NeuroImage 273 (2023) 120083

Contents lists available at ScienceDirect

NeuroImage

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

Naturalistic viewing increases individual identifiability based on connectivity within functional brain networks

Jean-Philippe Kröll[a,b,*], Patrick Friedrich[a,b], Xuan Li[a,b], Kaustubh R. Patil[a,b], Lisa Mochalski[a,b], Laura Waite[a], Xing Qian[c], Michael WL Chee[c,e], Juan Helen Zhou[c,d,e], Simon Eickhoff[a,b], Susanne Weis[a,b]

[a] Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Centre Jülich, Jülich 52428, Germany
[b] Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany
[c] Yong Loo Lin School of Medicine, Centre for Sleep and Cognition and Centre for Translational MR Research, National University of Singapore, Singapore
[d] Department of Electrical and Computer Engineering, National University of Singapore, Singapore
[e] Integrative Sciences and Engineering Programme (ISEP), National University of Singapore, Singapore, Singapore

A B S T R A C T

Naturalistic viewing (NV) is currently considered a promising paradigm for studying individual differences in functional brain organization. While whole brain functional connectivity (FC) under NV has been relatively well characterized, so far little work has been done on a network level.

Here, we extend current knowledge by characterizing the influence of NV on FC in fourteen meta-analytically derived brain networks considering three different movie stimuli in comparison to resting-state (RS). We show that NV increases identifiability of individuals over RS based on functional connectivity in certain, but not all networks. Furthermore, movie stimuli including a narrative appear more distinct from RS. In addition, we assess individual variability in network FC by comparing within- and between-subject similarity during NV and RS. We show that NV can evoke individually distinct NFC patterns by increasing inter-subject variability while retaining within-subject similarity. Crucially, our results highlight that this effect is not observable across all networks, but rather dependent on the network-stimulus combination. Our results confirm that NV can improve the detection of individual differences over RS and underline the importance of selecting the appropriate combination of movie and cognitive network for the research question at hand.

1. Introduction

Understanding functional brain organization is a major goal of human neuroscience. Typically, researchers have focused on commonalities between individuals and used group-averages to reveal the shared neural underpinnings of certain brain functions. In recent years, the interest in individual functional brain architecture has grown. At the same time, neuroimaging has shifted from mapping brain functions towards investigating interactions within distributed brain networks by considering functional brain connectivity. Specifically, functional connectivity studies yielded insight into the foundation of individual brain organization (Biswal et al., 1995; Greicius et al., 2003; Fox et al., 2006; Damoiseaux et al., 2006). However, it is yet unclear which paradigms are best suited to study individual differences.

Most research on FC has been done on connectivity patterns occurring during resting state (RS), where participants lie in the scanner without any particular task or any external stimulation (Damoiseaux et al., 2006; Amft et al., 2015; Langner and Eickhoff, 2013; Binder et al., 2009; Buhle et al., 2014; Shehzad et al., 2009; Schaefer et al., 2018). In contrast to task-based studies, RS is thought to reveal the intrinsic brain organization (Smith et al., 2009). In addition, the ease of implementation of RS data allows for the relatively quick acquisition of large healthy and clinical samples due to low demands on participants. Although the RS paradigm has provided a variety of insights into the organization of the human brain, it also comes with limitations: In the absence of a task, RS is likely influenced by spontaneous thoughts of the participant (Christoff et al., 2004; Gonzalez-Castillo et al., 2021). Furthermore, experimental decisions such as instructing participants to keep their eyes open or closed can affect the measurement (Patriat et al., 2013). Finally, various studies have shown that individual FC during RS is heavily influenced by state effects (Geerligs et al., 2015).

To address these limitations, naturalistic viewing (NV) has been suggested as a promising tool for the study of individual differences (Finn et al., 2017; Finn et al., 2020). During NV, participants are instructed to watch a movie clip without any additional task. Therefore, NV reduces the variability induced by spontaneous thought content of the subject, because all participants are presented with the same stimulus (Hasson et al., 2004). By more closely mimicking conditions under which the brain naturally operates, NV promises to capture more ecologically valid neuronal responses. Despite NV increasing the similarity of FC across participants, individual differences still persist. Using “fingerprinting” (Finn et al., 2015) or identifiability as a proxy for individual

* Corresponding author at: Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Centre Jülich, Jülich 52428, Germany.
E-mail address: j.kroell@fz-juelich.de (J.-P. Kröll).

https://doi.org/10.1016/j.neuroimage.2023.120083.
Received 26 October 2022; Received in revised form 7 March 2023; Accepted 31 March 2023
Available online 2 April 2023.

differences, Vanderwal et al., (2007) demonstrated that NV shows better identification accuracy than RS (Vanderwal et al., 2017). Furthermore, Finn et al. (2020) showed that the implementation of NV data outperforms RS in predicting trait-like phenotypes, thus suggesting that individual variability might be enhanced during NV (Finn and Bandettini, 2020). Different attempts have been made to explain why NV might enhance FC variability. For instance, Geerligs et al. (2015) argued that the differences in interpretation of a given movie content might promote individual FC variability (Geerligs et al., 2015). Van de Meer and colleagues (der et al., 2020) suggested that NV might impose richer brain state dynamics and therefore more distinct connectivity profiles, which in turn might better reflect phenotypes of interest than brain states during RS. Naturalistic Viewing paradigms provide further advantages over conventional RS: By increasing participant engagement, NV reduces fatigue and head movement during the measurement (Finn and Bandettini, 2020; Vanderwal et al., 2019). In addition, movie-watching can increase scanner tolerability for cohorts which might either struggle with staying still (e.g. ADHD patients) or completing demanding tasks (subjects with cognitive impairments) (Eickhoff et al., 2020).

Current literature evinces the potential for naturalistic viewing as a useful paradigm for studying individual brain architecture. So far, most studies primarily focused on whole-brain connectivity reflecting a holistic view on brain functions. However, brain architecture is commonly seen as segregated into modular clusters of spatially distinct areas constituting functional networks (Sporns and Betzel, 2016). These networks represent specific cognitive domains, such as memory (Spreng et al., 2009), social cognition (Bzdok et al., 2012) and executive function (Rottschy et al., 2012). Therefore, investigating networks functional connectivity (NFC) increases the interpretability of findings over whole-brain connectivity. Furthermore, connectivity in different networks likely yields distinct patterns of variance in reaction to NV stimuli. For example, a functional network related to the processing of emotions should react differently to a movie scene with strong emotional content, as compared to the motor network.

The most commonly used method to define functional networks is to estimate them from FC under resting-state (Damoiseaux et al., 2006; Schaefer et al., 2018; Thomas Yeo et al., 2011). RS-networks have shown good reproducibility and seem to generally converge well with studies on task-evoked networks (Smith et al., 2009; Mennes et al., 2010; Dosenbach et al., 2007). However, there are several other methods for defining functional networks (Schaefer et al., 2018; Smith et al., 2009; Power et al., 2011), one of which are meta-analytically defined networks (Eickhoff et al., 2012). The latter have the advantage of representing the most likely core nodes involved in a given cognitive function, because they incorporate convergent information from a multitude of studies (Eickhoff et al., 2020). Thus, studying NFC in meta-analytical networks might grant robust insights into the effects of naturalistic viewing on individual variability, which has not been studied yet.

The present study aims to investigate the influence of NV on individual variability in NFC by use of three different movie stimuli and RS. There is a plethora of NV stimuli available. Depending on the research question at hand, studies have suggested to use stimuli that are disease-specific (e.g. a movie evoking suspicion to study paranoia) (Eickhoff et al., 2020; Finn et al., 2018), emotionally or socially engaging (Finn and Bandettini, 2020; Saarimäki, 2021; Mishra et al., 2022; Schaefer et al., 2010) or as neutral as possible (Vanderwal et al., 2015). Previous studies on individual variability under NV employed stimuli that the researchers deemed to be the most engaging, thus resorting to movies with high social and emotional content (Finn and Bandettini, 2020; Saarimäki, 2021; Mishra et al., 2022; Schaefer et al., 2010). We employ stimuli with different levels of social content, ranging from the neutral movie *Inscapes*, over the silent movie *The Circus,* to the most social movie *Indiana Jones and the Temple of Doom.* Understanding how different levels of social and emotional content influence individual variability on a network level might aid researchers in choosing adequate stimuli for future studies.

We compare several measures of individual variability (e.g. identifiability and inter- and intra-subject variability) between the three different movie stimuli and RS across three scanning sessions on the basis of various meta-analytical networks covering affective (Amft et al., 2015; Buhle et al., 2014; Liu et al., 2011; Sabatinelli et al., 2011), social (Amft et al., 2015; Bzdok et al., 2012; Caspers et al., 2010), executive (Langner and Eickhoff, 2013; Rottschy et al., 2012; Camilleri et al., 2018; Cieslik et al., 2015), memory (Binder et al., 2009; Spreng et al., 2009) and motor (Witt et al., 2008) functions. Furthermore, we validate our results in RS-derived networks by Thomas Yeo et al. (2011), and on a whole-brain atlas by Shen et al. (2013). As a first step, we examined the similarity of connectivity profiles evoked by different movies and RS. Secondly, we assessed the identifiability of subjects based on NFC-patterns evoked by NV or RS. Subsequently, we investigated to what extent identifiability is affected by network size. Finally, we compared the effect of different movies and RS on inter- and intra-subject variability.

2. Material and methods

2.1. Participants

36 healthy right-handed and ambidextrous adults were scanned at the centre for Translational MR Research, National University of Singapore. Two subjects were excluded for having incomplete sessions, leaving a final cohort at 34 (19 females, mean age 27 + / - 2.7 years). Exclusion criteria were neurological or psychiatric diagnoses, significant visual or hearing impairment, alcohol or caffeine consumption 6 h prior to the scan and self-reporting of bad sleep the night before the scan days. All participants underwent three identical testing sessions within a one-week interval. Subjects gave written, informed consent and were compensated for their participation. The study was approved by the institutional review board of the National University of Singapore.

2.2. Data acquisition

The data was acquired on a Siemens Magnetom PrismaFit 3-Tesla with a 20-Channel head coil. Structural images were collected using an MP-RAGE sequence (TR=2300 ms, TE =2,28 ms, TI=900 ms, flip-angle=8°) and 1 mm voxel size. All RS and NV runs used the same echo planar imaging sequence (TR=719 ms, TE=30 ms, flip-angle=52°, slices=44, FOV=225×225 mm^2) resulting in 2.96×2.96×3 mm voxel size. Data were retrieved from collaborators at the National University of Singapore, and structured in the form of a DataLad dataset, a research data management solution providing data versioning, data transport, and provenance capture Halchenko et al. (2021). Each of the three testing sessions per participant, which were conducted within a seven day period, comprised three NV runs and two RS scans. The order of scans was identical on all three days, starting with a structural scan, followed by 5 functional scans in the order of RS 1, *Inscapes, Circus, Indiana Jones* and RS 2, with each functional scan lasting for 10 min. All movies had been cut to the same length. For RS scans, participants were asked to lay as still as possible and think of nothing in particular, while keeping their eyes open. Instructions for the NV scans were to watch the movies while staying as still as possible. For all scans, participants were asked to not fall asleep during the measurement. The movie clips were presented via a mirror that was mounted on the head coil and the sound was played through headphones. *Inscapes* is a non-verbal, non-social series of animated abstract shapes created by Vanderwal et al. which was looped to match the 10 min duration (original length 7 min) (Vanderwal et al., 2015). *The Circus* (United Artists Digital Studios, 1928, directed by Charlie Chaplin) is a silent black-and-white film which depicts the protagonist being chased by the police and unintentionally causing comic situations during his escape. *Indiana Jones and the Temple of Doom* (Paramount Pictures, 1984, directed by Steven Spielberg) shows the opening scene of the movie during which the protagonist has to fight off several hitmen who are trying to kill him. Foam wedges were fitted around each subject's head for comfort and to decrease movement. For all subsequent analyses, only the first RS scan (RS1) was used.

2.3. Data preprocessing

Preprocessing of MRI data was performed using fMRIPrep, version 20.1.1 (Esteban et al., 2019). In brief, the T1-weighted volumes were corrected for intensity non-uniformity and skull-stripped. The extracted brain images were then transformed into Montreal Neurological Institute (MNI) space and motion corrected using Advanced Normalization Tools (ANTS) (Avants et al., 2009). The functional data was motion-corrected with MCflirt (Jenkinson et al., 2002) and subsequently co-registered to the native T1-weighted image using boundary based registration with six degrees of freedom from Freesurfer (Greve and Fischl, 2009). Subsequently, ICA-AROMA (Pruim et al., 2015) was used on the MNI-aligned BOLD images to remove motion artifacts and applied an isotropic Gaussian kernel of 6 mm FWHM (full-width half-maximum) for spatial smoothing. Global signals were extracted within the CSF, the WM, and the whole-brain masks and regressed from the preprocessed fMRI data for each subject.

2.4. Network functional connectivity

For each subject, NFC matrices were constructed for each of the 14 meta-analytical networks, comprising nine to 23 nodes (a detailed description of the networks can be found in the supplements). Isotropic 5 mm spheres were created around the local maxima of each meta-analytical network node and the mean time series were subsequently extracted. Only gray matter voxels were included. In addition, NFC matrices were constructed for the seven RS derived networks created by Thomas Yeo et al. (2011), comprising the Default, Control, Dorsal Attention, Salience, Visual, Somatomotor and Limbic networks, and the whole-brain atlas created by Shen et al. (2013). Pearson's correlation coefficient (PCC) between all node pairs was calculated to generate a n-times-n connectivity matrix per subject and condition, where n denotes the number of nodes of the respective network.

2.5. Representational dissimilarity matrix (RDM) analysis

To investigate how patterns of inter-individual differences in NFC vary across conditions (RS and three different NV conditions), we applied a RDM analysis. The present analysis closely followed the methods described by Kriegeskorte (2008). The procedure can be summarized in three steps. First, the correlation between the FC patterns of every possible pair of subjects is calculated for each condition and network. Second, to generate a measure of dissimilarity, the correlation distance (1-r) is computed. Third, the dissimilarity values for all subject pairs are assembled into an RDM (as a subjects * subjects size matrix) that serves as the signature of the given condition.

To visually compare RDMs, we employed Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018), a technique for dimensionality reduction and visualization. Instead of preserving large-scale structures, UMAP seeks to preserve local neighborhood distances. To this end, a pre-set number of nearest neighbors (NN) are specified and the distances to these neighbors is represented as a weighted graph, with the NN being assigned with higher weights. UMAP then finds a low-dimensional representation of the data that best preserves these neighborhoods. The NN parameter controls whether UMAP focuses on the local or global structure of the data. Large values force UMAP to consider a larger number of neighbors and therefore focus on the broader structure of the data. In contrast, low values of NN force UMAP to focus on the local structure of the data. We here applied UMAP to the previously described RDMs. To account for the small number of data points (fifteen RDMs per network) the NN parameter was set to two. Considering more than 4 NN led to a more global clustering of RDMs that partly obscured differences between conditions. By grouping closely related RDMs together, UMAP allowed us to visualize which conditions evoked similar responses. Of note, distance metrics in UMAP are non-linear and not necessarily the same for each dimensionality reduction. Therefore, the results are suited to compare the similarity of condition evoked responses within, but not across networks. An analysis of the RDMs on a whole-brain level is reported in the supplementary material (Fig. S1).

2.6. Assessment of identifiability

Assessment of identifiability was closely based on the methods described by previous papers (Finn et al., 2015; Vanderwal et al., 2017). The FC matrices belonging to the same session and condition were grouped, resulting in 12 databases (three sessions times the four conditions). For every combination of two databases, Pearson's correlation between the FC matrix of one subject from the first database and every other FC matrix from the second database was calculated. The two FC matrices with the highest correlation were considered to be from the same subject. Identification accuracy was defined as the frequency of correctly identified subjects divided by the total number of subjects. Afterwards, the accuracies were averaged across session pairs to quantify the identifiability per condition and network. An analysis of identifiability on a whole-brain level is reported in the supplementary material (Table S1).

2.7. Influence of network size

To ensure that the differences in identification accuracy between networks were not just reflections of network size, we systematically compared identifiability in artificially created networks, constituting up to 50 nodes. Artificial networks were created by randomly choosing coordinates from the MNI152 gray matter mask. Around each coordinate, an isotropic sphere was created, which was matched to the node size of the meta-analytical networks (5 mm). The mean Euclidean distance between nodes from the meta-analytically defined networks was calculated (14.62 mm) and set as the minimal distance between nodes for the artificial networks. Thereby, the randomly chosen nodes were prevented from overlapping whilst preserving some degree of spatial comparability between artificial and meta-analytically defined networks. This process was repeated 100 times for each network size, creating a new random configuration of nodes during each repetition. Subsequently, identification accuracies for all networks and the different conditions were calculated to evaluate (1) how network size influences identification accuracy, (2) how identifiability between the different conditions behaves in artificial networks and (3) how the meta-analytically defined networks compare to the artificial networks.

2.8. Within- and between subject correlation

Within-subject correlations were calculated as Pearson's correlation between the FC matrices of the same subject across session pairs (e.g. Ses-1 to Ses-2, Ses-1 to Ses-3) and then averaged. This process was performed for each of the four conditions (RS and the three movie stimuli) separately. For each network or whole-brain atlas, a one-way ANOVA was computed with condition (RS, *Inscapes, Circus, Indiana Jones*) as within-subject factor to evaluate the effect of condition on within-subject correlations within the specific networks. Subsequently, Bonferroni correction was applied to account for Type 1 error and Tukey's HSD test was performed to reveal which of the conditions significantly differed. The between-subject correlations were defined as the mean PCC between the FC matrix of one subject and every other subject's FC matrix from the same session and condition. For each network, a one-way ANOVA was computed with condition (RS, movie1, movie2, movie3) as between-subject factor to evaluate the effect of condition on between-subject correlations within the specific networks. Subsequently, Bonferroni correction was applied to account for Type 1 error and Tukey's

HSD test was performed to reveal which of the conditions significantly differed. It is important to note that the between-subject comparisons in this study are based on correlations between static NFC of subjects, in contrast to an Inter-subject Correlation (ISC) approach that correlates the fMRI time series of subjects and is often used to analyze NV (Halchenko et al., 2021). As such, our results should not be interpreted as a measure of synchrony across subjects, but rather as their similarity in FC. The analysis of within- and between-subject correlations on a whole-brain level can be found in the supplementary material (Fig. S2).

3. Results

3.1. Similarity of different movies and RS connectivity profiles in meta-analytic networks

We investigated the similarity of different conditions by embedding the respective RDMs into a low dimensional space (UMAP).The UMAP representation showed that RS was embedded separately from all NV conditions in AM, CogAC, VigAtt and WM, and separately from most NV conditions in MNS and Motor networks. In eMDN, EmoSF, ER, eSAD, Rew and ToM networks, RS shows overlaps with the movie *Inscapes*. On the other hand, the movies *Circus* and *Indiana Jones* tended to cluster together in (AM, CogAC, eMDN, Empathy, ER, eSAD, MNS, Motor, Rew). We did not observe any evidence for a systematic session-effect, as RDMs of the same session (session 3) were only embedded together in the motor network (Fig. 1).

3.2. Similarity of different movies and RS connectivity profiles in RS derived networks

The UMAP representation of the different conditions in RS derived networks showed that RS was embedded separately from all NV conditions in the Control network and separately from most NV conditions in Limbic, SomatoMotor and Visual networks. In all networks except for the Control network, RS shows overlap with the movie *Inscapes. Indiana Jones* and *Circus* overlap in all networks (Fig. 2).

3.3. Identification accuracies in meta-analytic networks

Identifiability of subjects was assessed based on NFC-patterns evoked by NV or RS. Overall, individual FC matrices could be matched across sessions with moderate to high accuracy with identification accuracies ranging from 52% to 100%. The motor network represented an exception with low identification accuracies across conditions (27.5%–30.4%). In eleven out of 14 networks, identifiability was highest in either the *Circus* or *Indiana Jones* NV conditions. Among the naturalistic stimuli, *Indiana Jones* led to the highest identification accuracies in eight of the networks (SM, CogAc, EmoSF, eMDN, ER, VigAtt, MNS, and eSAD). The top three highest accuracies were achieved using NV, with FC matrices using the *Indiana Jones* movie reaching the highest accuracy (98%) in the SM network. Generally, networks with more nodes tended to achieve higher accuracies.

3.4. Identification accuracies in RS derived networks

In addition, identifiability of subjects was assessed based on NFC in RS derived networks. Generally, individual FC matrices could be matched with moderate to high accuracy with accuracies ranging from 43% to 91%. The limbic network represented an exception with low identification accuracies across conditions (9.3%–14.22%). In the control, dorsal attention and visual networks, *Indiana Jones* led to the highest identification accuracy. In the default, salience and somatomotor networks, RS led to the highest identifiability. The highest accuracy was achieved by RS in the default network (91%). Overall, accuracies in the RS derived networks were lower than in the majority of meta-analytically derived networks.

3.5. Identification accuracies for different network sizes

To evaluate the effect of network size on identification accuracy, we computed identifiability in random networks with sizes between 3 and 50 nodes. We then compared these to the accuracies achieved in meta-analytic networks, as the meta-analytic networks showed higher accuracy then the RS derived networks. Identifiability in artificial networks showed how network size influences identification accuracy for all modalities (Fig. 2). A continuous increase of identification accuracy can be seen for all conditions up until a network size of 20 nodes, where the increase rate stabilizes. All networks, apart from the Motor network, achieved higher accuracies than the artificially created networks of the same size, regardless of condition. Furthermore, identification accuracies for the *Indiana Jones* movie exceeded those of the other three conditions, regardless of network size (Fig. 3).

3.6. Within- and between-subject correlations in meta-analytic networks

We calculated within-subject correlations, as a measure of how similar subjects are to themselves across sessions, and between-subject correlations, as a measure of similarity between subjects. The average within-subject correlations for RS and NV ranged between 0.5 and 0.8, with the exception of the Motor network (0.1–0.6), indicating a high level of similarity of connectivity patterns across sessions. For multiple networks, most prominently the MNS network, within-subject correlations strengthened from RS < *Inscapes* < *Circus* < *Indiana Jones*.

RS state differed from one or more movie conditions in various networks: RS showed significantly higher within-subjects correlations compared to *Indiana Jones* (AM) and *Circus* (AM). In contrast, some movies showed significantly higher within-subject correlations than RS in emoSF (*Indiana Jones*), and MNS (*Indiana Jones* and *Circus*).

In several networks certain movies differed from one another, with significantly higher correlations in *Indiana Jones* compared to *Circus* in emoSF; and higher correlations in *Indiana Jones* compared to *Inscapes* in Empathy and MNS networks. *Circus* never showed significantly higher correlations compared to any other movie in any network.

RS and the movie *Inscapes* exhibited similar correlations across networks. Overall, the movie *Indiana Jones* tended to stand out in that it was the only condition that showed significantly higher within-subject correlations than RS in several networks (EmoSF and MNS). On the contrary, the movie *Circus* often led to decreased within-subject correlations in comparison to the other conditions.

Between-subject correlations were generally lower than those previously observed on a whole-brain level, ranging from below 0.1 to 0.75. In several networks, the opposite pattern of what was observed for within-subject correlations can be seen, such that increasingly complex stimuli weaken between-subject correlations (AM, ER, eSAD and SM). In other networks, the three movies made connectivity across subjects more similar, increasing between-subject correlations in comparison with RS (CogAc, EmoSF, Rew and VigAtt).

Comparing within- and between-subject correlations, it is evident that increased within-subject correlations did not automatically lead to decreased between-subject correlations (and vice versa), such that a subject's scan can be highly individual (or reliable) and still share substantial overlap with others.

RS differed from one or more movie conditions in various networks: RS showed significantly higher between-subjects correlations compared to *Indiana Jones* (AM, eSAD, SM, ToM), *Inscapes* (ToM) and *Circus* (AM, Motor, SM, ToM). In contrast, other networks showed significantly higher between-subject correlations than RS for *Indiana Jones* (CogAC, eMDN, EmoSF, Rew, VigAtt,WM), *Inscapes* (CogAC, emoSF, Rew, VigAtt, WM,) and *Circus* (CogAC, EmoSF, MNS, Rew, VigAtt).

In several networks certain movies differed from one another, with significantly higher between-subject correlations of *Inscapes* compared to *Circus* in the AM, CogAC, EmoSF, eSAD, Motor, SM, and ToM; and higher correlations in *Inscapes* compared to *Indiana Jones* in AM, EmoSF,

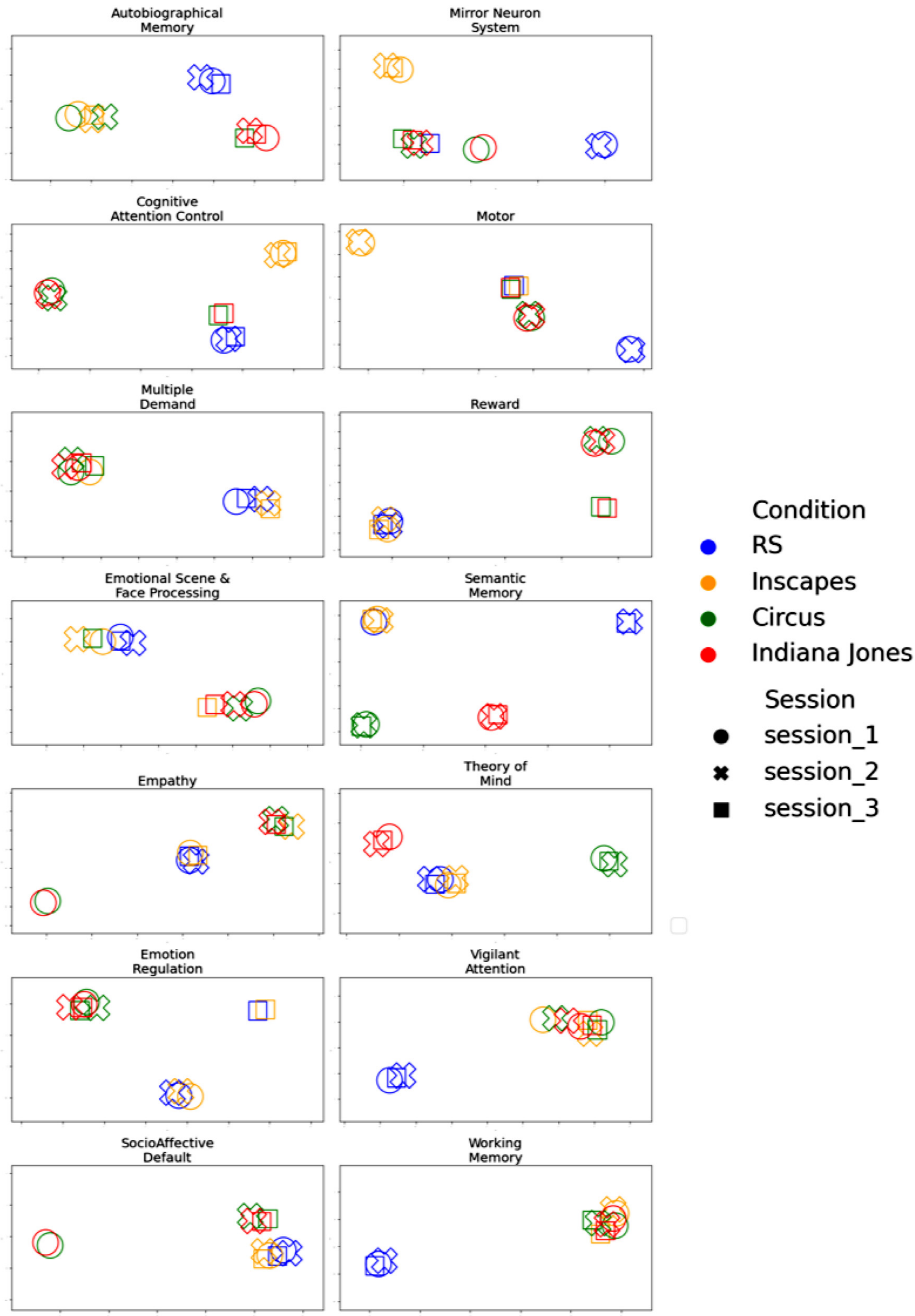

Fig. 1. UMAP representation of the RDMs of the different conditions in each meta-analytic network.

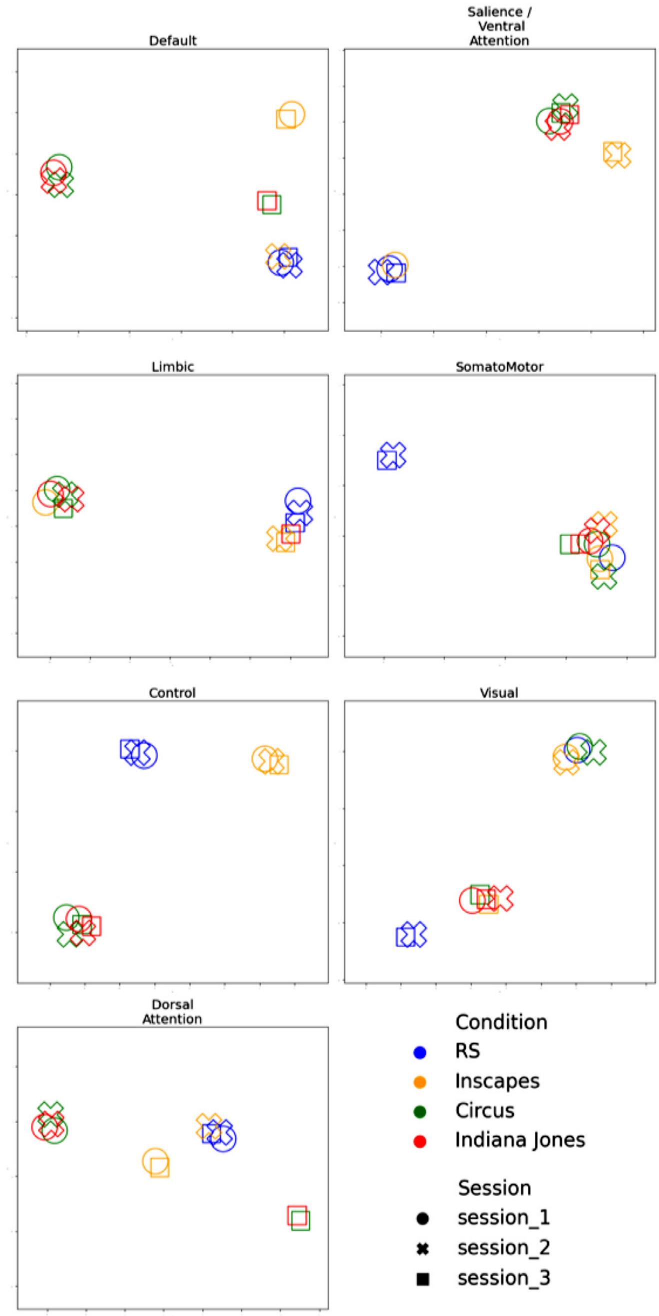

Fig. 2. UMAP representation of the RDMs of the different conditions in each RS derived network.

ER, eSAD and SM; and higher correlations in *Indiana Jones* compared to *Circus* in eDMN, Motor and ToM; and higher correlations in *Circus* compared to *Indiana Jones* in AM and SM networks (Figs. 4 and 5).

3.7. Within- and between-subject correlations in RS derived networks

We calculated within- and between-subject correlations for the RS derived networks. The average within-subject correlation for RS and NV ranged between 0.6 and 0.9, with the exception of the limbic network (0.1–0.8), indicating a high level of similarity of connectivity across sessions. The within-subject correlations in the RS derived networks were generally higher than the within-subject correlations in the meta-analytic networks. RS showed significantly higher within-subject correlations than *Circus* in the default network.The movie *Indiana Jones* showed significantly higher within-subject correlations than *Circus* in the Default network.

The average between-subject correlations ranged between 0.1 and 0.9 and were generally higher than the between-subject correlations in the meta-analytic networks. In five out of seven networks, at least one of the movie conditions led to higher between-subject correlations than for RS.

RS differed from one or movie conditions in various networks. RS showed significantly higher between-subject correlations compared to *Circus* (Cont) and *Indiana Jones* (Default). In contrast, other networks showed higher between-subject correlations than RS for *Inscapes* (DorsAtt), *Circus* (SalVentAtt, SomMot, Vis) and *Indiana Jones* (SalVentAtt, SomMot, Vis).

In several networks, certain movies differed from each other with significantly higher between-subject correlations for *Inscapes* than *Circus* in the Default and DorsAtt network; and higher correlations for *Inscapes* compared to *Indiana Jones* in the Default network; and higher correlations for *Circus* than *Inscapes* in the SalVentAtt, SomMot and Vis networks; and higher correlations for *Circus* than for *Indiana Jones* in the Vis network; and higher correlations for *Indiana Jones* than for *Inscapes* in the SalVentAtt and SomMot networks Figs. 6 and 7).

4. Discussion

In the current study we examined and compared the NFC evoked by different NV stimuli and RS with respect to similarity of connectivity profiles, individual identifiability, as well as within- and between-subject correlations. Our results showed that NV stimuli evoke connectivity profiles that are distinct from RS across meta-analytically defined and RS derived networks. NV stimuli, especially *Indiana Jones*, enhance the identifiability of individual subjects in the vast majority (10 of 14) of meta-analytic networks. Crucially, our results show that NFC analysis might reveal differences that are obscured on a whole brain basis. Lastly, our results emphasize that the similarity of individuals to themselves and to others is highly dependent on the combination of condition and network.

4.1. Comparison of connectivity profiles during NV and RS

In this study, we compared NFC evoked by three different NV stimuli and RS. A low-dimensional embedding of NFC similarity across subjects in meta-analytic networks showed that FC patterns during *Inscapes* are mostly similar to those during RS, while *Circus* and *Indiana Jones* exhibited distinct connectivity profiles across networks (Fig. 1). The relative similarity of connectivity patterns during Inscapes and RS has been reported before: For instance, based on Pearson's correlations between FC matrices, *Inscapes* was shown to be more similar to RS than to another movie condition (Vanderwal et al., 2017). These authors argued that due to the abstract nature of the movie, participants might not engage in temporally synchronized cognitive processes, which is similar to RS (Vanderwal et al., 2015). Furthermore, our embedding shows little similarity of NFC during *Inscapes* and either *Circus* or *Indiana Jones* in the majority of networks. This is in line with the previous argument, as both *Circus* and *Indiana Jones* contain a narrative that is likely to increase similarity across subjects, as has been shown for verbal narratives (e.g. emotional speeches (Nummenmaa et al., 2014; Schmälzle et al., 2015)). Accordingly, connectivity profiles during *Circus* and *Indiana Jones* overlap in the vast majority of networks. For the whole brain, similarity across conditions seemed more widespread and all conditions clustered together at least once (Fig. S1).

4.2. Identifiability

To assess the stability of individual patterns on the network level, we calculated the identifiability of NFC matrices across the three movies and RS (Table 1). Considering that NV has been shown to increase the

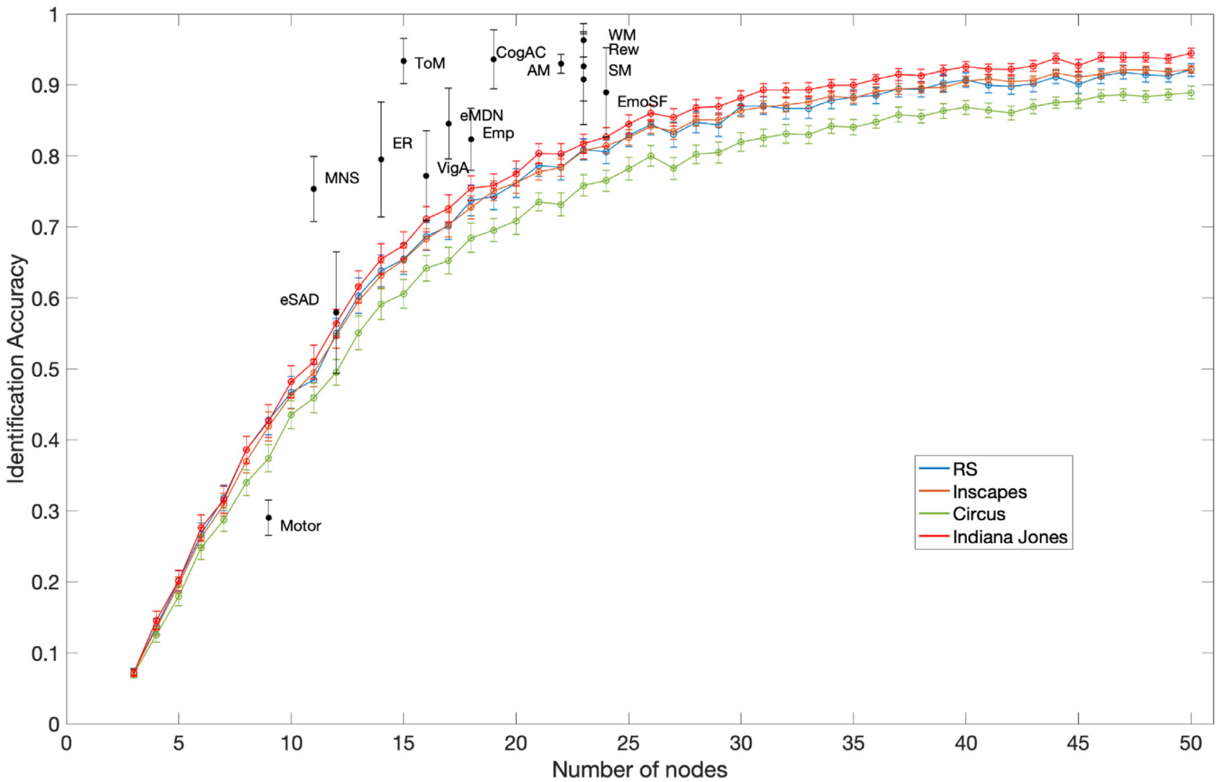

Fig. 3. Identification accuracies in artificial networks. The figure depicts the network size as the number of nodes (x-axis) against averaged identification accuracy (y-axis) for each of the four conditions (RS = blue; Inscapes = orange; Circus = green; Indiana Jones = red). Black dots denote the mean identification accuracy of meta-analytically defined networks, averaged across conditions and placed at their respective node count. (AM =Autobiographical Memory, CogAC = Cognitive Attention Control,eMDN=extended Multiple Demand Network, EmoSF= Emotional Scene and Face Processing, ER = Emotion Regulation, eSAD=Extended Social-affective Default, MNS = Mirror Neuron System, Rew = Reward, SM = Semantic Memory, ToM = Theory of Mind, VigAtt= Vigilant Attention, WM = Working memory,.

Table 1
Identification accuracies per network and modality, averaged across sessions. Networks are in order of highest average accuracy. The highest identification accuracy in each network is denoted in bold. (AM =Autobiographical Memory, CogAC = Cognitive Attention Control,eMDN=extended Multiple Demand Network, EmoSF= Emotional Scene and Face Processing, ER = Emotion Regulation, eSAD=Extended Social-affective Default, MNS = Mirror Neuron System, Rew = Reward, SM = Semantic Memory, ToM = Theory of Mind, VigAtt= Vigilant Attention, WM = Working memory, Shen = Shen atlas).

Network	RS	Inscapes	Circus	Indiana Jones	Node Number
Semantic Memory	95.1%	95.1%	97.1%	**98.0%**	23
Cognitive Attention Control	93.6%	90.2%	94.1%	**96.6%**	19
Theory of Mind	93.1%	90.7%	**95.1%**	94.6%	15
Autobiographical Memory	**94.1%**	92.2%	93.1%	92.6%	22
Working Memory	**96.1%**	93.6%	88.7%	92.2%	23
Reward	**96.1%**	90.7%	86.3%	90.2%	23
Emotional Scene & Face Perception	88.7%	85.8%	86.8%	**94.6%**	24
Multiple Demand Network	85.8%	85.8%	79.9%	**86.8%**	17
Empathy	**86.3%**	81.4%	79.9%	81.9%	18
Emotion Regulation	81.9%	80.9%	72.1%	**83.3%**	14
Vigilant Attention	80.4%	74.0%	73.5%	**80.9%**	16
Mirror Neuron System	**77.0%**	76.5%	71.1%	**77.0%**	11
Socio Affective Default	59.8%	52.9%	54.4%	**64.7%**	12
Motor	27.9%	30.4%	**30.4%**	27.5%	9

reliability of individual FC patterns (Geerligs et al., 2015; Hasson et al., 2010), we hypothesized that identifiability should be higher for movies as compared to RS. However, present results suggest that this is not the case for movies in general, but rather identification accuracy appears to highly depend on the specific movie as well as on the chosen network. Specifically, *Indiana jones* achieved the highest accuracy in 8 of 14 networks (SM, CogAC, EMOSF, eMDN, ER, VigAtt, MNS, eSAD), whereas *Inscapes* and *Circus* produced highest accuracies in two networks (*Inscapes*: Motor; *Circus*: ToM, Motor). RS, on the other hand, achieved the highest accuracies in 5 networks (AM, WM, ReW, Empathy, MNS). Notably, the connectivity profiles within the Motor network yielded low identification accuracies in comparison with the other networks across all stimuli. Lower-level cognitive structures such as the motor network show low variance between participants (Croxson et al., 2018). Further-

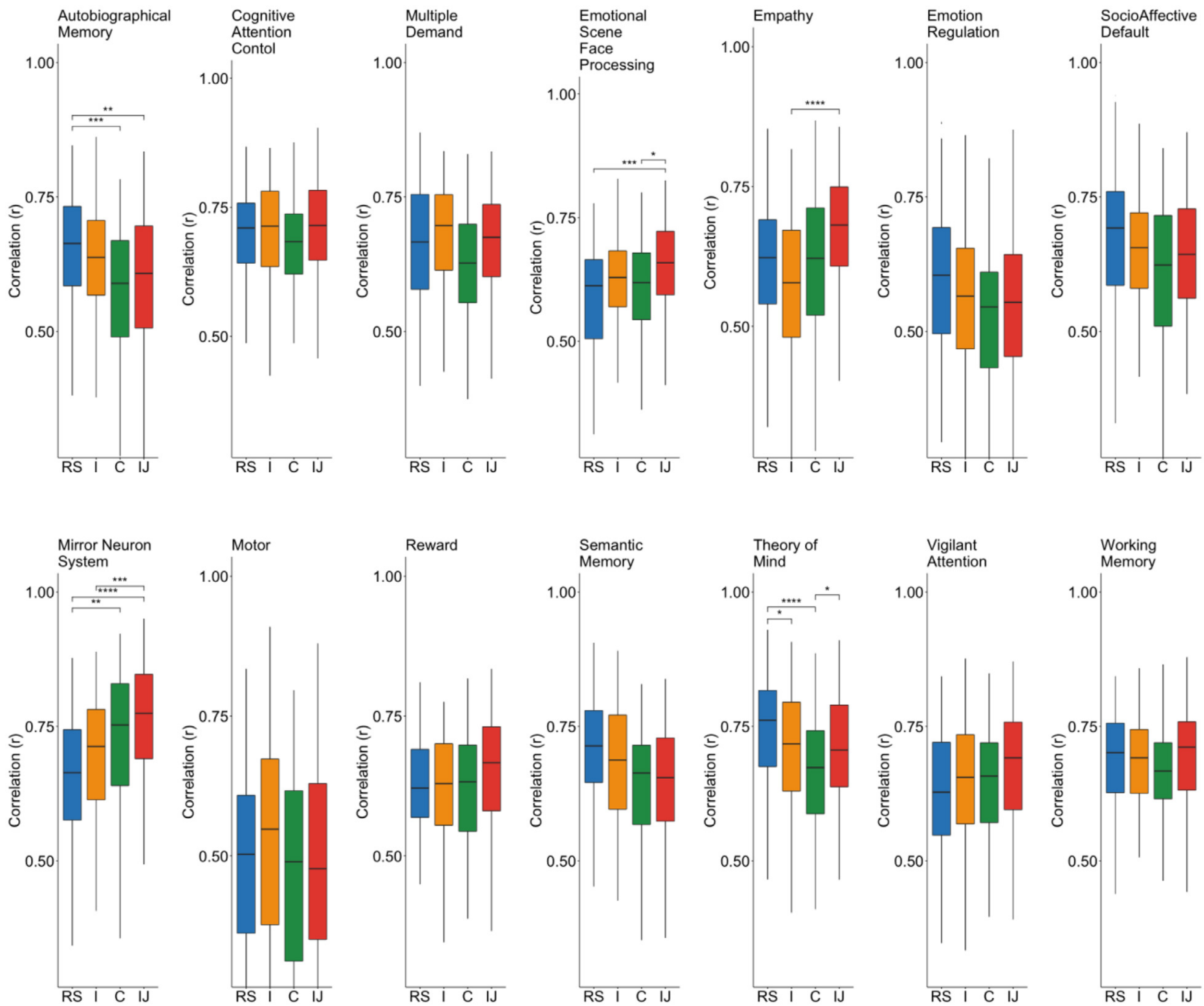

Fig. 4. Within-subject correlations for the meta-analytically defined networks. Correlations across all session pairings are depicted. (RS= Resting State, *I* = Inscapes, C = Circus, IJ = Indiana Jones).

more, as the motor network was created solely based on fingertapping tasks, it seems reasonable to assume that activation was low in this network. Therefore, connectivity patterns are expected to be rather similar across participants.

Indiana Jones was the stimulus that achieved the highest identification accuracy in the majority of networks. Previous studies have argued that the major driving factor for improvement of individual identifiability is the social content of a stimulus (Nummenmaa et al., 2014; Schmälzle et al., 2015; Dmochowski et al., 2014)., which in the present study was most pronounced for *Indiana Jones*. In comparison, neither *Circus* nor *Inscapes* reach the level of social content depicted in *Indiana Jones. Circus'* complete lack of speech might have taken away from the social component whereas *Inscapes* does not depict any human interaction at all.

4.3. Identification accuracies for different network sizes

Since we observed an increase of identification accuracy with network size such that bigger networks tended to show higher accuracies, we investigated the influence of network size on identifiability in artificially created networks (Fig. 3). The results show the same tendency that was observed in the meta-analytically defined networks, such that identification accuracy was highest for *Indiana Jones*, followed by RS, *Inscapes* and *Circus*. Confirming our observation, identification accuracy in artificial networks increased with network size, regardless of condition. Notably, all meta-analytical networks, except the motor network, outperformed artificial networks of the same size, supporting their biological validity. Following our previous line of argument, the motor network might not be suitable for subject identification based on FC, which might explain the underperformance compared to artificial networks.

4.4. Within- and between-subject correlations in meta-analytic networks

To better understand the differences in identifiability across stimuli and networks, we investigated within- and between-subject correlations (Figs. 4 and 5). Our results showed that in the majority of networks, within- and between-subject correlations were significantly altered during NV in comparison to RS. It is generally assumed that NV should increase between-subject similarity, given that all subjects are presented with the same stimuli, in comparison to no stimuli at all during RS (Hasson et al., 2004; Hasson et al., 2010; Kauppi, 2010). On the other hand, it is unclear whether NV can evoke unique and reliable patterns across sessions, as measured by within-subject correlations. Vanderwal and colleagues investigated FC variability in NV and RS and showed that naturalistic paradigms increased within- and between-subject correlations on a whole brain level (Vanderwal et al., 2017). However, our results showed no significant differences for either within- or between-

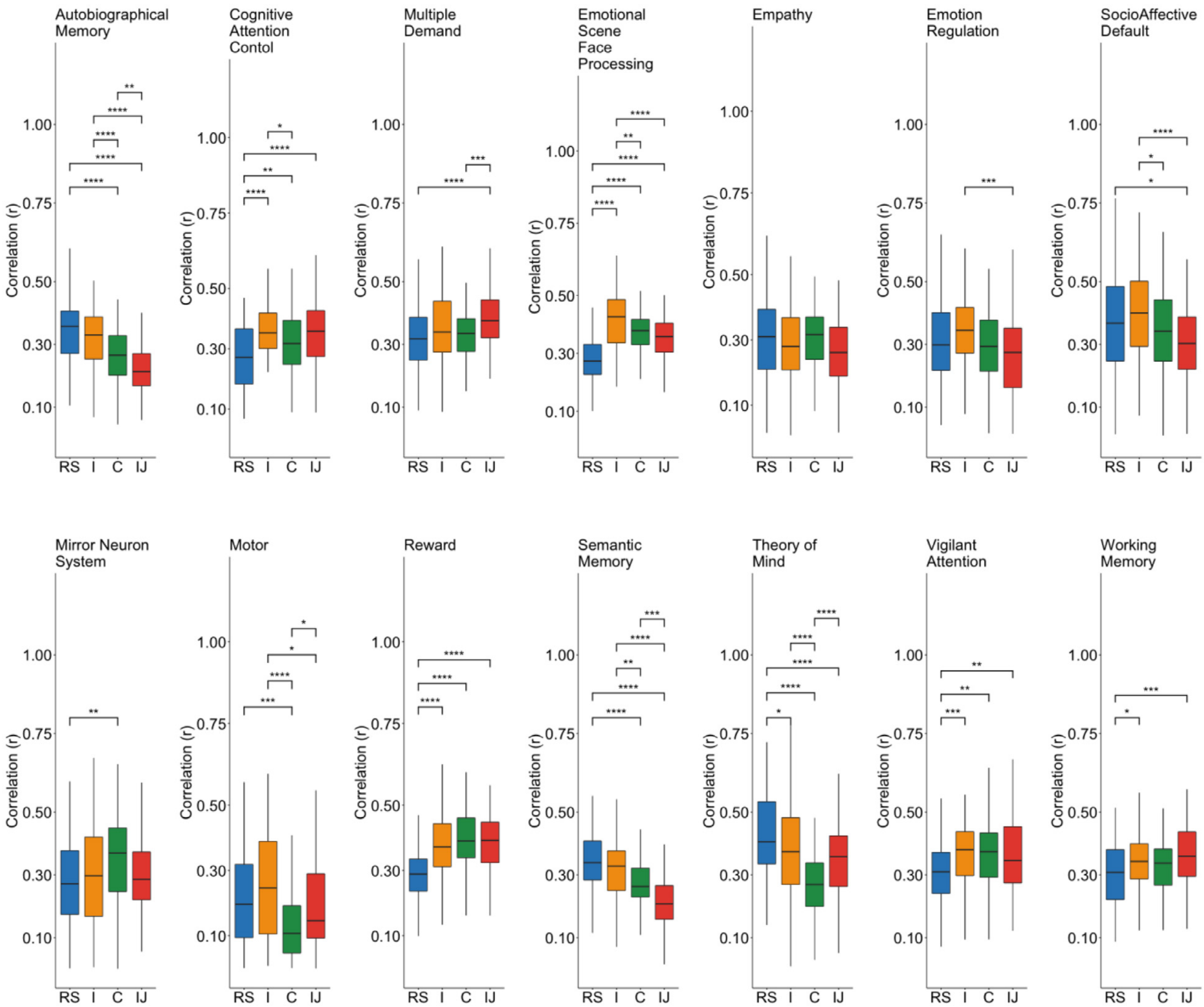

Fig. 5. Between-subject correlations for the meta-analytically defined networks. Correlations for all sessions are depicted. (RS= Resting State, *I* = Inscapes, C = Circus, IJ = Indiana Jones).

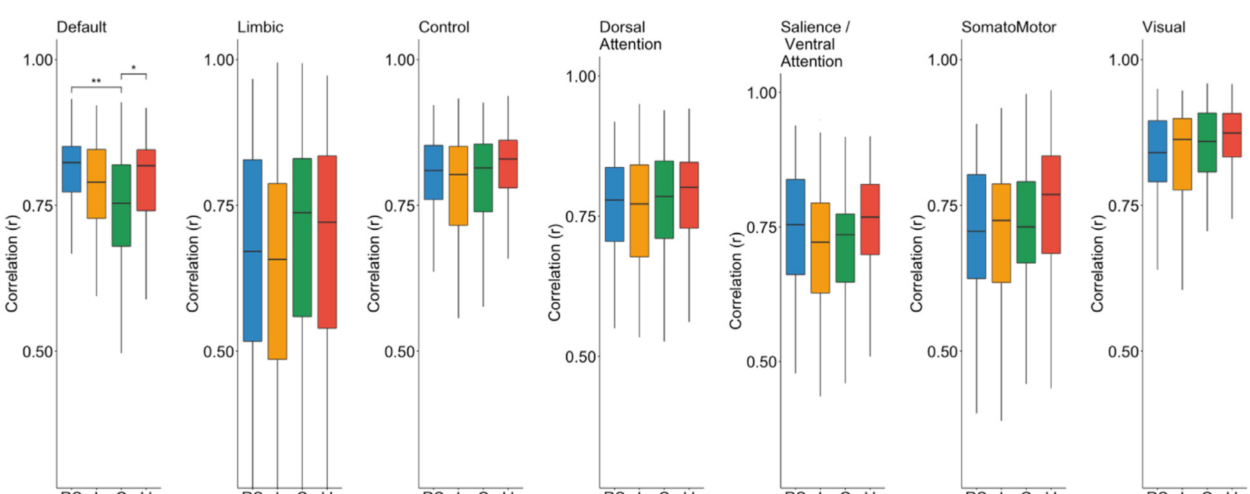

Fig. 6. Within-subject correlations for the RS derived networks. Correlations across all session pairings are depicted.

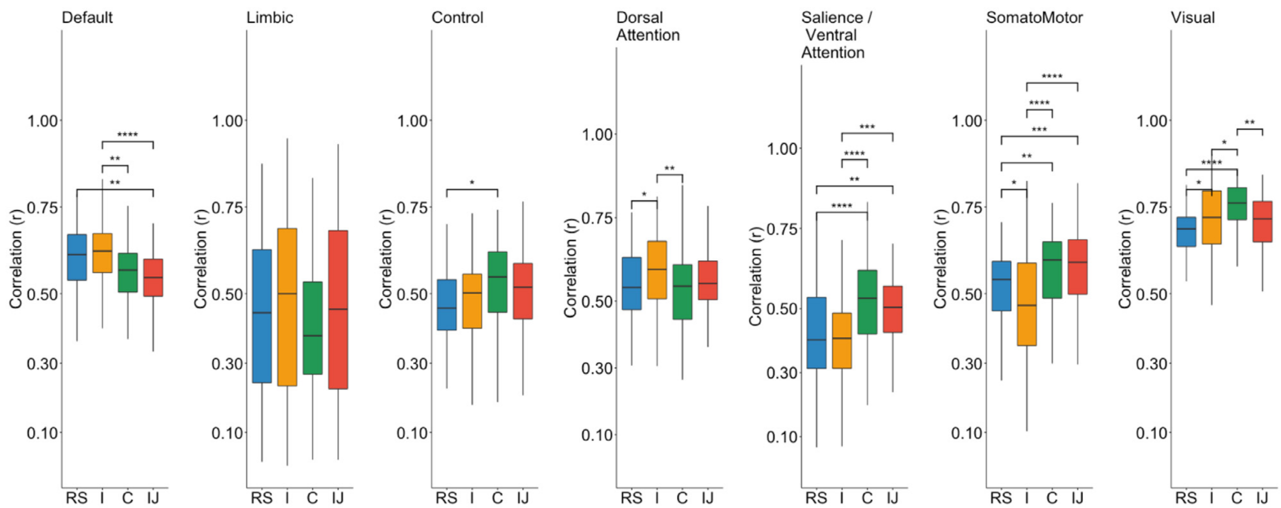

Fig. 7. Between-subject correlations for the RS derived networks. Correlations for all sessions are depicted.

subject correlations on a whole-brain level (supplementary Fig. S1 and S2). On the other hand, our analysis revealed varying effects on a network basis. Increased within-subject correlations were mainly observed in meta-analytic networks that are essential for perception and processing of action, behavior and emotions, namely EmoSF, Empathy and MNS. In other networks, NV resulted in more similar patterns between subjects (CogAC, eMDN, Rew, VigAtt and WM). Notably, multiple networks showed decreased within- and between-subject correlations during NV (AM, ER, eSAD, SM, ToM). We will discuss these three groups of networks subsequently.

4.4.1. Networks with higher within-subject correlations in movies

NV showed significantly higher within-subject correlations in networks that are essential for perception and processing of action, behavior and emotions (EmoSF and MNS). In a recent publication by Finn and Bandettini (2020) it was shown that NV outperformed RS in FC-based prediction of behavioral scores. In their study, movies with strong social content led to the more accurate predictions, regardless of whether the predicted score was social or cognitive. The authors hypothesize that social movies are not only more engaging, but also most likely to evoke divergent interpretations and reactions across participants. In agreement with this assumption, several studies have shown that social movies induce different neural responses across subjects (Finn et al., 2018; Rikandi et al., 2017; Salmi et al., 2020) and that shared interpretation of a narrative or movie is associated with similarity in neural responses (Nguyen et al., 2019; Gruskin et al., 2020). Assuming that the social aspect of a movie stimulus induces stable individual connectivity patterns, it is reasonable to expect that this effect is more pronounced in networks that deal with the processing of social interactions.

In the EmoSF network for example, which deals with the visual and emotional processing of faces or scenes (Sabatinelli et al., 2011), all three movie stimuli led to higher within-subject correlations compared to RS. Notably, the movie *Indiana Jones*, during which the emotional processing of faces is a key aspect, shows highest within-subject correlations. Here, differences in the emotional assessment of the particular faces and scenes might have been the driving factor that evoked stable individual connectivity patterns during *Indiana Jones*.

In the MNS network, which is involved in the understanding of actions and their underlying intentions as well as the imitation of observed behavior, we observed an increase of individuality with increasingly complex stimuli (Caspers et al., 2010). Especially the two stimuli *Circus* and *Indiana Jones*, during which action and behavior of different characters are depicted, should engage the MNS network which in turn might have led to the increased within-subject correlations. The between-subject correlations were significantly stronger for *Circus* than for *RS*, but not different between the remaining conditions. Presumably, the movie *Circus* serves as the optimal stimulus for action observation since it shows moving characters, but (unlike *Indiana Jones*) does not include competing stimuli like speech.

Another network that showed similar patterns, although not reaching significance is the Empathy network, which deals with the emotional cognition of moral behavior (Bzdok et al., 2012). The within-subject correlations were increased during the movies *Circus* and *Indiana Jones*. During both movies, characters show varying emotions in response to different situations, which might have been experienced differently across subjects. *Inscapes* on the other hand performed similar to RS, likely because the depicted abstract shapes failed to engage the network.

4.4.2. Networks with higher between-subject correlations in movies

NV showed significantly higher between-subject correlations in networks that are associated with executive functions and/or stimulus evaluation (CogAC, eMDN, EmoSF, MNS, Rew, VigAtt and WM). Here, NV increased the similarity of FC across participants (i.e. higher-between subject correlation), but did not increase within-subject correlations. Several other studies have found NV to increase the similarity between subjects (Finn et al., 2020; Hasson et al., 2004; Vanderwal et al., 2017; Wang et al., 2017), which is likely caused by exposure to the same stimulus. Although these studies mostly agree that individual differences can exist on top of the shared response on a whole-brain level, they acknowledge two possible scenarios: On the one hand, the stimulus evoked similarity across subjects might enable better observation of individual differences (Vanderwal et al., 2017; Finn and Bandettini, 2020). On the other hand, strongly increased similarity across subjects' neuronal response might blur individual features (Finn et al., 2017). The same assumptions hold true from a network perspective, such that networks subjected to the same stimulus can either exhibit deviating patterns on top of the shared response, or highly similar patterns which conceal individual differences, depending on the specific network function.

Considering the main function of each respective network, none of the networks should be particularly engaged during RS or during any of the movies. The CogAC network is essential for the suppression of a predominant but inadequate response in favor of the contextually appropriate response (Cieslik et al., 2015). The eMDN consists of core regions that are active during most processes which involve executive or higher cognitive functions and a set of more task-specific regions extending these core regions (Camilleri et al., 2018). The Rew network is essential for reward-related decision making (Liu et al., 2011). The

Table 2
Identification accuracies per network and modality, averaged across sessions. Networks are in order of highest average accuracy. The highest identification accuracy in each network is denoted in bold.

Network	RS	Inscapes	Circus	Indiana Jones	Node Number
Default	**91.18%**	85.78%	87.75%	90.69%	24
Control	86.76%	87.75%	79.90%	**89.71%**	13
Dorsal Attention	75.49%	73.04%	66.18%	**75.98%**	15
Salience	**74.51%**	66.67%	56.37%	73.04%	12
Visual	68.63%	58.82%	57.84%	**73.04%**	17
Somatomotor	**62.75%**	60.29%	43.14%	57.84%	14
Limbic	14.22%	14.22%	9.31%	12.25%	5

VigAtt network is involved in vigilant attention, i.e. the continued focus on intellectually un-challenging tasks (Langner and Eickhoff, 2013). The WM network is fundamental for the storage and manipulation of short-term memory (Rottschy et al., 2012). Since individual differences are likely only enhanced in networks that are engaged during a certain condition, we assume that NV did not evoke stable individual connectivity patterns, as the processing of movies may not rely on the core network function. Therefore, subjects are less unique and more similar to themselves, increasing between-subject correlations especially in comparison with unconstrained RS where more heterogeneous responses are expected.

4.4.3. Networks with higher between- or within-subject correlations in RS

The vast majority of previous studies reported increased within- and between-subject correlations for NV in comparison with RS (Finn et al., 2020; Hasson et al., 2004; Vanderwal et al., 2017; Wang et al., 2017; Nastase et al., 2019). However, all of these studies employed analyses of whole brain connectivity, disregarding effects in single networks. While previous result patterns hold true in some networks, we also show that NV decreases within- and between-subject correlations in other networks (AM, ER, eSAD, SM and ToM).

The majority of these networks at least partially overlap with the default mode network, which is tied to intrinsically oriented functions, rather than the processing of external stimuli (Hasson et al., 2004; Golland et al., 2007). Therefore, it seems plausible that NV does not increase within- or between-subject correlations in these networks which are likely not engaged during movie watching. The AM network was the only network in which within-subject correlations for RS exceeded *Indiana Jones*. This network comprises brain regions engaged in processes concerning scene-construction and self-projection, or the ability to mentally project oneself from the present moment into another time, place, or perspective. Consequently we would expect the AM network to be more strongly activated during RS, when the mind is not occupied by the content of a movie. Our data indeed shows that participants during RS showed higher within-subject correlations than during the two narrative movie clips *Circus* and *Indiana Jones*, but not significantly different from the purely abstract animation *Inscapes*. Therefore, we conclude that in absence of a storyline, subjects divert to imagined situations instead of the external stimuli, thus engaging the AM network which leads to higher within-subject correlations for RS than for the narrative movies. We assume that the movie *Inscapes* is inbetween a narrative and the complete absence of a stimuli, thus it may fail to engage participants over a longer period of time, therefore letting the participant zone out eventually. In addition, RS and *Inscapes* also increased between-subject correlations in comparison to both narrative movies. Likely, increased between-subject correlations are driven by the joint activation of the AM network during RS and *Inscapes*. On the other hand, *Circus* and *Indiana Jones* likely engage the network to a lesser extent, thereby falling short of evoking coordinated activity which in turn reduces similarity between subjects.

The eSAD network was defined to comprise those brain regions that are part of the default mode network, but at the same time also involved in social or affective processing (Amft et al., 2015). Thus, the network is engaged in socio-affective processing including emotional processes, cognition, reward, introception, memory and theory of mind functions. Although not exclusively a “task-negative” network, the eSAD network is highly overlapping with the default mode network and generally presumed to be more active when participants can let their thoughts run free (Amft et al., 2015). RS showed higher within-subject correlations than *Circus* as well as higher between-subject correlations than *Indiana Jones*. In addition, *Inscapes*, which is arguably closer to RS than the other movies, also showed higher between-subject correlations as compared to *Circus* and *Indiana Jones*. Due to the default mode aspects of the eSAD network, it is perceivable that this network is more strongly engaged during RS and *Inscapes*. Thus, participants are more likely to express different connectivity patterns as compared to NV where the network is mostly unengaged. The movies *Circus* and *Indiana Jones* on the other hand might result in a less pronounced engagement of the network, thus failing to evoke similar patterns across participants.

The SM network is involved in retrieving semantic knowledge and is highly overlapping with the default mode network (Binder et al., 2009). The authors argue that task-unrelated thoughts are inherently semantic, because they require the manipulation of stored knowledge (Binder et al., 1999). Furthermore, semantic processing was reliably shown to be suppressed during demanding perceptual tasks (Binder et al., 2009), which is in accordance with our result pattern, showing increasingly complex stimuli to decrease within- and between-subject similarity (RS > *Inscapes* > *Circus* > *Indiana Jones*). We thus suggest that increasing complexity of the movie stimuli suppresses semantic processing and therefore leads to less engagement of the SM network. Presumably, due to a less pronounced engagement of the SM network during that *Circus* and *Indiana Jones*, participants show low between-subject correlations as well as low within-subject correlations.

The ToM network is fundamental for the understanding and contemplation of the behavior and intentions of others (Bzdok et al., 2012). Within- and between-subject correlations in the ToM network were generally higher during RS than during the NV conditions. We assume that movies evoke different interpretations of the intentions of the depicted characters and thus may have led to diverging connectivity profiles, in turn increasing differences between subjects. On the other hand, these differences seem to be unstable across sessions, thus decreasing within-subject correlations during NV.

4.5. Comparison with RS-derived networks

Identification accuracies in RS-derived networks confirm the assumption that identifiability is dependent on the network-stimulus combination (Table 2). Highest identification accuracy for RS was achieved in the Default, Salience and SomatoMotor networks, whereas highest accuracy for *Indiana Jones* was found in the Control, Dorsal Attention and Visual networks. For RS highest overall accuracy (91%) was achieved in the Default network, which is prominently active during RS (Long et al., 2008). However, the accuracies achieved in RS-derived networks were generally lower than those achieved in meta-analytic networks. Out of the 14 meta-analytic networks, eight yielded higher accuracies than the best performing RS derived network (Default).

In accordance with our results on meta-analytic networks, within- and between-subject correlations were also significantly altered during NV, in comparison to RS, in the RS-derived networks (Figs. 6 and 7). In the Control, Dorsal Attention, Salience, SomatoMotor and Visual networks NV resulted in more similar patterns between subjects. Only in the Default network, NV showed decreased between-subject correlations in comparison with RS.

Noticeably, differences in within-subject correlations between NV conditions and RS are less pronounced in the RS-derived networks than in the meta-analytic networks. This is further supported by the fact that RDMs of RS and NV stimuli tended to cluster together more often in RS derived networks (Fig. 2). Furthermore, within- and especially between-subject correlations are largely increased for the RS networks, resulting in reduced identifiability in RS derived networks compared to the meta-analytic networks. On the one hand, meta-analytic networks seem to be more sensitive to differences between NV stimuli and RS, likely because they best represent the core nodes of a given cognitive function. On the other hand, although within-subject correlations are increased in RS derived networks, the larger increase in between-subject similarity overshadows this effect and consequently leads to decreased identifiability. Taken together, the present results underline the viability of using specific meta-analytic networks for reliably identifying subjects' brain connectivity patterns under NV conditions.

4.6. Limitations

While the current study sheds new light onto individual differences in, and stability of, brain states elicited by movie watching, it comes with some limitations. Firstly, individual outliers might have biased identification accuracies, due to the small sample size. However, previous studies on RS and NV reported similar identification accuracies as those achieved in this study (Finn et al., 2015; Vanderwal et al., 2017). Nevertheless, future studies should be conducted on larger samples to confirm our results. Secondly, while we demonstrated enhanced individual difference and identifiability for certain stimulus-network combinations, our study did not include any phenotypes. Therefore this study is not suited to determine whether enhanced individual differences under NV can be used to more accurately predict phenotypes as compared to RS. Hence, future studies should investigate the interplay between increased identifiability and the accuracy of phenotype predictions. Thirdly, reliability of FC might at least partly be driven by structured noise such as vascular effects (Varikuti et al., 2017). Although we applied a number of denoising strategies, results might thus be confounded by non-neuronal signals. Additionally, only static FC was considered in the present study. Future studies investigating dynamic FC might shed more light on how individual variability in functional brain organization changes over the time course of a movie. A previous study on dynamic FC showed that NV improved test-retest reliability over RS, similar to the results in this study (Zhang et al., 2022). Finally, it was not assessed whether participants had seen any of the movie clips prior to participating in the study. Knowing the film beforehand could affect engagement of the participant and thereby modulate the effect of NV. In addition, previous studies have shown that expected stimuli can decrease the neuronal response (Alink et al., 2010; Koster-Hale and Saxe, 2013). Since the three sessions in our study were conducted within a week, participants are expected to be rather familiar with the movie content during the second and third session. Therefore, it is possible that the predictable content reduced the neuronal response and influenced our results. However, our results showed that connectivity patterns rather clustered according to stimulus than repetition, which suggests that the same movie stimulus can be used repeatedly to study FC of a subject across various time points. Similarly, a study by Wang et al. (2017) showed that movie fMRI increased reliability over RS across two sessions. The authors concluded that the effect that is achieved by increased engagement during movie watching, outweighs the impact of familiarity with a given movie. Taken together, these findings encourage the application of movie fMRI in clinical studies where it is necessary to monitor patients over a longer period of time.

5. Conclusions

NV has been suggested to show high potential for emphasizing individual differences, but effects have often been reported on a whole-brain level only. Our study extends the current knowledge by characterizing the influence of NV on FC in meta-analytically derived functional networks. We show that NV increases identifiability of individuals based on functional connectivity in certain networks. However, there is not one naturalistic stimulus that will enhance individual differences across the brain. Therefore it is crucial to select the appropriate stimulus and networks for the research question at hand.

Declaration of Competing Interest

The authors report no competing interests.

Credit authorship contribution statement

Jean-Philippe Kröll: Formal analysis, Software, Validation, Visualization, Writing – original draft. **Patrick Friedrich:** Writing – review & editing, Methodology. **Xuan Li:** Writing – review & editing. **Kaustubh R. Patil:** Writing – review & editing. **Lisa Mochalski:** Writing – review & editing. **Laura Waite:** Data curation, Writing – review & editing. **Xing Qian:** Writing – review & editing. **Michael WL Chee:** Writing – review & editing. **Juan Helen Zhou:** Conceptualization, Writing – review & editing, Funding acquisition. **Simon Eickhoff:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Susanne Weis:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition.

Data and code availability statement

The data that support the findings of this study are available upon request with data sharing agreement from the co-author Dr. Helen Zhou, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon request and with permission of the Institutional Review Board of the National University of Singapore.

Custom code used in this study is available upon request from the corresponding author JK.

Acknowledgements

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 945539 (HBP SGA3), and the Deutsche Forschungsgemeinschaft (491111487). We also acknowledge the funding support from Yong Loo Lin School of Medicine, National University of Singapore (J.H.Z), the Duke-NUS Medical School Signature Research Program Core Funding (J.H.Z.), and Ministry of Education, Singapore (MOE-T2EP40120–0007, J.H.Z), and Far East Organization (E-546–00–0398–01, MWLC.)

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120083.

References

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus predictability reduces responses in primary visual cortex. J. Neurosci. 30 (8), 2960–2966. doi:10.1523/JNEUROSCI.3730-10.2010.

Amft, M., Bzdok, D., Laird, A.R., Fox, P.T., Schilbach, L., Eickhoff, S.B., 2015. Definition and characterization of an extended social-affective default network. Brain Struct. Funct. 220 (2), 1031–1049. doi:10.1007/s00429-013-0698-0.

Avants, B., Tustison, N.J., Song, G., 2009. Advanced normalization tools: V1.0. Insight J. doi:10.54294/uvnhin, Published online July 29.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Rao, S.M., Cox, R.W., 1999. Conceptual processing during the conscious resting state: a functional MRI study. J. Cogn. Neurosci. 11 (1), 80–93. doi:10.1162/089892999563265.

Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 19 (12), 2767–2796. doi:10.1093/cercor/bhp055.

Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. 34 (4), 537–541. doi:10.1002/mrm.1910340409.

Buhle, J.T., Silvers, J.A., Wager, T.D., et al., 2014. Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. Cereb. Cortex 24 (11), 2981–2990. doi:10.1093/cercor/bht154.

Bzdok, D., Schilbach, L., Vogeley, K., et al., 2012. Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct. Funct. 217 (4), 783–796. doi:10.1007/s00429-012-0380-y.

Camilleri, J.A., Müller, V.I., Fox, P., et al., 2018. Definition and characterization of an extended multiple-demand network. Neuroimage 165, 138–147. doi:10.1016/j.neuroimage.2017.10.020.

Caspers, S., Zilles, K., Laird, A.R., Eickhoff, S.B., 2010. ALE meta-analysis of action observation and imitation in the human brain. Neuroimage 50 (3), 1148–1167. doi:10.1016/j.neuroimage.2009.12.112.

Christoff, K., Ream, J.M., Gabrieli, J.D.E., 2004. Neural basis of spontaneous thought processes. Cortex 40 (4–5), 623–630. doi:10.1016/S0010-9452(08)70158-8.

Cieslik, E.C., Mueller, V.I., Eickhoff, C.R., Langner, R., Eickhoff, S.B., 2015. Three key regions for supervisory attentional control: evidence from neuroimaging meta-analyses. Neurosci. Biobehav. Rev. 48, 22–34. doi:10.1016/j.neubiorev.2014.11.003.

Croxson, P.L., Forkel, S.J., Cerliani, L., Thiebaut de Schotten, M, 2018. Structural variability across the primate brain: a cross-species comparison. Cereb. Cortex 28 (11), 3829–3841. doi:10.1093/cercor/bhx244.

Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., et al., 2006. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. 103 (37), 13848–13853. doi:10.1073/pnas.0601417103.

van der Meer, J.N., Breakspear, M., Chang, L.J., Sonkusare, S., Cocchi, L, 2020. Movie viewing elicits rich and reliable brain state dynamics. Nat. Commun. 11 (1), 5004. doi:10.1038/s41467-020-18717-w.

Dmochowski, J.P., Bezdek, M.A., Abelson, B.P., Johnson, J.S., Schumacher, E.H., Parra, L.C., 2014. Audience preferences are predicted by temporal reliability of neural processing. Nat. Commun. 5 (1), 4567. doi:10.1038/ncomms5567.

Dosenbach, N.U.F., Fair, D.A., Miezin, F.M., et al., 2007. Distinct brain networks for adaptive and stable task control in humans. Proc. Natl. Acad. Sci. U. S. A. 104 (26), 11073–11078. doi:10.1073/pnas.0704320104.

Eickhoff, S.B., Bzdok, D., Laird, A.R., Kurth, F., Fox, P.T., 2012. Activation likelihood estimation meta-analysis revisited. Neuroimage 59 (3), 2349–2361. doi:10.1016/j.neuroimage.2011.09.017.

Eickhoff, S.B., Milham, M., Vanderwal, T., 2020. Towards clinical applications of movie fMRI. Neuroimage 217, 116860. doi:10.1016/j.neuroimage.2020.116860.

Esteban, O., Markiewicz, C.J., Blair, R.W., et al., 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods 16 (1), 111–116. doi:10.1038/s41592-018-0235-4.

Finn, E.S., Bandettini, P.A., 2020. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. Neuroscience doi:10.1101/2020.08.23.263723.

Finn, E.S., Shen, X., Scheinost, D., et al., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18 (11), 1664–1671. doi:10.1038/nn.4135.

Finn, E.S., Scheinost, D., Finn, D.M., Shen, X., Papademetris, X., Constable, R.T., 2017. Can brain state be manipulated to emphasize individual differences in functional connectivity? Neuroimage 160, 140–151. Accessed June 29, 2022 https://linkinghub.elsevier.com/retrieve/pii/S1053811917302872.

Finn, E.S., Corlett, P.R., Chen, G., Bandettini, P.A., Constable, R.T., 2018. Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. Nat. Commun. 9 (1), 2043. doi:10.1038/s41467-018-04387-2.

Finn, E.S., Glerean, E., Khojandi, A.Y., et al., 2020. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. Neuroimage 215, 116828. doi:10.1016/j.neuroimage.2020.116828.

Fox, M.D., Corbetta, M., Snyder, A.Z., Vincent, J.L., Raichle, M.E., 2006. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proc. Natl. Acad. Sci. U. S. A. 103 (26), 10046–10051. doi:10.1073/pnas.0604187103.

Geerligs, L., Rubinov, M., Cam-CAN, Henson, RN, 2015. State and trait components of functional connectivity: individual differences vary with mental state. J. Neurosci. 35 (41), 13949–13961. doi:10.1523/JNEUROSCI.1324-15.2015.

Golland, Y., Bentin, S., Gelbard, H., et al., 2007. Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. Cereb. Cortex 17 (4), 766–777. doi:10.1093/cercor/bhk030.

Gonzalez-Castillo, J., Kam, J.W.Y., Hoy, C.W., Bandettini, P.A., 2021. How to interpret resting-state fMRI: ask your participants. J. Neurosci. 41 (6), 1130–1141. doi:10.1523/JNEUROSCI.1786-20.2020.

Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. U. S. A. 100 (1), 253–258. doi:10.1073/pnas.0135058100.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48 (1), 63–72. doi:10.1016/j.neuroimage.2009.06.060.

Gruskin, D.C., Rosenberg, M.D., Holmes, A.J., 2020. Relationships between depressive symptoms and brain responses during emotional movie viewing emerge in adolescence. Neuroimage 216, 116217. doi:10.1016/j.neuroimage.2019.116217.

Halchenko, Y., Meyer, K., Poldrack, B., et al., 2021. DataLad: distributed system for joint management of code, data, and their relationship. JOSS 6 (63), 3262. doi:10.21105/joss.03262.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303 (5664), 1634–1640. doi:10.1126/science.1089506.

Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. Trends Cogn. Sci. (Regul. Ed.) 14 (1), 40–48. doi:10.1016/j.tics.2009.10.011.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841. doi:10.1006/nimg.2002.1132.

Kauppi, 2010. Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. Front. Neuroinform doi:10.3389/fninf.2010.00005, Published online.

Koster-Hale, J., Saxe, R., 2013. Theory of mind: a neural prediction problem. Neuron 79 (5), 836–848. doi:10.1016/j.neuron.2013.08.020.

Kriegeskorte, N., 2008. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci. doi:10.3389/neuro.06.004.2008, Published online.

Langner, R., Eickhoff, S.B., 2013. Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol. Bull. 139 (4), 870–900. doi:10.1037/a0030694.

Liu, X., Hairston, J., Schrier, M., Fan, J., 2011. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neurosci. Biobehav. Rev. 35 (5), 1219–1236. doi:10.1016/j.neubiorev.2010.12.012.

Long, X.Y., Zuo, X.N., Kiviniemi, V., et al., 2008. Default mode network as revealed with multiple methods for resting-state functional MRI analysis. J. Neurosci. Methods 171 (2), 349–355. doi:10.1016/j.jneumeth.2008.03.021.

McInnes, L., Healy, J., Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Published online 2018. doi:10.48550/ARXIV.1802.03426.

Mennes, M., Kelly, C., Zuo, X.N., et al., 2010. Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. Neuroimage 50 (4), 1690–1701. doi:10.1016/j.neuroimage.2010.01.002.

Mishra, S., Srinivasan, N., Tiwary, U.S., 2022. Dynamic functional connectivity of emotion processing in beta band with naturalistic emotion stimuli. Brain Sci. 12 (8), 1106. doi:10.3390/brainsci12081106.

Nastase, S.A., Gazzola, V., Hasson, U., Keysers, C., 2019. Measuring shared responses across subjects using intersubject correlation. Soc. Cogn. Affect. Neurosci. nsz037. doi:10.1093/scan/nsz037, Published online May 16.

Nguyen, M., Vanderwal, T., Hasson, U., 2019. Shared understanding of narratives is correlated with shared neural responses. Neuroimage 184, 161–170. doi:10.1016/j.neuroimage.2018.09.010.

Nummenmaa, L., Saarimäki, H., Glerean, E., et al., 2014. Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. Neuroimage 102, 498–509. doi:10.1016/j.neuroimage.2014.07.063.

Patriat, R., Molloy, E.K., Meier, T.B., et al., 2013. The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated. Neuroimage 78, 463–473. doi:10.1016/j.neuroimage.2013.04.013.

Power, J.D., Cohen, A.L., Nelson, S.M., et al., 2011. Functional network organization of the human brain. Neuron 72 (4), 665–678. doi:10.1016/j.neuron.2011.09.006.

Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F., 2015. ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage 112, 267–277. doi:10.1016/j.neuroimage.2015.02.064.

Rikandi, E., Pamilo, S., Mäntylä, T., et al., 2017. Precuneus functioning differentiates first-episode psychosis patients during the fantasy movie Alice in Wonderland. Psychol. Med. 47 (3), 495–506. doi:10.1017/S0033291716002609.

Rottschy, C., Langner, R., Dogan, I., et al., 2012. Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage 60 (1), 830–846. doi:10.1016/j.neuroimage.2011.11.050.

Saarimäki, H., 2021. Naturalistic stimuli in affective neuroimaging: a review. Front. Hum. Neurosci. 15, 675068. doi:10.3389/fnhum.2021.675068.

Sabatinelli, D., Fortune, E.E., Li, Q., et al., 2011. Emotional perception: meta-analyses of face and natural scene processing. Neuroimage 54 (3), 2524–2533. doi:10.1016/j.neuroimage.2010.10.011.

Salmi, J., Metwaly, M., Tohka, J., et al., 2020. ADHD desynchronizes brain activity during watching a distracted multi-talker conversation. Neuroimage 216, 116352. doi:10.1016/j.neuroimage.2019.116352.

Schaefer, A., Nils, F., Sanchez, X., Philippot, P., 2010. Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. Cogn. Emot. 24 (7), 1153–1172. doi:10.1080/02699930903274322.

Schaefer, A., Kong, R., Gordon, E.M., et al., 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb. Cortex 28 (9), 3095–3114. doi:10.1093/cercor/bhx179.

Schmälzle, R., Häcker, F.E.K., Honey, C.J., Hasson, U., 2015. Engaged listeners: shared neural processing of powerful political speeches. Soc. Cogn. Affect. Neurosci. 10 (8), 1137–1143. doi:10.1093/scan/nsu168.

Shehzad, Z., Kelly, A.M.C., Reiss, P.T., et al., 2009. The Resting Brain: unconstrained yet Reliable. Cereb. Cortex 19 (10), 2209–2229. doi:10.1093/cercor/bhn256.

Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403–415. doi:10.1016/j.neuroimage.2013.05.081.

Smith, S.M., Fox, P.T., Miller, K.L., et al., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. U. S. A. 106 (31), 13040–13045. doi:10.1073/pnas.0905267106.

Sporns, O., Betzel, R.F., 2016. Modular Brain Networks. Annu. Rev. Psychol. 67 (1), 613–640. doi:10.1146/annurev-psych-122414-033634.

Spreng, R.N., Mar, R.A., Kim, A.S.N., 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J. Cogn. Neurosci. 21 (3), 489–510. doi:10.1162/jocn.2008.21029.

Thomas Yeo, B.T., Krienen, F.M., Sepulcre, J., et al., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106 (3), 1125–1165. doi:10.1152/jn.00338.2011.

Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L.C., Castellanos, F.X., 2015. Inscapes : a movie paradigm to improve compliance in functional magnetic resonance imaging. Neuroimage 122, 222–232. doi:10.1016/j.neuroimage.2015.07.069.

Vanderwal, T., Eilbott, J., Finn, E.S., Craddock, R.C., Turnbull, A., Castellanos, F.X., 2017. Individual differences in functional connectivity during naturalistic viewing conditions. Neuroimage 157, 521–530. doi:10.1016/j.neuroimage.2017.06.027.

Vanderwal, T., Eilbott, J., Castellanos, F.X., 2019. Movies in the magnet: naturalistic paradigms in developmental functional neuroimaging. Dev. Cogn. Neurosci. 36, 100600. doi:10.1016/j.dcn.2018.10.004.

Varikuti, D.P., Hoffstaedter, F., Genon, S., Schwender, H., Reid, A.T., Eickhoff, S.B., 2017. Resting-state test–retest reliability of a priori defined canonical networks over different preprocessing steps. Brain Struct. Funct. 222 (3), 1447–1468. doi:10.1007/s00429-016-1286-x.

Wang, J., Ren, Y., Hu, X., et al., 2017. Test-retest reliability of functional connectivity networks during naturalistic fMRI paradigms: test-retest reliability of naturalistic fMRI. Hum. Brain Mapp. 38 (4), 2226–2241. doi:10.1002/hbm.23517.

Witt, S.T., Laird, A.R., Meyerand, M.E., 2008. Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. Neuroimage 42 (1), 343–356. doi:10.1016/j.neuroimage.2008.04.025.

Zhang, X., Liu, J., Yang, Y., et al., 2022. Test–retest reliability of dynamic functional connectivity in naturalistic paradigm functional magnetic resonance imaging. Hum. Brain Mapp. 43 (4), 1463–1476. doi:10.1002/hbm.25736.