See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/303467932

Quantifying the test-retest reliability of MEG resting state functional connectivity

Article in Brain Connectivity · May 2016

DOI: 10.1089/brain.2015.0416

CITATIONS		READS	- READS				
		114					
utho	rs:						
	Pilar Garcés		María Carmen Martín-Buro				
	Roche		Complutense University of Madrid				
	21 PUBLICATIONS 77 CITATIONS		6 PUBLICATIONS 16 CITATIONS				
	SEE PROEILE						
	Fernando Maestú						
Ş	Complutense University of Madrid						
	247 PUBLICATIONS 2,917 CITATIONS						
	SEE PROFILE						
	f the authors of this publication are also	working on those	related projects				
ine o	The autions of this publication are also	working on these	related projects.				
Project	Stochastic process and cognitive brain	n abilities View pro	ject				

Project Psicobiología View project

All content following this page was uploaded by Pilar Garcés on 28 June 2016.

The user has requested enhancement of the downloaded file. All in-text references <u>underlined in blue</u> are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

Quantifying the Test-Retest Reliability of Magnetoencephalography Resting-State Functional Connectivity

Pilar Garcés,^{1–3} María Carmen Martín-Buro,^{1,3,4} and Fernando Maestú^{1,3,5}

Abstract

The coordinated activity of the resting-state brain can be evaluated with magnetoencephalography (MEG) for distinct brain rhythms by performing source reconstruction to estimate the activities of target brain regions and employing one of the many existent functional connectivity (FC) algorithms. Although this procedure has been applied in a great amount of studies both with healthy and pathological populations, the reliability of such FC estimates is unknown, and this impairs the use of resting-state MEG FC at the individual level. In this study, the test-retest reliability of MEG resting FC was evaluated by exploring both within- and betweensubject variability in FC in 16 healthy subjects who underwent three resting-state MEG scans. FC was computed after beamforming source reconstruction with four popular FC metrics: phase-locking value (PLV), phase lag index (PLI), direct envelope correlation (d-ecor), and envelope correlation with leakage correction (lc-ecor). Then, test-restest reliability and within- and between-subject agreement were evaluated with the intraclass correlation coefficient (ICC) and Kendall's W, respectively. Reliability was found to depend on the FC metric, the frequency band, and the specific link. As a general trend, greater test-retest reliability was found for PLV in theta to gamma, and for lc-ecor and d-ecor in beta. Further inspection of the ICC distribution revealed that volume conduction effects could be contributing to high ICC in PLV and d-ecor. In addition, stronger links were found to be more reliable. Overall, this encourages the further use of resting-state MEG FC for individuallevel studies, especially with PLV or envelope correlation metrics.

Key words: amplitude synchronization; brain network; functional connectivity; MEG; phase synchronization; resting state; test-retest reliability

Introduction

S INCE BISWAL ET AL. (1995) observed a highly correlated brain activity between bilateral motor cortices, increasing focus has been shifted to the functioning of the restingstate brain (Deco et al., 2011; van den Heuvel and Hulshoff Pol, 2010). In fact, the resting-state brain activity is highly organized, and separate brain regions have been repeatedly shown to be spontaneously coordinated, accounting for sensory and cognitive functions, such as visual or motor processes, attention, or executive control (Rosazza and Minati, 2011).

This spontaneous brain coordination is usually characterized with functional connectivity (FC), which evaluates statistical dependencies between the activities of distinct brain regions (Pereda et al., 2005). FC can be seen as a mechanism of information transfer between brain areas, and resting FC has been proven to be indicative of an individual's brain functioning in a variety of studies (Baldassarre et al., 2012; Rosazza and Minati, 2011; Yamashita et al., 2015).

Resting-state FC has been predominantly examined with functional magnetic resonance imaging (fMRI) as correlations between blood oxygenation level-dependent (BOLD) signals (van den Heuvel et al., 2009). Although fMRI is of great use in the exploration of human brain function, slowly fluctuating BOLD signals constitute only an indirect measure of brain activity. Therefore, over the past few years, attention

¹Laboratory of Cognitive and Computational Neuroscience (UCM-UPM), Centre for Biomedical Technology, Pozuelo de Alarcón, Madrid, Spain.

²Department of Applied Physics III, Faculty of Physics, Universidad Complutense de Madrid, Madrid, Spain.

³Biomedical Research Networking Center in Bioengineering Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain.

⁴Psychology Division, Cardenal Cisneros University College, Universidad Complutense de Madrid, Madrid, Spain.

⁵Department of Basic Psychology II, Faculty of Psychology, Universidad Complutense de Madrid, Madrid, Spain.

has been drawn to the electrophysiological basis of FC. In fact, although magnetoencephalography (MEG) is often overshadowed by the high spatial resolution of fMRI, MEG provides an exceptional insight into the time-frequency dynamics of brain activity. It is indeed a direct measure of neuronal firing and offers an excellent temporal resolution. Resting-state networks (RSNs) have been successfully extracted with MEG over the past years using source-space FC (Brookes et al., 2011a,b; de Pasquale et al., 2010; Hall et al., 2013; Hipp et al., 2012; Luckhoo et al., 2012; Wens et al., 2014). Moreover, resting-state MEG FC has been proven sensible to detect abnormal brain functioning in a variety of diseases, including Alzheimer's disease (Engels et al., 2015; López et al., 2014), multiple sclerosis (Tewarie et al., 2015), and schizophrenia (Bowyer et al., 2015).

Overall, this indicates that FC can be robustly assessed with MEG. However, FC estimates could be affected by measurement noise and the subjects' cognitive or emotional state, and it is unclear to what extent FC estimates are representative of an individual. This prevents the potential use of MEG FC in the characterization of individual subjects or in the proper classification between healthy and pathological populations. Therefore, the reliability of FC estimates needs to be carefully examined. To date, only a handful of studies have tackled test-retest reliability of MEG/electroencephalography (EEG) FC (Deuker et al., 2009; Hardmeier et al., 2014; Jin et al., 2011), and they were restricted to sensor space, therefore failing to provide the spatial resolution necessary for the evaluation of whole-brain FC. To our knowledge, no study has quantified the test-retest reliability of source-space MEG FC estimates.

In this work, we investigated the test-retest reliability of MEG FC. For that, we computed whole-brain resting state FC for 16 subjects, who were scanned thrice with a 1-week test-retest interval. For each subject and session, source reconstruction of the sensor space data was performed with beamforming, and FC was computed between 66 brain regions. FC was evaluated with four phase and envelope synchronization metrics, both with and without correction for volume conduction: envelope correlation (d-ecor), envelope correlation with leakage correction (lc-ecor), phase locking value (PLV), and phase lag index (PLI). The test-retest reliability of the FC estimates was then assessed with the intraclass correlation coefficient (ICC). The effect of the volume conduction on the variability in ICC values across FC metrics and frequency bands was assessed through a linear mixed model (LMM), including the ICC of the beamformer weights correlation. In addition, within- and between-subject agreement of FC maps was computed with Kendall's W.

Materials and Methods

Subjects and MEG acquisition

Sixteen healthy subjects (age 30.4 ± 5.8 , 10 women and 6 men) underwent three resting-state MEG sessions (two conditions: eyes open and eyes closed) over 3 consecutive weeks. For each participant, scans were scheduled at the same day of the week and same time of the day (8 subjects in the morning, 8 subjects in the afternoon). The study was approved by the local Ethics Committee. MEG resting-state recordings were acquired at the Center for Biomedical Technology (Madrid, Spain) using an Elekta Vectorview

system with 306 sensors (102 magnetometers and 204 planar gradiometers), inside a magnetically shielded room (Vacuumschmelze GmbH). During the measurements, participants sat 4 min with their eyes open staring at a white cross at 2 m distance and 4 min with their eyes closed. They were instructed to remain calm and move as little as possible. A Fastrak Polhemus system digitized each subject's head and four coils were attached to the forehead and mastoids, so that the head position with respect to the MEG helmet was continuously determined. The activity in electrooculogram channels was also recorded to keep track of ocular artifacts. Signals were sampled at 1000 Hz with an online filter of bandwidth 0.1-300 Hz. Maxfilter software (version 2.2., Elekta Neuromag) was used to remove external noise with the temporal extension of the signal space separation (tsss) method with movement compensation (Taulu and Simola, 2006).

Source reconstruction

The continuous resting-state sensor data were scanned for jump, muscle, and ocular artifacts with FieldTrip (Oostenveld et al., 2011) artifact detection procedures, and nonoverlapping artifact-free 6-sec epochs were located in the continuous resting-state recordings. This yielded 24.3 ± 5.7 clean epochs for the eyes open condition and 26.8 ± 4.7 clean epochs for the eyes closed condition. Then, MEG data were filtered into delta (2–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–45 Hz) frequency bands with a finite impulse response filter of order 1000 and downsampled to a 250 Hz sampling rate.

For each subject, source models were built by spreading dipoles over the cortical surface, with 6 mm spacing. This was achieved by segmenting each subject's T1-weighted MRI with Freesurfer (version 5.1.0; Fischl et al., 2002; Ségonne et al., 2007) and using MNE software (Gramfort et al., 2014) to downsample the corresponding surfaces and realigning to the MEG coordinate system. The forward model was then solved with a three-shell boundary element method, using the inner skull, outer skull, and scalp surfaces extracted from the subject's MRI with NFT software (Acar and Makeig, 2010) and computing leadfield values with MNE.

Source time series were reconstructed with a linearly constrained minimum variance beamformer (Van Veen et al., 1997). For each subject, session, source, and frequency band beamforming filters were built with the magnetometer covariance matrix averaged over trials and a 5% regularization factor. These filters were then applied to the continuous resting-state data, yielding a 4-min time series per source location. To avoid edge artifacts, source time series were only split into clean trials after the FC computation.

FC computation

To compute FC between source time series, four different metrics were used: envelope correlation (d-ecor), envelope correlation with leakage correction (lc-ecor), PLV, and PLI. First, for all source time series x(t), the phase $\varphi_x(t)$ and the amplitude $A_x(t)$ were extracted as argument and modulus of the complex analytical signal (Niso et al., 2013; Rosenblum et al., 2001):

$$\zeta(t) = x(t) + ix_H(t) = A_x(t)e^{i\varphi_x(t)}.$$
(1)

TEST-RETEST RELIABILITY OF MEG RESTING STATE FC

Phase synchronization was computed by evaluating the distribution of phase differences extracted from two source time series x(t) and y(t), sampled simultaneously at $\{x_k\}_{k=1,2...,n}$ and $\{y_k\}_{k=1,2...,n}$. First, the PLV measures how much this phase difference differs from a uniform distribution (Lachaux et al., 1999):

$$PLV_{xy} = \frac{1}{n} \left| \sum_{k=1}^{n} \exp\left(i \left(\varphi_{x,k} - \varphi_{y,k} \right) \right) \right|.$$
(2)

 PLV_{xy} varies between 1 for a constant phase difference and 0 for a randomly distributed phase difference (within $]-\pi,\pi]$). Second, an alternative phase synchronization measure is the PLI, which has the advantage of escaping volume conduction bias, since it is insensitive to zero and π phase differences (Stam et al., 2007):

$$PLI_{xy} = \frac{1}{n} \left| \sum_{k=1}^{n} \operatorname{sign} \left[\sin \left(\varphi_{x,k} - \varphi_{y,k} \right) \right] \right|.$$
(3)

 PLI_{xy} also varies between 0 (no synchronization or synchronization centered around 0 or $\pm \pi$) and 1 (maximal synchronization). It quantifies the asymmetry in the phase difference over the upper and lower half-unit circle (]0, π] and]- π , 0], respectively). For instance, $PLI_{xy}=1$ if $\varphi_x - \varphi_y$ remains in the [0, π] interval.

Envelope synchronization can be directly assessed with the correlation coefficient between $A_x(t)$ and $A_y(t)$, which captures simultaneous power increases/decreases between x(t) and y(t).

$$d-ecor_{xy} = |corr(A_x(t), A_y(t))|.$$
(4)

Envelopes were smoothed with a 0.5 sec moving average filter. Envelope correlations are particularly relevant in MEG RSNs, since they seem to be closely related to fMRI BOLD fluctuations and have successfully reconstructed fMRI RSNs (Brookes et al., 2011a,b; Maldjian et al., 2014). As for the PLV, d-ecor estimates could be affected by volume conduction. The lc-ecor measure accounts for this effect by performing an orthogonalization of the source time series (Brookes et al., 2012) before envelope and correlation computation:

$$y_R(t) = y(t) - \beta x(t), \tag{5}$$

where β is the linear regression coefficient between x(t) and y(t).

To avoid edge artifacts, envelopes and phases were estimated in the whole 4-min time series, but correlations PLV and PLI were only computed for the previously determined 6-sec clean trials. FC values were averaged over trials to produce a final FC estimate per pair of source locations, condition, frequency band, subject, and session.

Reliability definition and estimation

Reliability is usually considered a relative quantity, which can be defined as follows (Weir, 2005):

$$reliability = \frac{{\sigma_r}^2}{{\sigma_r}^2 + {\sigma_w}^2},$$
(6)

where σ_r^2 represents the true variance of the magnitude of interest and σ_w^2 the error variance. Reliability ranges therefore

from 0 (no reliability) to 1 (perfect reliability). In this study, we focus on test-retest reliability, which is estimated by carrying out various measurement sessions for each subject. Let us call z_{ij} the measured (or observed) score for subject *i* in session *j*. For instance, z_{ij} can designate an FC estimate between two sources for subject *i* and session *j*. Then, reliability is defined as follows (Bartlett and Frost, 2008; Scholtes et al., 2011; Weir, 2005):

$$reliability = \frac{\sigma_{z,i}^2}{\sigma_{z,i}^2 + \sigma_w^2},$$
(7)

where $\sigma_{z,i}^2$ corresponds to the true between-subject variance, and the measurement error *w* is defined as follows:

$$z_{i,j} = z_{true,i} + w_{i,j}.$$
(8)

We assume thereby that a true quantity (or, in our case, coupling strength) $Z_{true,i}$ exists for each subject *i*, and we call error any deviation between $z_{i,j}$ and $z_{true,i}$.

The intraclass correlation (or ICC, Shrout and Fleiss, 1979) is the most commonly employed reliability estimate when dealing with interval and ratio values. It evaluates between-subject variance $\sigma_{z,i}^2$ and error variance σ_w^2 from a measurement sample, and then uses Equation (7) to quantify reliability. Various ICC types exist, depending on the model used to estimate $\sigma_{z,i}^2$ and σ_w^2 . In this text, we focus on the ICC type 1–1 described in McGraw and Wong (1996); Shrout and Fleiss (1979); and Weir (2005), which applies to our study of the test-retest reliability of MEG estimates. This model is based on a one-way random effects model, in which the ordering of *j* is assumed to be irrelevant. This is equivalent to assuming that there is no systematic error between observations or subjects do not change consistently across MEG sessions.

The type 1–1 ICC is:

$$ICC = \frac{MSR - MSW}{MSR + 2MSW},\tag{9}$$

where *MSR* and *MSW* are between-subject and withinsubject mean square values, respectively, and were computed as follows (McGraw and Wong, 1996):

$$MSR = \frac{1}{n-1} \sum_{i=1}^{n} (\bar{x}_{i.} - \bar{x}_{..})^2$$
(10)

$$MSW = \frac{1}{n(k-1)} \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\bar{x}_{ij} - \bar{x}_{i.} \right)^{2},$$
 (11)

where n = 16 is the number of subjects, k = 3 is the number of MEG sessions per subject, $\overline{x}_{i.}$ is the session average for subject *i*, and $\overline{x}_{..}$ is the subject and session average. We note that, although reliability is a nonnegative magnitude, in an experiment with no true test-retest reliability, ICC estimates fluctuate around zero. It is then possible to obtain MSR < MSW and ICC < 0.

Reliability of whole-brain FC

The test-retest FC reliability was first assessed for all links connecting 66 cortical regions of interest (ROIs). These ROIs were defined in the individual's T1 volume, using Freesurfer software (version 5.1.0) and all 33 cortical regions per hemisphere of the Desikan–Killiany atlas (Desikan et al., 2006). For each ROI, a representative time series was chosen: it corresponded to the source time series that presented the highest average correlation with the other sources within the ROI. FC between ROIs was then computed using all four FC metrics described in the FC Computation section, yielding a 66×66 connectivity matrix per FC metric and frequency band. ICC was then computed for each link, FC metric, and frequency band following Equation (9).

The ICC quantifies the test-retest reliability of our FC estimates, but cannot separate which fraction of the reliability is due to the underlying true FC strength and which fraction is due to artifacts in the FC estimate, such as volume conduction. In fact, we can expect that volume conduction presents lower within- than between-subject variability, and therefore may contribute to higher test-retest reliability in FC metrics such as d-ecor or PLV, which are sensible to volume conduction. As a further step, we attempted therefore to elucidate the effect of volume conduction in the ICC. We note that this is a tricky task, since there are no direct measures of volume conduction for a given MEG recording. The correlation between beamformer weights is, however, indicative of source leakage (Hillebrand et al., 2012). In fact, if source leakage exists between two sources, their beamfomer weights are similar (or highly correlated) and their reconstructed time series are highly coupled. In the opposite case, if the beamfomer weight correlation between two sources is low, it is unlikely that high FC between these two sources result from volume conduction. We accordingly computed 66×66 beamformer weight correlation matrices for each frequency band, subject, and session as follows:

$$BWC_{pq} = corr(w_p, w_q), \qquad (12)$$

where corr(...) denotes Pearson correlation and w_p and w_q are $1 \times Nchannels$ beamformer weights for the representative sources in ROIs p and q, respectively. Next, the test-retest reliability of *BWC*, which will be referred to as *ICCBWC*, was computed for each link and frequency band with the ICC.

To assess the effect of beamformer weight correlation (BWC) in the ICC of FC values, the following LMM was employed:

$$ICCFC = X\beta + z_0 + \varepsilon_0, \tag{13}$$

where X = [iscorrd, iscorrleak, isPLV, isPLI, iscorrd·ICCBWC, iscorrleak · ICCBWC, isPLV · ICCBWC, isPLI· $ICCBWC] is the fixed effect design matrix, <math>\beta = [a_{d-ecor}a_{lc-ecor}a_{PLV}a_{PLI}b_{d-ecor}b_{lc-ecor}b_{PLV}b_{PLI}]^T$ is the vector of fixed effect coefficients, z_0 is the random effect coefficient, and ε_0 is the error term. The coefficients *b* represent the strength of the effect of *ICCBWC* in *ICCFC* for the corresponding FC metric, and the coefficients *a* represent the strength of all other fixed effects in ICCFC different than ICCBWC, for the corresponding FC metric. The model was fit (for each frequency band separately) using Matlab's *fitlmematrix* function and a restricted maximum likelihood estimation method.

Finally, to explore a possible relation between ICCFC and FC strength, Spearman correlations were computed across links between both variables, for each FC metric and frequency band separately. Furthermore, partial correlation between ICCFC and FC, controlling for the effect of ICCBWC,

was computed to account for volume conduction effects. Bonferroni correction was applied to control for the multiple comparisons problem.

Within- and between-subject agreement of whole-brain FC maps

Reliability, as computed above, depends both on withinand between-subject variability. To inspect both effects separately, we evaluated the within- and between-subject agreement of FC maps. For a given condition, FC metric, and frequency band, FC maps were sorted across links (ROI pairs) *l* and transformed into ranks r_{ijl} ($1 \le r_{ijl} \le N_{links}$, Nlinks = 66 * 65/2 = 2145) for every subject *i* and session *j* separately. Then, the agreement between two sessions j_1 and j_2 (or within-subject agreement) was assessed with Kendall's W (Legendre, 2005):

$$W_{w} = \frac{12\sum_{l=1}^{N_{links}} (A_{il} - \overline{A}_{i})^{2}}{2^{3} (N_{links}^{3} - N_{links})},$$
(14)

where $A_{il} = \sum_{j=\{j_1, j_2\}} r_{ijl}$ and $\overline{A}_i = \frac{1}{N_{links}} \sum_{l=1}^{N_{links}} A_{il}$. Analogously, the agreement between two subjects i_1 and i_2 (or between-subject agreement) was assessed with:

$$W_{b} = \frac{12\sum_{l=1}^{N_{links}} \left(B_{jl} - \overline{B}_{j}\right)^{2}}{2^{3} \left(N_{links}^{3} - N_{links}\right)},$$
 (15)

where $B_{jl} = \sum_{i=\{i_1i_2\}} r_{ijl}$ and $\overline{B}_j = \frac{1}{N_{links}} \sum_{l=1}^{N_{links}} B_{jl}$. W ranges from 0 to 1 and reaches 1 when the ranks obtained in two sessions/subjects are in perfect agreement. Wilcoxon rank sum tests were performed to compare within- and betweensubject W values for all FC metrics and frequency bands, and Bonferroni correction was applied to deal with the multiple comparisons problem.

Reliability of RSNs

To further investigate the spatial characteristics of the reliability of resting-state FC, we focused on seven wellknown RSNs: visual, sensorimotor, auditory, DMN (default mode network), left FP (frontoparietal), right FP, and frontoinsula networks. Each RSN was defined as a set of 2 or 3 nodes, following their spatial extent described previously in the literature (Laird et al., 2011; Smith et al., 2009). The nodes were defined as a 1 cm-radius sphere around a central coordinate defined in MNI space (Table 1), which was then transformed into subject space with a homogeneous transformation extracted from Freesurfer's cortical segmentation process.

First, within-network reliability was quantified as the ICC of FC between all source pairs connecting given RSNs nodes, for each FC metric, RSN, and frequency band separately. Second, the reliability of seed-based FC maps was also quantified, using all nodes in Table 1 as seeds and computing the ICC of the FC between the seed and the remaining source locations. To enable within-subject comparisons, individual source meshes were transformed into a template mesh in MNI space with $N_{sources} = 4739$ sources generated with Freesurfer from the subjects' cortical segmentations. FC values were then spatially smoothed with a 15 mm moving average filter.

Network	Node	x	у	z	BA
Visual	Left visual cortex	-41	-77	3	19
	Right visual cortex	41	-72	1	19
Sensorimotor	Left primary somatosensory cortex	-38	-27	52	1
	Right primary somatosensory cortex	40	-25	50	1
Auditory	Left primary auditory cortex	-55	-21	7	41
·	Right primary auditory cortex	57	-20	7	41
DMN	Precuneus	1	-57	28	31
	Left inferior parietal	-45	-65	30	39
	Right inferior parietal	52	-60	26	39
	Anterior cingulate	2	42	7	32
Left FP	Left angular gyrus	-40	-58	56	39
	Left middle frontal gyrus	-42	28	23	9
Right FP	Right angular gyrus	53	-50	43	39
e	Right middle frontal gyrus	45	28	26	9
Frontoinsular	Left insula	-38	16	2	13
	Right insula	38	16	2	13
	Median cingulate	-2	12	40	32

TABLE 1. MNI COORDINATES OF THE NODES FORMING THE RESTING-STATE NETWORKS

MNI coordinates are obtained from the functional magnetic resonance imaging RSNs released by Laird et al. (2011) and Smith et al. (2009). Labels are assigned with FSL atlas tools (Desikan et al., 2006; Eickhoff et al., 2005; Jenkinson et al., 2012).

BA, Broadmann area; DMN, default mode network; FP, frontoparietal.

Results

Reliability of whole-brain FC maps

ICCFC distribution. The distribution of ICCFC values, which quantify the test-retest reliability of eyes closed resting-state FC, is represented in Figure 1A. ICCFC was greatly dependent both on the FC metric and the frequency band. Highest values were reached for the PLV in alpha, beta, and gamma bands (averages 0.74-0.82). For d-ecor and lc-ecor, ICC values were generally smaller, and peaked in the beta band (average 0.23-0.37). For PLI, average ICC was lower (<0.1) for all frequency bands. The same effects across frequency bands and FC metric were obtained for the eyes open condition: results are displayed in Supplementary



FIG. 1. Distribution of ICC values for whole-brain FC during restingstate eyes closed. (A) Distribution of ICCFC values across links, for each FC metric and frequency band separately. The colored boxes indicate the position of the first, second, and third quartile, and the central lines span the range from the 5th to the 95th percentile. (B) Estimated a coefficients in the LMM described in the Reliability of whole-brain FC section. (C) Estimated b coefficients in the LMM. For (B) and (C), the colored boxes span the coefficients' 99% confidence interval. Stars indicate that the corresponding coefficient is significantly larger than zero (p < 0.01, Bonferroni corrected). FC, functional connectivity; ICC, intraclass correlation coefficient; LMM, linear mixed model. Color images available online at www .liebertpub.com/brain

Figure S1 (Supplementary Data are available online at www .liebertpub.com/brain). Given that eyes open and closed conditions were also found to behave similarly in subsequent analyses, the following sections will for simplicity only describe the eyes closed results.

BWC effect. The contribution of BWC to the previously described ICCFC was explored by introducing both ICCFC and ICCBWC in an LMM, as detailed in the Reliability of whole-brain FC section. The 99% confidence intervals of the LMM coefficients are displayed in Figure 1B and C. On the one hand, the *a* coefficients, which weigh ICCBWC-unrelated fixed effects on ICCFC, were significantly greater than zero (p < 0.01, Bonferroni corrected) for PLV for all frequency bands, lc-ecor in alpha and beta, d-ecor in beta, and PLI in alpha. On the other hand, the *b* coefficients, which weigh the contribution of ICCBWC to ICCFC, were significantly greater than zero (p < 0.01, Bonferroni corrected) for d-ecor and PLV for all frequency bands, while there were close to zero for leakage-corrected FC metrics (PLI and d-ecor).

Reliability was, however, unevenly distributed across links. Figure 2 represents the 5% of links with highest $ICCFC - b \cdot ICCBWC$ (i.e., the links with highest ICCFC after subtracting the contribution of BWC). The topography of these links differed across FC metrics and frequency bands. As a general trend, the following areas were found to be of especial relevance: anterior cingulate and temporal ROIs in delta, temporal, occipital, and posterior ROIs in theta, temporal, occip ital, and anterior ROIs in alpha, central and temporal ROIs in beta, and parietal and frontal ROIs in gamma.

Dependence between ICC and FC strength. The dependence between ICC and FC was investigated with Spearman correlations. The results are listed in Table 2. ICC and FC were positively correlated, especially for d-ecor and PLV (r=0.50-0.78). Correlations for lc-ecor and PLI were significant in alpha and beta bands (r=0.14-0.46), and smaller (ABS|r| < 0.08) in delta, theta and gamma. When computing partial correlation between ICC and FC, while controlling for ICCBWC, correlation coefficients for d-ecor and PLV were much smaller, although remained significant for all frequency bands (r=0.29-0.57). On the other hand, for lc-ecor and PLI, as expected, correlation coefficients between ICCFC and FC remained unchanged after controlling for ICCBWC.

Within- and between-subject agreement in whole-brain FC maps

The within- and between-subject agreement of FC maps was computed with Kendall's W from pairwise maps of a single and distinct subjects. The distribution of W values for every frequency band and FC metric is displayed in Figure 3. First, highest within- and between-subject average W was found for PLV (0.59–0.98). Moreover, within-subject W was significantly higher than between-subject W for PLV (20% increment on average), and this difference was higher for high-frequency bands (22% increment in beta and gamma



	Correlation ICCFC-FC				Parti	ial correlation IC	CCFC-FC (ICC	CBWC)
	d-ecor	lc-ecor	PLV	PLI	d-ecor	lc-ecor	PLV	PLI
Delta	0.54*	-0.04	0.78*	-0.01	0.30*	-0.04	0.57*	-0.02
Theta	0.63*	0.00	0.71*	0.00	0.36*	0.00	0.48*	0.00
Alpha	0.70*	0.16*	0.62*	0.27*	0.40*	0.16*	0.43*	0.27*
Beta	0.70*	0.46*	0.50*	0.14*	0.56*	0.46*	0.30*	0.13*
Gamma	0.69*	0.08*	0.51*	-0.02	0.43*	0.08*	0.29*	-0.02

TABLE 2. CORRELATION BETWEEN ICC AND FC

In the left side, the Spearman correlation coefficients between ICC and FC across links are listed, for all FC metrics and frequency bands separately, for the resting-state eyes closed condition. In the right side, Spearman partial correlation coefficients between ICCFC and FC, controlling for BWC ICC are listed. Asterisks denote significant correlations (Bonferoni corrected, p < 0.01).

BWC, beamformer weight correlation; FC, functional connectivity; ICC, intraclass correlation coefficient; PLI, phase lag index; PLV, phase locking value.

vs. 16% in delta). Second, high average W values (0.56–0.82) were also found for d-ecor, and within-subject W was higher than between-subject W for all frequency bands (8.4% on average). Third, for lc-ecor, within- and between-subject W were generally smaller (0.42–0.63) and similar, although for the beta band, a higher W (0.42–0.79) was obtained, and within-subject W significantly exceeded between-subject W. Fourth, for PLI both between- and within-subject W values were smaller (0.48–0.60), and within-subject W was significantly greater than between-subject W for alpha and beta bands.

Reliability of RSNs

Within-network reliability. By exploring the test-retest reliability of whole-brain FC in the Reliability of Whole-Brain FC Maps section, it was found that ICC values have a great spatial variability and some of the links with higher testretest reliability bear some similarity with the RSNs described in the literature. The test-retest reliability of seven RSNs was therefore explored. First, the reliability of the FC of links connecting nodes of the same RSN was assessed with the ICC. Results are listed in Table 3. ICC values were dependent both on the FC metric and frequency band. For all RSNs, highest ICC values were found for PLV, usually ranging from 0.7 to 0.9 in alpha, beta, and gamma bands and smaller values from 0.4 to 0.7 in delta and theta bands. High ICC values (>0.7) were, however, also found in delta and theta bands for the sensorimotor and visual networks. d-ecor and lc-ecor presented greater variability across frequency bands than PLV. The ICC of both FC metrics topped in the beta band (ICC = 0.46-0.78 for d-ecor and ICC = 0.43-0.66 for lc-ecor), while low to medium ICC (<0.5) was found for delta, theta, and gamma bands. Although ICC was rather low in the alpha band, high values (>0.6) were found with d-ecor in the visual and sensorimotor networks. With PLV and envelope correlation metrics, higher ICC values were found for the sensory networks (especially sensorimotor and visual), than for DMN, FP, and frontoinsular networks. Finally, ICC values were medium to low for the PLI in all frequency bands and networks and did not exceed 0.44.

Reliability of seed-based FC maps. FC was computed from each seed to the remaining source location for each condition, frequency band, and FC metric separately. Then, the ICC of the FC values was computed for each link. The distributions of ICC values over all seed locations are displayed in Supplementary Figure S2. These ICC values were lower than the corresponding within-network ones for most seeds for PLV and d-ecor across frequency bands, lcecor in alpha, and beta and PLI in alpha indicating greater within- than between-network reliability. Moreover, the variability over frequency bands and FC metrics was similar than in the within-network analysis. First, highest values (ICC >0.6) were found for PLV, especially in alpha, beta, and gamma bands for all seed locations, although slightly higher ICC was found for sensory seeds and precuneus. Second, ICC values for d-ecor and lc-ecor were small for delta, theta, and gamma bands (third quartile <0.3), but higher values were obtained for alpha and especially beta bands (ICC 0.3-0.6). ICC values also varied somewhat across seeds, and were higher in the beta band in sensorimotor and left middle frontal seeds with lc-ecor and in the alpha band in the sensory and precuneus seeds for both lc-ecor and d-ecor. Third, ICC was poor for PLI (third quartile <0.2).



FIG. 3. Distribution of within- and between-subject agreement of FC maps. Each pair of same-color boxplots represents the within- (left) and between-subject (right) agreement for given a frequency band and FC metric, for the resting-state eyes closed condition. Stars indicate significantly higher within- than between-subject agreement (Wilcoxon rank sum test, p < 0.01, Bonferroni corrected). Color images available online at www .liebertpub.com/brain

Network		Visual	Sensorimotor	Auditory	DMN	Left FP	Right FP	Frontoinsular
d-ecor	Delta Theta	0.53 0.69	0.29 0.49	$0.23 \\ -0.03$	0.18 0.22	0.05 0.13	0.20 0.32	0.05 0.24
	Alpha Beta	0.61 0.69	0.73 0.76	0.64 0.58	0.32 0.42	0.38 0.46	0.66 0.78	0.29 0.59
lc-eco r	Gamma Delta	0.34 0.20	$0.46 \\ -0.03$	0.43 0.10	0.32 0.10	$\begin{array}{c} 0.20 \\ -0.08 \end{array}$	$0.10 \\ -0.22$	$\begin{array}{c} 0.15 \\ -0.07 \end{array}$
	Theta Alpha	0.07 0.51	$-0.15 \\ 0.25$	0.02 0.17	0.01 0.32	0.19 0.02	$-0.06 \\ 0.25$	$-0.02 \\ 0.11$
	Beta Gamma	$0.66 \\ -0.02$	0.54 0.15	0.52 0.20	$0.45 \\ 0.00$	0.43 0.00	$0.55 \\ -0.32$	$\begin{array}{c} 0.47 \\ -0.04 \end{array}$
PLV	Delta Theta	$0.76 \\ 0.85$	0.82 0.91	0.72 0.67	0.41 0.59	$0.56 \\ 0.67$	0.53 0.71	$0.49 \\ 0.72$
	Alpha Beta	$\begin{array}{c} 0.86\\ 0.89 \end{array}$	0.95 0.91	0.71 0.73	$0.75 \\ 0.84$	$\begin{array}{c} 0.87\\ 0.88\end{array}$	0.82 0.86	$0.67 \\ 0.78$
PLI	Gamma Delta	0.90 0.07	$\begin{array}{c} 0.87 \\ -0.09 \end{array}$	$0.68 \\ -0.03$	$\begin{array}{c} 0.88 \\ -0.06 \end{array}$	0.78 0.19	0.92 0.17	$\begin{array}{c} 0.84 \\ -0.08 \end{array}$
	Theta Alpha Beta	-0.10 0.40 -0.14	$0.05 \\ -0.10 \\ -0.10$	-0.19 0.03 0.22	$0.06 \\ 0.08 \\ 0.07$	-0.14 0.11 -0.02	-0.07 0.12 -0.10	-0.05 0.07 0.02
	Gamma	-0.05	-0.15	-0.01	-0.10	-0.21	0.11	-0.02

TABLE 3. ICC OF THE WITHIN-NETWORK FC FOR THE RESTING-STATE EYES CLOSED CONDITION

ICC values were computed for every frequency band, FC metric, and RSN separately. For RSNs with more than two nodes (DMN and frontoinsular), the average ICC across links within the network is listed. RSNs, resting-state networks.



FIG. 4. FC and ICC for beta band resting-state eyes closed and lc-ecor. For three selected seeds (left primary somatosensory, left primary auditory, and precuneus), average FC maps (first row), ICC (second row), and the normalized joint histo-gram of ICC and FC values (third row) are displayed. FC maps are averaged over all sessions and subjects. The seed location is indicated with a white circle. Color images available online at www.liebertpub.com/brain



FIG. 5. FC and ICC for beta band resting-state eyes closed and PLV. For three selected seeds (left primary somatosensory, left primary auditory, and precuneus), average FC maps (first row), ICC (second row), and the normalized joint histogram of ICC and FC values (third row) are displayed. FC maps are averaged over all sessions and subjects. The seed location is indicated with a white circle. PLV, phase locking value. Color images available online at www.liebertpub.com/brain

ICC values were unevenly distributed across brain regions. As an example, beta band ICC and average FC maps for three selected seeds are displayed for lc-ecor in Figure 5 and PLV in Figure 4 (corresponding figures for d-ecor and PLI can be found in the Supplementary Figures S3 and S4). Average FC was generally higher between regions belonging to the same network. For instance, seeds in sensory networks (visual, somatosensory, and auditory cortices) were strongly connected to their homologous contralateral areas. This was less evident for DMN, FP, and frontoinsular networks. ICC maps resembled average FC maps, showing high ICC for links between homologous sensory areas. DMN and FP networks presented lower ICC than sensory networks for all FC metrics. Moreover, FC increased along with ICC values for d-ecor, lc-ecor, and PLV, as revealed by the joint distribution of both magnitudes. This relation was, however, not linear, especially for FC metrics without zero-lag correction (PLV and d-ecor), for which the ICC increased first rapidly along with FC values and then saturated for medium to high FC (Fig. 5 and Supplementary Fig. S3).

Discussion

In this work, we assessed the test-retest reliability of resting-state FC by exploring within- and between-subject

variability of FC values in a cohort of 16 subjects. Following previous resting-state MEG literature (Brookes et al., 2011b; Hillebrand et al., 2005; Schoffelen and Gross, 2009), source time series were estimated with beamforming. Then, source space FC was computed between source locations using four distinct FC metrics: d-ecor, lc-ecor, PLV, and PLI. Withinand between-subject variability was employed to assess both reliability (with the ICC) and agreement (with Kendall's W) of the FC maps. First, overall trends over FC metrics and frequency bands were inspected for whole-brain connections, exploring all links connecting 66 cortical ROIs. Average ICCFC was highest in the PLV in alpha, beta, and gamma bands (0.74-0.82), medium for d-ecor and lc-ecor in the beta band (average 0.23–0.37), and considerably low (<0.1) for PLI across frequency bands. An LMM, including ICCBWC revealed that volume conduction, indirectly assessed through BWC, could be contributing to high testretest reliability in PLV and d-ecor. When accounting for this effect, the overall test-retest reliability of whole-brain FC was found to be significant for PLV for all frequency bands, lc-ecor in alpha and beta, d-ecor in beta, and PLI in alpha. When evaluating within- and between-subject agreement separately with Kendall's coefficient of concordance W, medium to high values were obtained (0.5-0.97), depending on the frequency band and the FC metric. W was highest with PLV and lowest with lc-ecor. In addition, within-subject W was significantly higher than between-subject FC for d-ecor and PLV throughout frequency bands, PLI in alpha and beta, and lc-ecor in beta.

In addition, FC metric and frequency band were not the only important factors influencing test-retest reliability. Instead, the FC test-retest reliability was highly dependent on the specific link considered. First, ICCFC correlated positively with the FC strength, especially in alpha and beta bands. Second, when focusing on well-known RSNs, withinnetwork connections were particularly reliable. Within each network, similar FC metrics and frequency band trends were found than for whole-brain FC: highest values found for PLV in alpha, beta, and gamma bands and for d-ecor and lc-ecor in beta band.

To our knowledge, this constitutes the first test-retest reliability study of MEG source space FC. Test-retest reliability of FC has, however, already been explored with both sensor space MEG/EEG and fMRI. Deuker et al. (2009) and Jin et al. (2011) employed the ICC to estimate the reliability of resting-state graph metrics derived from MEG sensor space mutual information. In the first study, Deuker et al. (2009) obtained high ICC (0.5-0.8) for strength, clustering, minimum path length, synchronizability, global efficiency, cost efficiency, and small world index in alpha band, and low to medium ICC (0–0.5) in other frequency bands. Interestingly, the ICC of these graph metrics during a working memory task was found to be higher than during resting state. In the latter study, Jin et al. (2011) found medium ICC (0.4-0.65) for nodal efficiency, nodal degree, and betweeness centrality in alpha and beta bands and lower ICC for theta and gamma bands. Using EEG and sensor space PLI, Hardmeier et al. (2014) obtained good reliability (ICC range: 0.5–0.8) for clustering coefficient, average path length, and small-world index in theta, alpha, and beta bands. Test-retest studies of sensor space graph metrics have therefore obtained divergent results, although they cannot be easily compared because of differences in the processing methodology, the choice of FC metric, and the test-retest interval. Using resting-state fMRI, Shehzad et al. (2009) could directly inspect the test-retest reliability of RSNs by computing FC between ROIs selected from the fMRI literature. They overall obtained low reliability (average ICC = 0.10– 0.40), although higher ICC was found when considering significant correlations exclusively. This is in agreement with our results and suggests that stronger connections are more reliable. Although not focusing on reliability, Wens et al. (2014) studied short-term within-subject variability in MEG FC with four subjects who underwent 20 scan sessions in a single day. After selecting three seeds in visual, auditory, and sensorimotor cortices, they evaluated the spatial similarity between FC maps. They finally recommended between 2 and 11 sessions to produce a robust subject average, depending on the network. Moreover, in line with our work, they obtained that the sensorimotor network presented the highest within-subject similarity and hence required the least sessions to produce good subject-representative average FC maps.

An interesting result in our work is the frequency specificity of the reliability of RSNs. With PLV, RSNs reached medium to good reliability for all frequency bands, but highest ICC values were consistently found in alpha, beta, and gamma bands. In the previously commented graph metric studies (Deuker et al., 2009; Hardmeier et al., 2014; Jin et al., 2011), reliability was generally highest for alpha, reasonably good for theta and beta, and quite low for gamma. Although we cannot establish any direct analogy between our approach and sensor space graph metrics, given that high frequency oscillations are supposed to be more local than low-frequency oscillations (Buzsáki and Draguhn, 2004), we could speculate that we found higher gamma band reliability because gamma network properties are not well captured with sensor space FC. When using d-ecor and lc-ecor, reliability varied greatly among frequency bands and was higher in beta, followed by alpha. This finding is not particularly surprising, since MEG RSNs have been predominantly studied with envelope correlation in the alpha and beta bands. In fact, although the seminal work of (Brookes et al., 2011a,b; de Pasquale et al., 2010; Hipp et al., 2012) investigated RSNs with envelope correlation at several frequency bands, they concluded that most networks were best extracted in the beta band and alpha band in de Pasquale et al. (2010) and in Brookes et al. (2011b) for the DMN. In this study, we found that the test-retest reliability was quite high across networks for alpha and beta band with envelope correlation measures, although higher values were generally found in sensory networks and DMN.

In addition, we found that the choice of FC metric greatly influenced reliability. ICC was highest for PLV, medium for d-ecor and lc-ecor, and lowest for PLI, indicating that PLV estimates are the most reliable ones. The most striking difference was found between PLV and PLI: PLV presented consistently higher test-retest reliability than PLI. This could be due to the fact that both metrics capture different aspects of phase synchronization: while PLV evaluates how much the distribution of phase differences between brain oscillations deviates from a uniform distribution, PLI quantifies asymmetries in the phase differences over $[0, \pi]$ and $[-\pi, 0]$. The differences in test-retest reliability could be therefore caused by very different effects. On the one hand, it is possible that zero-lag synchronization represents meaningful and reliable FC, since various animal studies with intracranial EEG have proven the existence of true zero-phase lag synchronization between distinct areas (Fell and Axmacher, 2011; Frien et al., 1994; Roelfsema et al., 1997). On the other hand, it is also possible that volume conduction, which presents higher within- than between-subject variability, spuriously increased the magnitude of the test-retest reliability of PLV (and d-ecor) estimates. Recent literature of MEG RSNs has, however, predominantly employed d-ecor and lc-ecor. Envelope correlation has in fact produced FC maps, which resemble the ones obtained with fMRI (Brookes et al., 2011b; de Pasquale et al., 2010, 2012). Furthermore, lcecor refines d-ecor by eliminating zero-lag effects, using either time domain (Brookes et al., 2012) or frequency domain (Hipp et al., 2012; Wens et al., 2014) orthogonalization. The average FC maps obtained in this study were similar to the ones obtained in the previously commented studies: with lc-ecor, homologous areas in sensory networks were strongly connected and, contrary to PLV or d-ecor, the neighborhood of the seed did not present artifactual high values. Our results suggest nonetheless, that, although lc-ecor produces meaningful and unbiased FC maps, PLV estimates present higher test-retest reliability. Furthermore, phase and amplitude

TEST-RETEST RELIABILITY OF MEG RESTING STATE FC

synchronization evaluate different mechanisms of information transfer, so that the choice of FC metric should depend on the underlying FC model. On the one hand, the phase of the oscillations is supposed to represent the degree of excitability of neuronal populations, and phase synchronization has been thought to contribute both to precise spike timing, especially in the gamma band, and to neural communication and plasticity (Fell and Axmacher, 2011). On the other hand, the amplitude of the oscillations represents the intensity of coherent neural activity, and amplitude correlation reflects the simultaneous activation of distinct brain areas.

Finally, it is important to note that test-retest reliability does not evaluate the goodness of the measurement (in this case, FC estimation). The test-retest reliability is instead a relative quantity, which depends on within- and betweensubject variability. Any within-subject variability produces a reliability decrease, independent of the origin: instrumentation (e.g., MEG sensors deviating from their ideal working point), measurement situation (e.g., varying magnetic fields interfering in the MEG measurement, acoustic noise that could induce interfering brain activity in the subject) or the subject itself (true change in the coupling strength between two regions within a single individual). We can therefore not conclude that an FC estimate with low test-retest reliability (small ICC) contains a great amount of error or is inaccurate: it could as well be possible that it captures a true withinsubject change. In this line, higher within-subject agreement in FC maps does not necessarily imply that FC is more stable within individuals: it could as well be due to volume conduction bias in the FC metric. Reliability relates instead on how well subjects can be distinguished in a given group. It is therefore indicative of whether FC estimates can be employed to characterize individual subjects within the group sample. We should also indicate that reliability is a population-specific quantity: for a fixed within-subject variability, the more heterogeneous the subject sample, the higher the reliability.

Conclusions

In conclusion, we provided the first source space test-retest reliability report of MEG resting-state FC. We computed both whole-brain FC between 66 ROIs and seed-based FC (focusing on RSNs nodes) using four commonly employed FC metrics: d-ecor, lc-ecor, PLV, and PLI. We then evaluated the test-retest reliability of the FC estimates with the ICC, and the within- and between-subject agreement of the FC maps with Kendall's W. ICC values were greatly dependent on the FC metric, the frequency band, and the source locations. Overall, highest ICC values were obtained for PLV, even reaching 0.8 or higher for broad regions, especially in alpha, beta, and gamma bands. D-ecor and lc-ecor presented similar reliability, and generally yielded low to medium ICC (<0.5), except for alpha and beta bands, where ICC values of 0.6 were reached. PLI estimates yielded the lowest ICC throughout frequency bands, and ICC values rarely exceeded 0.4. Joint inspection of the ICCFC (test-retest reliability of FC) and the ICCBWC (test-retest reliability of beamformer weights correlation) revealed that volume conduction (assessed through the BWC) could be partly responsible for the high ICCFC obtained in PLV and d-ecor estimates. Across FC metrics and frequency bands (but especially in alpha and beta bands), ICCFC values correlated positively with FC strength, and within-network connections presented higher ICCFC than between-network connections. Both findings point out a link between stronger connections and higher test-retest reliability. Reliability was defined in this study as a relative quantity (Weir, 2005), which represents the fraction of the measurement variability that is accounted for by between-subject variability. Our results are therefore encouraging for future studies, which would aim at using FC strength in RSNs to characterize individual subjects, especially when using PLV in theta, alpha, beta, and gamma bands, or envelope correlation measures in the alpha and beta bands.

Acknowledgments

The authors are grateful to the research participants for their participation in this study. Research by P.G. has been supported by a PICATA predoctoral contract of the Moncloa Campus of International Excellence (UCM–UPM).

Author Disclosure Statement

No competing financial interests exist.

References

- Acar ZA, Makeig S. 2010. Neuroelectromagnetic forward head modeling toolbox. J Neurosci Methods 190:258–270.
- Baldassarre A, Lewis CM, Committeri G, Snyder AZ, Romani GL, Corbetta M. 2012. Individual variability in functional connectivity predicts performance of a perceptual task. Proc Natl Acad Sci USA 109:3516–3521.
- Bartlett JW, Frost C. 2008. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol 31:466–475.
- Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34:537–541.
- Bowyer SM, Gjini K, Zhu X, Kim L, Moran JE, Rizvi SU, et al. 2015. Potential biomarkers of schizophrenia from MEG resting-state functional connectivity networks: Preliminary data. J Behav Brain Sci 05:1–11.
- Brookes MJ, Hale JR, Zumer JM, Stevenson CM, Francis ST, Barnes GR, et al. 2011a. Measuring functional connectivity using MEG: Methodology and comparison with fcMRI. Neuroimage 56:1082–1104.
- Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, et al. 2011b. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proc Natl Acad Sci USA 108:16783–16788.
- Brookes MJ, Woolrich MW, Barnes GR. 2012. Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. Neuroimage 63:910–920.
- Buzsáki G, Draguhn A. 2004. Neuronal oscillations in cortical networks. Science 304:1926–1929.
- Deco G, Jirsa VK, McIntosh AR. 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci 12:43–56.
- de Pasquale F, Della Penna S, Snyder AZ, Lewis C, Mantini D, Marzetti L, et al. 2010. Temporal dynamics of spontaneous MEG activity in brain networks. Proc Natl Acad Sci USA 107:6040–6045.
- de Pasquale F, Della Penna S, Snyder AZ, Marzetti L, Pizzella V, Romani GL, Corbetta M. 2012. A cortical core for

dynamic integration of functional networks in the resting human brain. Neuron 74:753–764.

- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.
- Deuker L, Bullmore ET, Smith M, Christensen S, Nathan PJ, Rockstroh B, Bassett DS. 2009. Reproducibility of graph metrics of human brain functional networks. Neuroimage 47:1460–1468.
- Engels MM, Stam CJ, van der Flier WM, Scheltens P, de Waal H, van Straaten EC. 2015. Declining functional connectivity and changing hub locations in Alzheimer's disease: An EEG study. BMC Neurol 15:145.
- Fell J, Axmacher N. 2011. The role of phase synchronization in memory processes. Nat Rev Neurosci 12:105–118.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. 2002. Whole brain segmentation. Neuron 33:341–355.
- Frien A, Eckhorn R, Bauer R, Woelbern T, Kehr H. 1994. Stimulus-specific fast oscillations at zero phase between visual areas V1 and V2 of awake monkey. Neuroreport 5:2273–2277.
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. 2014. MNE software for processing MEG and EEG data. Neuroimage 86:446–460.
- Hall EL, Woolrich MW, Thomaz CE, Morris PG, Brookes MJ. 2013. Using variance information in magnetoencephalography measures of functional connectivity. Neuroimage 67: 203–212.
- Hardmeier M, Hatz F, Bousleiman H, Schindler C, Stam CJ, Fuhr P. 2014. Reproducibility of functional connectivity and graph measures based on the phase lag index (PLI) and weighted phase lag index (wPLI) derived from high resolution EEG. PLoS One 9:e108648.
- Hillebrand A, Barnes GR, Bosboom JL, Berendse HW, Stam CJ. 2012. Frequency-dependent functional connectivity within resting-state networks: An atlas-based MEG beamformer solution. Neuroimage 59:3909–3921.
- Hillebrand A, Singh KD, Holliday IE, Furlong PL, Barnes GR. 2005. A new approach to neuroimaging with magnetoencephalography. Hum Brain Mapp 25:199–211.
- Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK. 2012. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat Neurosci 15:884–890.
- Jin S-H, Seol J, Kim JS, Chung CK. 2011. How reliable are the functional connectivity networks of MEG in resting states? J Neurophysiol 106:2888–2895.
- Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. 1999. Measuring phase synchrony in brain signals. Hum Brain Mapp 8: 194–208.
- Laird AR, Fox PM, Eickhoff SB, Turner JA, Ray KL, McKay DR, et al. 2011. Behavioral interpretations of intrinsic connectivity networks. J Cogn Neurosci 23:4022–4037.
- Legendre P. 2005. Species associations: The Kendall coefficient of concordance revisited. J Agric Biol Environ Stat 10:226– 245.
- López ME, Bruña R, Aurtenetxe S, Pineda-Pardo JÁ, Marcos A, Arrazola J, et al. 2014. Alpha-band hypersynchronization in progressive mild cognitive impairment: A magnetoencephalography study. J Neurosci 34:14551–14559.
- Luckhoo H, Hale JR, Stokes MG, Nobre AC, Morris PG, Brookes MJ, Woolrich MW. 2012. Inferring task-related networks using independent component analysis in magnetoencephalography. Neuroimage 62:530–541.

- Maldjian JA, Davenport EM, Whitlow CT. 2014. Graph theoretical analysis of resting-state MEG data: Identifying interhemispheric connectivity and the default mode. Neuroimage 96:88–94.
- McGraw K, Wong S. 1996. Forming inferences about some intraclass correlation coefficients. Psychol Methods.
- Niso G, Bruña R, Pereda E, Gutiérrez R, Bajo R, Maestú F, del-Pozo F. 2013. HERMES: Towards an integrated toolbox to characterize functional and effective brain connectivity. Neuroinformatics 11:405–434.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M. 2011. Field-Trip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869.
- Pereda E, Quiroga RQ, Bhattacharya J. 2005. Nonlinear multivariate analysis of neurophysiological signals. Prog Neurobiol 77:1–37.
- Roelfsema PR, Engel AK, König P, Singer W. 1997. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. Nature 385:157–161.
- Rosazza C, Minati L. 2011. Resting-state brain networks: Literature review and clinical applications. Neurol Sci 32:773– 785.
- Rosenblum M, Pikovsky A, Kurths J, Schäfer C, Tass PA. 2001. Neuro-Informatics and Neural Modelling, Handbook of Biological Physics. New York, NY: Elsevier.
- Schoffelen JM, Gross J. 2009. Source connectivity analysis with MEG and EEG. Hum Brain Mapp 30:1857–1865.
- Scholtes VA, Terwee CB, Poolman RW. 2011. What makes a measurement instrument valid and reliable? Injury 42:236–240.
- Ségonne F, Pacheco J, Fischl B. 2007. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Trans Med Imaging 26:518–529.
- Shehzad Z, Kelly AMC, Reiss PT, Gee DG, Gotimer K, Uddin LQ, et al. 2009. The resting brain: Unconstrained yet reliable. Cereb Cortex 19:2209–2229.
- Shrout PE, Fleiss JL. 1979. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 86:420–428.
- Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, et al. 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci USA 106:13040–13045.
- Stam CJ, Nolte G, Daffertshofer A. 2007. Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum Brain Mapp 28:1178–1193.
- Taulu S, Simola J. 2006. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys Med Biol 51:1759–1768.
- Tewarie P, Schoonheim MM, Schouten DI, Polman CH, Balk LJ, Uitdehaag BMJ, et al. 2015. Functional brain networks: Linking thalamic atrophy to clinical disability in multiple sclerosis, a multimodal fMRI and MEG Study. Hum Brain Mapp 36:603–618.
- van den Heuvel MP, Hulshoff Pol HE. 2010. Exploring the brain network: A review on resting-state fMRI functional connectivity. Eur Neuropsychopharmacol 20:519–534.
- van den Heuvel MP, Mandl RCW, Kahn RS, Hulshoff Pol HE. 2009. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. Hum Brain Mapp 30:3127–3141.
- Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. 1997. Localization of brain electrical activity via linearly constrained

TEST-RETEST RELIABILITY OF MEG RESTING STATE FC

minimum variance spatial filtering. IEEE Trans Biomed Eng 44:867–880.

- Weir JP. 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 19:231–240.
- Wens V, Bourguignon M, Goldman S, Marty B, Op de Beeck M, Clumeck C, et al. 2014. Inter- and intra-subject variability of neuromagnetic resting state networks. Brain Topogr 27:620– 634.
- Yamashita M, Kawato M, Imamizu H. 2015. Predicting learning plateau of working memory from whole-brain intrinsic network connectivity patterns. Sci Rep 5:7622.

Address correspondence to: Pilar Garcés Laboratory of Cognitive and Computational Neuroscience (UCM-UPM) Centre for Biomedical Technology (CTB) Campus de Montegancedo s/n 28223 Pozuelo de Alarcón Madrid Spain

E-mail: pilar.garces@ctb.upm.es