HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2023 September 27.

Published in final edited form as:

Neuroimage. 2023 August 15; 277: 120238. doi:10.1016/j.neuroimage.2023.120238.

Effects of phase encoding direction on test-retest reliability of human functional connectome

Hengyi Cao[a,b,c,*], **Anita D. Barber**[a,b,c], **Jose M. Rubio**[a,b,c], **Miklos Argyelan**[a,b,c], **Juan A. Gallego**[a,b,c], **Todd Lencz**[a,b,c], **Anil K. Malhotra**[a,b,c]

[a]Institute of Behavioral Sciences, Feinstein Institutes for Medical Research, Manhasset, NY, United States

[b]Division of Psychiatry Research, Zucker Hillside Hospital, 265-16 74th Avenue, Glen Oaks, NY 11004, United States

[c]Department of Psychiatry, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States

Abstract

The majority of human connectome studies in the literature based on functional magnetic resonance imaging (fMRI) data use either an anterior-to-posterior (AP) or a posterior-to-anterior (PA) phase encoding direction (PED). However, whether and how PED would affect test-retest reliability of functional connectome is unclear. Here, in a sample of healthy subjects with two sessions of fMRI scans separated by 12 weeks (two runs per session, one with AP, the other with PA), we tested the influence of PED on global, nodal, and edge connectivity in the constructed brain networks. All data underwent the state-of-the-art Human Connectome Project (HCP) pipeline to correct for phase-encoding-related distortions before entering analysis. We found that at the global level, the PA scans showed significantly higher intraclass correlation coefficients (ICCs) for global connectivity compared with AP scans, which was particularly prominent when using the Seitzman-300 atlas (versus the CAB-NP-718 atlas). At the nodal level, regions most strongly affected by PED were consistently mapped to the cingulate cortex, temporal lobe, sensorimotor areas, and visual areas, with significantly higher ICCs during PA scans compared with AP scans, regardless of atlas. Better ICCs were also observed during PA scans at the edge level, in particular when global signal regression (GSR) was not performed.

[*]Corresponding author: hcao2@northwell.edu (H. Cao).

Authors' statement

Study conceptualization: Hengyi Cao

Data analysis and Methodology: Hengyi Cao, Miklos Argyelan

Data collection: Juan Gallego

Data interpretation and critical comments: Hengyi Cao, Anita Barber, Jose Rubio, Miklos Argyelan, Todd Lencz, Anil Malhotra

Paper writing: Hengyi Cao

Paper review & editing: All authors

Funding acquisition: Anil Malhotra

Declaration of Competing Interest

Dr. Malhotra is a consultant for Genomind, Inc., InformedDNA, Janssen Pharma, and Acadia Pharma. The other authors report no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120238.

Further, we demonstrated that the observed reliability differences between PEDs may relate to a similar effect on the reliability of temporal signal-to-noise ratio (tSNR) in the same regions (that PA scans were associated with higher reliability of tSNR than AP scans). Averaging the connectivity outcome from the AP and PA scans could increase median ICCs, especially at the nodal and edge levels. Similar results at the global and nodal levels were replicated in an independent, public dataset from the HCP-Early Psychosis (HCP-EP) study with a similar design but a much shorter scan session interval. Our findings suggest that PED has significant effects on the reliability of connectomic estimates in fMRI studies. We urge that these effects need to be carefully considered in future neuroimaging designs, especially in longitudinal studies such as those related to neurodevelopment or clinical intervention.

Keywords

Test-retest reliability; Functional connectome; Phase encoding direction; fMRI

1. Introduction

The broad application of functional magnetic resonance imaging (fMRI) in the study of human functional connectome at the macroscale has considerably promoted the understanding of the brain function and organization in healthy individuals (Power et al., 2011), neurodevelopment (Somerville et al., 2018), aging (Bookheimer et al., 2019), and among different mental disorders (Baker et al., 2019; Cao et al., 2020; Cao et al., 2021b; Dong et al., 2018; Ilioska et al., 2022; Javaheripour et al., 2021). Albeit the fruitful discoveries so far, a key consideration in the connectomic research is test-retest reliability, which quantifies the consistency of the brain connectivity readouts across multiple assessments and therefore justifies the validity and practicality of the functional connectomic measures, in particular for studies with a longitudinal design. For this purpose, a great number of studies in the literature have sought to investigate test-retest reliability of the human functional connectome (Anderson et al., 2011; Cao et al., 2019; Cao et al., 2014; Noble et al., 2019; Noble et al., 2017a; Pannunzi et al., 2017; Shah et al., 2016; Shehzad et al., 2009), and to identify factors that may help to improve the reliability of the outcome (Birn et al., 2013; Cao et al., 2019; Noble et al., 2017b; Pannunzi et al., 2017; Yoo et al., 2019). While reliability results reported in these prior studies are variable, it has generally been accepted that factors such as scan length (Anderson et al., 2011; Birn et al., 2013; Gordon et al., 2017; Noble et al., 2017b), global signal regression (GSR) (Cao et al., 2019; Noble et al., 2019; Song et al., 2012), brain atlas (Cao et al., 2019; Cao et al., 2014; Noble et al., 2019), and type of outcome measure (multivariate vs univariate)(Noble et al., 2019; Noble et al., 2017b; Pannunzi et al., 2017; Yoo et al., 2019) may play a critical role in tuning the connectome reliability. Specifically, better reliability tends to be acquired from longer scan time, without GSR, more brain nodes with finer parcellation, and the assessment of multivariate or summarized connectivity scores compared with single connectivity strength.

Phase encoding (PE) is an important technique to pinpoint the spatial location of voxel signals along the y-axis during the fMRI scans, is achieved by applying a PE gradient to impose a specific phase angle to a transverse magnetization vector. The most frequently

used PE directions (PED) in the fMRI studies are either from anterior to posterior (hereafter "AP") or from posterior to anterior (hereafter "PA"). Different PEDs are known to have different impacts on susceptibility-induced distortion and signal loss in the brain, which may in turn influence image quality and the reliability of functional connectome constructed from these images. Specifically, previous studies have shown that AP scans are commonly associated with greater signal loss in the orbitofrontal cortex, anterior cingulate cortex (ACC), and temporal lobe (Mori et al., 2018; Wang et al., 2021; Weiskopf et al., 2006; Weiskopf et al., 2007). Less consistent findings are reported for the PA scans, which may relate to worse signals in the frontal pole, striatum, and cerebellum (Mori et al., 2018; Wang et al., 2021). Although field maps have increasingly been collected and used with the aim of correcting for PED-related distortion, the efficacy of such corrections has not been fully investigated. This raises concern as whether PED would have significant effects on test-retest reliability of the derived connectomic measures. Notably, recent studies have demonstrated significant PED-related effects on the connectivity outcome in comparison of patients with schizophrenia and healthy controls (Mori et al., 2018) as well as contrasting males and females (Wang et al., 2021), further strengthening the possibility of such concern.

When designing a neuroimaging study, investigators need to make decisions about which PED to use, despite uncertainty about how results may be affected by specific directions. As a result, many recent studies have acquired data with both directions in consecutive runs in the same session (Bookheimer et al., 2019; Demro et al., 2021; Somerville et al., 2018). Systematic differences in reliability between PEDs would impact the accuracy of the detected outcome, and therefore it might be advisable to recommend prioritizing one PED over the other when designing experiments. In this study, we examined the effects of PED on test-retest reliability of fMRI-based functional connectomic measures in healthy young adults, using two independent datasets with repeated fMRI scans. In the discovery dataset, participants received two sessions of scans with 12-week apart; while in the replication dataset as part of the Human Connectome Project Early Psychosis (HCP-EP) study, fMRI scans were acquired twice on the same day separated by approximately half an hour. During each scan session in both datasets, both AP and PA images were acquired that allows direct comparison of the reliability between the two PEDs. For each PED, we constructed connectomes using two state-of-the-art functional brain atlases (the Seitzman atlas including 300 parcels (Seitzman et al., 2020) and the CAB-NP atlas including 718 parcels (Ji et al., 2019)), both with and without GSR, and further tested if there could be any interactive effects between the PED and these different processing strategies. All data went through the Human Connectome Project (HCP) pipeline to correct for FE-related effects before entering the reliability analysis.

2. Methods and materials

2.1. Subjects

Two independent healthy young adult datasets with repeated fMRI scans were included in the study. The discovery dataset consisted of 32 healthy participants (age 28.1±3.9 years, 13 males). The exclusion criteria included: 1) lifetime history of any psychiatric disorders as determined by Structured clinical Interview for DSM-5 (SCID), non-patient

version; 2) lifetime history of neurological disorders; 3) mental retardation; and 4) MR imaging contraindications. All participants provided written informed consent for protocols approved by the Institutional Review Board of Northwell Health. The replication dataset was part of the HCP-EP study (https://www.humanconnectome.org/study/human-connectome-project-for-early-psychosis/), including 49 healthy subjects (age 24.9±4.3 years, 30 males) who provided their written consents for protocols approved by Harvard University and Indiana University.

2.2. Study design and data acquisition

In the discovery dataset, each participant underwent two sessions of multi-paradigm fMRI scans with a time interval of 12 weeks. During each session, resting-state scans and two task-fMRI scans were performed (an event-related cognitive control task and a block-designed reward processing task). For each paradigm, two runs of data were collected with different phase encoding directions (one with AP and the other with PA) in pseudorandom order, with the sequence of the two runs counterbalanced across the sample. An overall diagram for study design is presented in Fig. 1. Here, resting state included two 7-min eyes-closed runs. The cognitive control paradigm was based on the Multi-Source Interference Task (MSIT (Bush and Shin, 2006)), during which subjects were shown with three numbers (1, 2, or 3) at each trial. They were asked to select the number that is different from the other two, while ignoring the location of the number. The task lasted for a total of 8 min separated by two runs (4 min with AP and 4 min with PA). The reward processing task was similar to the one used in the Human Connectome Project (Barch et al., 2013), where subjects were asked to guess whether the number to emerge after a question mark was above or below five during each trial. They either won or lost money based on the correctness of their responses. The task lasted for 6 min with 3-min AP scans and 3-min PA scans. Notably, all of the 32 subjects completed the resting-state scans for both sessions with 12-week apart, while only 20 and 23 subjects completed the cognitive control and reward processing tasks for both sessions, respectively. Due to differences in scan length and sample size between the three paradigms, we did not directly compare reliability outcomes between fMRI paradigms but rather treated paradigm as a random-effect variable throughout the entire study.

In the replication dataset, two sessions of resting-state scans were conducted on the same day (separated by approximately half an hour). During each session, two runs of data were collected (one with AP, the other with PA), each lasting for 5.6 min. Notably, we used our own dataset as the discovery sample and the HCP-EP as the replication sample due to following reasons: 1) Our own data were scanned 12-week apart, which resembles a typical clinical longitudinal study to investigate interventional effects; 2) Our own data had longer resting-state scan time (7 min per PED) compared with the HCP-EP (5.6 min per PED); and 3) Except for resting state, we also collected data from task-based paradigms, while the HCP-EP only had resting-state data. Therefore, the HCP-EP was used as the replication sample to further validate whether a similar effect could be observed with a much shorter scan interval using only resting-state data.

Imaging data from both datasets were scanned on 3T SIEMENS Prisma scanners following the HCP protocol (Van Essen et al., 2012). Specifically, the BOLD images were acquired

with multi-band echo-planar imaging (EPI) sequence. Except for PED, all other parameters remained the same across all runs and sessions. The scan parameters are: 1) discovery sample: TR = 720 ms, TE = 33 ms, FA = 52 degree, slice thickness = 2 mm, 72 continuous slices, FOV = 231*231 mm, voxel size = 2.2*2.2*2 mm, multi-band factor = 8; 2) replication sample: TR = 800 ms, TE = 37 ms, FA = 52 degree, slice thickness = 2 mm, 72 continuous slices, FOV = 208*208 mm, voxel size = 2*2*2 mm, multi-band factor = 8.

2.3. Data preprocessing

All image data were preprocessed with the HCP pipeline (Glasser et al., 2013), including a total of five major steps (PreFreeSurfer, FreeSurfer, PostFreeSurfer, fMRI Volume, fMRI Surface). Briefly, structural images were corrected for gradient- and phase encoding-related distortions, aligned at the native space, and further registered to the standard MNI space. The distortion-corrected images were submitted to the FreeSurfer *recon-all* command to segment the volume into predefined structures, reconstruct white and pial cortical surfaces, and perform FreeSurfer's standard folding-based surface registration to their surface atlas. Similarly, functional images were first corrected for gradient and phase encoding distortions, realigned to reduce head motion, registered to the native space, and then normalized to the MNI space.

The preprocessed images were further scrutinized for head motion. In particular, we calculated the frame-wise displacements (FD) for each participant based on Power et al. (Power et al., 2012). Subjects with an average FD either > 0.5 mm or larger than the group mean plus three times the standard deviation during each paradigm were excluded for further analysis. This led to the rejection of one subject for the cognitive control task and one subject for the reward processing task in the discovery sample, as well as one subject in the replication sample.

2.4. Construction of whole-brain connectome

Two state-of-the-art atlases covering the entire brain were used to construct functional connectome matrices (the Seitzman-300 (Seitzman et al., 2020) and the CAB-NP-718 (Ji et al., 2019)). In brief, the Seitzman-300 atlas is an updated version of the widely used Power-264 atlas (Power et al., 2011) to include the subcortex and cerebellum, and the CAB-NP-718 atlas is an extension of the Glasser's HCP atlas (Glasser et al., 2016) to include the subcortex and cerebellum. In the Seitzman-300 atlas, a node was represented by a 5-mm radius sphere around a given coordinate defined by the atlas; while nodes in the CAB-NP-718 atlas were defined by brain parcels. For each node, time series were extracted and corrected for the effects of task-evoked coactivations (for task data), white matter and cerebrospinal fluid signals, 24 head motion parameters (i.e. 6 translational and rotational parameters, their first derivatives, and the squares of these 12 parameters), and FD. Here, the task-evoked coactivations were removed by using a linear regression model, in which each task condition was included as a regressor. The regressors were generated by convolving the task stimulus function with the standard hemodynamic response function. These noise-corrected time series were then temporally filtered (rest data: band pass 0.01–0.1 Hz; task data: high pass 0.01 Hz) and subsequently used to compute brain connectome matrices based on pairwise Pearson correlations. Notably, since whether to apply global

signal regression (GSR) is still an open question (Murphy and Fox, 2017), we calculated the connectome matrices both with and without GSR in this study. The above processing pipeline was kept the same for all scan sessions and runs (Fig. 1).

2.5. Reliability assessments

The reliability of the derived connectome matrices was assessed at three levels. At the global level, we calculated the measure of global functional connectivity, which is the grand mean of the whole connectome matrices. At the nodal level, reliability was evaluated for the connectivity strength of each single node (i.e. mean of connectivity between a given node and all other nodes in the matrices). At the edge level, reliability was computed for each single connection in the brain connectome matrices. Here, following the previous work, test-retest reliability was quantified by intraclass correlation coefficient (ICC(2,1)) based on the following formula, which gauges the absolute agreement of the measurements between two sessions.

$$ICC(2,1) = (BMS - EMS)/(BMS + (k-1)*EMS + k*(JMS - EMS)/n)$$

Here, BMS is the between-subject mean square, EMS is the residual mean square, JMS is the between-session mean square, n is the number of subjects, and k is the number of sessions. A larger positive ICC value (closer to 1) indicates a better agreement of measurements between two sessions and thereby higher test-retest reliability. Based on common definition (Cao et al., 2021a; Cao et al., 2019; Cao et al., 2014), an ICC of 0.4 and above indicates fair reliability of the examined measurement, and ICC > 0.6 indicates good reliability.

2.6. Statistics

The estimated ICC values were subsequently entered into linear mixed-effect models to determine potential effects of PED on these values. For global connectivity, PED (AP vs PA), atlas (Seitzman 300 vs CAB-NP 718), and GSR (with GSR vs without GSR) were included in the model as fixed-effect variables, with paradigm as random-effect variable. Effects of PED and interactions between PED and other fixed-effect variables (PED * atlas, PED * GSR) were estimated to decide how PED would affect the reliability of global connectivity and whether such effect could be influenced by different data processing strategies. For nodal and edge connectivity, a similar model was employed but separated by atlas, with PED and GSR modelled as fixed-effect variables and paradigm as random-effect variable. This was primarily applied to estimate the effects on median ICCs across all nodes or edges but also on single nodal and edge ICCs. Similar models were also used in the replication sample, except that no random-effect variable was modeled (only one paradigm).

3. Results

3.1. Reliability of global, nodal, and edge connectivity

The derived ICC values for the functional connectome at the global, nodal, and edge levels are present in Figs. 2, 3, and 4, respectively. At the global level, the ICCs of global functional connectivity ranged between 0.02 and 0.65, depending on PED, atlas, paradigm,

and whether GSR was applied (Fig. 2). In general, the lowest ICC was observed with AP scans and GSR when using the Seitzman-300 atlas (with the only exception of the cognitive control task), while relatively higher ICCs (>0.4) were detected with PA scans, especially when GSR was not used. An ICC of 0.4 and above was also evident for PA scans when GSR was applied, again with the cognitive control task as the only exception. In addition, when using the CAB-NP-718 atlas, we also observed fair to good reliability of global connectivity with AP scans and GSR, suggesting that there may be an atlas by PED interaction on the derived reliability measures.

In terms of functional connectivity for single nodes, relatively consistent findings were shown for both atlases and all three paradigms. Specifically, PA scans without GSR had the highest reliability, regardless of atlas and paradigm (Seitzman-300: median ICC > 0.41 for all paradigms; CAB-NP-718: median ICC > 0.40 for all paradigms), followed by AP scans without GSR (median ICC > 0.25 with Seitzman-300 and median > 0.27 with CAB-NP-718), while both AP and PA scans with GSR had relatively low reliability (median ICC > 0.16 for all paradigms with both atlases). The ICC distributions for all nodes in the constructed connectomes are present in Fig. 3A, 3C, and 3E. The median, first and third quartile ICCs across all nodes are present in Table 1.

In Fig. 3B, 3D, and 3F, we present the distribution of nodes with fair to good reliability (ICC > 0.4) in both atlases and all three paradigms. We found that during resting state, approximately 30% of total nodes in both atlases had fair reliability (ICC > 0.4) and only <1% of total nodes had good reliability (ICC > 0.6) with the AP scans. These numbers increased to >50% for nodes with fair reliability and ~10% for nodes with good reliability with the PA scans. During the cognitive control task, fairly reliable nodes comprised of 20% and 34% of total nodes in the Seitzman-300 and CAB-NP-718 atlases respectively with the AP scans and ~55% total nodes in both atlases with the PA scans. Similarly, proportion of nodes with good reliability increased from <7% with AP to ~10% with PA. In high consistency, the PA scans continued to show larger proportion of nodes with fair (>60%) and good (>10%) reliability during the reward processing task, compared with 18% and <1% with the AP scans. Spatially, nodes with highest reliability during the AP scans tended to be concentrated at the posterior part of the brain, chiefly the subcortex and cerebellum; while the reliable nodes during the PA scans had a more widespread distribution across the entire brain.

The Fig. 4 presents the ICC distribution of all edges in the examined connectomes. Consistently, for all paradigms and both atlases, PA without GSR had the highest overall reliability across all edges (median ICCs > 0.25, 0.23, and 0.20 for resting state, cognitive control, and reward processing, respectively, also see Table 1). When using the Seitzman-300 atlas without GSR, more than 30% of edges in the PA scans had fair reliability and more than 5% of edges had good reliability, regardless of paradigm. In contrast, only slightly above 20% and 3% of edges had fair and good reliability respectively in the AP scans. Similarly, when using the CAB-NP-718 atlas without GSR, more than 20% of edges demonstrated fair and good reliability during PA, compared to ~15% during AP.

3.2. Effects of PED on reliability measures

At the global level, significant effect of PED was detected on the ICCs of global functional connectivity ($P = 0.042$), with higher reliability for PA compared with AP (Fig. 5). Moreover, there was also a significant PED by atlas interactive effect on the ICC values ($P = 0.038$), where significantly higher reliability in PA was particularly observed with the Seitzman-300 atlas ($P = 0.028$) but not with the CAB-NP-718 atlas ($P = 0.96$), suggesting that the effect of PED on reliability of global connectivity is atlas-dependent. In contrast, no significant PED by GSR interactive effect was observed.

Similarly, at the nodal level, significant PED effect was observed on the median nodal ICCs with the Seitzman-300 atlas (PA > AP, $P = 0.026$) but not with the CAB-NP-718 atlas ($P = 0.33$, Fig. 6A). In terms of single nodes, ICCs of connectivity strength in 63 nodes within the Seitzman-300 atlas and in 107 nodes within the CAB-NP-718 atlas were significantly affected by PED at $P < 0.05$. While none of these effects survived false discovery rate (FDR) correction, nodes with largest effect size primarily mapped to the cingulate cortex, temporal cortex, sensorimotor areas, and visual cortex in both atlases. The top ten nodes ranked by P value in each atlas are present in Fig. 6B. Notably, for all of the ten top nodes in both atlases, their ICCs were significantly higher in PA compared with AP (Seitzman-300: $P = 0.001$; CAB-NP-718: $P < 0.001$), suggesting that PA scans may generally boost the reliability of functional connectivity at the nodal level.

The significant PED effect observed with the Seitzman-300 atlas also applied to the median edge ICCs ($P = 0.01$, Fig. 7A). Moreover, the effect was also close to significant with the CAB-NP-718 atlas at the edge level ($P = 0.06$). For both atlases, PA scans were found to be more reliable than AP scans, with a significant PED by GSR interaction also detected ($P < 0.035$), in which the PED effect was most prominent without GSR. For single edges, the top ten affected edges in the Seitzman-300 atlas involved connections between the frontal cortex and the subcortex, sensorimotor area, and temporal cortex; while the top ten affected edges in the CAB-NP-700 atlas were largely connections of the subcortex and cerebellum. Similarly, for those most affected edges, PA scans showed significantly higher reliability compared with AP scans ($P < 0.001$, Fig. 7B).

3.3. Are PED effects on reliability of connectivity related to PED effects on reliability of signal-to-noise ratio and head motion?

Based on these findings, we further asked the questions as whether the observed PED effects on reliability of functional connectivity may relate to 1) a similar PED effect on reliability of temporal signal-to-noise ratio (tSNR) in the same regions; and 2) a similar PED effect on reliability of head motion. To this end, we calculated node-specific tSNRs and FDs for the most affected nodes as shown in Fig. 6. The node-specific FDs were generated based on previous publications (Satterthwaite et al., 2013; Yan et al., 2013) using the DPARSF toolbox (Chao-Gan and Yu-Feng, 2010). Briefly, the position of each voxel in the brain at each time point was estimated by applying the motion transformation matrix to the original position of each voxel derived from the reference image during realignment (here the single-band images during the scan), resulting in a series of maps quantifying the distance change of each voxel relative to its preceding time point. The node-specific FDs were further

calculated by averaging all voxels within a given node. Similarly, in terms of prior definition (Chen et al., 2020; Demetriou et al., 2018), node-specific tSNRs were evaluated as the mean of the time series divided by its standard deviation in each node.

Using the same measure of ICC(2, 1), we subsequently estimated the reliability of FD and tSNR for each of the top ten nodes with the strongest PED effect. The derived ICC values were then entered into a similar linear mixed model as dependent variable, with PED and atlas as fixed variables and paradigm as random variable. Here, we observed highly significant effect of PED on the reliability of tSNR across the same regions ($P < 0.001$, Fig. 8A), with much higher ICCs during the PA scans compared with the AP scans. By contrast, the reliability of FD in these regions did not significantly differ between the AP and PA scans ($P = 0.47$, Fig. 8B), suggesting that the detected PED effect on reliability of nodal connectivity may to certain degree relate to a similar effect on reliability of signal intensity in the same regions.

In addition, we also tested if the FD and tSNR themselves (rather than their reliability) in these nodes would differ between PEDs. We did not observe any significant PED effects on either tSNR ($P = 0.45$) or FD ($P = 0.32$), suggesting that the reliability of tSNR, rather than the tSNR itself, may contribute to the detected reliability differences in functional connectivity (Fig. 8).

3.4. Does averaging AP and PA scans boost the reliability?

One commonly used approach to combining data acquired from different PEDs, as recommended by the HCP, is to average the derived AP and PA functional connectivity matrices (Smith et al., 2013). This raises the question as whether test-retest reliability could be boosted by using these averaged connectivity measures. To answer this question, we further calculated the ICC(2,1) values for the averaged connectome at the global, nodal, and edge levels, and compared the derived ICCs with those from the PA scans. As presented in Fig. 9, at the global level, no significant reliability differences were shown between the PA and the averaged connectomes, regardless of atlas ($P = 0.69$). However, at the nodal level, we found a significant increase of median nodal reliability for the averaged connectome compared with the PA connectome, which was only present with the CAB-NP-718 atlas ($P = 0.017$) but not the Seitzman-300 atlas ($P = 0.42$). At the edge level, both atlases showed significantly increased median edge ICCs in the averaged connectome compared with those in the PA connectome ($P < 0.001$). These results suggest that averaging AP and PA scans may boost the reliability at the nodal and edge levels, in particular with the CAB-NP-718 atlas.

3.5. Replication of reliability findings in an independent dataset

When analyzing data from the HCP-EP project, we found overall similar effects for global ICC and median nodal ICCs (see ICC distributions in Figure S1). As presented in Fig. 10A, at the global level, significant PED by atlas interactive effect was shown on the reliability of global connectivity, with significantly higher ICCs during PA scans (vs AP scans) with the Seitzman-300 atlas ($P = 0.02$). No significant difference was observed in ICCs between the PA scans and the average of AP and PA scans ($P > 0.14$), despite an upward trend in

the reliability of averaged scans. At the nodal level (Fig. 10B), compared with AP scans, PA showed significantly higher median ICCs with the CAB-NP-718 atlas ($P = 0.02$) and trend-level higher median ICCs with the Seitzman-300 atlas ($P = 0.09$). With both atlases, averaging AP and PA yielded significantly higher median ICCs compared with the PA scans ($P < 0.04$). Consistent with the discovery sample, the top ten nodes most affected by PED showed significantly higher ICCs in PA scans compared with AP scans (Seitzman-300: $P = 0.026$; CAB-NP-718: $P = 0.002$, Figure S2). These nodes were similarly distributed in the cingulate cortex, temporal cortex, sensorimotor areas, and visual cortex in the Seitzman-300 atlas and in the cingulate cortex, temporal cortex, sensorimotor areas, and subcortex in the CAB-NP-718 atlas. At the edge level (Fig. 10C), however, we did not detect a significant difference for median edge ICCs between AP and PA scans with both atlases ($P > 0.52$), although significantly higher reliability was still observed when contrasting the averaged connectivity with the AP or PA connectivity ($P < 0.04$). These findings suggest that the detected PED effects are largely preserved in the replication sample at the global and nodal levels.

4. Discussion

This study for the first time systematically examined the effects of PED on test-retest reliability of human functional connectome. In two independent datasets with repeated scans using both AP and PA directions, we showed that PA scans were associated with significantly higher reliability of global connectivity compared with the AP scans, in particular using the Seitzman-300 atlas with relatively fewer brain nodes (versus the CAB-NP-718 atlas). At the nodal level, while reliability differed node by node, regions most strongly affected by PED were consistently mapped to the cingulate cortex, temporal cortex, sensorimotor areas, and visual areas in both atlases, with almost uniformly increased reliability during PA scans in these regions. Further, we found that differences in test-retest reliability between different PEDs may relate to a similar PED effect on the reliability of image tSNR, and averaging AP and PA scans may increase the reliability of functional connectivity, in particular at the nodal and edge levels. These findings suggest that PED has significant effects on the reliability of human functional connectome and should be taken into consideration during the design of imaging protocol, especially for longitudinal studies.

For reliability assessed at different levels, they all varied between different PEDs, atlases, and global signal approaches. While the effects of PED are a new finding in the present study, the other factors have been well-known to play an important role in modulating the reliability outcome in the functional connectome data. In particular, data without GSR have been found to be more reliable than data with GSR (Cao et al., 2019; Noble et al., 2019; Song et al., 2012; Tozzi et al., 2020), and data constructed from a more fine-grained atlas with more functional parcels are reported to be more reliable than those constructed from an atlas with smaller number of parcels (Cao et al., 2021a; Cao et al., 2019; Cao et al., 2014; Noble et al., 2019; Tozzi et al., 2020). Our present findings are highly consistent with these prior observations: when using the CAB-NP-718 atlas without GSR, fair to good reliability (ICC > 0.4) was detected at the global level regardless of PED and paradigm; at the nodal and edge levels, averagely ~40% of total nodes and ~25% of edges showed fair to good reliability, across both PEDs and all paradigms. While the effects of atlas and GSR are not

the research focus in the present study, we nevertheless tested the potential interactive effect between PED and these factors on the functional connectivity measures. The significant PED by atlas and PED by GSR interactions observed in the present work implies that certain PED effects on the functional connectome may be atlas- and GSR-dependent.

During the AP scans, nodes with good reliability (ICC > 0.6) were largely distributed at the posterior part of the brain, chiefly the subcortex and cerebellum; while the distribution of reliable nodes was more widespread across the whole brain during the PA scans. The more constrained distribution of the reliable nodes during AP scans may relate to artifacts and signal distortions induced by the sinuses and eyes, which render relatively lower reliability in image quality for the anterior part of the brain compared with the posterior part. This explanation is supported by the subsequent finding that AP and PA images were associated with different reliability in tSNR, in particular regions whose reliability of connectivity was most affected by PED. One possibility that the subcortical regions are more resilient to this effect is that these regions have extensive connections with the entire cerebral cortex (both anterior and posterior part), and therefore the effects of less reliable connections at the anterior part are compensated by the effects of more reliable connections at the posterior part, rendering the reliability of overall subcortical connectivity relatively stable. Statistically, regions consistently influenced by PED in both atlases were primarily the cingulate cortex, temporal cortex, sensorimotor area, and visual area. This is similar to results in previous studies that both functional connectivity strength and amplitude of low-frequency fluctuation (ALFF) in these regions are strongly affected by PED (Mori et al., 2018; Wang et al., 2021). Notably, these regions are not only along the PE axis during imaging acquisition but also close to air/tissue interface, which possibly suffer magnetic susceptibility induced image distortion and propagated flow or motion artifacts (De Panfilis and Schwarzbauer, 2005; Weiskopf et al., 2006). Tentatively, such effects may be more variable and less stable with the AP scans, leading to lower reliability of functional connectivity in these regions.

In this study, we tested two possibilities that may underlie the observed reliability differences between PEDs. Parallel to the interpretations above, we found that image tSNR was also more reliable during PA scans compared with the AP scans, in particular in regions with the largest PED effects on reliability of functional connectivity. This finding suggests that even after careful correction of PE effects during data processing (such as the standard pipeline implemented in the HCP), the PE effects are still not able to be fully removed and could impact the image signal intensity and stability. By contrast, no significant differences were observed between AP and PA scans in terms of reliability of head motion. As head motion is a critical consideration in the calculation of functional connectivity (Power et al., 2012; Satterthwaite et al., 2013; Yan et al., 2013) and may affect the reliability of connectivity measure (Cao et al., 2019; Noble et al., 2019; Yan et al., 2013), this result mitigates the possibility that the PED effects on reliability of human functional connectome are due to differences in reliability of head motion. This is not surprising since head motion has been proposed to reflect an individualized trait (Couvy-Duchesne et al., 2014; Engelhardt et al., 2017; Zeng et al., 2014) and relate to psychopathology (Couvy-Duchesne et al., 2016; Kong et al., 2014). Therefore, head motion assessments should theoretically be associated with inter-subject variation rather than variation in scan protocol.

Our results also demonstrated that averaging data collected from both PEDs may increase the reliability of functional connectivity, chiefly at the nodal and edge levels and particularly with the CAB-NP-718 atlas. This finding suggests that the use of averaged data, as commonly applied in the neuroimaging studies, may indeed be beneficial, at least from a reliability perspective. An explanation is that the averaged data may mitigate state-dependent and scan/PED-related noise, rendering the connectivity measures more accurate. Therefore, if multiple scans with different PEDs are available (such as the data from the HCP family), the approach to averaging AP and PA scans is generally preferred to achieve more reliable measures, compared with using data from a single run. It remains to be determined, however, whether acquisition of multiple runs of only PA scans would further boost the reliability (versus multiple runs with different PEDs).

One major limitation of this study, as we clearly acknowledge here, is that the scan length is relatively short for each run (7 min for the discovery sample and 5.6 min for the replication sample). As prior work has shown that at least 20–30 min scan time is required to achieve a stable result (Gordon et al., 2017; Laumann et al., 2015), we cannot exclude the possibility that the findings reported here are influenced by scan length. It is important to further investigate whether longer scans would compensate for the observed effects of PED in future studies. The second limitation relates to the sample size, which is relatively small in both samples. Although similar PED effects were observed across both samples, future studies with larger test-retest datasets are still warranted to verify these findings. Thirdly, our results were purely based on the PE correction approach implemented in the HCP pipeline using field maps, which may not be generalizable to other PE correction approaches (Gu and Eklund, 2019). We employed the HCP pipeline since this is considered as a state-of-the-art approach for imaging preprocessing and has widely been used in the neuroimaging community. However, our results may not necessarily imply a lack of efficacy for the employed approach in PE correction but may reflect an intrinsic problem for the MRI scans. Fourthly, all data in both samples were collected from scanners with the same make and model (SIEMENS Prisma) using multiband EPI scans, and therefore whether the detected effects are specific to scanner and/or scan sequence needs to be examined. Finally, as both scan length and sample size differed between fMRI paradigms, we treated paradigm as a random-effect variable in this study. However, since paradigm itself may affect reliability of functional connectivity (Cao et al., 2019; Cao et al., 2014; Shah et al., 2016), direct comparison of PE effects between different paradigms should also be performed in the future.

In sum, this study for the first time reveals the effects of PED on test-retest reliability of human connectomic measures derived from fMRI data, at global, nodal, and edge levels. Overall, the PA scans appear to show superiority in reliability of connectivity measures, suggesting that PA scans may be preferred if the research is not focused on a specific region of interest. We urge that the PED effects need to be carefully considered in future neuroimaging designs, especially in longitudinal studies such as those related to neurodevelopment or clinical intervention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by fundings from the National Institute of Mental Health (NIMH) grants P50MH080173 and R01MH108654 to Dr. Malhotra. Dr. Cao acknowledges funding from the Feinstein Institutes for Medical Research. The Human Connectome Project for Early Psychosis (HCP-EP) was supported by NIMH funding U01MH109977.

Data and code availability statement

The discovery dataset will be made available to other research groups upon request. Due to ethics constraints, data will be shared on a project-specific basis. The replication dataset is part of the HCP-EP study that can be downloaded at the Connectome Coordination Facility (https://nda.nih.gov/ccf/) upon request. The HCP pipeline used for preprocessing is publicly available at https://github.com/Washington-University/HCPpipelines. Other in-house Matlab codes used in this study for computation of ICC are available to share when the manuscript is accepted.

References

Anderson JS, Ferguson MA, Lopez-Larson M, Yurgelun-Todd D, 2011. Reproducibility of single-subject functional connectivity measurements. AJNR Am. J. Neuroradiol 32, 548–555. [PubMed: 21273356]

Baker JT, Dillon DG, Patrick LM, Roffman JL, Brady RO, Pizzagalli DA, Öngür D, Holmes AJ, 2019. Functional connectomics of affective and psychotic pathology. Proc. Natl. Acad. Sci. U. S. A 116, 9050–9059. [PubMed: 30988201]

Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith S, Johansen-Berg H, Snyder AZ, Van Essen DC, Consortium W-MH, 2013. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 80, 169–189. [PubMed: 23684877]

Birn RM, Molloy EK, Patriat R, Parker T, Meier TB, Kirk GR, Nair VA, Meyerand ME, Prabhakaran V, 2013. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. Neuroimage 83, 550–558. [PubMed: 23747458]

Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL, Burgess GC, Curtiss SW, Diaz-Santos M, Elam JS, Fischl B, Greve DN, Hagy HA, Harms MP, Hatch OM, Hedden T, Hodge C, Japardi KC, Kuhn TP, Ly TK, Smith SM, Somerville LH, Uğurbil K, van der Kouwe A, Van Essen D, Woods RP, Yacoub E, 2019. The lifespan human connectome project in aging: an overview. Neuroimage 185, 335–348. [PubMed: 30332613]

Bush G, Shin LM, 2006. The multi-source interference task: an fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network. Nat. Protoc 1, 308–313. [PubMed: 17406250]

Cao H, Chen OY, McEwen SC, Forsyth JK, Gee DG, Bearden CE, Addington J, Goodyear B, Cadenhead KS, Mirzakhanian H, Cornblatt BA, Carrión RE, Mathalon DH, McGlashan TH, Perkins DO, Belger A, Thermenos H, Tsuang MT, van Erp TGM, Walker EF, Hamann S, Anticevic A, Woods SW, Cannon TD, 2021a. Cross-paradigm connectivity: reliability, stability, and utility. Brain Imaging Behav 15, 614–629. [PubMed: 32361945]

Cao H, Chung Y, McEwen SC, Bearden CE, Addington J, Goodyear B, Cadenhead KS, Mirzakhanian H, Cornblatt BA, Carrión R, Mathalon DH, McGlashan TH, Perkins DO, Belger A, Seidman LJ, Thermenos H, Tsuang MT, van Erp TGM, Walker EF, Hamann S, Anticevic A, Woods SW, Cannon TD, 2020. Progressive reconfiguration of resting-state brain networks as psychosis

develops: preliminary results from the North American Prodrome Longitudinal Study (NAPLS) consortium. Schizophr. Res 226, 30–37. [PubMed: 30704864]

Cao H, McEwen SC, Forsyth JK, Gee DG, Bearden CE, Addington J, Goodyear B, Cadenhead KS, Mirzakhanian H, Cornblatt BA, Carrión RE, Mathalon DH, McGlashan TH, Perkins DO, Belger A, Seidman LJ, Thermenos H, Tsuang MT, van Erp TGM, Walker EF, Hamann S, Anticevic A, Woods SW, Cannon TD, 2019. Toward leveraging human connectomic data in large consortia: generalizability of fMRI-based brain graphs across sites, sessions, and paradigms. Cereb. Cortex 29, 1263–1279. [PubMed: 29522112]

Cao H, Plichta MM, Schäfer A, Haddad L, Grimm O, Schneider M, Esslinger C, Kirsch P, Meyer-Lindenberg A, Tost H, 2014. Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. Neuroimage 84, 888–900. [PubMed: 24055506]

Cao H, Zhou H, Cannon TD, 2021b. Functional connectome-wide associations of schizophrenia polygenic risk. Mol. Psychiatry 26, 2553–2561. [PubMed: 32127647]

Chao-Gan Y, Yu-Feng Z, 2010. DPARSF: a MATLAB toolbox for “pipeline” data analysis of resting-state fMRI. Front. Syst. Neurosci 4, 13. [PubMed: 20577591]

Chen JCC, Forsyth A, Dubowitz DJ, Muthukumaraswamy SD, 2020. On the quality, statistical efficiency, and safety of simultaneously recorded multiband fMRI/EEG. Brain Topogr 33, 303–316. [PubMed: 32144628]

Couvy-Duchesne B, Blokland GA, Hickie IB, Thompson PM, Martin NG, de Zubicaray GI, McMahon KL, Wright MJ, 2014. Heritability of head motion during resting state functional MRI in 462 healthy twins. Neuroimage 102 (2), 424–434 Pt. [PubMed: 25132021]

Couvy-Duchesne B, Ebejer JL, Gillespie NA, Duffy DL, Hickie IB, Thompson PM, Martin NG, de Zubicaray GI, McMahon KL, Medland SE, Wright MJ, 2016. Head motion and inattention/hyperactivity share common genetic influences: implications for fMRI studies of ADHD. PLoS One 11, e0146271. [PubMed: 26745144]

De Panfilis C, Schwarzbauer C, 2005. Positive or negative blips? The effect of phase encoding scheme on susceptibility-induced signal losses in EPI. Neuroimage 25, 112–121. [PubMed: 15734348]

Demetriou L, Kowalczyk OS, Tyson G, Bello T, Newbould RD, Wall MB, 2018. A comprehensive evaluation of increasing temporal resolution with multiband-accelerated protocols and effects on statistical outcome measures in fMRI. Neuroimage 176, 404–416. [PubMed: 29738911]

Demro C, Mueller BA, Kent JS, Burton PC, Olman CA, Schallmo MP, Lim KO, Sponheim SR, 2021. The psychosis human connectome project: an overview. Neuroimage 241, 118439. [PubMed: 34339830]

Dong D, Wang Y, Chang X, Luo C, Yao D, 2018. Dysfunction of large-scale brain networks in schizophrenia: a meta-analysis of resting-state functional connectivity. Schizophr. Bull 44, 168–181. [PubMed: 28338943]

Engelhardt LE, Roe MA, Juranek J, DeMaster D, Harden KP, Tucker-Drob EM, Church JA, 2017. Children’s head motion during fMRI tasks is heritable and stable over time. Dev. Cogn. Neurosci 25, 58–68. [PubMed: 28223034]

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC, 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. [PubMed: 27437579]

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M, Consortium W-MH, 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. [PubMed: 23668970]

Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, Hampton JM, Coalson RS, Nguyen AL, McDermott KB, Shimony JS, Snyder AZ, Schlaggar BL, Petersen SE, Nelson SM, Dosenbach NUF, 2017. Precision functional mapping of individual human brains. Neuron 95, 791–807 e797. [PubMed: 28757305]

Gu X, Eklund A, 2019. Evaluation of six phase encoding based susceptibility distortion correction methods for diffusion MRI. Front. Neuroinform 13, 76. [PubMed: 31866847]

Ilioska I, Oldehinkel M, Llera A, Chopra S, Looden T, Chauvin R, Van Rooij D, Floris DL, Tillmann J, Moessnang C, Banaschewski T, Holt RJ, Loth E, Charman T, Murphy DGM, Ecker C, Mennes M, Beckmann CF, Fornito A, Buitelaar JK, 2022. Connectome-Wide Mega-Analysis Reveals Robust Patterns of Atypical Functional Connectivity in Autism. Biological Psychiatry

Javaheripour N, Li M, Chand T, Krug A, Kircher T, Dannlowski U, Nenadić I, Hamilton JP, Sacchet MD, Gotlib IH, Walter H, Frodl T, Grimm S, Harrison BJ, Wolf CR, Olbrich S, van Wingen G, Pezawas L, Parker G, Hyett MP, Sämann PG, Hahn T, Steinsträter O, Jansen A, Yuksel D, Kämpe R, Davey CG, Meyer B, Bartova L, Croy I, Walter M, Wagner G, 2021. Altered resting-state functional connectome in major depressive disorder: a mega-analysis from the PsyMRI consortium. Transl. Psychiatry 11, 511. [PubMed: 34620830]

Ji JL, Spronk M, Kulkarni K, Repovš G, Anticevic A, Cole MW, 2019. Mapping the human brain's cortical-subcortical functional network organization. Neuroimage 185, 35–57. [PubMed: 30291974]

Kong XZ, Zhen Z, Li X, Lu HH, Wang R, Liu L, He Y, Zang Y, Liu J, 2014. Individual differences in impulsivity predict head motion during magnetic resonance imaging. PLoS One 9, e104989. [PubMed: 25148416]

Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen MY, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NU, Schlaggar BL, Mumford JA, Poldrack RA, Petersen SE, 2015. Functional system and areal organization of a highly sampled individual human brain. Neuron 87, 657–670. [PubMed: 26212711]

Mori Y, Miyata J, Isobe M, Son S, Yoshihara Y, Aso T, Kouchiyama T, Murai T, Takahashi H, 2018. Effect of phase-encoding direction on group analysis of resting-state functional magnetic resonance imaging. Psychiatry Clin. Neurosci 72, 683–691. [PubMed: 29774625]

Murphy K, Fox MD, 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. Neuroimage 154, 169–173. [PubMed: 27888059]

Noble S, Scheinost D, Constable RT, 2019. A decade of test-retest reliability of functional connectivity: a systematic review and meta-analysis. Neuroimage 203, 116157. [PubMed: 31494250]

Noble S, Scheinost D, Finn ES, Shen X, Papademetris X, McEwen SC, Bearden CE, Addington J, Goodyear B, Cadenhead KS, Mirzakhanian H, Cornblatt BA, Olvet DM, Mathalon DH, McGlashan TH, Perkins DO, Belger A, Seidman LJ, Thermenos H, Tsuang MT, van Erp TGM, Walker EF, Hamann S, Woods SW, Cannon TD, Constable RT, 2017a. Multisite reliability of MR-based functional connectivity. Neuroimage 146, 959–970. [PubMed: 27746386]

Noble S, Spann MN, Tokoglu F, Shen X, Constable RT, Scheinost D, 2017b. Influences on the test-retest reliability of functional connectivity MRI and its relationship with behavioral utility. Cereb. Cortex 27, 5415–5429. [PubMed: 28968754]

Pannunzi M, Hindriks R, Bettinardi RG, Wenger E, Lisofsky N, Martensson J, Butler O, Filevich E, Becker M, Lochstet M, Kühn S, Deco G, 2017. Resting-state fMRI correlations: from link-wise unreliability to whole brain stability. Neuroimage 157, 250–262. [PubMed: 28599964]

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE, 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142–2154. [PubMed: 22019881]

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE, 2011. Functional network organization of the human brain. Neuron 72, 665–678. [PubMed: 22099467]

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH, 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. [PubMed: 22926292]

Seitzman BA, Gratton C, Marek S, Raut RV, Dosenbach NUF, Schlaggar BL, Petersen SE, Greene DJ, 2020. A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum. Neuroimage 206, 116290. [PubMed: 31634545]

Shah LM, Cramer JA, Ferguson MA, Birn RM, Anderson JS, 2016. Reliability and reproducibility of individual differences in functional connectivity acquired during task and resting state. Brain Behav 6, e00456. [PubMed: 27069771]

Shehzad Z, Kelly AM, Reiss PT, Gee DG, Gotimer K, Uddin LQ, Lee SH, Margulies DS, Roy AK, Biswal BB, Petkova E, Castellanos FX, Milham MP, 2009. The resting brain: unconstrained yet reliable. Cereb. Cortex 19, 2209–2229. [PubMed: 19221144]

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Uğurbil K, Van Essen DC, Glasser MF, Consortium W-MH, 2013. Resting-state fMRI in the Human Connectome Project. Neuroimage 80, 144–168. [PubMed: 23702415]

Somerville LH, Bookheimer SY, Buckner RL, Burgess GC, Curtiss SW, Dapretto M, Elam JS, Gaffrey MS, Harms MP, Hodge C, Kandala S, Kastman EK, Nichols TE, Schlaggar BL, Smith SM, Thomas KM, Yacoub E, Van Essen DC, Barch DM, 2018. The lifespan Human Connectome Project in development: a large-scale study of brain connectivity development in 5–21 year olds. Neuroimage 183, 456–468. [PubMed: 30142446]

Song J, Desphande AS, Meier TB, Tudorascu DL, Vergun S, Nair VA, Biswal BB, Meyerand ME, Birn RM, Bellec P, Prabhakaran V, 2012. Age-related differences in test-retest reliability in resting-state brain functional connectivity. PLoS One 7, e49847. [PubMed: 23227153]

Tozzi SL, Fleming SL, Taylor ZD, Raterink CD, Williams LM, 2020. Test-retest reliability of the human functional connectome over consecutive days: identifying highly reliable portions and assessing the impact of methodological choices. Netw. Neurosci 4, 925–945. [PubMed: 33615097]

Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, Della Penna S, Feinberg D, Glasser MF, Harel N, Heath AC, Larson-Prior L, Marcus D, Michalareas G, Moeller S, Oostenveld R, Petersen SE, Prior F, Schlaggar BL, Smith SM, Snyder AZ, Xu J, Yacoub E, Consortium W-MH, 2012. The Human Connectome Project: a data acquisition perspective. Neuroimage 62, 2222–2231. [PubMed: 22366334]

Wang Y, Chen X, Liu R, Zhang Z, Zhou J, Feng Y, Jiang C, Zuo XN, Zhou Y, Wang G, 2021. Effect of phase-encoding direction on gender differences: a resting-state functional magnetic resonance imaging study. Front. Neurosci 15, 748080. [PubMed: 35145372]

Weiskopf N, Hutton C, Josephs O, Deichmann R, 2006. Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. Neuroimage 33, 493–504. [PubMed: 16959495]

Weiskopf N, Hutton C, Josephs O, Turner R, Deichmann R, 2007. Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. MAGMA 20, 39–49. [PubMed: 17268781]

Yan CG, Cheung B, Kelly C, Colcombe S, Craddock RC, Di Martino A, Li Q, Zuo XN, Castellanos FX, Milham MP, 2013. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. Neuroimage 76, 183–201. [PubMed: 23499792]

Yoo K, Rosenberg MD, Noble S, Scheinost D, Constable RT, Chun MM, 2019. Multivariate approaches improve the reliability and validity of functional connectivity and prediction of individual behaviors. Neuroimage 197, 212–223. [PubMed: 31039408]

Zeng LL, Wang D, Fox MD, Sabuncu M, Hu D, Ge M, Buckner RL, Liu H, 2014. Neurobiological basis of head motion in brain imaging. Proc. Natl. Acad. Sci. U. S. A 111, 6058–6062. [PubMed: 24711399]

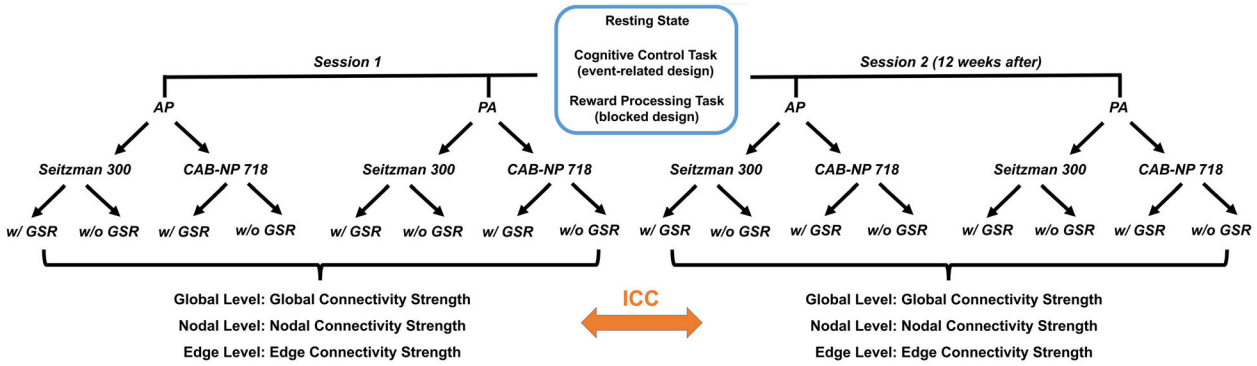

Fig. 1.

Data analysis pipeline in the discovery sample. A similar pipeline was also used for the replication sample, with data acquired from a single paradigm (resting state) and much shorter scan interval (two sessions on the same day).

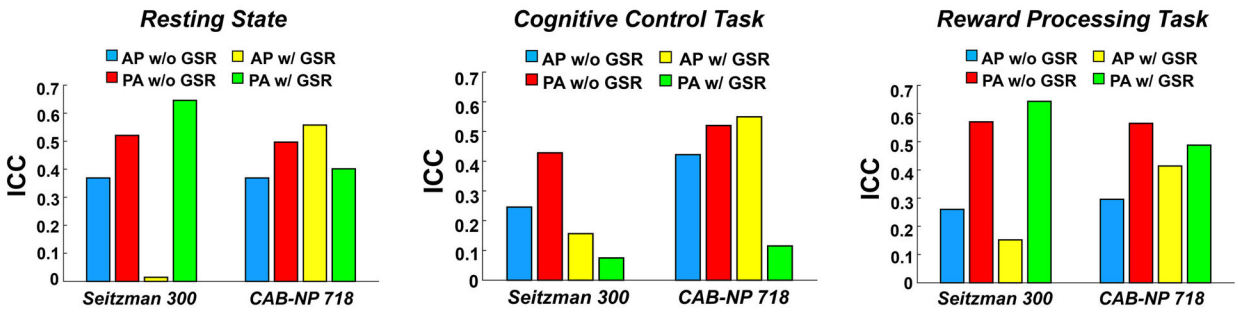

Fig. 2.

Intraclass correlation coefficients (ICCs) of global connectivity across different PEDs, atlases, and global signal approaches in the discovery sample. The left, middle, and right panels represent ICCs during resting state, cognitive control, and reward processing, respectively.

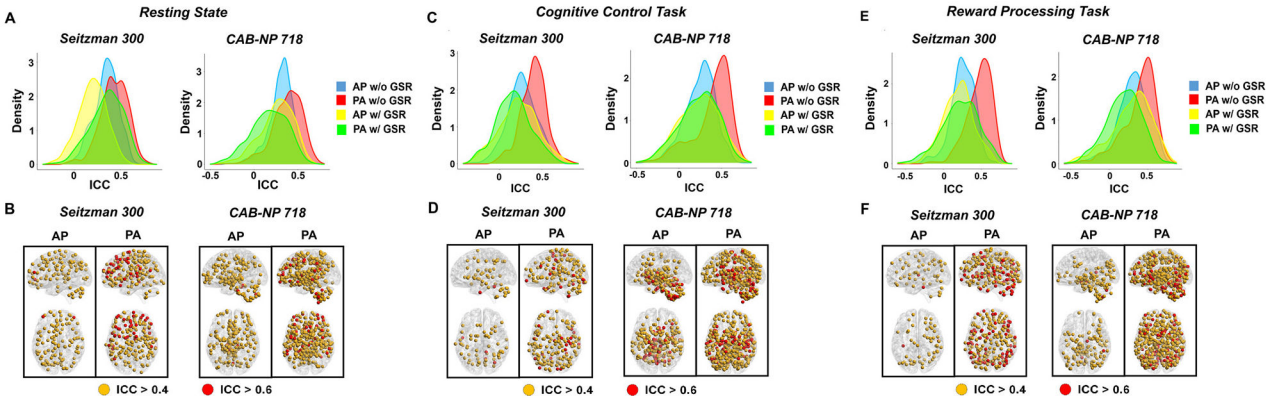

Fig. 3.

ICCs of nodal connectivity across different PEDs, atlases, and global signal approaches in the discovery sample. The upper panels present the ICC distributions across all nodes in the brain, and the lower panels present the nodes with fair (ICC > 0.4, in orange) and good (ICC > 0.6, in red) test-retest reliability. The left, middle, and right panels show results during resting state, cognitive control, and reward processing, respectively. Across all paradigms and both atlases, PA scans without global signal regression had the highest reliability across the whole brain. Nodes shown on the lower panels were therefore based on ICCs without global signal regression.

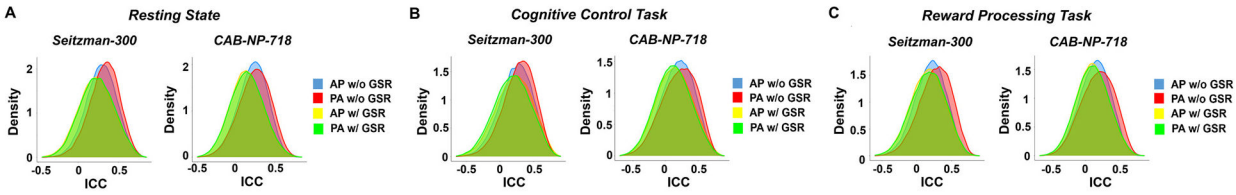

Fig. 4.

ICCs of edge connectivity across different PEDs, atlases, and global signal approaches in the discovery sample. The left, middle, and right panels show results during resting state, cognitive control, and reward processing, respectively.

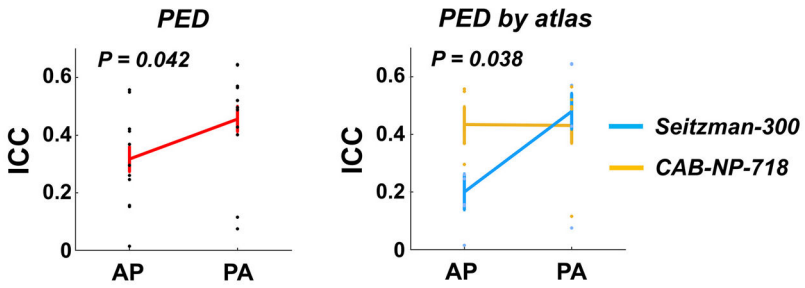

Fig. 5.

Effects of PED on the reliability measures of global connectivity. Significant PED effect and PED by atlas interactive effect were shown on the reliability of global connectivity, with reliability in PA scans significantly higher than that in AP scans, in particular using the Seitzman-300 atlas.

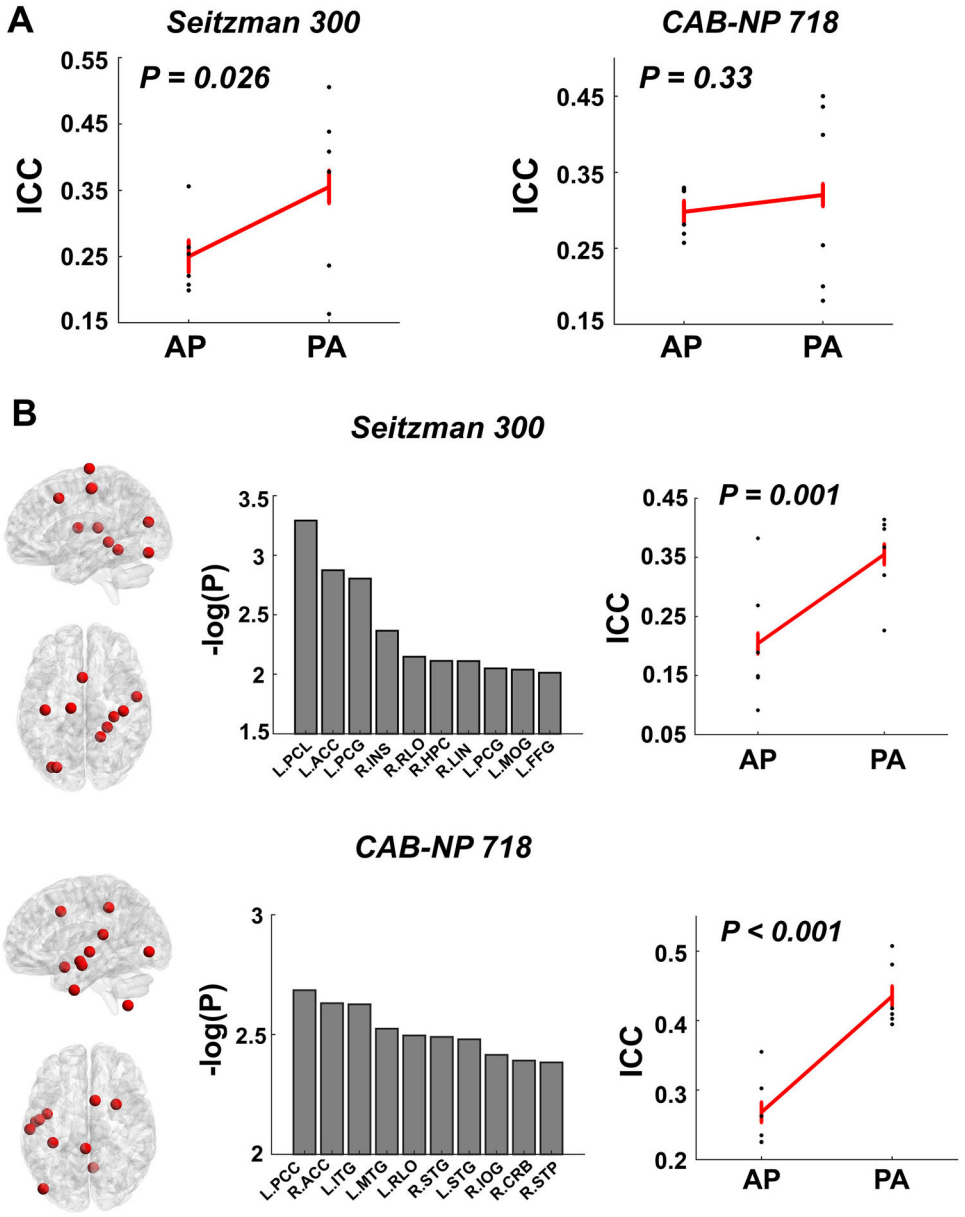

Fig. 6.

Effects of PED on the reliability measures of nodal connectivity. (A) Significantly higher median reliability across all nodes was shown for the PA scans compared with AP scans with the Seitzman-300 atlas. (B) Top ten nodes ranked by the *P* value in both atlases showed higher reliability during PA scans. These nodes were largely distributed in the cingulate cortex, temporal cortex, sensorimotor area, and visual area with both atlases. Error bars indicate standard error.

Abbreviations: ACC = anterior cingulate cortex; PCL = paracentral lobule; PCG = postcentral gyrus; INS = insula; RLO = Rolandic operculum; HPC = hippocampus; LIN = lingual gyrus; MOG = middle occipital gyrus; FFG = fusiform gyrus; PCC = posterior cingulate cortex; ITG = inferior temporal gyrus; MTG = middle temporal gyrus; STG =

superior temporal gyrus; IOG = inferior occipital gyrus; CRB = cerebellum; STP = superior temporal pole.

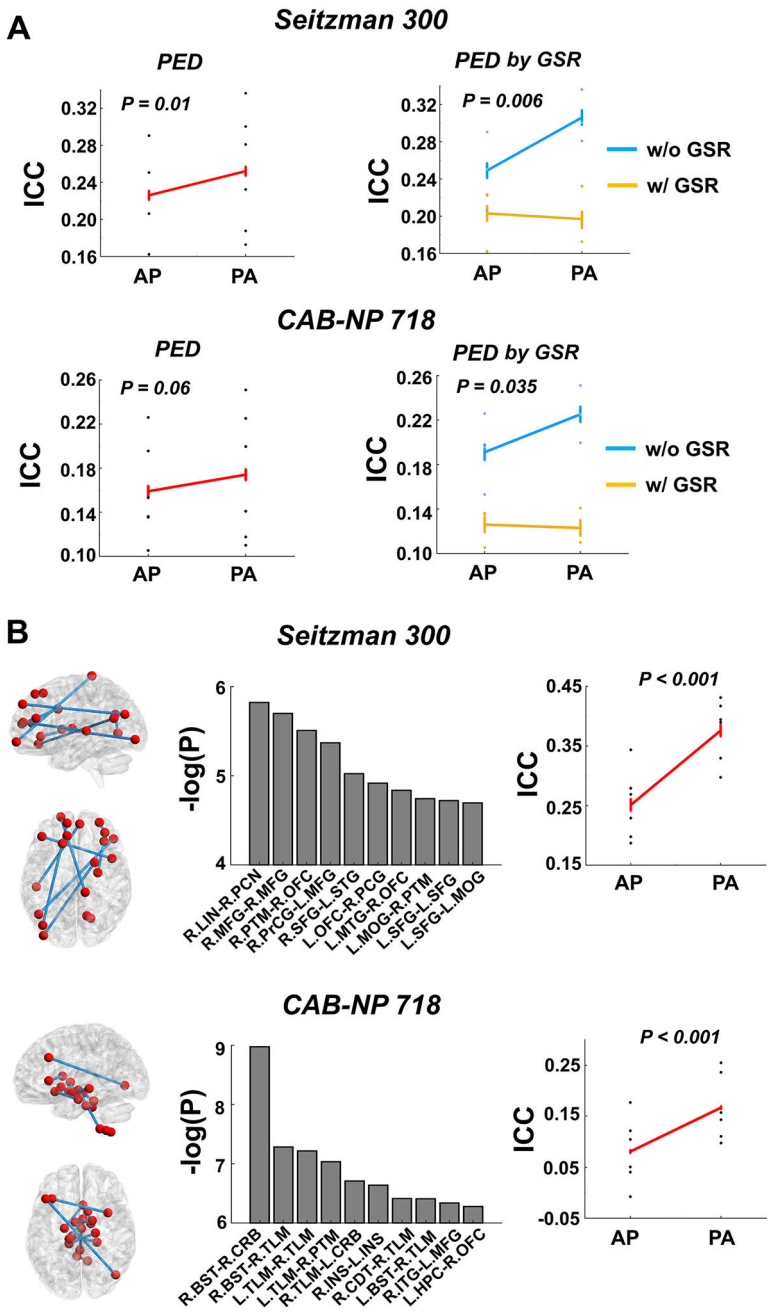

Fig. 7.

Effects of PED on the reliability measures of edge connectivity. (A) Significantly higher median reliability across all edges was detected for the PA scans compared with AP scans. Moreover, significant PED by GSR interaction was also observed, in which higher reliability was particularly detected without GSR. (B) Top ten edges ranked by the *P* value in both atlases showed higher reliability during PA scans. In the Seitzman-300 atlas, these edges chiefly included connections between the frontal cortex and the subcortex, sensorimotor area, and temporal cortex; while in the CAB-NP-718 atlas, these edges were largely connections of the subcortex and cerebellum. Error bars indicate standard error.

Abbreviations: PCG = postcentral gyrus; INS = insula; LIN = lingual gyrus; MOG = middle occipital gyrus; MTG = middle temporal gyrus; STG = superior temporal gyrus; PCN = precuneus; MFG = middle frontal gyrus; PTM = putamen; OFC = orbitofrontal cortex; PrCG = precentral gyrus; SFG = superior frontal gyrus; BST = brainstem; CRB = cerebellum; TLM = thalamus; CDT = caudate; ITG = inferior temporal gyrus; HPC = hippocampus.

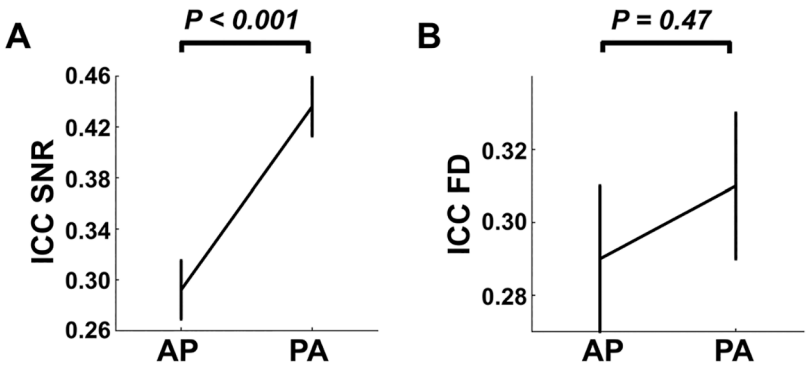

Fig. 8.

Effects of PED on reliabilities of temporal signal-to-noise ratio (tSNR, Panel A) and head motion (Panel B). For the same regions shown in Fig. 6B, their SNR reliability was significantly higher during PA scans compared with AP scans. No significant effect was shown for the reliability of head motion. Error bars indicate standard error.

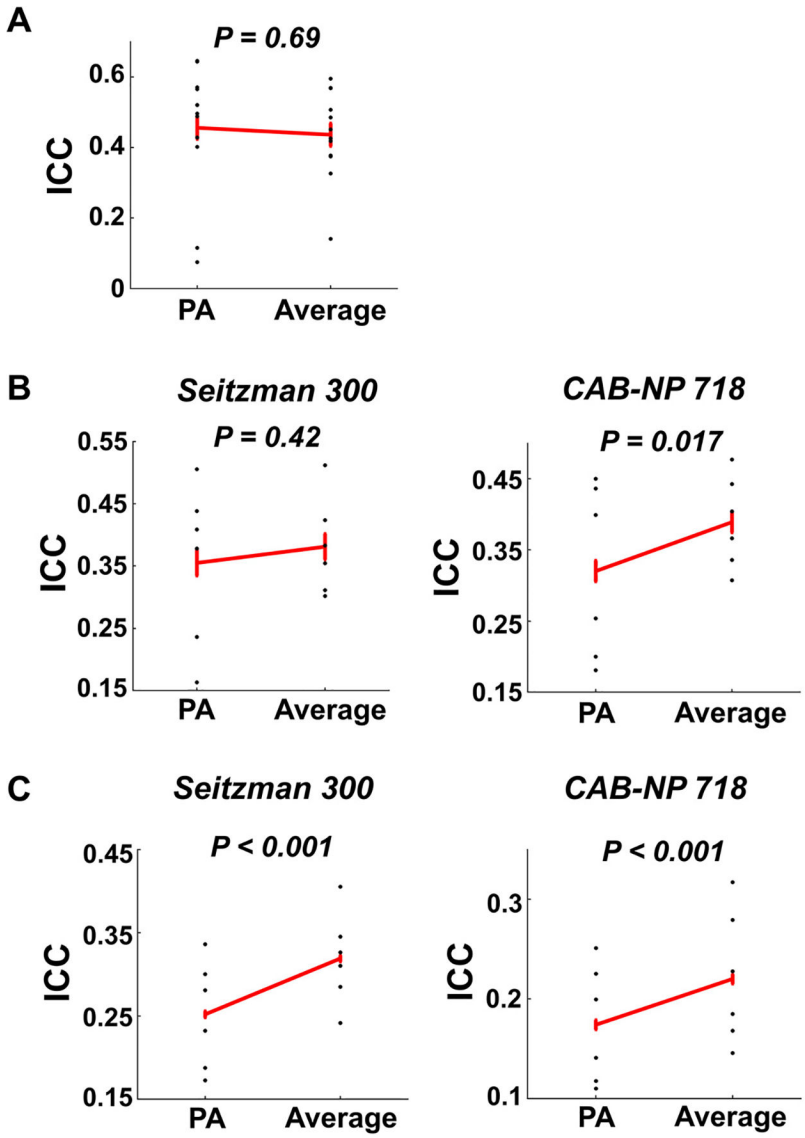

Fig. 9.

Reliability of global (Panel A), nodal (Panel B), and edge (Panel C) connectivity for data averaged by AP and PA scans. No significant reliability differences were shown between the PA and the averaged data on global connectivity. However, significant differences in median reliability of nodal and edge connectivity were demonstrated. Error bars indicate standard error.

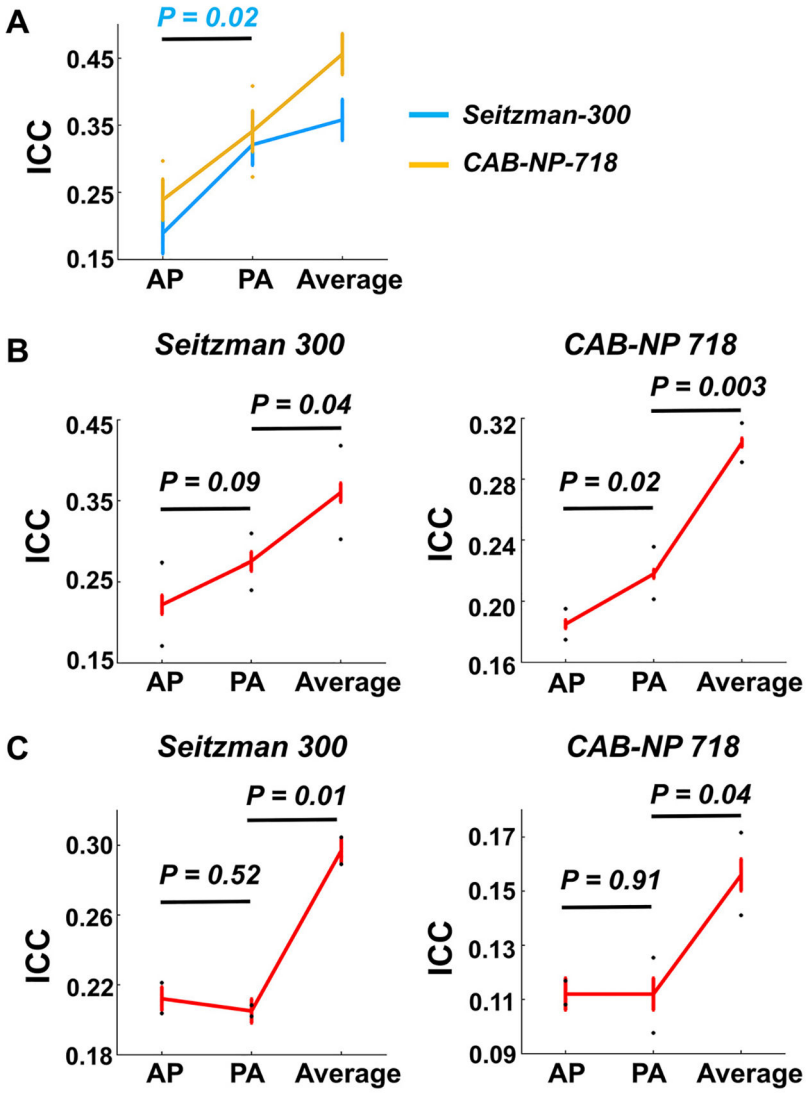

Fig. 10.

Replication of findings in the HCP-EP dataset. (A) Similar PED and PED by atlas effects were shown for reliability of global connectivity. (B) Similar PED effects on median reliability of nodal connectivity were observed. (C) No significant difference was observed in median reliability of edge connectivity between AP and PA scans, although better reliability was still shown for the averaged connectome. Error bars indicate standard error.

Table 1

Median, first and third quartiles of nodal and edge-level ICCs in the discovery sample, separated by PED, global signal approach, atlas, and paradigm.

	Node-level Median (Q1, Q3)	Edge-level Median (Q1, Q3)
Resting state, Seitzman-300		
AP without GSR	0.36 (0.26, 0.43)	0.29 (0.16, 0.42)
PA without GSR	0.44 (0.35, 0.54)	0.34 (0.21, 0.46)
AP with GSR	0.20 (0.10. 0.30)	0.22 (0.08, 0.37)
PA with GSR	0.38 (0.24, 0.48)	0.23 (0.09, 0.38)
Resting state, CAB-NP-718		
AP without GSR	0.33 (0.24, 0.40)	0.23 (0.09, 0.35)
PA without GSR	0.40 (0.28, 0.50)	0.25 (0.10, 0.39)
AP with GSR	0.28 (0.15, 0.40)	0.14 (0, 0.28)
PA with GSR	0.18 (0.03, 0.33)	0.14 (0, 0.29)
Cognitive control, Seitzman-300		
AP without GSR	0.26 (0.15, 0.38)	0.25 (0.08, 0.41)
PA without GSR	0.41 (0.31, 0.50)	0.30 (0.13, 0.45)
AP with GSR	0.22 (0.07, 0.38)	0.22 (0.03, 0.41)
PA with GSR	0.16 (0.03, 0.29)	0.19 (0, 0.37)
Cognitive control, CAB-NP-718		
AP without GSR	0.33 (0.19, 0.45)	0.20 (0.01, 0.37)
PA without GSR	0.44 (0.28, 0.54)	0.23 (0.02, 0.41)
AP with GSR	0.33 (0.14, 0.47)	0.14 (0, 0.33)
PA with GSR	0.20 (0.04, 0.34)	0.12 (0, 0.31)
Reward processing, Seitzman-300		
AP without GSR	0.25 (0.17, 0.37)	0.21 (0.05, 0.35)
PA without GSR	0.51 (0.38, 0.60)	0.28 (0.11, 0.44)
AP with GSR	0.21 (0.07, 0.32)	0.16 (0, 0.33)
PA with GSR	0.24 (0.08, 0.39)	0.17 (0, 0.34)
Reward processing, CAB-NP-718		
AP without GSR	0.27 (0.14, 0.38)	0.15 (0, 0.31)
PA without GSR	0.45 (0.30, 0.55)	0.20 (0.01, 0.38)
AP with GSR	0.26 (0.06, 0.41)	0.11 (0, 0.27)
PA with GSR	0.25 (0.09, 0.40)	0.11 (0, 0.28)