Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging

Emily S. Finn [a],[*], Enrico Glerean [b], Arman Y. Khojandi [a], Dylan Nielson [c], Peter J. Molfese [a], Daniel A. Handwerker [a], Peter A. Bandettini [a]

[a] Section on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, MD, USA
[b] Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland
[c] Mood Brain & Development Unit, National Institute of Mental Health, Bethesda, MD, USA

ARTICLE INFO

ABSTRACT

Two ongoing movements in human cognitive neuroscience have researchers shifting focus from group-level inferences to characterizing single subjects, and complementing tightly controlled tasks with rich, dynamic paradigms such as movies and stories. Yet relatively little work combines these two, perhaps because traditional analysis approaches for naturalistic imaging data are geared toward detecting shared responses rather than between-subject variability. Here, we review recent work using naturalistic stimuli to study individual differences, and advance a framework for detecting structure in idiosyncratic patterns of brain activity, or "idiosynchrony". Specifically, we outline the emerging technique of inter-subject representational similarity analysis (IS-RSA), including its theoretical motivation and an empirical demonstration of how it recovers brain-behavior relationships during movie watching using data from the Human Connectome Project. We also consider how stimulus choice may affect the individual signal and discuss areas for future research. We argue that naturalistic neuroimaging paradigms have the potential to reveal meaningful individual differences above and beyond those observed during traditional tasks or at rest.

1. Introduction

At present, there are two exciting movements afoot in cognitive neuroscience. First, the field is shifting focus from the group to the individual: instead of averaging data across a population, studies are isolating brain function in single subjects and determining how it relates to behavioral phenotypes (Bartolomeo et al., 2017; Dubois and Adolphs, 2016; Seghier and Price, 2018). Second, researchers are embracing the complexity of so-called "naturalistic" stimuli—e.g., film clips, spoken narratives—as experimental paradigms, to complement and extend the tightly controlled, parametric tasks that form the pillars of classical psychology work (Sonkusare et al., 2019).

While each of these movements has brought promising discoveries on its own, studies at the intersection—i.e., that use naturalistic stimuli to study individual differences—are relatively few and far between. Why might this be? Naturalistic stimuli evoke patterns of brain activity that are, by and large, highly consistent across subjects (Hasson et al., 2004, 2010), and typically, naturalistic imaging data are collected and analyzed in ways geared toward detecting similarities between subjects rather

than differences (Nastase et al., 2019). Adapting naturalistic tasks to study individual differences and brain-behavior relationships raises several challenges, demanding new approaches to both experimental design and data analysis. The potential payoff, however, is high. Naturalistic tasks offer a "happy medium" between the extremes of highly controlled cognitive tasks, which often lack ecological validity, and resting-state acquisitions, which are entirely unconstrained, making them vulnerable to confounds and difficult to interpret (Vanderwal et al., 2019). They allow experimenters to probe intermingled signals throughout the hierarchy of neural systems—from low-level sensory processing up to social cognition—using data from a single acquisition, and while the brain is engaged in activities more similar to everyday life. Thus, they may offer more "bang for the buck" than either rest or traditional tasks for characterizing spatiotemporal patterns of brain activity in individual subjects.

Here, our goals are threefold. In the first section, we briefly describe the inter-subject correlation (ISC) family of approaches to analyzing naturalistic imaging data, which exploit the time-locked nature of the task across subjects to isolate brain activity driven by the stimulus

(Hasson et al., 2004; Nastase et al., 2019). Also in this section, we review existing literature using ISC to study individual differences. (For comprehensive reviews on naturalistic tasks and inter-subject correlation more generally, we refer the reader to Sonkusare et al. (2019) and Nastase et al. (2019), respectively. Our focus here is specifically on how these intersect with the study of individual differences.) In the second section, we advance a framework for studying individual differences in spatio-temporal patterns of brain activity during naturalistic stimulation, a phenomenon we call "idiosynchrony". We focus specifically on the emerging technique of inter-subject representational similarity analysis (IS-RSA), which adapts ISC to highlight stimulus-driven responses that are idiosyncratic rather than shared. We present the theory behind IS-RSA, then apply it to data from the Human Connectome Project to demonstrate how brain responses during movie viewing share structure with behavioral traits, including working memory and personality. In the final section, we consider directions for future work, including how we might choose stimuli to amplify behaviorally relevant individual signals.

2. Inter-subject correlation: stimulus-locked, fewer assumptions

Because naturalistic paradigms sit somewhere between resting state and traditional tasks, researchers have at their disposal a variety of analysis approaches. The presence of a stimulus permits classic techniques for detecting activation to specific features of the stimulus, such as the general linear model (GLM) and other regression techniques. At the same time, thanks to their continuous, non-parametric nature, these paradigms also lend themselves to approaches developed primarily for rest—for example, functional connectivity, which considers coactivations between regions rather than magnitude of activation in any single region.

We place these approaches in a landscape with two major axes: how time-locked to the stimulus they are, and how many assumptions they require. Functional connectivity (FC)—especially static FC, which collapses across the whole time series—demands relatively few assumptions, but is impossible to map back onto the stimulus, and therefore fails to separate stimulus-driven from stimulus-independent sources of neural activity. Dynamic FC gets closer to a time-resolved signal, but in most cases the need for windows of a minimum length makes it difficult to map connectivity fluctuations to specific events within the stimulus, and resulting signals still contain a mixture of stimulus-related and stimulus-unrelated variation within each subject.

GLM-based analyses do capitalize on the "ground truth" of a stimulus with known timing, but in doing so make two critical assumptions: one, that the experimenter knows which features of the stimulus are important for driving brain activity, and two, that they have modeled these features accurately. Non-optimal assumptions at either step will impair sensitivity.

While both families of approaches can and have been applied to naturalistic data (regression-based activation: (Bartels and Zeki, 2004; Lahnakoski et al., 2012a; Lahnakoski et al., 2012b; Russ and Leopold, 2015); functional connectivity: (Betti et al., 2013; Geerligs et al., 2015; Guo et al., 2016; Vanderwal et al., 2017), to name a few), neither was designed specifically for these paradigms. The inter-subject correlation (ISC) family of approaches (Hasson et al., 2004; Nastase et al., 2019; Simony et al., 2016), which was developed specifically for naturalistic paradigms, maximize sensitivity to stimulus-driven activity with fewer assumptions. These approaches use one subject's brain activity as a model for a second subject's brain activity, reasoning that as long as two subjects receive the same input at the same time, any shared variance must be due to stimulus processing.

In brief, ISC is defined as the Pearson correlation of the activity timecourse in a spatial location (i.e., voxel, parcel) across different subjects. It is typically computed using either a leave-one-out framework, in which one subject's timecourse is correlated with the average of all other subjects, or a pairwise framework, in which correlation is performed between every possible pair of subjects. For our purposes, this latter approach is more relevant, since it preserves information about which specific subject pairs show the highest and lowest correlations, which can later be related to behavioral scores.

Unlike FC, ISC isolates stimulus-driven signal and can be interpreted with respect to events in the stimulus itself. And unlike GLM-based approaches, even if the experimenter doesn't know—or can't model—the most important features of the stimulus, as long as there is some consistent signal across subjects, ISC will recover it (Pajula et al., 2012), in some cases with more sensitivity than deconvolution/GLM-based analyses (Hejnar et al., 2007). Thus, ISC is a powerful, data-driven technique for detecting shared responses, anticipated or otherwise.

2.1. Adapting ISC to an individual-differences framework

The classic formulation of ISC assumes that the signal observed at each voxel x and timepoint t reflects a mixture of three components. Loosely following the notation of Nastase et al. (2019), these components are:

c, a stimulus-evoked response that is consistent across subjects;
id, a stimulus-evoked response that is idiosyncratic to each subject;
ε, a noise component (which may reflect both neural activity that is unrelated to the stimulus, i.e., mind-wandering, as well as noise from non-neural physiological and scanner sources).

Thus, for a given subject i, the response in voxel x at timepoint t can be described as follows:

$$x_i(t) = c(t) + id_i(t) + \varepsilon_i(t)$$

In this formulation, it is not possible to distinguish id from ε, unless subjects are exposed to the same stimulus multiple times. Even then, results must be interpreted with caution, because repetition can change how a stimulus is processed. Another approach is to anchor id to some other known information about each subject, like a trait score (b), and search for structure in these responses across subjects:

$$x_i(t) = c(t) + \mathbf{b_i}(id(t)) + \varepsilon_i(t)$$

$$x_j(t) = c(t) + \mathbf{b_j}(id(t)) + \varepsilon_j(t)$$

…

$$x_n(t) = c(t) + \mathbf{b_n}(id(t)) + \varepsilon_n(t)$$

Notice that now the id term has lost its subject subscript, since we are assuming that there is some canonical response associated with a given trait, and each subject's trait score acts as a sort of bias term governing to what degree a subject expresses that response.[1]

This framework lets us distinguish stimulus-related responses in individual subjects from stimulus-unrelated noise, and does so in a way that facilitates interpretation, since we are linking idiosyncratic responses to a known behavioral measure. Generally, we might predict that the influence of individual differences on $x(t)$—that is, the ratio of id to c—grows as one moves up the cortical processing hierarchy, such that the shared, behavior-independent signal dominates in unimodal cortex, while the idiosyncratic, behavior-dependent signal becomes stronger in higher-order association cortex. This may explain why ISC is traditionally high in primary visual and auditory cortex, and drops off (but doesn't necessarily disappear completely) in areas of prefrontal cortex, for example: it is not the case that these regions are not responding to the stimulus, but rather that they respond with different spatiotemporal signatures across

---

[1] In this formulation, we assume that b is static within individuals, since it represents a trait-level measure. However, this formulation could be extended to support an individual behavioral measure that changes over the course of the stimulus—for example, attentional state or affective experience—by postulating that b is also a function of time (i.e., $b_i(t)$).

subjects (Chang et al., 2018; Hasson et al., 2010). Some of this variance might be explained by trait-level phenotypes.

3. Existing work on ISC & individual differences

Despite the traditional focus on shared responses, in recent years, researchers have begun to investigate how ISC varies with both state-level factors related to either the stimulus or experimental instructions, and—more relevant for our purposes—with trait-like factors intrinsic to the subjects themselves.

Several studies have shown that ISC is sensitive to features of the stimulus (Dmochowski et al., 2014; Hasson et al., 2010; Nummenmaa et al., 2014; Schmälzle et al., 2015). Other studies have kept the stimulus constant and used priming to show that ISC is sensitive to explicit manipulations of attention or prior beliefs about a stimulus (Cooper et al., 2011; Lahnakoski et al., 2014; Yeshurun et al., 2017). While informative, these studies do not examine why different individuals often spontaneously—i.e., with no explicit priming—show different neural or behavioral responses to the same stimulus.

A handful of studies have reported spontaneous individual differences in time-locked activity to a stimulus that relate to behavioral appraisal of the stimulus. Using emotional clips, comedy videos, and moral dilemma scenarios, respectively, Nummenmaa et al. (2012), Jaaskelainen et al. (2016) and Tei et al. (2019) found that ISC in certain brain regions correlated with similarity of post-hoc behavioral ratings (of dynamic valence/arousal, humor, and moral conflict) across pairs of subjects. Using an abstract video of animated shapes and an audio story, respectively, Nguyen et al. (2019) and Saalasti et al. (2019) found that subject pairs with higher ISC in certain cortical regions during the stimulus ultimately had more similar interpretations of the stimulus. These studies provide a fascinating window into how individual differences in brain activity during stimulus exposure relate to subsequent differences in reactions and interpretations. But because both brain responses and behavior are tied to the specific stimulus at hand, the extent to which these state-like differences reflect trait-like predispositions—i.e., intrinsic individual characteristics—remains unclear.

Recent studies take this additional step and link idiosyncratic responses during naturalistic stimulation to stable, trait-like factors. Finn et al. (2018) report that individuals with higher trait paranoia show stronger ISC in cortical regions involved in social processing during an ambiguous social narrative. Bacha-Trams et al. (2018) report that individual differences in cognitive style (i.e., holistic versus analytical thinking) relate to strength of ISC in several cortical regions during a drama movie. In a sample of college-age males, Chen et al. (2020) report that variations in sociosexual desire and self-control preferences are linked to variations in activity in several higher-order brain networks, during erotic (but not neutral) films. In a sample of children and adolescents spanning ages 7–21 years, Gruskin et al. (2020) report that patterns and severity of depression symptoms are linked to brain responses during an emotionally evocative animated clip, but only in adolescents, suggesting that these idiosyncratic responses emerge over the course of development. Finally, in a study of individuals with dyslexia, ISC calculated from MEG envelope time series correlated with similarity in phonological processing, technical reading, and working memory (Thiede et al., 2019). Encouragingly, these studies show that there is meaningful—i.e., behaviorally relevant—structure in idiosyncratic responses to naturalistic stimuli.

Other studies have addressed variability by dichotomizing subjects into diagnostic groups. Several studies have reported differences in ISC during naturalistic paradigms between healthy controls and populations with mental illnesses and disorders—most commonly, autism (Bolton et al., 2018; Byrge et al., 2015; Hasson et al., 2009; Salmi et al., 2013), but also depression (Guo et al., 2015) and first-episode psychosis (Mäntylä et al., 2018; Yang et al., 2020). In general, these studies report reduced cross-subject synchrony in the patient group, an effect that sometimes scales with symptom severity (Guo et al., 2015; Salmi et al., 2013).

Group contrasts between patients and controls can be useful to assess population-level differences in broad strokes. But rather than distinct canonical responses for each group, patients are typically characterized by increased heterogeneity of responses (Bolton et al., 2018). This suggests that a better framework would approach the question as an individual-differences problem. Indeed, Byrge et al.'s post-hoc analyses revealed that their ISC effect was driven entirely by five individuals from the ASD group with particularly idiosyncratic responses that could not be easily explained by other factors (i.e., data quality, symptom severity); when these five individuals were removed, the ASD and neurotypical groups were indistinguishable (Byrge et al., 2015). Interestingly, Hasson et al. (2009) reported that while the responses of subjects with autism were more variable, these idiosyncratic responses were reliable within single individuals across repeated presentations of the same stimulus, suggesting a trait-like component.
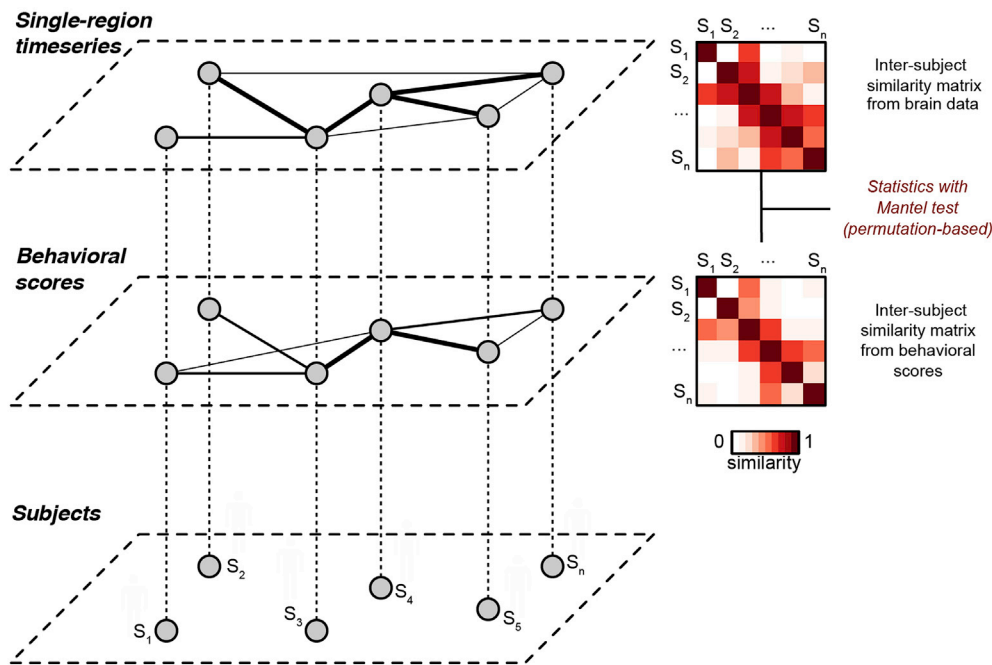
Rather than a dichotomy between health and disease, most mental illnesses are likely better conceptualized as the extreme end of a phenotypic spectrum (Cuthbert and Insel, 2013; Insel et al., 2010). Instead of stratifying subjects into patients and controls, we can improve sensitivity by using continuous measures—task performance, symptom severity, genetic load—as our independent variables (Finn and Constable, 2016). To then take advantage of the richness of information embedded in these continuous spaces, we need analysis frameworks that can appropriately handle the intricate interdependencies of dyads rather than individuals. Recent work has proposed using multilevel modeling and other expanded statistical formulations, some of which offer the ability to include subject-level covariates into ISC analyses (Chen et al., 2017; Chen et al., 2020). In the next section, we outline another promising approach, used by several of the studies cited above, that is flexible, intuitive and can be applied to detect relationships between any type of brain and behavioral data.

4. Inter-subject representational similarity analysis: Theory

By definition, inter-subject correlation cannot be calculated for a single subject. How, then, can we relate ISC, which operates at the level of subject pairs, to traits and behaviors, which operate at the level of single subjects? We can triangulate between these two levels of measurement using a framework we call "idiosynchrony", which leverages each subject's unique pattern of synchrony with other subjects to reveal a covariance structure that also reflects a known behavioral measure—for example self-report questionnaires, demographics, task performance, clinical assessments, or genotypes, among others. The intuition is that individuals who are more similar in behavior should also be more similar in their neural responses.

We operationalize this by computing two subject-by-subject distance matrices: one for the brain data (using, for example, ISC), and one for the behavioral data. We can then compare the geometry of these two matrices by correlating them, a procedure known as distance correlation or representational similarity analysis (Mantel, 1967; Kriegeskorte et al., 2008). We and others call this inter-subject representational similarity analysis (Chen et al., 2020; van Baar et al., 2019). The advantage of comparing similarity matrices over a typical first-level analysis is that instead of directly linking two physically different quantities like brain data and behavior, we use a second-order isomorphism to compare the geometry of brain data with the geometry of behavioral data (Fig. 1) (Kriegeskorte and Kievit, 2013).

This sounds straightforward enough. But one critical question is, how do we measure behavioral similarity? In choosing a distance metric, particularly when our behavior is one-dimensional (e.g., age (Moraczewski et al., 2018; Richardson et al., 2018), a trait score (Finn et al., 2018), accuracy on a cognitive task), we imbue our analysis with some fundamental assumptions about the structure of the brain-behavior representational similarity that affect the ultimate results and how we interpret them. To get a feel for some potential structures, imagine

arranging the rows and columns of the ISC matrix such that subjects are ordered by their behavioral score. What would we expect the resulting matrix to look like?

If we use Euclidean distance or another relative distance metric, we implicitly assume that subjects with closer scores should be more similar to one another, regardless of where they fall on the scale. In other words, for a behavior that is measured on a scale from 0 to 100, a pair of subjects scoring 0 and 1 should be just as similar as a pair of subjects scoring 99 and 100 (since in both cases, the Euclidean distance is 1). We call this the Nearest Neighbors (NN) model, since it assumes that a subject should always look most similar to his or her immediate neighbors, regardless of their absolute position on the scale (Fig. 2a).

The NN model may be appropriate for certain behaviors, but we could imagine an equally if not more plausible scenario: that similarity between subjects increases or decreases as one moves up or down the scale, in an absolute rather than relative sense. For example, perhaps high-scoring subjects are more similar to other high scorers, while low-scoring sub-jects are less similar both to high scorers and other low scorers. In other words, brain responses cluster together for subjects at one end of the behavioral spectrum, while variability increases as one moves toward the opposite end of the spectrum. We call this the *Anna Karenina* (or *AnnaK*) model, after the famous opening line of Leo Tolstoy's novel, which reads "All happy families are alike; each unhappy family is unhappy in its own way" (or, in this context, "all high [low] scorers are alike; each low [high] scorer is different in his or her own way"). In this case, Euclidean distance would not be the most appropriate choice. Instead, we would want to model similarity using a metric that reflects absolute position on the scale—for example, mean: $(i + j)/2$, minimum: $\min(i, j)$, or the product of

the mean and minimum (Fig. 2b–d).[2]

In the case of traits and behaviors that consist of a vector of responses per subject (e.g., self-report questionnaires) rather than a single scalar number (e.g., age), we have the option to calculate item-wise similarity using any number of potential distance metrics—for example, correla-tion, Euclidean distance, cosine distance and many others. In this case, we are assuming that it is the pattern of individual responses, and not the composite score, that should determine the inter-subject similarity structure. This approach is likely best suited to assessments consisting of unique items that are not interchangeable—personality questionnaires, for example.[3] As a general heuristic, quantitative assessments—those with clear "better" and "worse" ends of the scale—are likely more suited to distance models based on single composite scores, while qualitative scales could be suited to either composite or item-wise models. We can also imagine scenarios where both models are theoretically appropriate, and each might capture different effects. For example, questionnaires assessing symptom type and severity might show some effects that scale with overall score (Anna Karenina model), and others that scale with item-wise similarity (nearest neighbors)—and each of these effects might be present in different brain regions.

5. Inter-subject RSA: application

To investigate if and how brain similarity reflects behavioral simi-larity during naturalistic stimulation, and more specifically, how the choice of distance model affects results, we applied inter-subject RSA to an empirical dataset from the Human Connectome Project (HCP) (Van Essen et al., 2013). Subjects (all healthy volunteers aged 22–35 years) engaged in a movie-watching paradigm during high-resolution (voxel size = 1.6 mm^3, TR = 1s) functional MRI scanning at 7 T. The sample used here (n = 184) reflects all available data for this paradigm. This

---

[2] While in theory these three potential formulations of the AnnaK model have somewhat different interpretations, in practice, they tend to yield highly correlated distance matrices, making it difficult to select the most appropriate formulation via model comparison. In the empirical section of this paper, we elected to test only the first formulation (Fig. 2b, mean(i,j)), because it is the simplest, and because it has the advantage that any inverse relationships (i.e., lower behavioral scores associated with higher ISC) will be captured using the same model (see also Footnote 4). Future work could attempt to disentangle these and other formulations using advanced model comparison techniques and/or simulations.

[3] Since most cognitive performance tests consist of more than one item, we could in theory calculate item-wise distance on these tests as well. But because trials in these tests are generally interchangeable, it is more straightforward and interpretable to consider the similarity of two subjects' composite score rather than the similarity of their performance on individual trials (unless one has a specific hypothesis about learning rates, attention fluctuations, or other effects with a dynamic component).
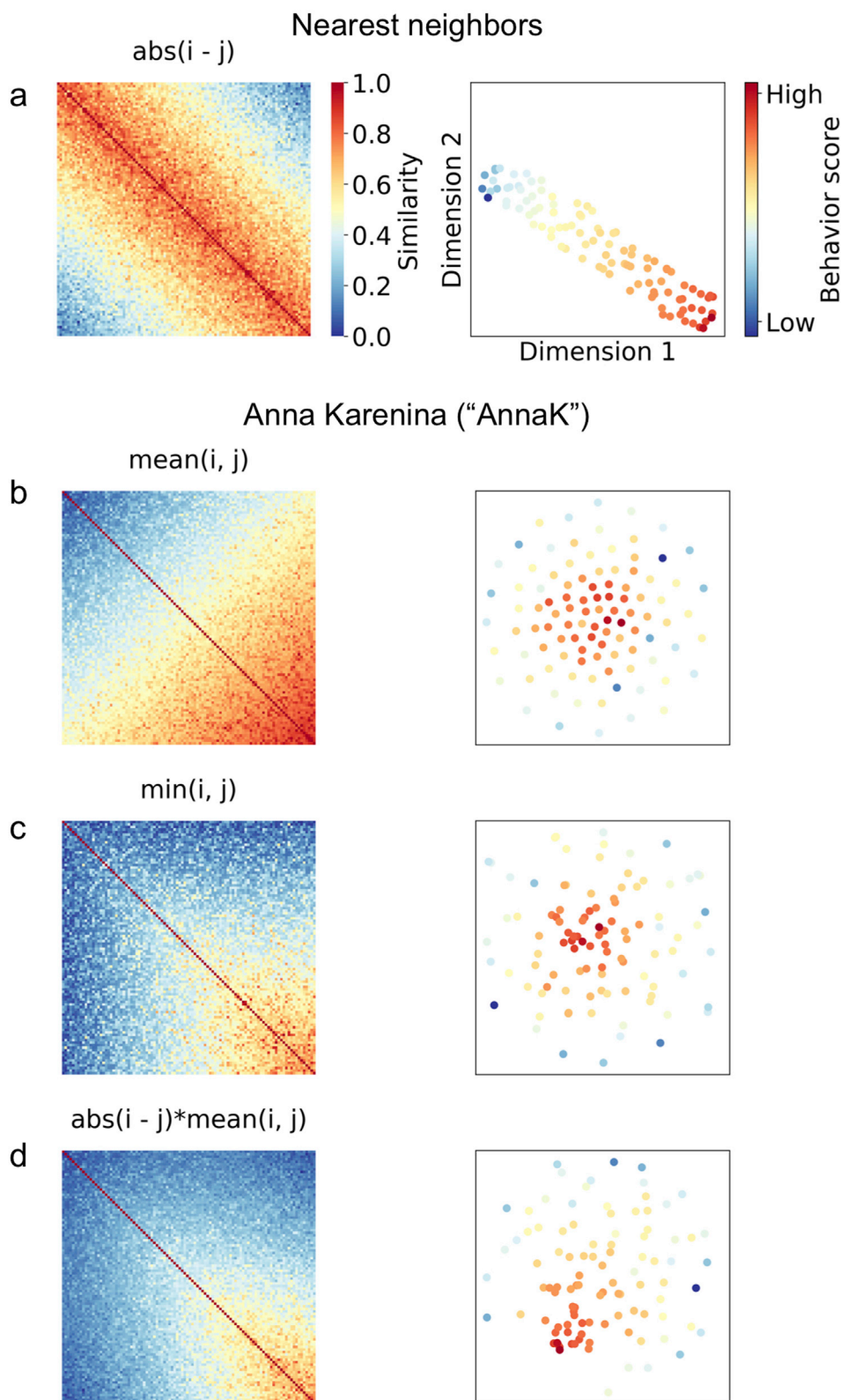
Fig. 2. Simulated potential structures for brain-behavior representational similarity matrices.

For each row a-d, the left panel depicts a simulated pairwise brain similarity matrix in which subjects are ordered along both rows i and columns j by their behavioral score (from low to high), and each cell $\{i, j\}$ reflects the correlation between subjects i and j of the timeseries of a given brain region (pairwise inter-subject correlation). The right panel depicts a two-dimensional embedding of the corresponding distance matrix (i.e., 1 – similarity matrix) using t-SNE (t-Distributed Stochastic Neighbor Embedding), in which each dot represents a subject, and subjects are colored according to their behavioral score. Under the t-SNE solution, similar observations (in this case, subjects) appear nearby, while dissimilar observations appear further away.

dataset contains many sets of twins (both mono- and dizygotic) and siblings (the 184 subjects come from n = 93 unique families). Details of data acquisition and basic preprocessing are published elsewhere (Glasser et al., 2013a; Van Essen et al., 2012; Vu et al., 2017). Each subject watched four 15-min movie runs; data from the first run (MOVIE1_7T_AP) are used here. This run comprised five video clips presented in a fixed order. Four clips were from independent films and documentaries, all with some degree of social and affective content, and one was a montage of brief (1.5s) moving scenes depicting people and places. All fMRI analyses began with the FIX-denoised data, which includes standard preprocessing (motion correction, distortion correction, highpass filtering, and nonlinear alignment to MNI template space (Glasser et al.,

2013b)) plus regression of 24 framewise motion estimates (six rigid-body motion parameters and their derivatives and the squares of those 12) and regression of confound timeseries identified via independent components analysis (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014).

Each subject also completed a battery of self-report and behavioral measures outside the scanner (Barch et al., 2013). We focused on two trait-level measures from the cognitive and emotional domains, respectively: working memory (as measured by a list-sorting task) and personality (as measured by the NEO Five-Factor Inventory). We chose these two traits because we hypothesized that the structure of brain-behavior similarity would manifest differently for each one, and therefore that each trait would be best modeled by a different distance function in IS-RSA, as detailed below. The primary outcome measure from the list-sorting task is a single scalar measure of working memory span, or how many items can be accurately stored and manipulated in working memory at one time (higher scores indicate higher performance). The NEO-Five Factor Inventory consists of 60 items, and subsets of these are summed to yield scores for five dimensions: agreeableness, extraversion, conscientiousness, neuroticism, and openness.

For both working memory and all five personality traits, we tested two models: 1) a nearest-neighbor model based on overall score (abs(i-j)), to test the prediction that people that score more similarly on a behavioral measure look more similar, regardless of whether they score high or low; and 2) an AnnaK model based on the mean(i,j) formulation, to test the prediction that people who score high on a given trait share similar patterns of brain activity, while people that score low show more variability (or vice versa).[4] Within the personality domain, we tested a third model: nearest-neighbor based on itemwise responses to the personality questionnaire, to test the prediction that people who fill out the questionnaire in more similar ways, regardless of their summary trait scores, would show more similar brain activity. We hypothesized that working memory would be best captured by the AnnaK model in most brain regions. We did not have strong hypotheses about the best model for personality, as any of the above scenarios were plausible *a priori*.

Because we might expect both behavioral phenotypes and brain activity to be more similar between siblings—and especially twins—due to any number of genetic and environmental influences, we avoided performing IS-RSA on related individuals. We split the dataset into two cohorts of unrelated subjects (n = 93 and n = 89, respectively), which had the benefit of giving us a natural replication sample to help guard against false positives. We performed all analyses on each cohort separately, and leveraged this test-retest framework to correct for multiple comparisons, described further below.

For each subject, we extracted activity timecourses from every node in a 268-node functional parcellation (Shen atlas; Shen et al. (2013)) by averaging signal across all voxels for each volume. (Because activity is expected to be smooth across neighboring voxels, this step reduces the dimensionality of the data, thus avoiding the computational cost of a voxelwise analysis.) For each node *n* and each subject pair {*i, j*} within a cohort, brain similarity was calculated as the Pearson correlation (ISC) of activity timecourses across the whole run. Behavioral similarity was calculated according to either an NN model (*i* – *j*; Fig. 2a) or an AnnaK model (mean(*i, j*); Fig. 2b). Representational similarity was assessed by calculating Spearman's rank correlation between the vectorized upper

triangles of the brain and behavioral similarity matrices.

For each node, the significance of the brain-behavior representational similarity was assessed non-parametrically using a Mantel test (Mantel, 1967), in which subject labels (i.e., rows and columns) are randomly permuted for one of the two similarity matrices a large number of times (in this case, 10,000) and the correlation between the two matrices is recalculated to form a null distribution of surrogate correlation values. The observed correlation coefficient is then compared to this null distribution to obtain a p-value for each node for each cohort. These p-values were then corrected for multiple comparisons using two parallel approaches, both of which leveraged the two-cohort framework: (1) Bonferroni-style, in which a corrected p-threshold was calculated to reflect the probability of obtaining a p < 0.05 result in both Cohort 1 and Cohort 2 for any given node (in this case, p = 0.0136[5]); and (2) familywise error control, in which we used permutation testing to compute a null distribution for how many nodes across the whole brain would be expected to survive an initial p-threshold (in this case 0.05) in both cohorts, and compared the observed number of nodes to this null distribution. Note that the second approach was not designed to test significance of any particular node, but rather to assess and compare overall detection power across the brain for our two models.

For working memory, as we hypothesized, brain-behavior representational similarity was best captured by the AnnaK model (Fig. 3, bottom row). High-scoring subjects were similar to other high scorers across much of the brain, while low-scoring subjects had more idiosyncratic responses—in other words, they were less similar to both the high scorers and other low scorers. The AnnaK model fit the data better than the nearest-neighbors model in that it yielded stronger effect sizes (compare distributions of RSA values between the two scatter plots in Fig. 3) that were more replicable across the two subject cohorts ($r_{cohort\ 1,\ cohort\ 2}$ = 0.60 for the AnnaK model versus 0.18 for the NN model). The AnnaK model captured significant relationships between brain and behavioral similarity in 52 nodes (familywise p < 0.0001) across parietal, temporal, occipital and cerebellar cortex (of these, 16 survived Bonferroni-style correction). This pattern is consistent with previous findings that working memory and other high-level cognitive abilities depend on distributed networks of regions mostly in association cortex, and that individual differences in the functional organization of these regions are to some degree intrinsic, i.e., they can be observed even in the absence of task states designed specifically to probe these processes (Cole et al., 2012; Finn et al., 2015; Hampson et al., 2006; Song et al., 2008). In contrast, the NN model captured significant brain-behavior relationships in only 2 nodes (familywise p = 0.14; neither survived Bonferroni correction).

For personality, results were more complicated. Using trait summary scores, the nearest-neighbor model based on trait summary scores did not yield any significant nodes across any of the five dimensions (Fig. 4b, top scatter plots), and the AnnaK model yielded only two significant nodes at an uncorrected threshold across all five dimensions (Fig. 4b, bottom scatter plots). However, the nearest-neighbor model based on item-wise responses captured 16 nodes across the brain (familywise p < 0.0001; though only one individual node survived Bonferroni-style correction), including fusiform, right inferior frontal gyrus, and several nodes in the

---

[4] One advantage of the mean(i,j) formulation of the AnnaK model is that the same model can detect effects in both directions, based on the sign of the resulting r-value between the brain and behavioral similarity matrices. If high scorers are alike and low scorers different, the resulting r-value would be positive; if low scorers are alike and high scorers different, it would be negative. (Note that the other two formulations—min(i,j) and abs(i-j)*mean(i,j)—would also be expected to yield negative r-values in a case where low scorers were alike, but because these models are not symmetric about the counterdiagonal [top-right to bottom-left], they would be less precise at detecting inverse relationships, and would require two different models to accurately detect inverse relationships.).

[5] If each node is tested at an alpha of α = 0.05, the probability of a false positive for the same node in both cohorts is 0.05^2, or 0.0025. If there are 268 tests, the probability that any one of them is significant in both cohorts is 0.0025*268, or 0.67. We can set this outcome probability to a desired alpha level, in this case 0.05: $268*α^2$ = 0.05, which gives 0.0136. Note that this approach assumes that Cohorts 1 and 2 are independent, which is not strictly true since there are twins and non-twin siblings split across cohorts. However, in the absence of clear data as to how genetic relatedness influences brain activity during naturalistic stimulation, we believe this approach, akin to using discovery and replication samples, is a step toward ensuring statistical rigor and generalizability (and is still preferable to including pairs of siblings within the same cohort, which would violate independence assumptions in the first-level analyses).

Working Memory

NN

Cohort 1

Cohort 2

AnnaK

Cohort 1

Cohort 2
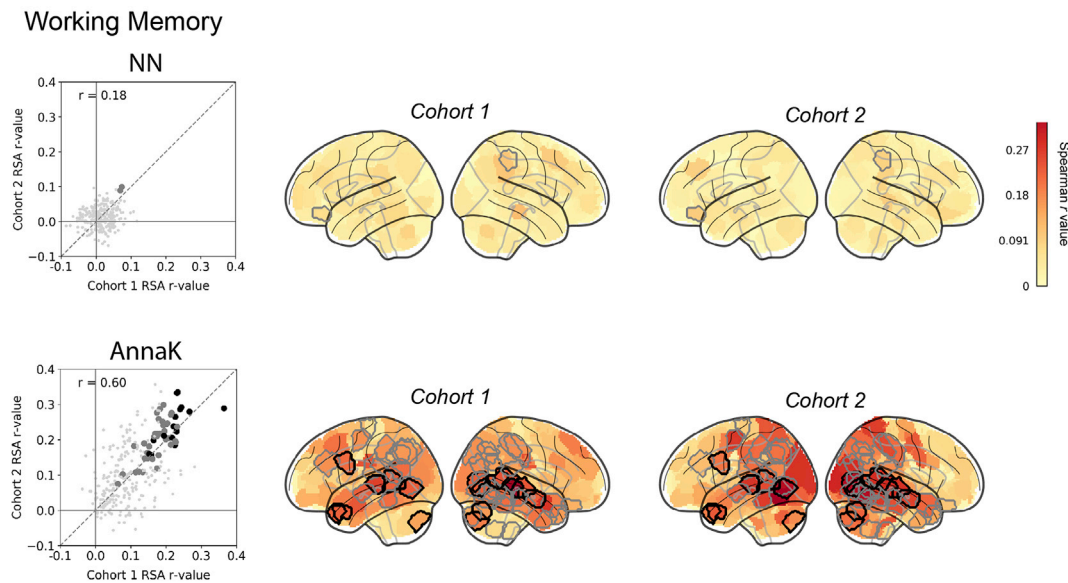
Fig. 3. Inter-subject RSA: Working Memory.

Do pairs of subjects that score more similarly on a test of working memory (Human Connectome Project: ListSort_Unadj) also show stronger ISC in certain brain regions during naturalistic viewing? Two models for behavioral similarity are tested: a nearest-neighbor model (top row; cf. Fig. 2a) where the behavioral similarity matrix is constructed as |i − j|, and an "Anna Karenina" model (bottom row, cf. Fig. 2b) where the behavioral similarity matrix is constructed as mean(i,j). In the scatter plots, each dot represents one node in the Shen atlas (268 total), plotted according to its representational similarity (Spearman correlation between brain similarity and behavioral similarity matrix, r) in cohort 1 (x-axis) versus its representational similarity in cohort 2 (y-axis). Large gray dots are nodes that show significant representational similarity (p < 0.05, uncorrected) after permutation testing in both cohorts (no. permutations = 10,000 for each cohort); large black dots are nodes that show significant representational similarity (p < 0.0136) after Bonferroni-style correction at α < 0.05. The dashed diagonal line represents the identity line y = x (not the regression line), to facilitate visual inspection of replicability—if the results are replicable across cohorts, the RSA r-values should fall close to this line. Glass brains show nodes colored by IS-RSA value. Nodes outlined in gray and black show significant representational similarity after familywise and Bonferroni correction, respectively (corresponding to the large gray and black dots in the scatterplots).

cerebellum (Fig. 4a). Interestingly, this suggests that brain responses during naturalistic stimulation may depend less on overall levels of certain personality traits, and more on personality "fingerprints" captured by specific patterns of item-wise responses to the personality questionnaire. The relative overrepresentation of results in the cerebellum adds to a growing body of evidence for its role in social cognition, especially processes that involve a high level of abstraction (Van Overwalle et al., 2014); the present results suggest that individual-specific traits may be encoded in patterns of cerebellar activity during complex stimulation of a largely social nature.

In sum, we observed that different traits and behaviors are best modeled by different distance functions in different brain regions. These results raise several questions to be addressed in future work. For example, which model(s) perform best on other traits and behaviors? Can we use dynamic analyses (Glerean et al., 2012; Simony et al., 2016) to reveal particular time windows where the brain-behavior signal is strongest? Do these windows relate to known features of the stimulus? For now, we can conclude that IS-RSA is a promising framework to detect brain-behavior relationships during naturalistic imaging, and that choosing the appropriate model not only improves sensitivity, but also offers the flexibility to test multiple hypotheses with fundamentally different interpretations.

6. Inter-subject RSA: Future directions

6.1. Boosting individual signal by removing common variance

Our intent here was to demonstrate IS-RSA using a straightforward, whole-brain method without an overly burdensome computational load. However, ISC may also suffer from the reliability paradox (Hedge et al., 2018): the stronger the shared response, the harder it is to identify individual differences. This is clear especially in sensory areas with high ISC values reflecting a strong shared time-locked response, possibly masking inter-individual differences.

A refined pipeline for IS-RSA might incorporate additional processing steps to boost individual signal and improve detection power for brain-behavior relationships. For example, simply regressing out the group-average timecourse from each individual's nodewise data before computing IS-RSA might highlight differences of interest. Alternatively, stimulus features (e.g., luminance, visual flow, acoustic properties (Lahnakoski et al., 2012b)) could be used as confounds to regress out strongly time-locked components from brain signals.

Yet another approach would be to first estimate a common set of latent features at the group level via hyperalignment (Haxby et al., 2011) or shared-response modeling (Chen et al., 2015), then calculate IS-RSA based on either the individual-subject data transformed into feature space or on the residuals after factoring out group variance. More sophisticated techniques using wavelet decomposition or mutual information can pull out shared signal components that are not time locked between individuals or with the stimulus, conceptually similarly to the evoked and induced oscillations in electrophysiology (Tallon-Baudry and Bertrand, 1999) (see also subsection "*Increasing spatiotemporal complexity*"). All of these techniques seek to decompose spatiotemporal signals into a set of features that model the shared component of subjects' responses, thereby minimizing the impact of misaligned anatomical or functional topographies in favor of emphasizing "true" functional differences between individuals. Shared-response modeling and hyperalignment have been shown to increase sensitivity to group differences (Chen et al., 2015) and individual differences (Feilong et al., 2018), respectively; future work should investigate whether one or more of these approaches increases sensitivity to brain-behavior relationships in the IS-RSA framework.
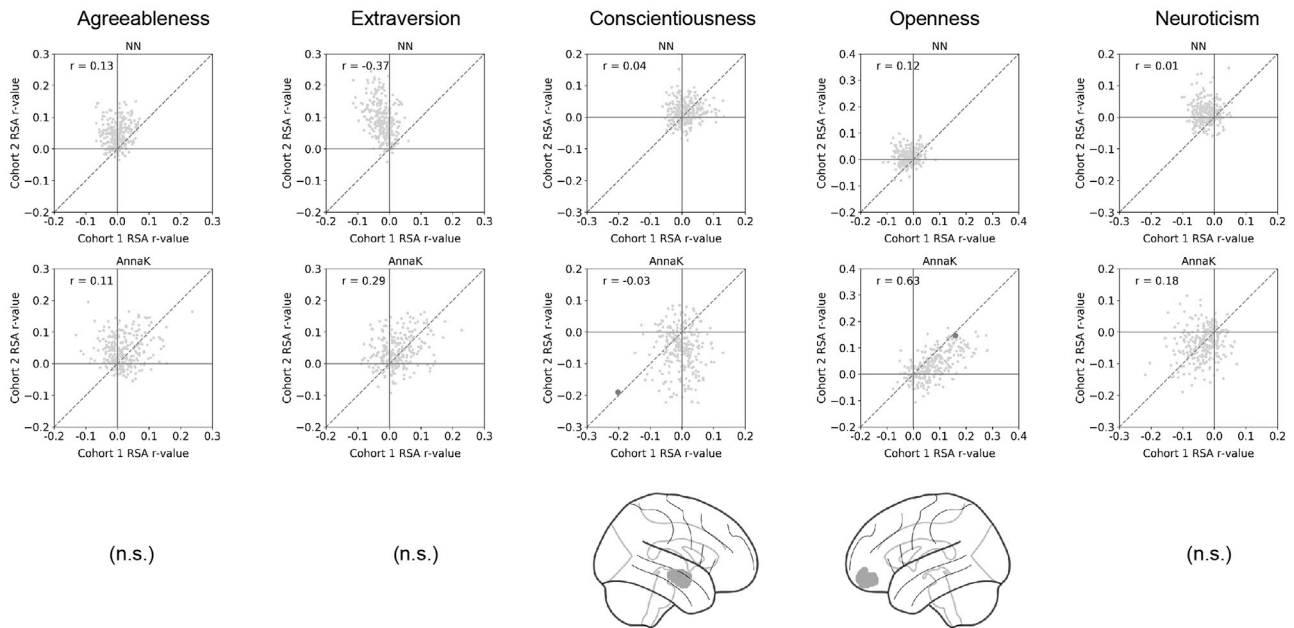
6.2. Extending IS-RSA to a predictive framework

We can readily extend IS-RSA to a predictive framework (Bzdok and Ioannidis, 2019; Gabrieli et al., 2015), in which a model is trained to take in patterns of brain activity during naturalistic stimulation and generate a

a. Personality: Itemwise
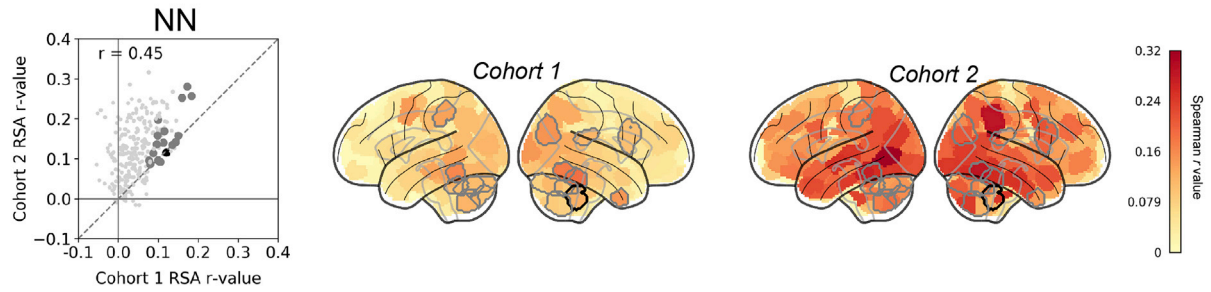
b. Personality: Trait scores

(n.s.) (n.s.) (n.s.)

Fig. 4. Inter-subject RSA: Personality.

Do pairs of subjects with more similar personalities (as measured with the Five-Factor Inventory) also show stronger ISC in certain brain regions during naturalistic viewing? a) IS-RSA, where personality similarity is calculated as the Pearson correlation between item-wise responses of each pair of subjects ("NN-itemwise"). b) IS-RSA where personality similarity is calculated based on summary scores for each of the five traits. For each trait, two models for behavioral similarity are tested: a nearest-neighbor model (top graph in each column; cf. Fig. 2a) where the behavioral similarity matrix is constructed as |i − j|, and an "Anna Karenina" model (bottom graph in each column, cf. Fig. 2b) where the behavioral similarity matrix is constructed as mean(i,j). In all scatter plots, each dot represents one node in the Shen atlas (268 total), plotted according to its representational similarity (Spearman correlation between brain similarity and behavioral similarity matrix) in cohort 1 (x-axis) versus its representational similarity in cohort 2 (y-axis). Large gray dots are nodes that show significant representational similarity (p < 0.05, uncorrected) after permutation testing in both cohorts (no. permutations = 10,000 for each cohort); large black dots are nodes that show significant representational similarity (p < 0.0136) after Bonferroni-style correction at α < 0.05. The dashed diagonal line represents the identity line y = x (not the regression line), to facilitate visual inspection of replicability—if the results are replicable across cohorts, the RSA r-values should fall close to this line. Glass brains show nodes colored by IS-RSA value. Nodes outlined in gray and black show significant representational similarity after familywise and Bonferroni correction, respectively (corresponding to the large gray and black dots in the scatterplots).

predicted behavioral score for never-before-seen individuals (or vice versa). In this context, too, the success of the model will depend crucially on knowing whether the brain-behavior relationship shows an *AnnaK* or *NN*-type structure.

The success of the AnnaK model for at least one behavior investigated here has interesting implications for not only IS-RSA, but any approach to predicting behavior from brain features: it suggests that the relationship between imaging-derived features and behavior may be linear (or at least monotonic) only at one end of the behavior spectrum, while the other end is associated with increased variability but not necessarily in a consistent direction. One implication may be to consider radial kernels for regression, which measure distance from a central point (as opposed

to linear kernels, which measure absolute position on a set of axes). However, the *AnnaK* structure does predict a linear or monotonic relationship between a subject's behavioral score and their mean (or median) ISC value with all other subjects, a fact that could be leveraged in a predictive analysis.

On the other hand, the NN models do not predict any relationship between behavior and overall mean or median ISC value, but rather between behavior and ISC with specific partners. In this case, nonparametric methods such as *k*-nearest neighbors would be more appropriate. In either case, feature selection could be applied to uncover the brain regions where similarity is most strongly related to behavioral similarity, to improve model performance.

6.3. Unsupervised IS-RSA

As we amass larger neuroimaging datasets that include naturalistic tasks, it may become possible to perform unsupervised analyses. In other words, instead of using known behavioral scores to pre-label the data, we could cluster inter-subject brain similarity matrices to detect natural categories or continua in a data-driven fashion, then see if these relate to present or future behavioral outcomes (Cerliani et al., 2017). In this way, we may be able to leverage naturalistic neuroimaging to organize individuals along axes that are more biologically valid than current diagnostic and self-report measures, potentially shedding new light on how individual variability is reflected in nuanced brain function.

6.4. Increasing spatiotemporal complexity

Thus far we have been using inter-subject correlation, or simple Pearson correlation between two subjects' timeseries, as our primary measure of brain similarity, largely because it is straightforward to compute, visualize, and interpret. However, in theory we could calculate a brain similarity matrix (cf. Fig. 1) based on any type of information extracted from single subjects' neuroimaging data. For example, we could use functional connectivity (Glerean et al., 2016), or compare subjects' temporal trajectories over the course of a stimulus using low-dimensional topological embeddings (e.g., Gonzalez-Castillo et al., 2019; Saggar et al., 2018), latent state discovery (Chang et al., 2018), or projection into a higher-order space using recurrent neural networks (Venkatesh et al., 2019). Furthermore, the brain and behavioral inter-subject similarity matrices (cf. Fig. 1) can also be interpreted as networks, meaning they could be analyzed using geometry-aware methods such as geodesic distance (Venkatesh et al., 2020). Rather than mass univariate tests for how behavior is reflected in single regions, these approaches allow for linking behavior to multivariate patterns of activity across multiple regions, and may prove even more powerful for uncovering brain-behavior relationships during naturalistic imaging.

7. Stimulus selection: How much synchrony is enough?

Thus far, we have discussed how to optimize sensitivity to stimulus-evoked individual differences from an analysis perspective. What about from an acquisition perspective? While naturalistic imaging experiments are growing in popularity, currently, there is no principled way to choose stimuli. Common wisdom is that for studying shared responses, we should choose something maximally engaging, to "drive" as much of the brain as possible. Indeed, previous work has shown that features of the stimulus affect ISC levels: more rhetorically powerful speeches (Schmälzle et al., 2015), emotionally arousing narratives (Nummenmaa et al., 2014), and highly rated television programs (Dmochowski et al., 2014) all evoke higher synchrony than their less engaging counterparts.

But, when the goal of a study is specifically to investigate individual differences, considerations for choosing a stimulus may be different. In the theoretical limit, a stimulus that evoked perfect synchrony across subjects would be useless for studying individual differences—since there would be no neural variability to relate to behavioral variability (Hedge et al., 2018). Practically, however, we are quite far from that theoretical limit, since individual BOLD responses are "noisy" both in terms of uninteresting variability (scanner noise, non-neural BOLD signals, stimulus-unrelated neural activity) and the stimulus-driven idiosyncratic responses that constitute the "signal" of interest here.

Is there a "sweet spot" where a stimulus evokes enough synchrony to build a successful cross-subject model, but not enough to saturate the individual signals of interest? To test their hypothesis that social network proximity predicts increasingly similar neural responses to movies, Parkinson et al. (2018) chose videos that might differentially appeal to those with different tastes (reasoning that friends would be more likely to have similar tastes; e.g., styles of humor, opinions on controversial topics). A handful of studies have created bespoke stimuli that were ambiguous by

design, such that different individuals might arrive at different interpretations of the same material. Finn et al. (2018) created a narrative describing a complex social scenario that seemed highly suspicious or nefarious to some individuals, but less so to others; Nguyen et al. (2018) created a Heider-Simmel-esque video (Heider and Simmel, 1944) depicting an interaction among animated shapes in which the relationships between the shapes were open to interpretation. These studies found that individuals who were more similar on either trait-level (i.e., intrinsic) or state-level (i.e., stimulus-driven) measures, respectively, showed increased inter-subject correlation during stimulus presentation in regions of higher-order association cortices, especially those linked to social cognition.

None of these studies, however, directly assessed how the degree of ambiguity or so-called "taste-dependence" of a stimulus affects its utility for drawing out meaningful individual signal. If the goal is to study idiosyncratic responses, how much synchrony is optimal? If we imagine plotting stimuli by the strength of the ISCs they evoke versus their sensitivity to individual differences, several potential scenarios emerge (Fig. 5.). One possibility is that idiosyncratic signals are quickly saturated by a powerful stimulus, such that the optimal stimulus would evoke only minimally correlated responses when averaged across all subjects, leaving room for specific subject pairs to be more or less correlated with one another according to variations in behavior (Scenario 1). Alternatively, the optimal level of average synchrony could be higher, such that sensitivity to individual differences benefits from a stronger foundational shared response at the group level (Scenario 2). Another possibility is that the theoretical "saturation point" is much higher, such that stimuli that evoke very strong responses are the ones that are most sensitive to individual differences (Scenario 3). This third scenario would be consistent with the observation that stimuli and task states that make subjects look more similar to one another can actually boost signal-to-noise for individual differences, since even though these states reduce
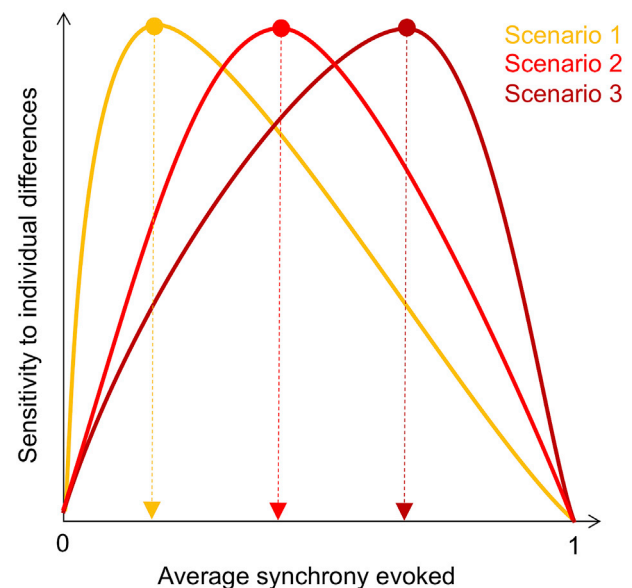

Fig. 5. Theoretical stimulus tuning curves for sensitivity to individual differences.
In the upper limit, as the degree of cross-subject synchrony evoked by a stimulus approaches 1, that stimulus will lose sensitivity to individual differences, since there will be no brain variability left to relate to behavioral variability. However, in the lower limit, if a stimulus evokes no correlation across subjects, there will be no meaningful structure in brain similarity to relate to behavioral similarity. Therefore, the optimal tuning curve likely follows an inverted-U shape. Determining where this curve peaks—in other words, the optimal degree of synchrony for extracting meaningful individual differences in a certain behavioral domain—should be a goal for future work.

overall cross-subject variation, the remaining variation is presumably more stable and trait-like (Finn et al., 2017; Vanderwal et al., 2017).

Understanding how stimulus choice affects sensitivity to individual differences will be an important area for future research, especially as large-scale data collection efforts begin to incorporate naturalistic scans into their protocols (Alexander et al., 2017; Van Essen et al., 2012). We believe acquiring naturalistic scans should be encouraged for several reasons: beyond improving subject compliance (and therefore data quality) compared to resting-state scans (Greene et al., 2018; Huijbers et al., 2017; Vanderwal et al., 2015), naturalistic paradigms yield data that can be mined in any number of ways, expanding their potential to generate insights into a variety of open questions in human neuroscience. Yet, to the extent that stimulus choice affects the individual signals we observe, it will behoove us to be as principled as possible in choosing stimuli.

Of course, there may not be a single "best" stimulus for studying individual differences; rather, the most appropriate stimulus may depend on the specific behavior(s) of interest. For example, a threatening or suspenseful stimulus—e.g., the opening scene in a horror film—might evince neural responses that share structure with trait anxiety, as compared to one depicting positive emotion—e.g., the happily-ever-after final scene of a romantic comedy. Conversely, the positive stimulus might yield better predictions of trait anhedonia. At the same time, both of these stimuli may yield better predictions of either trait than a neutral stimulus with little to no emotional content. Testing these hypotheses will require datasets with a range of both stimuli and behavioral measures per subject. However, there is already some supporting evidence: for example, differences between controls and patients with melancholic depression are more pronounced during a negative film clip than a positive one (Guo et al., 2015).

8. Limitations

Despite the many clear advantages and—we believe—great potential of using naturalistic paradigms to study individual differences, there are several outstanding challenges. For one, the test-retest validity of these paradigms is unclear. Repetition is known to alter neural processing, and this is likely especially true for emotionally evocative and memorable stimuli. Thus, naturalistic paradigms that rely on a single stimulus may be less appropriate for longitudinal studies, since it would be near impossible to disentangle within-subject changes from repetition effects.

There are computational challenges associated with using inter-subject approaches, since pairwise techniques mean that the number of observations increases as n^2 rather than n. Especially in the case of inter-subject functional connectivity (ISFC; Simony et al. (2016)), reducing dimensionality using pre-defined atlases, and/or selecting regions of interest based on *a priori* hypotheses or a first-pass analysis of which regions surpass some minimal threshold for response consistency (cf. Fig. 5.), may help in this regard.

Another challenge, albeit not one unique to naturalistic paradigms, is understanding which traits and behaviors we should target. Predicting a performance-based or self-report score acquired close in time to the imaging data itself, which is the goal of the vast majority of the current literature on brain-behavior relationships, is an important proof-of-concept, but ultimately the imaging data will always be simply a noisier version of whatever "ground truth" we are trying to predict. Ultimately, the goal should be to determine the value of baseline brain responses to naturalistic stimuli as predictors of follow-up measures such as learning rates (Cantlon and Li, 2013), illness trajectory, or response to intervention.

9. Concluding remarks

Here, we have advanced an emerging framework for studying individual differences during naturalistic neuroimaging, a phenomenon we call "idiosynchrony". Inter-subject representational similarity analysis (IS-RSA) combines the time-locked nature of the stimulus with known phenotypic information to move from shared responses to activity in individual subjects that is idiosyncratic, yet structured and interpretable. Unlike traditional approaches that rely on explicit models of the task, inter-subject approaches promise to capture as much nuance and variance of the evoked activity as possible. And, unlike functional connectivity approaches that treat naturalistic neuroimaging data akin to rest, inter-subject approaches afford near certainty that the observed signals are both neural in origin and driven by the stimulus. Above and beyond a boost in signal-to-noise, these paradigms open up exciting opportunities to link individual patterns of brain activity to specific events within the stimulus that unfold over various timescales, from its low-level sensory properties up to high-level narrative features that may evoke different memories, associations and emotions for each individual. Linking naturalistic patterns of brain activity to trait- and state-related variability across subjects will deepen our understanding of how individual brains give rise to individual behaviors, and may eventually lead to imaging-based tools for real-world applications.

Data and code availability

Raw data for the empirical results presented here come from the Human Connectome Project (http://www.humanconnectomeproject.org/). Code for all IS-RSA analyses—both simulations (cf. Fig. 2) and empirical application to HCP data (cf. Figs. 3 and 4)—can be found in the following Github repository: https://github.com/esfinn/intersubj_rsa, which also contains the processed HCP data (nodewise time series) that formed the input to the empirical analyses.

CRediT authorship contribution statement

Emily S. Finn: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Visualization, Writing - original draft. **Enrico Glerean:** Conceptualization, Formal analysis, Methodology, Visualization, Writing - review & editing. **Arman Y. Khojandi:** Data curation, Software. **Dylan Nielson:** Software, Visualization, Validation. **Peter J. Molfese:** Data curation, Resources. **Daniel A. Handwerker:** Methodology, Validation, Writing - review & editing. **Peter A. Bandettini:** Funding acquisition, Supervision, Writing - review & editing.

Acknowledgements

References

Alexander, L.M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., et al., 2017. An open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci. Data 4, 170181.

Bacha-Trams, M., Alexandrov, Y.I., Broman, E., Glerean, E., Kauppila, M., Kauttonen, J., Ryyppö, E., Sams, M., Jääskeläinen, I.P., 2018. A drama movie activates brains of holistic and analytical thinkers differentially. Soc. Cognit. Affect Neurosci. 13, 1293–1304.

Barch, D.M., Burgess, G.C., Harms, M.P., Petersen, S.E., Schlaggar, B.L., Corbetta, M., Glasser, M.F., Curtiss, S., Dixit, S., Feldt, C., et al., 2013. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 80, 169–189.

Bartels, A., Zeki, S., 2004. Functional brain mapping during free viewing of natural scenes. Hum. Brain Mapp. 21, 75–85.

Bartolomeo, P., Seidel Malkinson, T., de Vito, S., 2017. Botallo's error, or the quandaries of the universality assumption. Cortex 86, 176–185.

Betti, V., Della Penna, S., de Pasquale, F., Mantini, D., Marzetti, L., Romani, Gian L., Corbetta, M., 2013. Natural scenes viewing alters the dynamics of functional connectivity in the human brain. Neuron 79, 782–797.

Bolton, T.A.W., Jochaut, D., Giraud, A.L., Ville, D.V.D., 2018. Brain dynamics in ASD during movie-watching show idiosyncratic functional integration and segregation. Hum. Brain Mapp. 39, 2391–2404.

Byrge, L., Dubois, J., Tyszka, J.M., Adolphs, R., Kennedy, D.P., 2015. Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. J. Neurosci. 35, 5837–5850.

Bzdok, D., Ioannidis, J.P.A., 2019. Exploration, inference, and prediction in neuroscience and biomedicine. Trends Neurosci. 42, 251–262.

Cantlon, J.F., Li, R., 2013. Neural activity during natural viewing of Sesame Street statistically predicts test scores in early childhood. PLoS Biol. 11, e1001462.

Cerliani, L., Thomas, R.M., Aquino, D., Contarino, V., Bizzi, A., 2017. Disentangling subgroups of participants recruiting shared as well as different brain regions for the execution of the verb generation task: a data-driven fMRI study. Cortex 86, 247–259.

Chang, L.J., Jolly, E., Cheong, J.H., Rapuano, K., Greenstein, N., Chen, P.-H.A., Manning, J.R., 2018. Endogenous Variation in Ventromedial Prefrontal Cortex State Dynamics during Naturalistic Viewing Reflects Affective Experience. bioRxiv, p. 487892.

Chen, P.-H.A., Jolly, E., Cheong, J.H., Chang, L.J., 2020. Inter-subject representational similarity analysis reveals individual variations in affective experience when watching erotic movies. NeuroImage. In this issue.

Chen, G., Taylor, P.A., Qu, X., Molfese, P.J., Bandettini, P.A., Cox, R.W., Finn, E.S., 2020. Untangling the relatedness among correlations, part III: Inter-subject correlation analysis through Bayesian multilevel modeling for naturalistic scanning. NeuroImage. In this issue.

Chen, G., Taylor, P.A., Shin, Y.-W., Reynolds, R.C., Cox, R.W., 2017. Untangling the relatedness among correlations, Part II: inter-subject correlation group analysis through linear mixed-effects modeling. Neuroimage 147, 825–840.

Chen, P.-H.C., Chen, J., Yeshurun, Y., Hasson, U., Haxby, J., Ramadge, P.J., 2015. A reduced-dimension fMRI shared response model. In: Paper Presented at: Adv Neural Inf Process Syst.

Cole, M.W., Yarkoni, T., Repovš, G., Anticevic, A., Braver, T.S., 2012. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. J. Neurosci. 32, 8988–8999.

Cooper, E.A., Hasson, U., Small, S.L., 2011. Interpretation-mediated changes in neural activity during language comprehension. Neuroimage 55, 1314–1323.

Cuthbert, B.N., Insel, T.R., 2013. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Med. 11, 126.

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, 4567.

Dubois, J., Adolphs, R., 2016. Building a science of individual differences from fMRI. Trends Cognit. Sci. 20, 425–443.

Feilong, M., Nastase, S.A., Guntupalli, J.S., Haxby, J.V., 2018. Reliable individual differences in fine-grained cortical functional architecture. Neuroimage 183, 375–386.

Finn, E.S., Constable, R.T., 2016. Individual variation in functional brain connectivity: implications for personalized approaches to psychiatric disease. Dialogues Clin. Neurosci. 18, 277–287.

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, 2043.

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.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671.

Gabrieli, John D.E., Ghosh, Satrajit S., Whitfield-Gabrieli, S., 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. Neuron 85, 11–26.

Geerligs, L., Rubinov, M., Cam-Can, Henson, R.N., 2015. State and trait components of functional connectivity: individual differences vary with mental state. J. Neurosci. 35, 13949–13961.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., 2013a. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al., 2013b. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124.

Glerean, E., Pan, R.K., Salmi, J., Kujala, R., Lahnakoski, J.M., Roine, U., Nummenmaa, L., Leppämäki, S., Nieminen-von Wendt, T., Tani, P., 2016. Reorganization of functionally connected brain subnetworks in high-functioning autism. Hum. Brain Mapp. 37, 1066–1079.

Glerean, E., Salmi, J., Lahnakoski, J.M., Jääskeläinen, I.P., Sams, M., 2012. Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. Brain Connect. 2, 91–101.

Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D., Pereira, F., Bandettini, P., 2019. Imaging the Spontaneous Flow of Thought: Distinct Periods of

Cognition Contribute to Dynamic Functional Connectivity during Rest. bioRxiv, p. 527804.

Greene, D.J., Koller, J.M., Hampton, J.M., Wesevich, V., Van, A.N., Nguyen, A.L., Hoyt, C.R., McIntyre, L., Earl, E.A., Klein, R.L., et al., 2018. Behavioral interventions for reducing head motion during MRI scans in children. Neuroimage 171, 234–245.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., et al., 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage 95, 232–247.

Gruskin, D.C., Rosenberg, M.D., Holmes, A.J., 2020. Relationships between depressive symptoms and brain responses during emotional movie viewing emerge in adolescenc. NeuroImage. In this issue.

Guo, C.C., Hyett, M.P., Nguyen, V.T., Parker, G.B., Breakspear, M.J., 2016. Distinct neurobiological signatures of brain connectivity in depression subtypes during natural viewing of emotionally salient films. Psychol. Med. 46, 1535–1545.

Guo, C.C., Nguyen, V.T., Hyett, M.P., Parker, G.B., Breakspear, M.J., 2015. Out-of-sync: disrupted neural activity in emotional circuitry during film viewing in melancholic depression. Sci. Rep. 5, 11605.

Hampson, M., Tokoglu, F., Sun, Z., Schafer, R.J., Skudlarski, P., Gore, J.C., Constable, R.T., 2006. Connectivity–behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. Neuroimage 31, 513–519.

Hasson, U., Avidan, G., Gelbard, H., Vallines, I., Harel, M., Minshew, N., Behrmann, M., 2009. Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. Autism Res. 2, 220–231.

Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. Trends Cognit. Sci. 14, 40–48.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303, 1634–1640.

Haxby, James V., Guntupalli, J.S., Connolly, Andrew C., Halchenko, Yaroslav O., Conroy, Bryan R., Gobbini, M.I., Hanke, M., Ramadge, Peter J., 2011. A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron 72, 404–416.

Hedge, C., Powell, G., Sumner, P., 2018. The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. Behav. Res. Methods 50, 1166–1186.

Heider, F., Simmel, M., 1944. An experimental study of apparent behavior. Am. J. Psychol. 57, 243–259.

Hejnar, M.P., Kiehl, K.A., Calhoun, V.D., 2007. Interparticipant correlations: a model free FMRI analysis technique. Hum. Brain Mapp. 28, 860–867.

Huijbers, W., Van Dijk, K.R.A., Boenniger, M.M., Stirnberg, R., Breteler, M.M.B., 2017. Less head motion during MRI under task than resting-state conditions. Neuroimage 147, 111–120.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am. J. Psychiatr. 167, 748–751.

Jääskeläinen, I.P., Pajula, J., Tohka, J., Lee, H.-J., Kuo, W.-J., Lin, F.-H., 2016. Brain hemodynamic activity during viewing and re-viewing of comedy movies explained by experienced humor. Sci. Rep. 6, 27741.

Kriegeskorte, N., Kievit, R.A., 2013. Representational geometry: integrating cognition, computation, and the brain. Trends Cognit. Sci. 17, 401–412.

Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2.

Lahnakoski, J., Glerean, E., Salmi, J., Jääskeläinen, I., Sams, M., Hari, R., Nummenmaa, L., 2012a. Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. Front. Hum. Neurosci. 6.

Lahnakoski, J.M., Glerean, E., Jääskeläinen, I.P., Hyönä, J., Hari, R., Sams, M., Nummenmaa, L., 2014. Synchronous brain activity across individuals underlies shared psychological perspectives. Neuroimage 100, 316–324.

Lahnakoski, J.M., Salmi, J., Jääskeläinen, I.P., Lampinen, J., Glerean, E., Tikka, P., Sams, M., 2012b. Stimulus-related independent component and voxel-wise analysis of human brain activity during free viewing of a feature film. PloS One 7, e35215.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Canc. Res. 27, 209–220.

Mäntylä, T., Nummenmaa, L., Rikandi, E., Lindgren, M., Kieseppä, T., Hari, R., Suvisaari, J., Raij, T.T., 2018. Aberrant cortical integration in first-episode psychosis during natural audiovisual processing. Biol. Psychiatr. 84, 655–664.

Moraczewski, D., Chen, G., Redcay, E., 2018. Inter-subject synchrony as an index of functional specialization in early childhood. Sci. Rep. 8, 2252.

Nastase, S.A., Gazzola, V., Hasson, U., Keysers, C., 2019. Measuring shared responses across subjects using intersubject correlation. Soc. Cognit. Affect Neurosci. 14, 667–685.

Nguyen, M., Vanderwal, T., Hasson, U., 2019. Shared understanding of narratives is correlated with shared neural responses. Neuroimage 184, 161–170.

Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I.P., Hari, R., Sams, M., 2012. Emotions promote social interaction by synchronizing brain activity across individuals. Proc. Natl. Acad. Sci. U.S.A. 109, 9599–9604.

Nummenmaa, L., Saarimäki, H., Glerean, E., Gotsopoulos, A., Jääskeläinen, I.P., Hari, R., Sams, M., 2014. Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. Neuroimage 102, 498–509.

Pajula, J., Kauppi, J.-P., Tohka, J., 2012. Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. PloS One 7, e41196.

Parkinson, C., Kleinbaum, A.M., Wheatley, T., 2018. Similar neural responses predict friendship. Nat. Commun. 9, 332.

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., Saxe, R., 2018. Development of the social brain from age three to twelve years. Nat. Commun. 9, 1027.

Russ, B.E., Leopold, D.A., 2015. Functional MRI mapping of dynamic visual features during natural viewing in the macaque. Neuroimage 109, 84–94.

Saalasti, S., Alho, J., Bar, M., Glerean, E., Honkela, T., Kauppila, M., Sams, M., Jääskeläinen, I.P., 2019. Inferior parietal lobule and early visual areas support elicitation of individualized meanings during narrative listening. Brain Behav. 9, e01288.

Saggar, M., Sporns, O., Gonzalez-Castillo, J., Bandettini, P.A., Carlsson, G., Glover, G., Reiss, A.L., 2018. Towards a new approach to reveal dynamical organization of the brain using topological data analysis. Nat. Commun. 9, 1399.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90, 449–468.

Salmi, J., Roine, U., Glerean, E., Lahnakoski, J., Nieminen-von Wendt, T., Tani, P., Leppämäki, S., Nummenmaa, L., Jääskeläinen, I.P., Carlson, S., et al., 2013. The brains of high functioning autistic individuals do not synchronize with those of others. Neuroimage: Clin. 3, 489–497.

Schmälzle, R., Häcker, F.E.K., Honey, C.J., Hasson, U., 2015. Engaged listeners: shared neural processing of powerful political speeches. Soc. Cognit. Affect Neurosci. 10, 1137–1143.

Seghier, M.L., Price, C.J., 2018. Interpreting and utilising intersubject variability in brain function. Trends Cognit. Sci. 22, 517–530.

Shen, X., Tokoglu, F., Papademetris, X., Constable, R., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403–415.

Simony, E., Honey, C.J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., Hasson, U., 2016. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 7, 12141.

Song, M., Zhou, Y., Li, J., Liu, Y., Tian, L., Yu, C., Jiang, T., 2008. Brain spontaneous functional connectivity and intelligence. Neuroimage 41, 1168–1176.

Sonkusare, S., Breakspear, M., Guo, C., 2019. Naturalistic stimuli in neuroscience: critically acclaimed. Trends Cognit. Sci. 23, 699–714.

Tallon-Baudry, C., Bertrand, O., 1999. Oscillatory gamma activity in humans and its role in object representation. Trends Cognit. Sci. 3, 151–162.

Tei, S., Kauppi, J.-P., Fujino, J., Jankowski, K.F., Kawada, R., Murai, T., Takahashi, H., 2019. Inter-subject correlation of temporoparietal junction activity is associated with conflict patterns during flexible decision-making. Neurosci. Res. 144, 67–70.

Thiede, A., Glerean, E., Kujala, T., Parkkonen, L., 2019. Atypical Brain-To-Brain Synchronization during Listening to Continuous Natural Speech in Dyslexia. bioRxiv, p. 677674.

van Baar, J.M., Chang, L.J., Sanfey, A.G., 2019. The computational and neural substrates of moral strategies in social decision-making. Nat. Commun. 10, 1483.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., 2013. The Wu-Minn human connectome project: an overview. Neuroimage 80, 62–79.

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., et al., 2012. The Human Connectome Project: a data acquisition perspective. Neuroimage 62, 2222–2231.

Van Overwalle, F., Baetens, K., Mariën, P., Vandekerckhove, M., 2014. Social cognition and the cerebellum: a meta-analysis of over 350 fMRI studies. Neuroimage 86, 554–572.

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

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.

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.

Venkatesh, M., Jaja, J., Pessoa, L., 2019. Brain dynamics and temporal trajectories during task and naturalistic processing. Neuroimage 186, 410–423.

Venkatesh, M., Jaja, J., Pessoa, L., 2020. Comparing functional connectivity matrices: a geometry-aware approach applied to participant identification. Neuroimage 207, 116398.

Vu, A.T., Jamison, K., Glasser, M.F., Smith, S.M., Coalson, T., Moeller, S., Auerbach, E.J., Uğurbil, K., Yacoub, E., 2017. Tradeoffs in pushing the spatial resolution of fMRI for the 7T Human Connectome Project. Neuroimage 154, 23–32.

Yang, Z., Wu, J., Xu, L., Deng, Z., Tang, Y., Gao, J., Hu, Y., Zhang, Y., Qin, S., Li, C., et al., 2020. Individualized psychiatric imaging based on inter-subject neural synchronization in movie watching. Neuroimage. In this issue.

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C.J., Hasson, U., 2017. Same story, different story: the neural representation of interpretive frameworks. Psychol. Sci. 28, 307–319.