Resting-state fMRI can reliably map neural networks in children

[1] MORIAH E. THOMASON, [2] EMILY L. DENNIS, [2,3] ANAND A. JOSHI, [2,3] SHANTANU H. JOSHI, [2,3] IVO D. DINOV, [4] CATIE CHANG, [5] MELISSA L. HENRY, [5] REBECCA F. JOHNSON, [2,3] PAUL M. THOMPSON, [2,3] ARTHUR W. TOGA, [4] GARY H. GLOVER, [2,3] JACK D. VAN HORN, AND [5] IAN H. GOTLIB

[1] Department of Pediatrics, Wayne State University School of Medicine, Detroit, MI;
[2] Laboratory of Neuro Imaging (LONI), Department of Neurology, David Geffen School of Medicine, UCLA, Los Angeles, California;
[3] Center for Computational Biology at UCLA, Los Angeles, California;
[4] Department of Radiology, Stanford University School of Medicine, Stanford, California.
[5] Department of Psychology, Stanford University, Stanford, CA;

Correspondence:
Moriah Thomason, Ph.D.
Department of Pediatrics
Wayne State University
Merrill Palmer Skillman Institute
Detroit, MI 48202
moriah@stanford.edu
650-670-0717

Running Title: rs-fMRI reliability in children
Prepared for: NeuroImage

Abstract

Resting-state MRI (rs-fMRI) is a powerful procedure for studying whole brain neural connectivity. In this study we provide the first empirical evidence of the longitudinal reliability of rs-fMRI in children. We compared rest-retest measurements across spatial, temporal, and frequency domains for each of six cognitive and sensorimotor intrinsic connectivity networks (ICNs) both within and between scan sessions. Using Kendall's W, concordance of spatial maps ranged from .60 to .86 across networks, for various derived measures. The Pearson correlation coefficient for temporal coherence between networks across all Time one - Time two (T1/T2) z-converted measures was .66 ($p<.001$). There were no differences between T1/T2 measurements in low-frequency power of the ICNs. For the visual network, within-session T1 correlated with the T2 low-frequency power, across participants. These measures from resting-state data in children were consistent across multiple domains (spatial, temporal, and frequency). Resting-state connectivity is therefore a reliable method for assessing large-scale brain networks in children.

Introduction

Efforts in functional magnetic resonance imaging (fMRI) are shifting from research focused on specific cognitive domains such as vision, language, memory, and emotion, to assess individual differences in neural connectivity across multiple whole-brain networks. Resting-state fMRI (rs-fMRI) is being used increasingly to examine functional connections in the resting human brain. Rs-fMRI examines temporal correlations between segregated brain regions during unconstrained intrinsic activity, or task-free rest, periods. This reveals coherence within and between multiple whole-brain networks, and can be used to develop a more comprehensive model of human brain connectional architecture (Van Dijk et al., 2009).

Rs-fMRI is dramatically increasing our understanding of neural development, including the sequence of development, and the extent of neural system connectivity in normally and abnormally developing infants, children, and adolescents ((reviewed by Uddin et al., 2010)). Much more work needs to be done, however, to elucidate processes involved in brain network maturation and to relate them to behavioral development. Linking measures of neural and behavioral development is complex, as is using deviations in these measures to examine developmental disease processes. This undertaking is further complicated by non-neural contributions to network development, including physiological influences, environmental factors, and genes. Nevertheless, using rs-fMRI, and with advances in imaging science, researchers have begun to obtain a more comprehensive picture of neural network development.

In a short period of time, rs-fMRI has advanced our understanding of organizational principles of central nervous system development. For example, Kelly et al. and Fair et al. have documented less diffuse local connectivity and increased long-range connectivity with maturation (Fair et al., 2009; Kelly et al., 2009). The ordering of network maturation appears to parallel the ordering of behavioral maturation[3]. As is true of myelination, sensorimotor development precedes the development of systems underlying higher cognition (Kelly et al., 2009). One noteworthy movement in the field is to utilize advanced information-processing techniques to

identify brain regions that drive change in one or more neural networks. Gao et al., for example, used a graph theoretical measure of node importance to determine that the posterior cingulate/retrosplenial cortex plays a central role in the development of the default mode network in infants (Gao et al., 2009).

If rs-fMRI is to be used to advance our theories of neural network development, we must establish that the neural network signals that form the basis of rs-fMRI are stable across repeat measures. If these signals can be measured robustly despite the variance introduced by other sources (e.g., participant state, systemic physiological process such as breathing and cardiovascular function and scanner variance), and if these signals are reproducible, we are likely measuring reliable indicators of the status of neural networks. Resting-state fMRI measures have been shown to be reliable in adults (Meindl et al., 2009; Shehzad et al., 2009; van de Ven et al., 2004; Zuo et al., 2010a; Zuo et al., 2010b), but not yet in children or adolescents. Previous research has demonstrated that intra-individual variability is greater in children than in adults for measures of both behavioral (Williams et al., 2005) and blood oxygenation level dependent (BOLD) fMRI signal change (Thomason et al., 2005), rendering a systematic study of resting state reliability essential to further investigations in this area.

Our first goal was to identify reliable, testable intrinsic connectivity networks (ICNs) in a large sample of youth from whom peak network locations could be derived for future developmental rs-fMRI studies. The second goal was to examine the reliability of rs-fMRI in children both within and across sessions. We chose to test reliability in the most widely examined cognitive [i.e., default mode (DMN), executive (EN), and salience (SN)], and sensorimotor [i.e., auditory (AN), motor (MN), and visual (VN)] networks. We scanned 65 children and adolescents, about one-third of whom contributed multiple rest scans on either one or two different scan dates. Based on our prior experience with rs-fMRI data in children and based on adult studies of ICN reliability, we hypothesized that despite greater overall variance, children would demonstrate significant test-retest reliability in rs-fMRI ICNs. We compared

consistency in spatial topography, temporal connectivity between networks, and frequency content of the network time courses, and examined reliability estimates obtained for these measures.

Methods

Participants

Sixty-five children and adolescents aged 9-15 (mean=12.5 years, SD= 2.0) were scanned at least once, and 21 were scanned either two or three times. For the primary analyses of the present study, scans were partitioned into either within-session or between-sessions comparisons. For the between-sessions comparison, 15 children were scanned on the same rs-fMRI protocol using the same hardware (e.g., coil, scanner, etc) separated by 2-3 years. Two of these children were removed from analysis due to excess movement (> 1mm) and a technical problem identified during data reconstruction, leaving 13 children in this group. For the within-session comparison, 15 children were scanned with 2 consecutive resting-state scans within one scan session. Table 1 presents the number of participants in each comparison group. The within-session and between session comparison groups were matched for age, $t(26) = 1.2$, $p = .24$) and gender (χ^2 (1) = .19, p = .66).

Participants were recruited through their mothers via the Craigslist website and other online advertisements and parent networks, and each mother-child pair was compensated $25/hour. All participants had no reported history of brain injury, no behavioral indications of possible mental impairment, no past or present DSM-IV Axis I disorder, were right-handed, fluent in English, and had no reported learning disorders. Parents and children gave informed consent and assent, respectively, as approved by the Stanford Institutional Review Board.

MRI acquisition

Magnetic resonance imaging was performed on a 3.0 T GE whole-body scanner. Participants were positioned in a purpose-built single channel T/R head coil and stabilized by clamps and a bite bar formed with dental impression wax (made of Impression Compound Type

I, Kerr Corporation, Romulus, MI) to reduce motion-related artifacts during scanning. During the resting-state experiment, participants completed a six-minute scan during which they were instructed to lay still with their eyes closed. All resting-state scans were conducted following the anatomical localizer, a field inhomogeneity shim, and a 4-minute perfusion scan. For participants in the within-session comparison group, the two rest scans were performed back-to-back in the same scanning session.

For this study, 29 axial slices were acquired with 4mm slice thickness (no skip). High-resolution T2-weighted fast spin-echo structural images (TR = 3000ms, TE = 68ms, ETL=12) were acquired for anatomical reference. A T2*-sensitive gradient echo spiral in/out pulse sequence was used for all rs-fMRI imaging (TR = 2000ms, TE = 30ms, flip angle = 77°, FOV = 22 cm, 64 x 64). An automated high-order shimming procedure, based on spiral acquisitions, was used to reduce B0 heterogeneity (Kim et al., 2002). Spiral in/out methods have been shown to increase signal-to-noise ratio and BOLD contrast-to-noise ratio, and have also been shown to reduce signal loss in regions compromised by susceptibility-induced field gradients generated near air-tissue interfaces, such as PFC (Glover and Law, 2001). Compared to traditional spiral imaging techniques, spiral in/out methods result in less signal dropout and greater task-related activation in PFC regions (Preston et al., 2004). A high-resolution volume scan (140 slices, 1mm slice thickness) was collected for every participant using a spoiled gradient-recalled (SPGR) sequence for T1 contrast (TR = 3000ms, TE = 68ms, TI = 500ms, flip angle = 11°, FOV = 25 cm, 256 x 256). During the resting-state scan, participants' heart-rate (HR) and respiration waveform were recorded.

Physiological correction in reconstruction

Previous research has demonstrated that physiological noise can confound detection of neural activation during rs-fMRI. As a result, methods to model and correct for physiological effects of noise have been proposed by our group and by others (Birn et al., 2006; Chang et al., 2009). For the present study, rs-fMRI images were preprocessed using a correction that

diminishes BOLD signal fluctuations contributed by respiratory and HR variations. Using the method developed by Chang et. al. 2009, this correction reduces the effect of low-frequency respiratory variations (i.e., the "envelope" of the respiratory belt waveform) and heart rate (average rate in a 6-sec sliding window) by first convolving those independently measured signals with appropriate filters and then regressing them out of the time series for each voxel [15].

Image processing

Our first goal was to identify rs-fMRI network peaks using a data-driven independent components analysis (ICA) approach in a large youth sample. The second goal was to use consequent peaks from the ICA analysis to derive seed-based ICNs to study test-retest reliability. Thus, participant data were processed along two separate paths. The first used Statistical Parametric Mapping software (SPM8; www.fil.ion.ucl.ac.uk/spm/software/spm8/) to perform realignment, normalization, and smoothing (6 mm^3) for N = 65 rest scans from 65 different children (summarized in Table 1). These SPM8-preprocessed images were used for the group ICA analysis, described below. The second processing path was used to perform ROI-based connectivity analyses for only those children scanned more than once as part of either the within- or between-session groups. Beginning again from the raw-physiology-corrected image data, this second path of processing was implemented in AFNI (http://afni.nimh.nih.gov/afni) (Cox, 1996). Preprocessing of these data included slice-timing correction, volume registration, smoothing (4mm^3), bandpass filtering ($0.008 < f < 0.15$ Hz), and co-registration of functional and anatomical images. It is worth noting that the work could have been carried out within either software package (SPM or AFNI) and would generate the same results. The choice to use both SPM and AFNI for the different paths was primarily based on the convenience of the output file types. For example, SPM produces ANALYZE format images useful for subsequent group ICA, while AFNI produces time-course files useful for temporal analyses.

Identification of rs-fMRI networks

We used a three-step method to classify rs-fMRI networks. To address the first goal of the present study, Steps 1 and 2 were performed to generate network spatial maps in the sample of 65 children and adolescents**.** Each child contributed only one rest scan to the ICA analysis; in cases where more than one scan was available, the first was used. These maps were used both to visually represent the 6 *a priori* rs-fMRI networks of interest and to extract peak spatial coordinates for the subsequent ROI-based connectivity analysis in Step 3. Details of each step follow.

Step 1

Movement was plotted and visually inspected for every participant. Time frames corresponding to brief movement spikes (> 0.8mm) and lasting less than 5 frames were removed. This correction for excursions resulted in low average movement across participants (< 0.5 mm). In addition, the first three time frames were removed for all participants to allow for signal stabilization. In total, no more than 10% of time frames were removed. Remaining time frames for all 65 participants (32 females; mean age = 12.5) were concatenated in a group independent component analysis (ICA) implemented in Matlab (http://icatb.sourceforge.net) using GIFT (Calhoun et al., 2001). Infomax was used to estimate 34 components, after which binary spatial templates were used to automatically identify components corresponding to the 6 networks of interest. A spatial template-matching technique was used (as described in (Greicius et al., 2007)). The templates used in the present study were used in previously published work (Seeley et al., 2007); these were merely resampled (i.e., changed to the 3D spatial resolution of the target ICC and Kendall's W maps) for the present study. Finally, after an automated template-matching algorithm was used to determine those individual components corresponding to the default-mode, left and right executive, salience, motor, auditory, and visual resting-state networks, single-participant spatial maps were back reconstructed.

Step 2

To visualize the components and extract component peaks for subsequent ROI-based connectivity analyses, SPM8 was used to create statistical parametric maps for each network. In order to generate single subject spatial maps for each of the 65 participants that were included in the group ICA, we used GIFT to implement back reconstruction of single subject spatial maps from the raw data. Back-reconstructed single subject spatial maps that corresponded to each network of interest (default, executive, salience, motor, visual, auditory) were entered into network-specific one-sample t-tests, a method that has been used in a similar manner in previous studies (Stevens et al., 2009). Thus, t-tests were based on significance testing for the spatial overlap of 65 single-subject spatial maps, with the exception of the executive network, in which single-subject spatial maps corresponding to the right (N = 65) and left (N = 65) executive network ICA components were combined for a total of 110 spatial maps for that analysis. These networks were used to derive maximally significant peaks for seed-based analysis, and were otherwise excluded from further consideration in the present analysis. Peak locations from the resulting random effects maps are summarized in Table 2.

Step 3

Using a method that is now well established for analyzing rs-fMRI data (Fox et al., 2005), in step 3, we generated correlation maps in the subset of participants who were scanned multiple times. This method involves extracting timecourse data from a seed region and computing the correlation coefficient between that time course and the timecourses of all other brain voxels. Following previous methodology (Fox et al., 2005), we extracted our average time-course data from the 2-3 most significant foci (from network maps resulting from step 2) within each network and averaged these. The trace from each participant's seed regions (3D spheres with a radius of 3 mm, centered on the coordinates summarized in Table 2, and averaged across peaks summarizing each network) was de-trended for 3 translational and 3 rotational motion regressors (AFNI 3dDetrend). This trace was used to calculate the correlation between the seed region and the time-course data in all of the other voxels in the brain. Correlation estimates were controlled

for estimated translational and rotational motion as well as a white matter nuisance time-course (sphere with radius of 3mm, centered at the Talairach coordinate (-27, 9, 24)). In the present analysis we did not apply global signal correction because of the concern that it forces the presence of anticorrelated networks (Chang and Glover, 2009; Murphy et al., 2009; Weissenbacher et al., 2009), which could lead to false interpretation of the observed effects. After correlation coefficients were calculated for each voxel in the brain, correlation coefficients were converted to a normal distribution by Fischer's z transformation. We submitted the resulting z-converted correlation maps to concordance and mutual information analyses. We used time-course data for each network for the subsequent temporal and frequency analyses.

Voxelwise concordance measures

Intra-class correlation coefficient (ICC (Ruckert et al.)) and Kendall's W were applied to rs-fMRI maps. ICC and Kendall's W statistics are frequently used to measure test-retest reliability in fMRI data (Meltzer et al., 2009; Shehzad et al., 2009). Such concordance measures address the likelihood that regions of high group activation in a first scan session would be preserved within the participant in a second session, but discriminate between different participants. Large (approaching 1) or small (approaching 0) values of the Kendall's W statistic and the ICC indicate stability of inter-participant variability (i.e., participants' scans are highly stable and unique), or instability of the inter-participant variability (i.e., within-participant scans are highly variable and there is little between-participant differentiability), respectively. Intermediate values of these statistics indicate a greater or lesser degree of between-participant differentiability.

Intra-class correlation

ICC has been used to assess measurement reproducibility in fMRI (Caceres et al., 2009). It is defined as the ratio of the between-subjects variance to the total observed variance (Shrout and Fleiss, 1979),(McGraw and Wong, 1996). As Bland and Altman (1996) explain, ICC may be

understood as a measure of discrimination between subjects(Bland and Altman, 1996). We calculated reliability maps for the third ICC defined by Shrout and Fleiss (1979):

$$ICC(3,1) = \frac{BMS - WMS}{BMS + (k-1)WMS} \quad [1]$$

where *BMS* is the between-subjects mean squared variance, *WMS* is the within-subject mean squared variance, and *k* is the number of scan repetitions for each participant. Eq. (1) estimates the correlation of the subject signal intensities (z-scores) between sessions, modeled by a two-way analysis of variance (ANOVA), with random subject effects and fixed session effects. The ANOVA analysis splits the total mean squared intensity values into *BMS* and *WMS* components (Koerten et al.). In our analysis, we used k=2, which is the number of scans being considered for comparison. In this context, *WMS* is due to different scans. The measure, ICC(3,1) (Eq. (1)) was computed for each voxel.

Kendall's W

Kendall's W is a statistic based on ranks rather than values and is a robust, nonparametric descriptive statistic, with allowable values ranging from 0 (no agreement) to 1 (complete agreement). If $r_{i,j}$ indicates the rank of the *j*th subject in the *i*th scan, then W is given by the equations:

$$R_j = \sum_{i=1}^{n} r_{i,j}, \bar{R} = \frac{1}{2}m(n+1),$$
$$S = \sum_{j=1}^{m} (R_j - \bar{R})^2, W = \frac{12S}{m^2(n^3 - n)} \quad [2]$$

where m is number of subjects and n is number of scans, which in our case is 2, and ½[m(n+1)] is used as it represents the mean value of all the ranks. The measure W was also computed for each voxel.

Network masks for concordance measures

Results of the one-sample t-tests for each network were used to generate network masks. A masking procedure was used only to summarize mean concordance statistics. This allowed us to report results from a whole brain voxel-wise approach as well as a mean computed across voxels that fell within the network (similar to the approach taken by Shehzad et al., 2009). To generate masks, network results were thresholded at a level that kept the Family-Wise Error (FWE) corrected at $p<.05$ (i.e., minimum $Z = 7.75$) with a cluster extent of 30 voxels. Using SPM8 smoothing and imcalc functions, the resulting images were smoothed at 6mm and binarized whereby all voxels with values > 1 were assigned to a value of 1, and all other voxels were assigned a value of zero. Table 3 presents Kendall's W means for the whole brain and within-network masks.

Temporal analysis of correlations across networks

BOLD signal time-courses from the seed regions used for seed-based FC analysis (described in step 3, above) were extracted from each participant for temporal analysis of correlations between the 6 ICNs. Using custom Matlab routines, Pearson correlations were computed between all pairs of network time-courses within each scan. In addition, for each scan, a single measure of global network coherence was defined as the Fisher z-transformed average value of all such pair-wise correlations. The global network coherence reflects the overall relatedness of networks within a subject (see Stevens et al., 2009 for an expanded discussion of what this parameter might mean in development). Within- and between-session consistency of global network coherence was examined using the correlation between T2 and T1 measurements. Separate Pearson's correlations were computed to examine stability across scans for the following three comparisons: (1) T1/T2 within-session (N = 15); (2) T1/T2 between-session (N = 13); and (3) all T1 data correlated with all T2 data for within-and between-session comparisons (N = 28). The goal of this analysis was to examine whether the temporal relations between networks are stable across repeated measurements. Interpretations of the correlations between specific networks are beyond the scope of the present study, but were analyzed in

detail by Stevens and colleagues in a recent study of rs-fMRI networks in children (Stevens et al., 2009).

Temporal frequency analysis

An analysis of low-frequency power was conducted on the ICN time-courses for each participant and scan. For each time-course, the low-frequency power was computed as the percentage of the total spectral power lying in the range $0.008 < f < 0.08$ Hz. Independent samples t-tests were then used to test possible differences in low-frequency power across scans for each network, and Pearson correlation was used to examine the relationship between T1/T2 frequency measures for each network.

Results

Group ICA

Figure 1 shows well-established networks that we identified by group ICA analysis of 65 children/adolescents aged 9-15 years. Peaks for each network are summarized in Table 2 along with the nuisance white matter region of interest (ROI) used for seed-based analyses that followed the ICA analyses.

Seed-based connectivity

In order to show the level of individual variation in each network, we show data from a random selection of six individual participants in Figure 2: three from the within-session and three from the between-session group. Visual comparison of side-by-side network maps suggests that there was greater variance across participants than across scans. To quantify this observation, we used Kendall's W and ICC to apply discriminability and reproducibility statistics to spatial maps.

Rest-retest reliability in children

Intersession reproducibility was assessed using statistics that compared variance between scans versus variance across participants. All test-retest analyses were conducted in a sample of 15 children who performed two rs-fMRI scans within one scan session (the within-

session group), and 13 children who were scanned 2-3 years apart on the same scan system and imaging protocol (the between session group). Figure 3 shows measured Kendall's W and ICC values for both groups projected in the axial plane in a whole brain, voxel-wise analysis.

To provide greater detail on the distribution of ICC values across all brain voxels, ICC values are plotted in histograms in Figure 4 for both groups and all networks. We obtained positive ICC values for the majority of brain voxels in all comparisons, an indication of stability within participants across measurements.

Mean values for Kendall's W are listed in Table 3 for both within-network masked ICNs, and for all brain voxels. Within-session Kendall's W mean concordance measures range from .71 to .78 across a whole brain ROI, and from .71 to .86 in within-network ROIs. Between-sessions Kendall's W mean concordance measures range from .60 to .65 across a whole brain ROI, and from .60 to .66 in within network ROIs. One sample t-tests conducted for all networks on Kendall's W and ICC values across the brain were significant at $p < .0001$. Thus, these observed values within and between sessions reflect greater consistency within participants than between participants.

Correlations between network time-courses

The temporal dynamics among networks also changes with age (Stevens et al., 2009). In the present study we tested whether the global network coherence (defined above as the mean pairwise correlation between all 6 ICN time courses) of an individual would be consistent across time. We obtained a significant correlation across all Time one - Time two (T1/T2) z-converted global network coherence measures, $r(28)=.66$, $p<.001$, collapsing within- and between-session measurements, indicating consistency in time-course data over repeat measurements. Correlation statistics were also significant within each comparison group. That is, for the within-session T1/T2 comparisons between network z-converted correlation measures, $r(15)=.74$, $p=.002$, and for the between-session comparisons, $r(13)=.64$, $p=.02$. These results are plotted in panel D of Figure 5.

Figure 5 also presents the mean correlations separately for each pair of networks and for each of the T1 and T2 scans that comprise the within- and between-session comparisons. The correlation matrix in Figure 5 is organized into superordinate squares that show the results of each pair-wise network comparison. Consistency within superordinate squares indicates that the correspondence between pairs of networks in the temporal domain is consistent across repeated measures.

Frequency results

For all six networks, the proportion of data in the low-frequency power (range $0.008 < f < 0.08$ Hz) across participants did not change significantly across time for within-session measurements (all $t(15)$ values < 1.6, all $ps > .1$) or for between-session measurements (all $t(13)$ values < 1.4, all $ps > .1$; see Fig. 6). However, significant stability across T1/T2 measurements within a network was observed only within the visual network for the T1/T2 within-session comparison, $r(15)=.54$, $p=.037$.

Discussion

Our study examined test-retest reliability for a relatively new method, rs-fMRI, which may prove invaluable for assessing large-scale, functional neural networks in children and adolescents. Considerable anatomical evidence shows that human brain maturation is gradual and continuous, characterized by steadily increasing white matter, a general reduction in grey matter, extensive synaptic pruning, and elaboration through dendritic arborization (Changeux and Danchin, 1976; Giedd et al., 1999; Huttenlocher, 1990; Paus et al., 1999). Less is known, however, about human *functional* brain development.

By measuring function and connectivity in multiple large-scale brain networks, concurrently and without requiring task-compliance, rs-fMRI studies differ from task-based fMRI studies that have been the foundation of systems neuroscience imaging research over the past two decades. Rs-fMRI studies sample activity in spatially segregated brain regions. This activity (measured by BOLD signal) apparently occurs spontaneously, but with a coordinated temporal

pattern. Rs-fMRI may be useful for generating data relevant to the development of functional neural systems, and may increase our knowledge of developmental neuropsychiatric disorders. Already, rs-fMRI has been applied to infants to measure early neural network function (Fransson et al., 2007; Lin et al., 2008), to children with developmental disorders ((reviewed by Uddin et al., 2010)) and has recently been proposed as a measure for predicting brain developmental age (Dosenbach et al., 2010). Indeed, analysis of resting-state brain activity has already been useful in other populations in which task compliance is not possible, including individuals with disorders of consciousness (Vanhaudenhuyse et al., 2010) and chimpanzees (Rilling et al., 2007). This tool is made even more powerful as it can be linked to complementary data about the anatomy of the brain assessed in the same scan session (Honey et al., 2009).

The early and extraordinary success of rs-fMRI as a useful method for measuring systems-level brain organization has occurred sufficiently quickly that some of the assumptions currently being made have yet to be tested. We have experience examining hemodynamic responses pertaining to BOLD confounds (e.g., physiology, blood flow, breathing-rate) and have contributed to the methodological literature in that area (Chang et al., 2009; Chang and Glover, 2009; Thomason et al., 2007). Given our past work showing that BOLD in children is inherently noisier than it is in adults (Thomason et al., 2005), and given that the best practices for acquiring and analyzing resting-state data are still being developed (Van Dijk et al., 2009), it is critical that we determine whether resting-state network measurements are stable in children. The present data set would have been useful for a developmental study, but there have already been important contributions in that area (Dosenbach et al., 2010; Fair et al., 2009; Fair et al., 2007; Kelly et al., 2009; Supekar et al., 2009). The present study contributes to this literature by being the first to examine temporal stability in rs-fMRI measurements in children. Our results indicate that rs-fMRI is likely to primarily reflect features of the underlying biology (i.e., is stable within individual even over 2-3 years), with some lesser contributions from aspects of the acquisition process (i.e., is even more stable when within session).

Using group ICA in a large sample of youth, we identified rs-fMRI networks that have been found in previous research on adults. We reported the peak locations for networks composed of regions important for executive functioning, salience processing, motor, visual and auditory processing, and the default mode. These peaks may be useful for seed-based connectivity analyses in future studies of children.

ICC and Kendall's W values were predominantly positive across the whole brain volume, indicating that participant differentiability outweighed scan variance for most brain areas. The general pattern was one of moderately high concordance across spatial ICNs (Table 3), but there were some small areas of non-concordance (i.e., where these statistics were negative). These could reflect genuine neural developmental changes across the scan interval. For rs-fMRI to be effective, it should reliably measure stable features of the underlying biology, but still be sensitive to true biological differences. Investigators who conduct longitudinal rs-fMRI studies of children may find it useful to assess change using a mutual information approach that would quantify change across the interval and test its correspondence to behavioral measures, time, or developmental age, for example.

Concordance measured for spatial maps was greater for within-session than for between-session comparisons (measured by Kendall's W and ICC statistics). The distribution ICC coefficients across all brain voxels within each network are presented in Fig. 4. It is apparent that concordance was higher within sessions than between sessions in this study. Both distributions are significantly different than zero, indicating that the networks are stable within individuals, but the consistency is greater within sessions. Differences in these distributions could be driven by a number of factors that cannot be distinguished within this experimental design, including scan session specific biological factors (e.g., temperature), MRI technology factors (e.g., machine SNR, field shim), developmental maturation, and psychological factors (e.g., mood).

In this study we extended the investigation of the spatial reliability of ICNs to examine stability in temporal and frequency domains. We obtained a significant correlation across all

T1/T2 z-converted time-course correlation measures (i.e., scatterplot in Figure 5D, N = 28). The correlation measure used here (global network coherence) may be interpreted as the amount of total relatedness between networks, computed by averaging the pairwise correlations between networks. We found that measures of the relatedness of network time-course data are reliable for individual participants across time. This is among the first work to demonstrate empirically that network dynamics are stable reflections of individual differences, indicating that the study of network dynamics is a key area for future investigation.

Resting state low-frequency fluctuations are thought to reflect cyclic modulation of gross cortical excitability and network neuronal synchronization (Balduzzi et al., 2008). Here, we examined stability of the computed low-frequency power (range $0.008 < f < 0.08$ Hz) for each ICN at each measurement time. Time one to time two comparisons within network showed smaller differences than those observed across ICNs; this relation was significant, however, only in the visual network for the T1/T2 within-session comparisons where across participants the frequency from time 1 to time 2 was correlated ($p < .05$). Prior work in adults has shown frequency oscillations in visual cortex are impacted by eyes-open versus eyes-closed scanning (Yang et al., 2007), and also shown that coherent low-frequency fluctuations are particularly strong in visual cortex and posterior midline structures (Zuo et al., 2010a). Having obtained a significant result in the frequency domain in the visual cortex could therefore reflect aspects of development (i.e. early maturation in sensorimotor cortical networks), or could relate to qualities inherent to the visual network that persist across the life-span. It will be useful for future work in large samples of children to measure the regional specificity and developmental timing of BOLD-derived low-frequency fluctuations to refine what is understood about frequency dynamics within large-scale brain networks across development. Consistent with what has been observed in adults (Zuo et al., 2010a), we obtained significant results in both the temporal and frequency domains, further supporting stability in rs-fMRI data.

The present results indicate that, if motion is restrained and physiologically generated noise is appropriately controlled, rs-fMRI data in children are robust, and reflect meaningful characteristics of the underlying neurobiology. This is consistent with adult studies (Shehzad et al., 2009; van de Ven et al., 2004; Zuo et al., 2010a; Zuo et al., 2010b), and is the first indication that ICN maps are relatively stable in children and adolescents. We found that rs-fMRI measurements across spatial, temporal, and frequency domains were reproducible in children. This work provides an important demonstration that rs-fMRI measures are viable for studying developmental progress and of disease. We provide a critical foundation for using the resting state as a marker of large-scale neural network development, and as a basis to compare clinical and healthy population samples.

Acknowledgements

This project was partially supported by awards from the National Institute of Health [F32-MH081583 to MET, F31-AG032168 to CC, P41-RR009874 to GHG, RO1-MH074849 to IHG, U54-RR021813 and P41-RR013642 to AWT], National Science Foundation [0716055 and 0442992 to IDD], and by a NARSAD Young Investigator Award to MET. PT is supported by HD050735, AG016570, EB008432, EB008281, EB007813, and AG036535. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank Sarah Victor and Hannah Kang for their assistance in acquiring the scan data, and Yamanda Wright, Hannah Burley, and Lindsey Sherdell for their assistance in participant recruitment, screening, and conducting structured behavioral interviews. Finally, the authors thank the participants and their families for contributing their time and perspectives to this study.

Figure Legends

Figure 1. Spatial renderings of the components examined for reliability. These were identified by applying group ICA implemented using GIFT (Calhoun et al., 2001) to data from 65 children and adolescents aged 9-15. Components were selected using an automatic template matching technique developed by M. Greicius ((for description of the method, see Greicius et al., 2007)). Peaks from these networks are reported in Table 2. These may be useful for future studies using ROI seed-based connectivity approaches to rs-fMRI data in development. Peaks from these components were used for reliability analysis.

Figure 2. Individual seed-based connectivity maps for 6 participants illustrating each comparison (within-session, and between-session). Connectivity maps are more uniform within participants (rows) than across participants (columns).

Figure 3. To generate concordance maps, voxelwise ICN z-maps were compared for scans within session and between sessions. Resulting ICC and Kendall's W statistics are displayed as color maps on axial slices. Concordance analysis (verified in both measures) indicates that there is higher participant differentiability within session than across sessions. Differentiability across participants is high. Maps are coded using different color scales, to distinguish the two complementary statistical approaches.

Figure 4. Histogram of ICC coefficients. Blue bars indicate distribution of correlations for within-session comparisons across all brain voxels. Red bars indicate the distribution of ICC values for between-session, longitudinal (2-3 years) within subject comparisons. Positive values indicate consistency within participants over time.

Figure 5. Correlation matrix and individual participant scatterplot. In the correlation matrix on the left, temporal relations between networks were computed for each resting state scan. The resulting correlation coefficients are plotted within superordinate squares of the matrix. As an example, superordinate square (A) demonstrates that auditory and default mode network time-courses are more highly correlated than are motor and salience networks (presented in superordinate square B). It is noteworthy that for this data set, the cool and warm colors tend to be consistent within superordinate matrix squares, indicating a degree of consistency in the observed temporal relations between networks. Square (C) provides the legend for the ordering of the comparisons. The scatterplot on the right (D) demonstrates that participants with high correlations between network timecourses remain high across repeat measurement. In contrast, those with little correspondence between timecourses remain low.

Figure 6. Consistency in low frequency power. Low frequency power was computed in terms of a ratio with respect to the total power for each of the resting networks and within each set of scans that went into within and between group comparisons. Column graphs of means are clustered within network to illustrate similarities; within-session comparisons are shaded in blue and between-session comparisons shaded in green. Error bars show the standard deviation of the mean. The asterisk highlights a significant correlation observed between time1-time2 (T1/T2) scans for the visual network within-session comparison, $r(15)=.54$, $p=.037$.

References

Balduzzi, D., Riedner, B.A., Tononi, G., 2008. A BOLD window into brain waves. Proc Natl Acad Sci U S A 105, 15641-15642.

Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536-1548.

Bland, J.M., Altman, D.G., 1996. Measurement error and correlation coefficients. BMJ 313, 41-42.

Caceres, A., Hall, D., Zelaya, F., Williams, S., Mehta, M., 2009. Measuring fMRI reliability with the intra-class correlation coefficient. Neuroimage 45, 758-768.

Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp 14, 140-151.

Chang, C., Cunningham, J.P., Glover, G.H., 2009. Influence of heart rate on the BOLD signal: the cardiac response function. Neuroimage 44, 857-869.

Chang, C., Glover, G.H., 2009. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. Neuroimage 47, 1448-1459.

Changeux, J.P., Danchin, A., 1976. Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. Nature 264, 705-712.

Cox, R., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research 29, 162-173.

Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett, J.R., Jr., Barch, D.M., Petersen, S.E., Schlaggar, B.L., 2010. Prediction of individual brain maturity using fMRI. Science 329, 1358-1361.

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U., Church, J.A., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2009. Functional brain networks develop from a "local to distributed" organization. PLoS Comput Biol 5, e1000381.

Fair, D.A., Dosenbach, N.U.F., Church, J.A., Cohen, A.L., Brahmbhatt, S., Miezin, F.M., Barch, D.M., Raichle, M.E., Petersen, S.E., Schlaggar, B.L., 2007. Development of distinct control networks through segregation and integration. Proceedings of the National Academy of Sciences of the United States of America 104, 13507-13512.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102, 9673-9678.

Fransson, P., Skiold, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., Aden, U., 2007. Resting-state networks in the infant brain. Proc Natl Acad Sci U S A 104, 15531-15536.

Gao, W., Zhu, H., Giovanello, K.S., Smith, J.K., Shen, D., Gilmore, J.H., Lin, W., 2009. Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. Proc Natl Acad Sci U S A 106, 6790-6795.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. Nat Neurosci 2, 861-863.

Glover, G., Law, C., 2001. Spiral-in/out BOLD fMRI for increased SNR and reduced suceptibility artifacts. Magn Reson Med 46, 515-522.

Greicius, M.D., Flores, B.H., Menon, V., Glover, G.H., Solvason, H.B., Kenna, H., Reiss, A.L., Schatzberg, A.F., 2007. Resting-State Functional Connectivity in Major Depression: Abnormally Increased Contributions from Subgenual Cingulate Cortex and Thalamus. Biol Psychiatry.

Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences of the United States of America 106, 2035-2040.
Huttenlocher, P., 1990. Morphometric study of human cerebral cortex development. Neuropsychologia 28, 517-527.
Kelly, A.M., Di Martino, A., Uddin, L.Q., Shehzad, Z., Gee, D.G., Reiss, P.T., Margulies, D.S., Castellanos, F.X., Milham, M.P., 2009. Development of anterior cingulate functional connectivity from late childhood to early adulthood. Cereb Cortex 19, 640-657.
Kim, D., Adalsteinsson, E., Glover, G., Spielman, D., 2002. Regularized higher-order in vivo shimming. Man Reson Med 48, 715-722.
Koerten, H.K., de Bruijn, J.D., Daems, W.T., 1990. The formation of asbestos bodies by mouse peritoneal macrophages. An in vitro study. Am J Pathol 137, 121-134.
Lin, W., Zhu, Q., Gao, W., Chen, Y., Toh, C.H., Styner, M., Gerig, G., Smith, J.K., Biswal, B., Gilmore, J.H., 2008. Functional connectivity MR imaging reveals cortical functional connectivity in the developing brain. AJNR Am J Neuroradiol 29, 1883-1889.
McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychological Methods 1, 30-46.
Meindl, T., Teipel, S., Elmouden, R., Mueller, S., Koch, W., Dietrich, O., Coates, U., Reiser, M., Glaser, C., 2009. Test-retest reproducibility of the default-mode network in healthy individuals. Hum Brain Mapp.
Meltzer, J., Postman-Caucheteux, W., McArdle, J., Braun, A., 2009. Strategies for longitudinal neuroimaging studies of overt language production. Neuroimage 47, 745-755
Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., Bandettini, P.A., 2009. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 44, 893-905.
Paus, T., Zijdenbos, A., Worsley, K., Collins, D., Blumenthal, J., Giedd, J., Rapoport, J., Evans, A., 1999. Structural maturation of neural pathways in children and adolescents: in vivo study. Science 283, 1908-1911.
Preston, A.R., Thomason, M.E., Ochsner, K.N., Cooper, J.C., Glover, G.H., 2004. Comparison of spiral-in/out and spiral-out BOLD fMRI at 1.5 and 3 T. Neuroimage 21, 291-301.
Rilling, J.K., Barks, S.K., Parr, L.A., Preuss, T.M., Faber, T.L., Pagnoni, G., Bremner, J.D., Votaw, J.R., 2007. A comparison of resting-state brain activity in humans and chimpanzees. Proc Natl Acad Sci U S A 104, 17146-17151.
Ruckert, R., Inderbitzi, R., Picco, C., Schwarz, H., 1989. [Inflammatory aneurysm of the abdominal aorta and ureteral obstruction]. Helv Chir Acta 56, 629-632.
Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27, 2349-2356.
Shehzad, Z., Kelly, A.M.C., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X., Milham, M.P., 2009. The Resting Brain: Unconstrained yet Reliable. Cerebral Cortex 19, 2209-2229.
Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86, 420-428.
Stevens, M.C., Pearlson, G.D., Calhoun, V.D., 2009. Changes in the interaction of resting-state neural networks from adolescence to adulthood. Hum Brain Mapp 30, 2356-2366.
Supekar, K., Musen, M., Menon, V., 2009. Development of Large-Scale Functional Brain Networks in Children. Plos Biology 7, 15.
Thomason, M.E., Burrows, B.E., Gabrieli, J.D., Glover, G.H., 2005. Breath holding reveals differences in fMRI BOLD signal in children and adults. Neuroimage 25, 824-837.
Thomason, M.E., Foland, L.C., Glover, G.H., 2007. Calibration of BOLD fMRI using breath holding reduces group variance during a cognitive task. Hum Brain Mapp 28, 59-68.

Uddin, L., Supekar, K., Menon, V., 2010. Typical and atypical development of functional human brain networks: insights from resting-state fMRI. Front Syst Neurosci.
van de Ven, V.G., Formisano, E., Prvulovic, D., Roeder, C.H., Linden, D.E., 2004. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum Brain Mapp 22, 165-178.
Van Dijk, K.R.A., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2009. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J Neurophysiol 103, 297-321.
Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L.J., Bruno, M.A., Boveroux, P., Schnakers, C., Soddu, A., Perlbarg, V., Ledoux, D., Brichant, J.F., Moonen, G., Maquet, P., Greicius, M.D., Laureys, S., Boly, M., 2010. Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. Brain 133, 161-171.
Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., Windischberger, C., 2009. Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. Neuroimage 47, 1408-1416.
Williams, B.R., Hultsch, D.F., Strauss, E.H., Hunter, M.A., Tannock, R., 2005. Inconsistency in reaction time across the life span. Neuropsychology 19, 88-96.
Yang, H., Long, X.Y., Yang, Y., Yan, H., Zhu, C.Z., Zhou, X.P., Zang, Y.F., Gong, Q.Y., 2007. Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional MRI. Neuroimage 36, 144-152.
Zuo, X.N., Di Martino, A., Kelly, C., Shehzad, Z.E., Gee, D.G., Klein, D.F., Castellanos, F.X., Biswal, B.B., Milham, M.P., 2010a. The oscillating brain: complex and reliable. Neuroimage 49, 1432-1445.
Zuo, X.N., Kelly, C., Adelstein, J.S., Klein, D.F., Castellanos, F.X., Milham, M.P., 2010b. Reliable intrinsic connectivity networks: test-retest evaluation using ICA and dual regression approach. Neuroimage 49, 2163-2177.

4. Table

Click here to download high resolution image

	ICA-derived network maps	Repeat measure comparisons	
		between-session	within-session
N participants	65	13	15
age, T1/T2	12.5(2.0)	11.2(1.5)/13.7(1.4)	13.1
gender, M/F	33/32	8:5	8:7
time between sessions	n/a	2.5(0.4) years	consecutive

Table 1. Group demographic data summary. Repeat measure comparison groups were matched for age, $t(26)=1.2$, $p=.24$, and gender, $\chi^2 = .19$, $df = 1$, $p=.66$. Means are reported with (st.dev.). Abbreviations: T1/T2 = time1/time2 scans; M/F = male/female; n/a = not applicable.

4. Table

Click here to download high resolution image

Table 2. Network peaks derived from ICA perfomed on 65 children ages 9-15

NETWORK	PEAK			
MOTOR	L-BA4	-57	-13	36
MOTOR	R-BA4	62	-13	36
AUDITORY	R-BA41	52	-18	10
AUDITORY	L-BA41	-47	-28	14
VISUAL	R-BA17	8	-82	6
VISUAL	L-BA17	-13	-85	6
DEFAULT	R-BA29	5	-51	12
DEFAULT	L-BA10	-2	55	-8
DEFAULT	R-BA39	49	-61	23
EXECUTIVE	R-BA46	49	32	15
EXECUTIVE	R-BA7	36	-71	46
EXECUTIVE	R-BA8	1	30	41
EXECUTIVE	L-BA40	-50	-49	52
SALIENCE	R-BA13	39	18	1
SALIENCE	L-ACC	-2	17	34
SALIENCE	L-BA13	-43	12	1
WHITE MATTER		-27	9	24

Coordinates given in Talairach and Tournoux convention. For seed-based connectivity analysis, 3mm spheres were centered on these coordinates and timecourse was averaged across the peaks listed above for each network.

4. Table

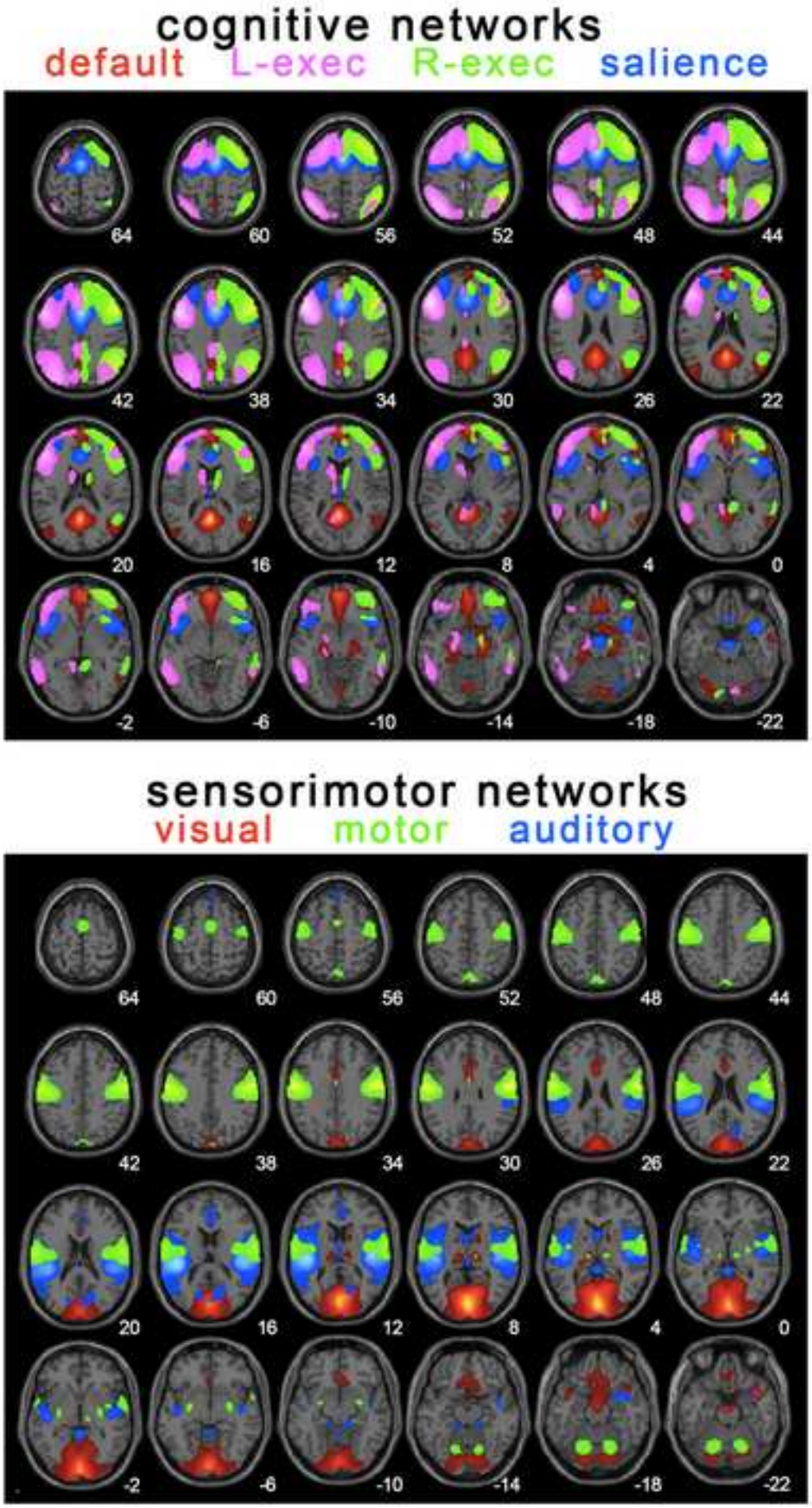
Table 3. Mean concordance measures

Kendall's W concordance		Within-session		Between-session	
		whole brain	within network	whole brain	within network
cognitive	default	.77(.12)	.78(.10)	.62(.14)	.63(.13)
	executive	.77(.12)	.81(.09)	.62(.14)	.65(.11)
	salience	.78(.12)	.86(.07)	.65(.14)	.62(.10)
sensorimotor	visual	.71(.13)	.72(.11)	.61(.14)	.60(.13)
	motor	.72(.12)	.71(10)	.60(.14)	.63(.12)
	auditory	.76(.12)	.85(.07)	.63(.14)	.66(.13)

Voxelwise mean and standard deviation are given for each comparison and each network.Whole brain results provide mean across all brain voxels, while the within network means are averages taken across a network-masked ROI.

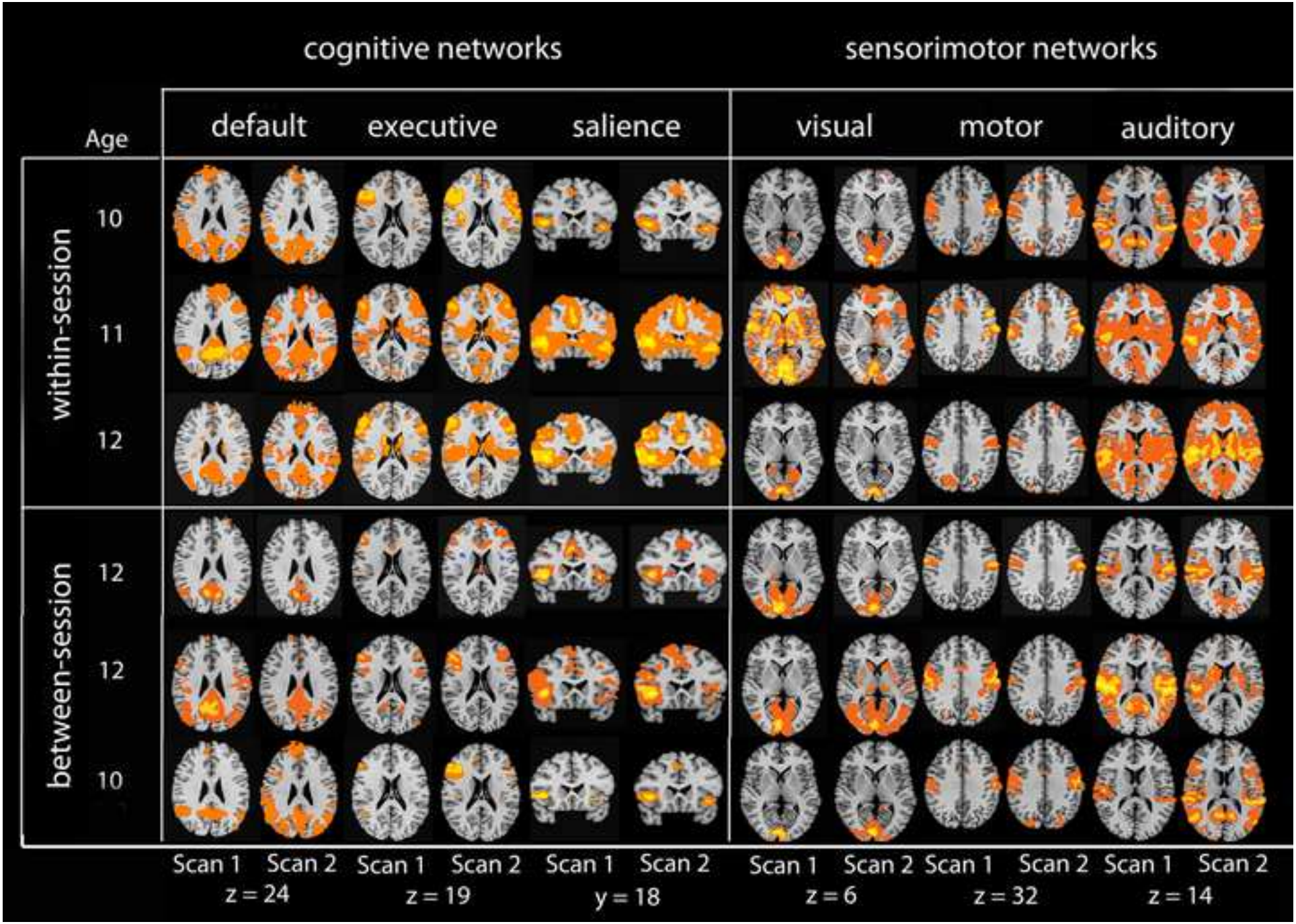
5. Figure

Click here to download high resolution image



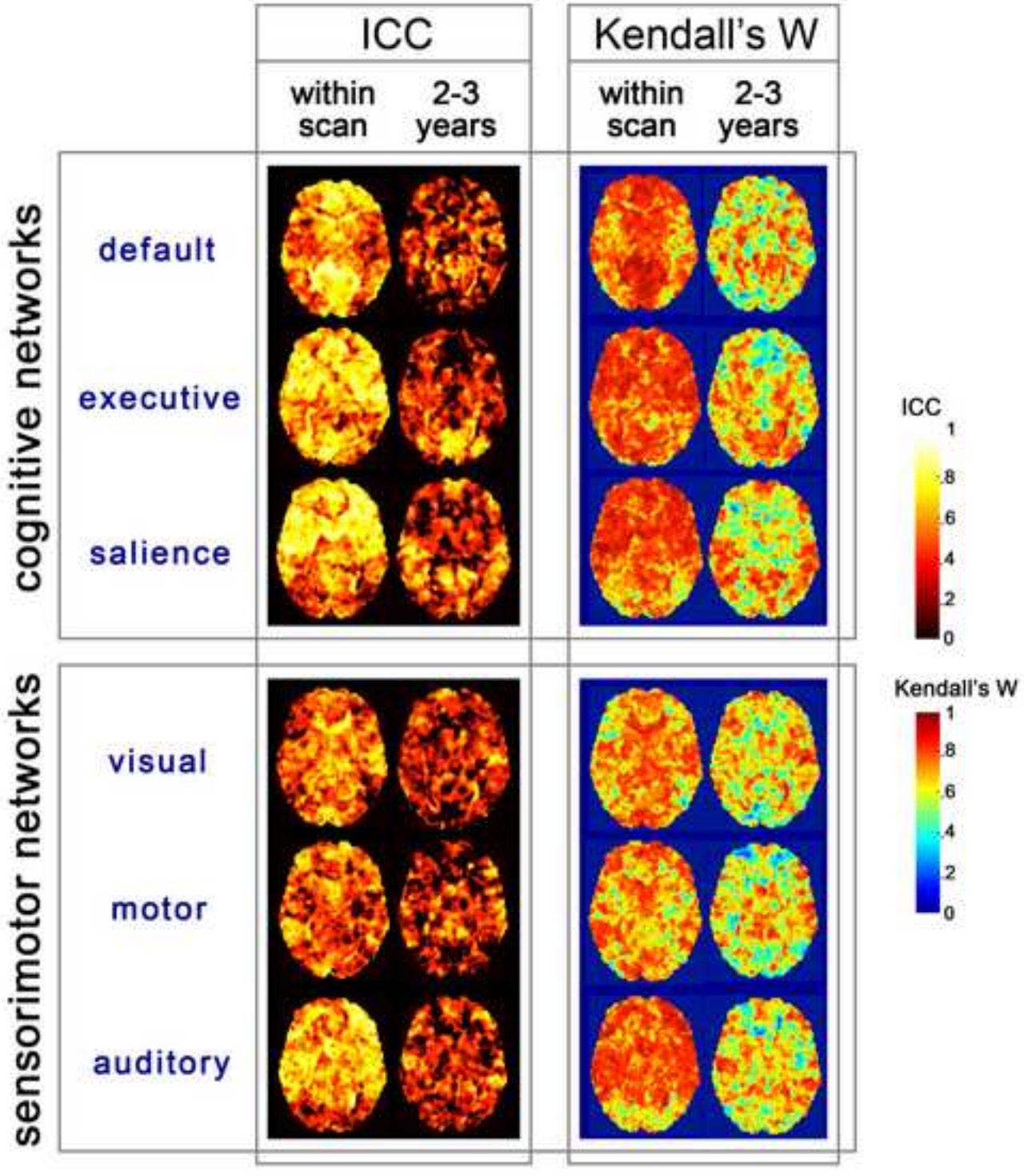
5. Figure

Click here to download high resolution image



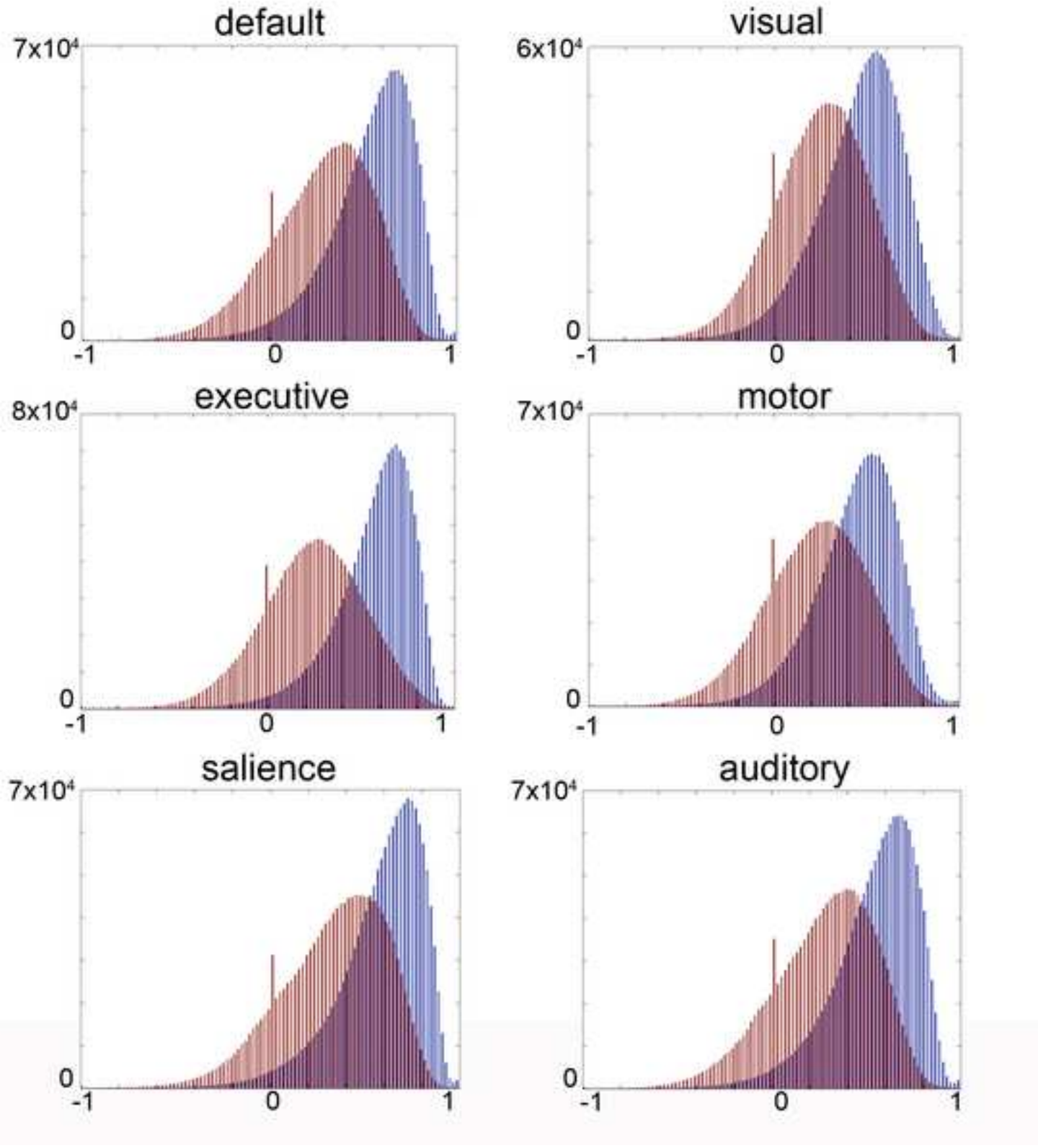
5. Figure

Click here to download high resolution image



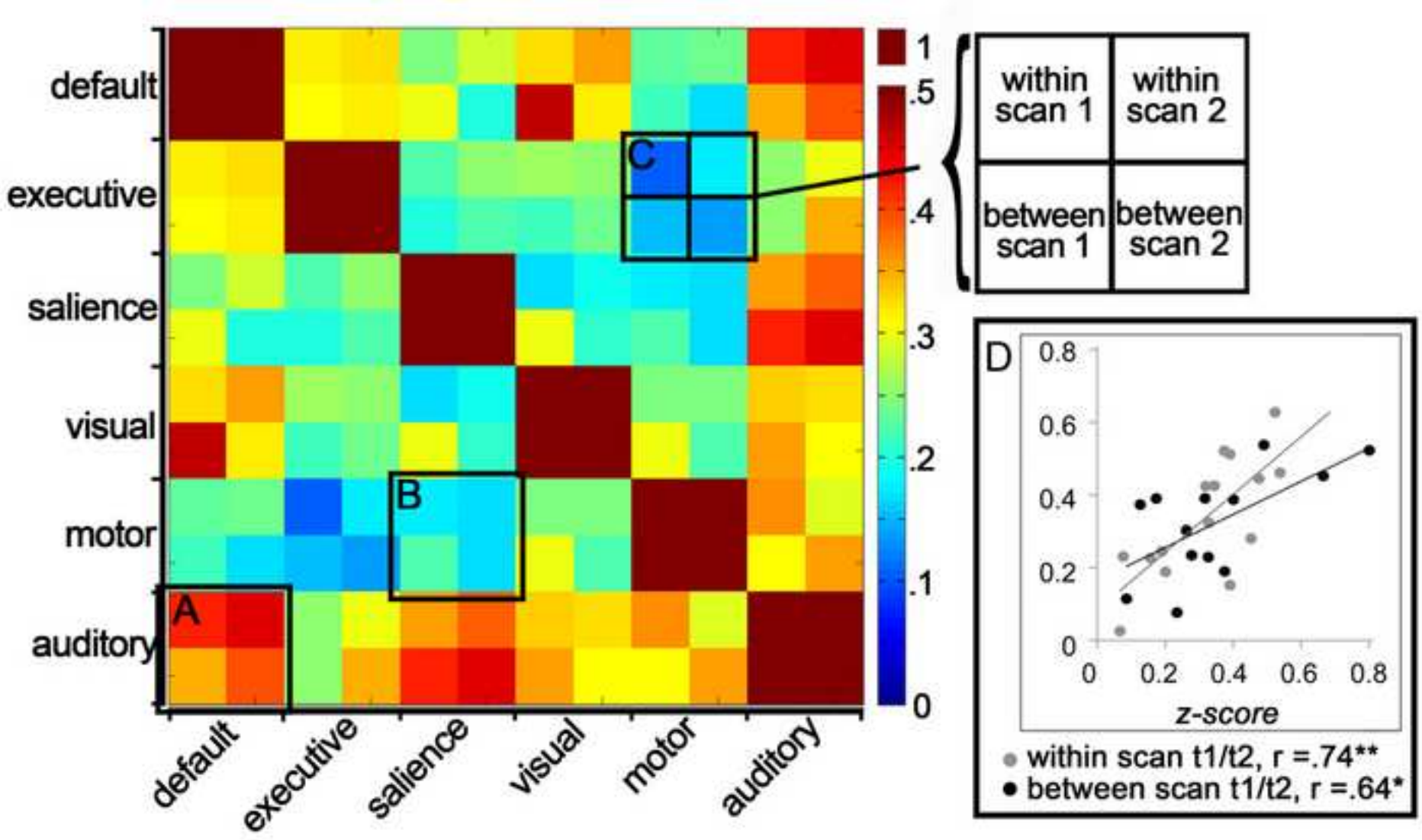
5. Figure

Click here to download high resolution image



5. Figure

Click here to download high resolution image



5. Figure

Click here to download high resolution image