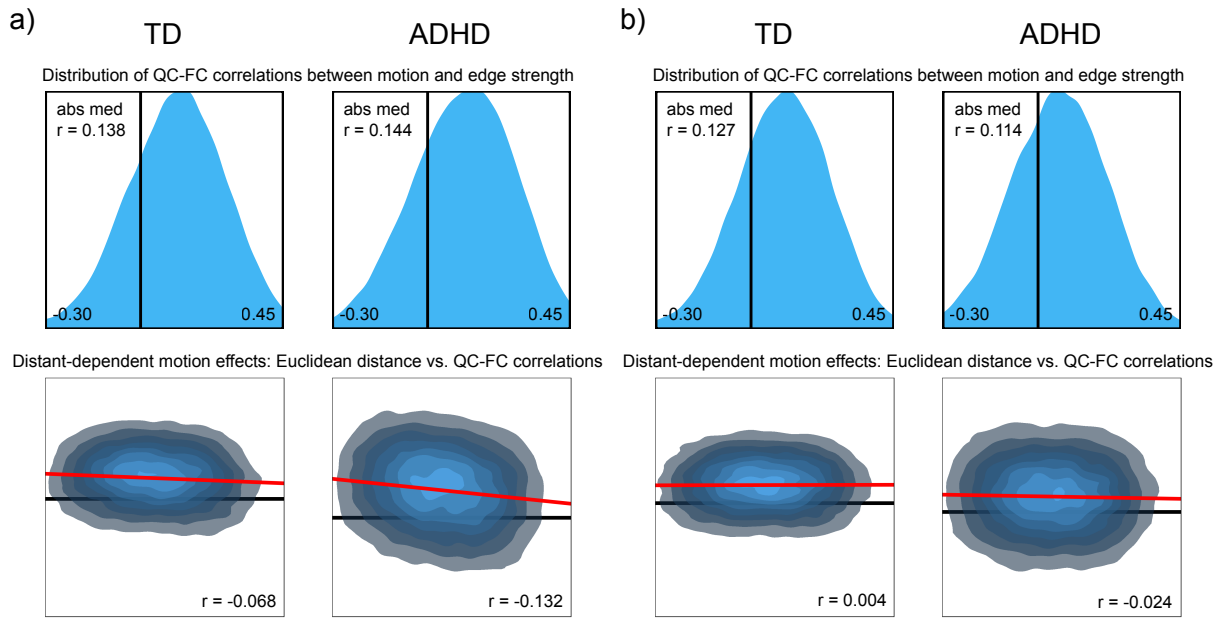

Figure S1: Residual effects of motion in the TD and ADHD groups after post-processing using quality metrics from Ciric et al. (2017) for a) scrubbed and b) unscrubbed data. The top panel depicts the distribution and absolute median correlation of quality control (QC)-functional connectivity (FC) correlations between participant motion (framewise displacement; FD) and edge strength (*r*). The bottom panel depicts distant-dependent effects of motion using density plots (x-axis: Euclidean distance between each pair of regions of interest; y-axis = QC-FC correlation of that edge pair). Plots were generated using eXtensible Connectivity Pipeline (XCP) software (https://github.com/PennBBL/xcpEngine; Ciric et al., 2018). Values of quality metrics in the current dataset are comparable to those reported in Ciric et al. (2017).

In addition to calculating the above quality control metrics, we ran several sensitivity analyses accounting for motion in different ways to ensure that the results reported in the *Main Text* were not driven by our particular methodological decisions. Specifically, we: (1) included mean FD as an additional covariate in all analyses; (2) included DVARS as an additional covariate in all analyses; (3) conducted analyses with unscrubbed, as opposed to scrubbed, data; and (4) used a motion-matched sample of participants. All analyses were corrected for multiple comparisons using procedures identical to those specified in the *Analyses* section of the *Main Text*. Only results of analyses reported as significant in the *Main Text* are included below. No results were found to be significant in the sensitivity analyses that were not significant in the *Main Text*.

See Tables S1 and S2 for comparison of results across sensitivity analyses and below sections for a detailed description of the results for each sensitivity analysis.

3.1. Replicating Results Controlling for Mean FD

First, we repeated analyses controlling for head motion (mean FD) in addition to mean FC and age. Specifically, we ran one-way ANCOVAs with group as the factor and mean FC, age, and mean FD as the covariates to compare static mean participation coefficient (PC) and within-module degree (WD), as well as time-varying coefficient of variation of PC (CVPC) and of WD (CVWD), across groups. Pearson's correlations (partial correlations controlling for mean FC, age, and mean FD) were conducted to relate graph metrics to performance on the GNG task, separately for each group. With regard to whole-network static FC analyses, there were significant differences between

ADHD and TD groups for mean whole-network PC ($F(1,70) = 4.05$, $p = .048$) and mean whole-network WD ($F(1,70) = 6.16$, $p = .016$). When assessing PC of specific network pairs, there was a significant group difference in PC between the DMN and FP networks ($F(1,70) = 8.71$, corrected $p = .017$). All static FC group differences were consistent with analyses reported in the *Main Text*. Correlations with ComErr in participants with ADHD were in the same direction as those reported in the *Main Text*, although did not reach our corrected significance threshold of FDR-corrected $p < .05$. Specifically, there was a trend toward whole-network PC being correlated with ComErr ($r = 0.31$, corrected $p = .099$), as well as the DMN-SM network pair PC being correlated with ComErr ($r = 0.38$, corrected $p = .066$). The DMN-SAL network pair PC was not correlated with ComErr ($r = 0.28$, corrected $p = .16$), although the relationship was in the same direction. Finally, the relationship between time-varying whole-network CVPC and ComErr was consistent with that reported in the *Main Text* ($r = -0.36$, corrected $p = .024$). Notably, no results from this set of analyses changed significantly when conducting direct comparisons between those reported in the *Main Text* and here (all z-values ≤ 0.37, all p-values $\geq .71$, uncorrected).

3.2. Replicating Results Controlling for DVARS

Next, we repeated analyses controlling for DVARS in addition to mean FC and age. DVARS is a measure of the change in BOLD signal across timepoints. DVARS is high during periods of high motion, thus it is another measure of motion artifact (Power et al., 2012). Specifically, we ran one-way ANCOVAs with group as the factor and mean FC, age, and DVARS as the covariates to compare static mean PC and WD, as well as time-varying CVPC and CVWD, across groups. Pearson's correlations (partial correlations controlling for mean FC, age, and DVARS) were conducted to relate graph metrics to performance on the GNG task, separately for each group. With regard to whole-network static FC analyses, there were significant differences between ADHD and TD groups for mean whole-network PC ($F(1,70) = 4.02$, $p = .049$) and mean whole-network WD ($F(1,70) = 6.04$, $p = .016$). When assessing PC of specific network pairs, there was a significant group difference in PC between the DMN and FP networks ($F(1,70) = 8.46$, corrected $p = .020$). All static FC group differences were consistent with analyses reported in the *Main Text*. Correlations with ComErr in participants with ADHD were in the same direction as those reported in the *Main Text*, although they did not reach our corrected significance threshold of FDR-corrected $p < .05$. Specifically, there was a trend toward whole-network PC being correlated with ComErr ($r = 0.34$, corrected $p = .068$), as well as DMN-SAL and DMN-SM network pair PC values being correlated with ComErr ($r = 0.34$, corrected $p = .068$ and $r = 0.38$, corrected $p = .065$ respectively). Finally, the relationship between time-varying whole-network CVPC and ComErr was consistent with that reported in the *Main Text* ($r = -0.39$, corrected $p = .012$). Notably, no results from this set of analyses changed significantly when conducting direct comparisons between those reported in the *Main Text* and here (all z-values ≤ 0.11, all p-values $\geq .91$, uncorrected).

3.3. Replicating Results Using Unscrubbed Data

Third, we used unscrubbed, instead of scrubbed, data as input into our analyses. For this analysis, we used models identical to those in the *Main Text* that controlled for mean FC and age for both the group difference ANCOVAs and the partial correlations with behavior. We only conducted static FC analyses for this method, as the time-varying FC analyses reported in the *Main Text* used unscrubbed data. The whole-network results were not significant, although they were in the same direction as those reported in the *Main Text*. That is, participants with ADHD had

numerically higher whole-network PC and whole-network WD than TD participants (PC: ADHD = 0.713, TD = 0.707 [as compared to 0.717 and 0.709 with scrubbed data]; WD: ADHD = 8.45, TD = 8.21 [as compared to 8.79 and 8.36 with scrubbed data]). Similar to the network pair analysis in the *Main Text*, DMN-FP PC was significantly higher in participants with ADHD as compared to TD participants ($F(1,78) = 7.80$, corrected $p = .026$). Additionally, correlations between brain metrics and ComErr in participants with ADHD were in the same direction as those reported in the *Main Text*, although they did not reach our significance threshold of FDR-corrected $p < .05$. Specifically, the correlation between whole-network PC and ComErr was $r = 0.29$ (corrected $p = .13$), as compared to $r = 0.35$ (corrected $p = .049$) with scrubbed data. For network pair analyses, the correlation between DMN-SAL PC and ComErr was $r = 0.30$ (corrected $p = .12$), as compared to $r = 0.36$ (corrected $p = .043$) with scrubbed data; and the correlation between DMN-SM PC and ComErr was similarly $r = 0.30$ (corrected $p = .12$), as compared to $r = 0.38$ (corrected $p = .043$) with scrubbed data. Notably, no results from this set of analyses changed significantly when conducting direct comparisons between those reported in the *Main Text* and here (all z-values ≤ 0.61, all p-values $\geq .54$, uncorrected).

3.4. Replicating Results in a Motion-Matched Sample

Finally, we used data processing and models identical to those in the *Main Text* with a subsample of our participants who were matched for motion. This analysis included 29 participants with ADHD (mean age 10.07, 12 girls, mean FD = 0.121) and 29 TD participants (mean age 10.17, 13 girls, mean FD = 0.121). Motion was not significantly different across the groups (Welch's $t(55.74) = 0.009$, $p = .99$). As expected, given the small sample size, results in general were not statistically significant, although as with the other sensitivity analyses they were in the same direction as those reported in the *Main Text*. Specifically, while there were no significant group differences in static FC brain metrics (whole-network PC, whole-network WD, or DMN-FP PC; all corrected p-values $> .30$), they were all numerically higher in ADHD, as compared to TD, participants (whole-network PC: ADHD = 0.713, TD = 0.711; whole-network WD: ADHD = 8.57, TD = 8.54; DMN-FP PC: ADHD = 0.370, TD = 0.358). Notably, the magnitude of correlations between brain metrics and ComErr in the participants with ADHD were numerically larger than those of the full sample, although with the smaller sample size they did not reach statistical significance. Specifically, the correlation between whole-network static PC and ComErr was $r = 0.41$ (corrected $p = .065$), as compared to $r = 0.35$ (corrected $p = .049$) with the full sample. For static network pair analyses, the correlation between DMN-SAL PC and ComErr was $r = 0.37$ (corrected $p = .080$), as compared to $r = 0.36$ (corrected $p = .043$) with the full sample; and the correlation between DMN-SM PC and ComErr was $r = 0.42$ (corrected $p = .080$), as compared to $r = 0.38$ (corrected $p = .043$) with the full sample. The correlation between whole-network time-varying CVPC and ComErr was was statistically significant as in the *Main Text* (r = -0.51, corrected $p = .007$). Notably, no results from this set of analyses changed significantly when conducting direct comparisons between those reported in the *Main Text* and here (all z-values ≤ 1.21, all p-values $\geq .23$, uncorrected).

Table S1: Impact of different methods of controlling for motion on significant group differences reported in the *Main Text*.

Analysis	Main Text	FD	DVARS	Unscrubbed	Matched
	F (p); df=1,78	F (p); df=1,70	F (p); df=1,70	F (p); df=1,78	F (p); df=1,50
Static FC					
Whole-network PC	4.08 (.047)	4.05 (.048)*	4.02 (.049)*	2.16 (.146)	0.12 (.728)
Whole-network WD	6.45 (.013)	6.16 (.016)*	6.04 (.016)*	1.73 (.192)	0.01 (.911)
DMN-FP PC	8.81 (.016)	8.71 (.017)*	8.46 (.020)*	7.80 (.026)*	3.25 (.309)

Results of significant group difference ANCOVAs covarying for mean functional connectivity (FC) and age as reported in the *Main Text*, when additionally controlling for mean FD (FD) or DVARS (DVARS), when inputting unscrubbed, instead of scrubbed, data (Unscrubbed), and when using a motion-matched subgroup of subjects (Matched). Whole-network includes DMN, FP, SAL, SM and SUB networks. All p-values are FDR-corrected for multiple comparisons as in the *Main Text*. * = consistent significant group difference when compared to main analysis (corrected $p < .05$).

Table S2: Impact of different methods of controlling for motion on significant correlations between brain metrics and behavior in participants with ADHD reported in the *Main Text*.

Analysis	Main Text	FD	DVARS	Unscrubbed	Matched
	r (p)	r (p)	r (p)	r (p)	r (p)
Static FC					
Whole-network PC	0.35 (.049)	0.31 (.099)$^{\sim}$	0.34 (.068)$^{\sim}$	0.29 (.126)	0.41 (.065)$^{\sim}$
DMN-SAL PC	0.36 (.043)	0.28 (.155)	0.34 (.068)$^{\sim}$	0.30 (.122)	0.37 (.080)$^{\sim}$
DMN-SM PC	0.38 (.043)	0.38 (.066)$^{\sim}$	0.38 (.065)$^{\sim}$	0.30 (.122)	0.42 (.080)$^{\sim}$
Time-Varying FC					
Whole-network CVPC	-0.41 (.008)	-0.36 (.024)*	-0.39 (.012)*	N/A	-0.51 (.007)*

Results of significant partial correlations between brain metrics and go/no-go (GNG) commission errors (ComErr) covarying for mean functional connectivity (FC) and age as reported in the *Main Text*, when additionally controlling for mean FD (FD) or DVARS (DVARS), when inputting unscrubbed, instead of scrubbed, data (Unscrubbed), and when using a motion-matched subgroup of subjects (Matched). Whole-network includes DMN, FP, SAL, SM and SUB networks. Unscrubbed sensitivity analysis not relevant to time-varying FC, as *Main Text* analyses are on unscrubbed data. All p-values are FDR-corrected for multiple comparisons as in the *Main Text*. * = consistent significant correlation when compared to main analysis (corrected $p < .05$); $^{\sim}$ = nonsignificant trend toward a correlation in the same direction when compared to main analysis (corrected $p < .10$).

3.5. Summary of Sensitivity Analysis Findings

As stated above, we conducted a series of sensitivity analyses to ensure that the results reported in the *Main Text* were not driven by head motion. Overall, results were remarkably similar when controlling for individual subject motion (using either FD or DVARS), and generally consistent when using unscrubbed timeseries or when including a motion-matched group of participants (Tables S1 and S2). While not all results were statistically significant, they were uniformly in the same direction and similar in magnitude to those reported in the primary analyses in the *Main Text*. As stated above in each individual section, no results from the sensitivity analyses significantly changed from those reported in the Main Text. Thus, nonsignificant results are not likely to have been due to significant effects of head motion artifacts in the analyses reported in the Main Text. Instead, they are more likely to have been due to a combination of reduced degrees of freedom in the case of controlling for additional variables, increased noise in the unscrubbed data,

and reduced power in the motion-matched subsample. As motion was correlated with symptoms of ADHD in our subjects (hyperactivity/impulsivity: r = 0.29, corrected p = .016; inattention: r = 0.30, corrected p = .016), it is likely that by covarying for mean FD or DVARS we removed signal in addition to noise from our models. Moreover, restricting our ADHD group to one with reduced motion in our motion-matched analysis likely resulted in a non-representative sample of children with ADHD, which could explain the reduction in group differences we observed. Non-significant results using the unscrubbed data could have been due to the increased noise inherent in unscrubbed data, particularly as related to motion artifacts (Ciric et al., 2017; Parkes et al., 2018). Critically, as it was necessary to use unscrubbed data for the time-varying FC analyses to ensure temporal contiguity, all sensitivity analyses of the time-varying analyses that included additional methods to control for motion remained statistically significant (all corrected p-values $< .025$ for the correlation between whole-network time-varying PC and ComErr in participants with ADHD). In other words, even though unscrubbed data is noisier and more susceptible to motion artifacts, when carefully controlling for motion the time-varying results remained consistent. Given the consistency between the above sensitivity analyses and the results reported in the *Main Text*, we are confident that our findings are not spurious as a result of increased head motion in the ADHD, as compared to the TD, participants

4. Static Functional Connectivity of non-DMN Pairs of Networks

Our main analyses focused on integration between the DMN and other networks included in our analyses (FP, SAL, SM, and SUB). To determine whether our results were specific to the DMN, we conducted exploratory post-hoc analyses quantifying group differences in PC, our measure of network integration, for all non-DMN network pairs (FP-SAL, FP-SM, FP-SUB, SAL-SM, SAL-SUB, SM-SUB). We used the same methods as described in the *Analyses* section of the *Main Text* (one-way ANCOVAs covarying for mean FC and age to assess group differences, and partial correlations controlling for mean FC and age to relate graph metrics to behavior). We corrected the following tests for six comparisons (one per network pair) using an FDR correction.

We found significantly higher PC in ADHD as compared to TD participants for the FP-SM and FP-SUB network pairs (FP-SM: mean ADHD = 0.447, mean TD = 0.428, $F(1,78) = 6.37$, corrected p = .048; FP-SUB: mean ADHD = 0.416, mean TD: 0.403, $F(1,78) = 6.07$, corrected p = .048), with a trend toward higher PC in ADHD for the SAL-SM network pair (mean ADHD = 0.460, mean TD = 0.449, $F(1,78) = 4.46$, corrected p = .076). All other corrected p-values $>$ .29. While we observed no significant correlations between PC and ComErr for any of the non-DMN network pairs, we did observe trends toward correlations between FP-SUB PC and ComErr (r = 0.35, corrected p = .077) and between SAL-SUB PC and ComErr (r = 0.36, corrected p = .077) in the participants with ADHD. All other corrected p-values $>$ .32. The correlation between SAL-SUB PC and ComErr was significantly stronger in the ADHD than in the TD participants (z = 2.67, p = .008), while there was no significant difference across groups in the correlation between FP-SUB PC and ComErr (z = 1.25, p = .21). Finally, we observed a trend toward a correlation in TD participants between FP-SAL PC and ComErr (r = 0.39, corrected p = .072), with all other corrected p-values $>$ .36. Although this relationship was not observed in the participants with ADHD, the magnitude of the relationship did not differ between groups (z = 0.83, p = .41). Results are summarized in Table S3.

Table S3: Mean participation coefficient (PC) and correlations with commission errors (ComErr) for each non-DMN network pair for ADHD and TD participants.

Networks	ADHD	TD	Group difference		Correlations	
			F-value (df=1,78)	p-value (FDR-corr)	ADHD ComErr r (FDR-corr p)	TD ComErr r (FDR-corr p)
FP-SAL	0.458	0.457	0.05	.969	0.22 (.330)	0.39 (.072)˜
FP-SM	0.447	0.428	6.37	.048*	-0.04 (.797)	-0.04 (.803)
FP-SUB	0.416	0.403	6.07	.048*	0.35 (.077)˜	0.08 (.767)
SAL-SM	0.460	0.449	4.46	.076˜	-0.19 (.362)	-0.21 (.370)
SAL-SUB	0.391	0.385	1.71	.292	0.36 (.077)˜	-0.22 (.370)
SM-SUB	0.369	0.369	0.001	.969	-0.17 (.362)	0.07 (.767)

Group difference ANCOVAs and Pearson correlations covarying for mean functional connectivity (FC) and age. All p-values are FDR-corrected for multiple comparisons. * = significant ($p < .05$); ˜ = nonsignificant trend ($p < .10$)

5. Static Functional Connectivity Using DMN Subnetworks

Some literature has demonstrated that the DMN may consist of several subnetworks that interact differentially with other brain networks (Andrews-Hanna et al., 2010; Buckner and DiNicola, 2019; Dixon et al., 2017). To explore this possibility in our participants, we used community detection to identify subnetworks of the predefined DMN nodes used in the main analysis (Power et al., 2011). Consensus clustering was performed on an average network consisting of all participants (TD and ADHD combined; Lanchichinetti and Fortunato, 2012). To perform the consensus clustering, first connectivity matrices of DMN nodes were averaged across all participants and thresholded at 0 (i.e., only positive connections were retained). The Louvain community detection algorithm was implemented 150 times on the average, weighted graph using a resolution parameter (γ) ranging from 1 to 1.5 in steps of 0.25. Next, an agreement matrix was constructed, with each cell containing the proportion of times a given pair of nodes was assigned to the same network. The weighted agreement matrix was thresholded at 0.5, indicating that a given pair of nodes was assigned to the same network at least 50% of the time. Finally, a consensus partition was obtained from the agreement matrix by implementing the Louvain community detection algorithm 100 times on the agreement matrix to obtain a single, representative partition. The Brain Connectivity Toolbox was used to partition the DMN into subnetworks (www.brain-connectivity-toolbox.net; Rubinov and Sporns, 2010). A γ value of 1.25 returned 3 subnetworks, which is in line with what previous studies have reported (Andrews-Hanna et al., 2010; Dixon et al., 2017), and thus all analyses were conducted using subnetworks defined from a γ of 1.25. Following community detection, these three DMN subnetworks, along with the 4 predefined task-relevant networks described in the *Main Text* (FP, SAL, SM, and SUB), were used in analyses to probe whether different DMN subnetworks differentially interacted with each of the task-relevant networks. To do so, we repeated static FC analyses investigating inter-network integration between individual network pairs involving the DMN, separately for each DMN subnetwork. That is, we conducted one-way ANCOVAs with group as the factor and mean FC and age as the covariates to compare mean PC of each network pair across groups. Pearson's correlations (partial correlations controlling for mean FC and age) were conducted to relate PC to ComErr on the GNG task, separately for each group. All analyses were FDR-corrected for 12 comparisons, corresponding to the number of network pairs tested (3 DMN sub-networks x 4 task-relevant networks). Results are presented in Table S4.

Table S4: Mean participation coefficient (PC) and correlations with commission errors (ComErr) for each DMN subnetwork pair for ADHD and TD participants.

Networks	ADHD	TD	Group difference		Correlations	
			F-value (df=1,78)	p-value (FDR-corr)	ADHD ComErr r (FDR-corr p)	TD ComErr r (FDR-corr p)
DMN1-FP	0.426	0.408	6.42	.058˜	0.30 (.092)˜	-0.04 (.941)
DMN2-FP	0.449	0.424	13.08	.003*	0.35 (.058)˜	-0.03 (.941)
DMN3-FP	0.423	0.383	17.24	.001*	0.18 (.294)	0.05 (.941)
DMN1-SAL	0.348	0.347	0.01	.975	0.31 (.092)˜	-0.07 (.941)
DMN2-SAL	0.424	0.411	3.01	.173	0.48 (.016)*	-0.01 (.967)
DMN3-SAL	0.378	0.350	5.70	.058˜	0.36 (.058)˜	0.04 (.941)
DMN1-SM	0.363	0.372	0.72	.597	0.43 (.033)*	-0.03 (.941)
DMN2-SM	0.407	0.407	0.001	.975	0.37 (.058)˜	-0.09 (.941)
DMN3-SM	0.395	0.401	0.36	.733	0.27 (.112)	-0.05 (.941)
DMN1-SUB	0.408	0.403	0.22	.764	0.13 (.413)	-0.05 (.941)
DMN2-SUB	0.404	0.393	1.55	.372	0.30 (.092)˜	-0.06 (.941)
DMN3-SUB	0.398	0.377	4.33	.098˜	0.21 (.214)	0.04 (.941)

Group difference ANCOVAs and partial correlations covarying for mean functional connectivity (FC) and age. All p-values are FDR-corrected for multiple comparisons. * = significant ($p < .05$); ˜ = nonsignificant trend ($p < .10$)

While the three subnetworks of the DMN identified in our participants involved nodes that were largely intermixed, there were some differences in node distribution (Figure S2). For example, DMN2 involved mostly anterior and posterior medial nodes, including regions considered to constitute the core DMN (anterior medial prefontal cortex, posterior cingulate cortex; Andrews-Hanna et al., 2010). DMN1 and DMN3, on the other hand, involved more lateral nodes. The nodes of DMN1 were mostly posterior and included the posterior cingulate cortex, precuneus, lateral occipital cortex, lateral temporal cortex, and parahippocampal gyrus. These regions span those considered to be part of multiple DMN subsystems as defined in extant literature (Andrews-Hanna et al., 2010; Dixon et al., 2017). Similarly, the nodes of DMN3 nodes spanned multiple DMN subsytems. Nodes of DMN3 were more distributed throughout both anterior and posterior regions as compared to DMN1, and included lateral occipital cortex, lateral temporal cortex, temporal fusiform cortex, lateral parietal cortex, anterior medial prefrontal cortex, and ventromedial prefrontal cortex. Notably, while initial literature defining DMN subsystems observed three spatially distinct subsystems (medial core, dorsal medial subsystem, and medial temporal subsystem; Andrews-Hanna et al., 2010; Dixon et al., 2017), more recent literature has observed parallel, spatially similar subsystems (Buckner and DiNicola, 2019). The DMN subsystems we observed are more in line with parallel subsystems, although literature to date has used data from healthy young adult participants and used different regions of interest, thus future research is needed to determine whether the differences we observed are due to methodological differences or participant characteristics.

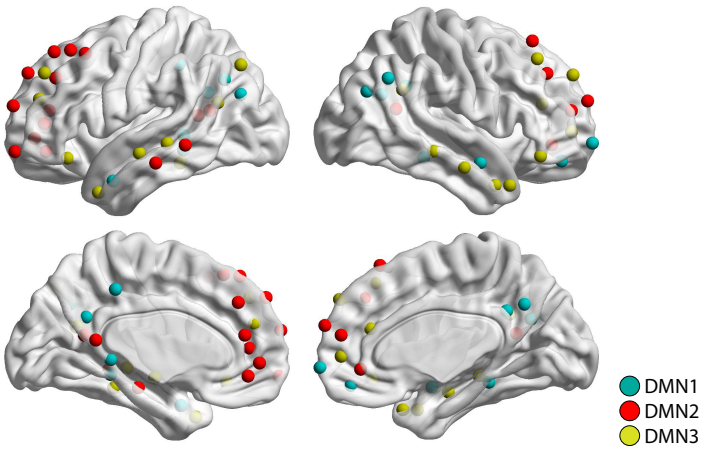

Figure S2: Regions of interest (ROIs) depicting the three default mode network (DMN) subnetworks. ROIs taken from Power et al., 2011.

We assessed group differences in PC between each DMN subnetwork and the four task-related networks. Consistent with the significant DMN-FP group difference observed when considering the DMN as a unitary network (corrected p-value = .016), we observed that participants with ADHD had significantly higher PC between the FP network and both DMN2 (mean ADHD: 0.449, mean TD: 0.424, $F(1,78) = 13.08$, corrected $p = .003$) and DMN3 (mean ADHD: 0.423, mean TD: 0.383, $F(1,78) = 17.24$, corrected $p = .001$), with a nonsignificant trend in the same direction between FP and DMN1 (mean ADHD: 0.426, mean TD: 0.408, $F(1,78) = 6.42$, corrected $p = .058$). Additionally, there were nonsignificant trends toward higher PC in participants with ADHD between DMN3 and both the SAL (mean ADHD: 0.378, mean TD: 0.350, $F(1,78) = 5.70$, corrected $p = .058$) and the SUB (mean ADHD: 0.398, mean TD: 0.377, $F(1,78) = 4.33$, corrected $p = .098$) networks.

Next we assessed relationships between inter-network PC and ComErr on the GNG task, separately for ADHD and TD participants. In analyses for which we considered the DMN a unitary network, we observed significant positive correlations in participants with ADHD between DMN-SAL PC and ComErr and between DMN-SM PC and ComErr (both corrected p-values = .043). Consistent with those results, in participants with ADHD we observed significant positive correlations between DMN2-SAL PC and ComErr ($r = 0.48$, corrected $p = .016$) and between DMN1-SM and ComErr ($r = 0.43$, corrected $p = .033$). There were nonsignificant trends in the same direction for DMN1-SAL and DMN3-SAL PC and ComErr ($r = 0.31$, corrected $p = .092$ and $r = 0.36$, corrected $p = .058$ respectively), as well as for DMN2-SM and ComErr ($r = 0.37$, corrected $p = .058$). The relationship between DMN3-SM PC and ComErr was nonsignificant but in the same direction ($r = 0.27$, corrected $p = .11$). Additionally, there were nonsignificant trends toward positive correlations with ComErr in participants with ADHD for DMN1-FP PC ($r = 0.30$, corrected $p = .092$), DMN2-FP PC ($r = 0.35$, corrected $p = .058$), and DMN2-SUB PC ($r = 0.30$, corrected $p = .092$). All other corrected p-values $> .21$. Finally, consistent with results from the *Main Text* in which we considered the DMN a unitary network, there were no correlations between PC of any network pairs and ComErr in the TD participants (all corrected p-values $> .94$).

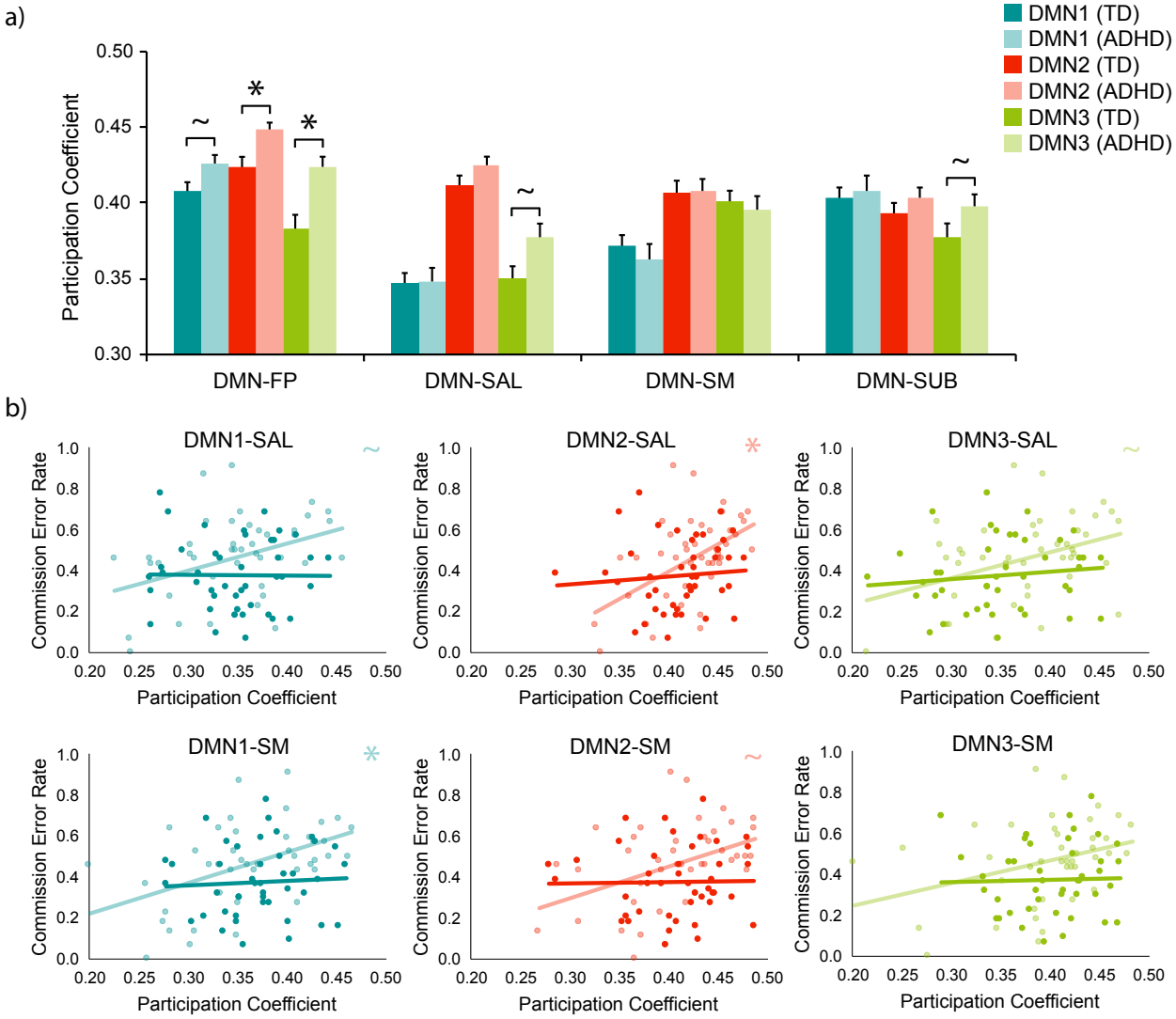

Figure S3: a) Group difference in PC separately for each DMN subnetwork pair. There was a significant group difference between the DMN2 and DMN3 subnetworks and the FP network (corrected $p = .003$ and .001 respectively), with a trend toward a group difference between the DMN1 subnetwork and the FP network (corrected $p = .058$). There were additionally trends toward group differences for the DMN3-SAL and DMN3-SUB network pairs (corrected $p = .058$ and .098 respectively). b) Relationships between PC and ComErr separately for the DMN-SAL and DMN-SM network pairs, separately for each DMN subnetwork. There were significant positive correlations between DMN2-SAL PC and DMN1-SM PC and ComErr in the participants with ADHD (corrected $p = .016$ and .033 respectively). There were additionally trends toward correlations between DMN1-SAL PC, DMN3-SAL PC, and DMN2-SM PC and ComErr in the participants with ADHD (corrected $p = .092$, .058, and .058 respectively). There were no relationships between PC and ComErr in the TD participants (all FDR-corrected p-values $> .94$). * indicates a relationship is significant at FDR-corrected $p < .05$. ~ indicates a nonsignificant trend at FDR-corrected $p < .10$.

To summarize, we observed that relationships between each DMN subnetwork and the task-relevant networks were largely consistent across subnetworks (Figure S3). While network structure is broadly in place by childhood, changes occur into young adulthood (Grayson and Fair, 2017), including increases in network segregation (Gu et al., 2015). Therefore, it is possible that the DMN has not yet differentiated into distinct subnetworks by middle childhood. Additionally, extant literature in adults has found that functional connectivity patterns of the DMN reconfigure between rest

and task contexts (Dixon et al., 2017; Fornito et al., 2012), with greater distinction between subnetworks during task. Thus, it is possible that our network analysis, assessed during the resting state, does not identify functionally dissociable subnetworks within the DMN. Future research quantifying network organization during both rest and task performance can clarify whether reliable DMN subnetworks can be identified in children.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8:14.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., and Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, 65(4):550–562.

Avants, B., Epstein, C., Grossman, M., and Gee, J. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41.

Buckner, R. L. and DiNicola, L. M. (2019). The brain's default network: updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience*, 20(10):593–608.

Ciric, R., Rosen, A. F. G., Erus, G., Cieslak, M., Adebimpe, A., Cook, P. A., Bassett, D. S., Davatzikos, C., Wolf, D. H., and Satterthwaite, T. D. (2018). Mitigating head motion artifact in functional connectivity MRI. *Nature Protocols*, 13(12):2801–2826.

Ciric, R., Wolf, D. H., Power, J. D., Roalf, D. R., Baum, G. L., Ruparel, K., Shinohara, R. T., Elliott, M. A., Eickhoff, S. B., Davatzikos, C., Gur, R. C., Gur, R. E., Bassett, D. S., and Satterthwaite, T. D. (2017). Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage*, 154:174–187.

Couvy-Duchesne, B., Ebejer, J. L., Gillespie, N. A., Duffy, D. L., Hickie, I. B., Thompson, P. M., Martin, N. G., de Zubicaray, G. I., McMahon, K. L., Medland, S. E., and Wright, M. J. (2016). Head motion and inattention/hyperactivity share common genetic influences: implications for fMRI studies of ADHD. *PLoS ONE*, 11(1):e0146271.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29(3):162–173.

Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis. *NeuroImage*, 9(2):179–194.

Dixon, M. L., Andrews-Hanna, J. R., Spreng, R. N., Irving, Z. C., Mills, C., Girn, M., and Christoff, K. (2017). Interactions between the default network and dorsal attention network vary across default subsystems, time, and cognitive states. *NeuroImage*, 147:632–649.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., and Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1):111–116.

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., and Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47(Suppl 1):S102.

Fornito, A., Harrison, B. J., Zalesky, A., and Simons, J. S. (2012). Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proceedings of the National Academy of Sciences*, 109(31):12788–12793.

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., and Ghosh, S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Frontiers in Neuroinformatics*, 5:13.

Gorgolewski, K. J., Esteban, O., Ellis, D. G., Notter, M. P., Ziegler, E., Johnson, H., Hamalainen, C., Yvernault, B., Burns, C., Manhães-Savio, A., Jarecka, D., Markiewicz, C. J., Salo, T., Clark, D., Waskom, M., Wong, J., Modat, M., Dewey, B. E., Clark, M. G., Dayan, M., Loney, F., Madison, C., Gramfort, A., Keshavan, A., Berleant, S., Pinsard, B., Goncalves, M., Clark, D., Cipollini, B., Varoquaux, G., Wassermann, D., Rokem, A., Halchenko, Y. O., Forbes, J., Moloney, B., Malone, I. B., Hanke, M., Mordom, D., Buchanan, C., Pauli, W. M., Huntenburg, J. M., Horea, C., Schwartz, Y., Tungaraza, R., Iqbal, S., Kleesiek, J., Sikka, S., Frohlich, C., Kent, J., Perez-Guevara, M., Watanabe, A., Welch, D., Cumba, C., Ginsburg, D., Eshaghi, A., Kastman, E., Bougacha, S., Blair, R., Acland, B., Gillman, A., Schaefer, A., Nichols, B. N., Giavasis, S., Erickson, D., Correa, C., Ghayoor, A., Küttner, R., Haselgrove, C., Zhou, D., Craddock, R. C., Haehn, D., Lampe, L., Millman, J., Lai, J., Renfro, M., Liu, S., Stadler, J., Glatard, T., Kahn, A. E., Kong, X.-Z., Triplett, W., Park, A., McDermottroe, C., Hallquist, M., Poldrack, R., Perkins, L. N., Noel, M., Gerhard, S., Salvatore, J., Mertz, F., Broderick, W., Inati, S., Hinds, O., Brett, M., Durnez, J., Tambini, A., Rothmei, S., Andberg, S. K., Cooper, G., Marina, A., Mattfeld, A., Urchs, S., Sharp, P., Matsubara, K., Geisler, D., Cheung, B., Floren, A., Nickson, T., Pannetier, N., Weinstein, A., Dubois, M., Arias, J., Tarbert, C., Schlamp, K., Jordan, K., Liem, F., Saase, V., Harms, R., Khanuja, R., Podranski, K., Flandin, G., Papadopoulos Orfanos, D., Schwabacher, I., McNamee, D., Falkiewicz, M., Pellman, J., Linkersdörfer, J., Varada, J., Pérez-García, F., Davison, A., Shachnev, D., and Ghosh, S. (2017). *Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. 0.13.1*. Zenodo.

Grayson, D. S. and Fair, D. A. (2017). Development of large-scale functional networks from birth to adulthood: a guide to the neuroimaging literature. *NeuroImage*, 160:15–31.

Greve, D. N. and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72.

Gu, S., Satterthwaite, T. D., Medaglia, J. D., Yang, M., Gur, R. E., Gur, R. C., and Bassett, D. S. (2015). Emergence of system roles in normative neurodevelopment. *Proceedings of the National Academy of Sciences*, 112(44):13681–13686.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841.

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Neto, E. C., and Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLoS Computational Biology*, 13(2):e1005350.

Lanchichinetti, A. and Fortunato, S. (2012). Consensus clustering in complex networks. *Scientific Reports*, 2:336.

Parkes, L., Fulcher, B., Yücel, M., and Fornito, A. (2018). An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage*, 171:415–436.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154.

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., and Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, 72(4):665–678.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, 84:320–341.

Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, 52(3):1059–1069.

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., and Gee, J. C. (2010). N4iTK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320.

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57.