Reviewed Preprint
v1 • May 7, 2025
Not revised

Reviewed Preprint
v2 • July 6, 2026
Revised by authors

Heritability of movie-evoked brain activity and connectivity

David C Gruskin[1] ✉, **Daniel J Vieira**[2], **Jessica K Lee**[2], **Gaurav H Patel**[2,3]

[1]Medical Scientist Training Program, Columbia University Irving Medical Center, New York, United States • [2]Division of Pathophysiology and Treatment Research, New York State Psychiatric Institute, New York, United States • [3]Department of Psychiatry, Columbia University Irving Medical Center, New York, United States

✉ **For correspondence:** david.gruskin@columbia.edu

Competing interests: No competing interests declared

Funding: See page 25

Reviewing editor: Emma Sprooten, Donders Institute for Brain, Cognition and Behaviour, Netherlands

eLife Assessment

This paper addresses a **valuable** research question on the modest heritability of the brain's response to movie watching, and how heritability varies under different parameters such as regional spatial hyperalignment and BOLD frequency bands. The topic of this paper is of interest to fMRI methodological experts, and potentially to a broader cognitive neuroscience audience, and those with an interest in understanding the heritable sources of individual differences in brain function. Although some of the conclusions could be strengthened by future cross validation studies in independent and larger family-based samples, and through complementary twin/family and SNP-based models, taken altogether, the analyses and results provide **convincing** evidence for the overall conclusions.

https://doi.org/10.7554/eLife.106081.2.sa3

Abstract

The neural bases of sensory processing are conserved across people, but no two individuals experience the same stimulus in exactly the same way. Recent work has established that the idiosyncratic nature of subjective experience is underpinned by individual variability in brain responses to sensory information. However, the fundamental origins of this individual variability have yet to be systematically investigated. Here, we establish a genetic basis for individual differences in sensory processing by quantifying (1) the heritability of high-dimensional brain responses to movies and (2) the extent to which this heritability is grounded in lower-level aspects of brain function. Specifically, we leverage 7T fMRI data collected from a twin sample to first show that movie-evoked brain activity is heritable across the cortex, and that this heritability is greater for information encoded in lower temporal frequencies, especially in more associative cortical areas. Next, we use hyperalignment to decompose this heritability into genetic similarity in *where* vs. *how* sensory information is processed. We also show that the heritability of brain activity patterns can be partially explained by the heritability of the neural timescale, a one-dimensional measure of local circuit functioning. Finally, we generalize our findings by illustrating a similar pattern of results for the heritability of movie-evoked functional connectivity. These results demonstrate that brain responses to complex stimuli are heritable, and that this heritability is due, in part, to genetic control over stable aspects of brain function.

Introduction

Although the neural machinery that allows us to process sensory information is broadly conserved across people, no two individuals experience the same sensory stimulus in exactly the same way. What underlying factors give rise to this person-to-person variability in subjective experience? Recent work has established that the idiosyncratic nature of subjective experience is reflected in idiosyncratic brain responses to sensory information, and that how an individual processes a stimulus is shaped by their psychosocial background and previous experiences. For example,

individuals who share more similar personality traits (Finn et al., 2018) and political orientations (van Baar et al., 2021) exhibit more similar interpretations of, and functional magnetic resonance imaging (fMRI) responses to, relevant audiovisual stimuli, as do individuals who are primed with more similar contextual information before listening to an ambiguous narrative (Yeshurun et al., 2017). Here, we extend this work by investigating a more fundamental source of variability in sensory-evoked brain responses and the experiences they represent: our genes.

Whether individual variability in a given trait is due to environmental or genetic factors is a central question in biology, and the extent to which this variability is underpinned by variation in genetics is captured by heritability (or h^2). Recent studies have quantified the heritability of various aspects of sensory brain function, revealing a genetic basis for patterns of brain activity elicited by auditory tones and visual gratings in sensory cortices (Alvarez et al., 2021; Renvall et al., 2012; van Pelt et al., 2012). However, the unimodal and low-dimensional nature of these stimuli may not capture the full complexity of real-life sensory experiences and the brain responses they evoke. Consequently, the extent to which genetic factors influence brain responses to more naturalistic stimuli remains unclear, especially for high-level (e.g., social and narrative) information encoded across longer timescales in association cortex.

In addition to activity patterns within individual brain areas, information can also be encoded in the functional connectivity (FC) between multiple areas or networks (Chen et al., 2014; Kohn et al., 2016). Research into the heritability of FC and related measures has largely focused on data acquired while subjects are at rest, during which an individual's unique FC profile (i.e., pattern of pairwise FC strengths), describes their brain's intrinsic functional architecture (Anderson et al., 2021; Burger et al., 2022; Busch et al., 2023; Dworetsky et al., 2024; Glahn et al., 2010; Sinclair et al., 2015; van den Heuvel & Hulshoff Pol, 2010). These studies have demonstrated that a range of resting state FC (rest FC)-derived measures are moderately heritable, and similar findings have resulted from work characterizing the heritability of FC during task performance, which additionally reflects the processing of information relevant to the task at hand (Cole et al., 2021; Elliott et al., 2019; Korgaonkar et al., 2014). Although this work has shed significant light on the genetic basis of FC during rest and cognition, the heritability of sensory-evoked FC patterns, which are known to encode stimulus features (Chen et al., 2014) and track individual differences in behavior (Finn & Bandettini, 2021), has yet to be investigated.

Finally, brain activity and connectivity patterns are complex phenomena that arise from a variety of physiological processes. Although the heritability estimates established by previous work could reflect emergent aspects of brain function, it might instead be possible to reduce them to genetic control over lower-level neural, vascular, and metabolic processes. For example, although the human cortex is topographically organized into areas that are specialized for processing specific kinds of information (e.g., facial features), the locations of these areas and the tuning patterns within them vary widely across individuals (Gordon et al., 2017; Haxby et al., 2020; Petersen et al., 2024). As such, these cortical topographies, or individual-specific maps of *where* stimulus features are processed, emerge over the course of development and constrain the activity and connectivity patterns an individual will exhibit during sensory processing. Independent of *where* stimuli are processed, stable aspects of brain function also shape high-dimensional activity and connectivity patterns by influencing *how* information is processed. For example, recent work from Shinn et al. (2023) showed that individual variability in higher-order aspects of brain function like FC profiles can be traced back to variability in simpler, low-level phenomena like temporal autocorrelation of the blood oxygen level-dependent (BOLD) signal. More specifically, a measure of temporal autocorrelation known as the neural timescale (NT) is thought to reflect the strength of local recurrent excitation (Cavanagh et al., 2020) but is also closely tied to the organization of brain-wide FC profiles (Shinn et al., 2023). Given that lower-level properties like functional topography (Alvarez et al., 2021; Anderson et al., 2021; Dworetsky et al., 2024) and BOLD temporal autocorrelation (Christova et al., 2022) are themselves heritable, it remains unclear (1)

to what extent high-dimensional brain responses to naturalistic stimuli are heritable and (2) how much of this heritability can be reduced to genetic control over these stable spatial and temporal aspects of brain function.

In the present work, we address these questions by analyzing 7T fMRI recordings of a twin sample acquired by the Human Connectome Project (Van Essen et al., 2013) to quantify the heritability of two distinct high-dimensional traits—stimulus-evoked BOLD time courses and functional connectivity profiles—across the cortex. Here, we focus on fMRI data acquired during movie-watching, as the rich and multimodal nature of movies engages multiple sensory and associative regions as well as the connections between them, making them well-suited for broadly assessing individual differences in sensory processing. Leveraging a multi-dimensional estimator of heritability (Anderson et al., 2021), we first show that movie-evoked BOLD time courses are heritable across the cortex. We extend this result by showing that BOLD time course heritability is greater in slower frequency bands, and especially in more associative parcels, suggesting that the neural processing of more abstract vs. lower-level sensory information is under greater genetic control. Next, we use hyperalignment to separate heritable differences in *where* information is processed from *how* it is processed by projecting voxel-level data into a common functional space, parsing the heritability of high-dimensional BOLD time courses into genetic control over stable spatial (e.g., functional cortical topography) versus temporal (e.g., neural timescale) aspects of brain function. Finally, we reveal a similar pattern of results for a different set of high-dimensional brain responses: functional connectivity profiles. Taken together, these results characterize the degree to which sensory processing is controlled by genetics and illustrate the benefits of a reductionist approach to studying the heritability of complex neurobiological phenomena, providing a foundation for future multi-scale studies of the mechanisms that underlie heritable differences in brain function.

Methods

Participants

Data used for this project come from the 178 subjects in the Human Connectome Project (HCP) Young Adult 7T release who completed every movie-watching run (Van Essen et al., 2013). All participants were healthy individuals between the ages of 22 and 36 (mean age = 29.4 years, standard deviation = 3.3) and provided informed written consent as part of their participation in the study. Self-reported racial identity in this sample was 87.6% White, 7.3% Black or African American, 3.9% Asian/Native Hawaiian/Other Pacific Islander, and 1.1% unknown/not reported, and 1.7% of the sample identified as Hispanic/Latino. HCP twin zygosity was determined by genotyping (168 subjects) or self-report (4 subjects), which identified 51 monozygotic (MZ) twin pairs and 34 dizygotic (DZ) twin pairs in the present sample, as well as 2 pairs of non-twin siblings and 4 singletons. All sibling pairs shared the same gender. Out of these 178 subjects, we identified 690 unrelated dyads who were matched in gender and age in years. Because two of these participants (from two separate MZ twin pairs) did not complete every resting state run, analyses involving resting state data use a sample size of $n = 176$.

fMRI data

All fMRI data were collected on a 7T Siemens Magnetom scanner across four sessions spanning multiple days. Each day involved two resting state (900 volumes) and two movie-watching scans (variable durations) across two sessions, all with the following sequence: repetition time (TR) = 1000 ms, echo time (TE) = 22.2 ms, number of slices = 85, flip angle = 45 degrees, spatial resolution = 1.6 mm^3. During movie runs, subjects passively watched short audiovisual clips from either independent films or major motion pictures as well as a montage of brief videos. All clips were only viewed once by each subject, with the exception of the brief montage which was included at the end of each of the four runs for test-retest purposes. These clips differed in their degree of narrative and social content, but many featured language and human characters; more information on the clips can be found at *https://db.humanconnectome.org*. Each video clip was

preceded by 20 seconds of rest, so we discarded all volumes that took place during these rest blocks as well as the first 20 volumes of each clip to prevent rest data and onset transients from biasing our intersubject correlation (ISC) measurements. Rest and movie data from the same day were normalized and concatenated, yielding one rest run (1800 volumes) and one movie run (1432 volumes Day 1, 1409 volumes Day 2) for each day of data collection.

Preprocessing and parcellation

The fMRI data used here were preprocessed as described in a previous publication (Gruskin & Patel, 2022). Briefly, we used ICA-FIX denoised data from which the global signal and its temporal derivative were removed. Our use of global signal regression (GSR) was motivated by work showing that this approach effectively reduces the impact of nuisance signals on FC measures (Parkes et al., 2018) and increases relationships between measures of brain function and behavior (Li et al., 2019) across individuals (presumably by making individual FC profiles more distinguishable), both of which should highlight heritable aspects of brain function. Still, it is important to consider that GSR remains a controversial preprocessing step and may affect our results. As such, we repeated our main BOLD time course heritability analysis without using GSR. To examine the effects of parcellation resolution, data were parcellated using the 10 resolutions of the Schaefer atlas (100 to 1000 parcels; Schaefer et al., 2018). The ICA-FIX data downloaded here were aligned across subjects using the HCP's Multimodal Surface Matching (MSMAll, henceforth "MSM") method, which registers data based on several multimodal properties in a topology-preserving manner (Feilong et al., 2021; Robinson et al., 2014).

Intersubject correlation (ISC)

Dyadic ISC analyses were used to quantify BOLD time course similarity between all pairs of participants. For each vertex or parcel, each participant's BOLD signal time course from a given day's movie-watching scan was normalized and (Pearson) correlated with the corresponding BOLD signal time courses from all other participants to yield an ISC matrix. We chose to use this pairwise ISC method over a leave-one-subject-out approach because it allows us to capitalize on the information contained in the n^2 pairwise ISC matrix (whereas the other approach averages out meaningful information to yield an $n \times 1$ ISC matrix). All Pearson r values in this and all other analyses were Fisher z-transformed before averaging (and converted back to Pearson r for visualization). This analysis was intended to illustrate group-level dyadic BOLD time course similarity in familiar units, not to test for significant group difference effects, as individuals may have contributed to multiple dyads such that each dyad was not independent. As such, we reserved hypothesis testing for our formal heritability analyses (below).

Functional connectivity (FC)

We constructed resting state functional connectivity (rest FC) and movie-watching functional connectivity (movie FC) matrices by (Pearson) correlating the time courses of all parcel pairings from the 400-parcel Schaefer atlas, using data from the two concatenated rest scans. These matrices were produced for each subject and for each day of data collection. We applied the same procedure to the movie-watching data, enabling us to compare the heritability of rest and movie FC profile similarities and strengths.

To evaluate the similarity of FC profiles for each combination of the 17 networks defined by Kong et al. (2019), we first vectorized the FC matrices of each subject. We then extracted the correlation coefficients corresponding to the parcel-level connections comprising each Kong network combination and computed correlations for these vectorized profiles across all subject pairs. For example, there are 16 parcels in the Kong et al. Auditory network and 17 parcels in the Language network, so the FC profile for a given subject's Auditory-Language network combination consists of the 272 (16*Auditoryparcels* × 17*Languageparcels*) correlation coefficients between all unique pairs of one parcel from each network. We also assessed subject-level FC strengths for each network combination by averaging the correlation coefficients for all parcel-level connections

within a network combination. As above, this analysis was intended to illustrate twin-twin similarity in familiar units, and given dyadic interdependence issues we reserved hypothesis testing for our formal heritability analyses (next section).

ISC and FC profile heritability analyses

BOLD time courses and FC profiles are high-dimensional variables, and reducing their dimensionality in order to use classical heritability analyses would sacrifice both statistical power and interpretability. As such, we quantified their heritability with a multidimensional estimator that has been used in similar studies (Anderson et al., 2021; Busch et al., 2023; Ge et al., 2016). This model (detailed in Anderson et al., 2021) takes as input a Subjects × Subjects kinship matrix describing the degree of genetic relatedness between individuals (1 for MZ twins, 0.5 for non-MZ siblings, 0 for all other pairs) as well as a Subjects × Subjects phenotypic similarity matrix. Here, each value of the phenotypic similarity matrix corresponded to a Pearson correlation coefficient describing either BOLD time course (per parcel or voxel) or FC profile similarity (per network combination) for a given subject pair. To estimate heritability, the variance in phenotypic similarity is then partitioned into a component attributable to genetic factors, represented by the kinship matrix, with age, gender, and per-scan head motion included as covariates. Significance testing of individual multidimensional heritability values and calculation of their standard errors were performed using the method established by Anderson et al. 2021. Specifically, the kinship matrix was shuffled 10,000 times to generate a null distribution against which the observed value could be compared. The resulting p-values were then false discovery rate (FDR) corrected using the Benjamini-Hochberg method (Benjamini & Hochberg, 1995) to adjust for multiple comparisons. FDR correction was always performed separately for each day of data collection, such that results reported as "significant at FDR-corrected $P < .05$ on both days" reflect a conservative criteria of $q < .05$ on two independent tests. Standard errors (SEs) were derived through a block jackknife method in which heritability was recalculated 90 times after leaving out all members of one of the 90 families in the dataset on each iteration. We then used these SEs to generate 95% confidence intervals (CIs).

To compare the heritability of movie and rest FC profiles, we used the following non-parametric permutation approach. For each day of data collection, we randomly shuffled each subject's movie and rest FC matrices (along with their corresponding framewise displacement [FD] covariates) and re-calculated FC profile heritability using the shuffled FC matrices. We then subtracted the two resulting heritability values for each of the 153 unique network combinations for the 17 Kong networks (e.g., Auditory-Language, Auditory-Auditory, etc.) and averaged these across networks to obtain 17 values reflecting the null difference in FC profile heritability for each network. We repeated this procedure 10,000 times and then used the two-sided test described above to generate a p-value for each network, and these p-values were then FDR corrected separately for each day of data collection. Because complete resting state datasets were not available for 2/178 movie-watching subjects, we only used movie-watching data from the 176 subjects who also had complete resting state data for this permutation test.

To determine the sample size necessary for stable multidimensional heritability results, we conducted our BOLD time course heritability analysis multiple times while systematically excluding between 5% and 90% of families. At each exclusion level, we performed 100 iterations, each time randomly removing a subset of families. After each iteration, we calculated the absolute difference between the heritability values obtained from the subsample and those from the full sample for each parcel. We then averaged these differences across all parcels and iterations to obtain the mean absolute error for that exclusion level. Similarly, to assess the stability of the spatial pattern of our results, we computed Spearman correlations between the subsample and full sample heritability values across parcels and averaged these correlations across all iterations. We included this analysis because it serves as a simple way to demonstrate the stability of our results at various sample sizes, and because this sort of subsampling approach has been used many times before in our field (e.g., Marek et al., 2022) and others (e.g., Manyara et al., 2024) to demonstrate the sample-size dependence of statistical effects.

Because the heritability of ISC is constrained by the degree of synchronization in a given area, we also sought to identify areas in which BOLD time courses were more/less heritable than would be expected based on ISC alone by fitting a linear model of the form $\text{heritability}_i = \beta_0 + \beta_1 \cdot \text{ISC}_i + \varepsilon_i$ and plotting the residuals.

Frequency-dependent ISC heritability analyses

To characterize the frequency-specific heritability of movie-evoked BOLD time courses, we performed a spectral analysis of the BOLD time series data across all subjects and parcels. First, we set a high-frequency cutoff equal to the Nyquist frequency of our data (0.5 Hz) and defined a low-frequency cutoff at 0.004 (1/238) Hz, corresponding to the length of the longest clip. We then computed the power spectrum for each parcel, subject, and day of data collection by applying a Fast Fourier Transform (FFT) to the time series data of each subject and then averaging these spectra across all subjects, parcels, and scanning days. To ensure consistent comparison across frequencies, we interpolated the power spectra onto a common frequency axis with a resolution of 0.001 Hz. The cumulative power distribution was then calculated from this averaged power spectrum. To partition the frequency range into bands containing equal fractions of the total power, we identified frequency cutoffs corresponding to quintiles of the cumulative power distribution. This resulted in five frequency bands: Band 5 (0.004–0.02 Hz), Band 4 (0.02–0.04 Hz), Band 3 (0.04–0.07 Hz), Band 2 (0.07–0.14 Hz), and Band 1 (0.14–0.50 Hz). For each of these bands, we applied fourth-order Butterworth bandpass filters to the concatenated BOLD time courses of each subject and parcel and recalculated ISC and BOLD time course heritability as described above. Standard errors of cortex-wide heritability were estimated for each band using a leave-one-family-out jackknife procedure. Heritability estimates were first averaged across all brain parcels within each jackknife fold, and the standard error of these parcel-averaged values was computed using the standard jackknife variance formula. To further characterize frequency-dependent changes in heritability at the parcel level, we then Spearman-correlated the spatial patterns of heritability with the sensorimotor-association hierarchy rankings from Sydnor et al. (2023). Significance testing for these sensorimotor-association results was performed using BrainSMASH (Brain Surrogate Maps with Autocorrelated Spatial Heterogeneity; Burt et al., 2020).

FC strength heritability analysis

Because FC strengths are inherently one-dimensional traits, their heritability was quantified with SOLAR's *polygenic* function (Almasy & Blangero, 1998), which generates estimates using variance-component models. Age, gender, and head motion were used as covariates in all FC strength analyses, and SEs were calculated using the block jackknife procedure described above. We note that because the heritability of univariate (vs. multivariate) traits requires larger sample sizes, these results should be considered preliminary and would benefit from further investigation in larger samples (Anderson et al., 2021). To test the significance of differences in rest vs. movie FC strength heritability, we compared the observed differences to null distributions generated by 1,000 permutations of the same procedure described above for FC profiles.

Hyperalignment

We used piecewise response and connectivity hyperalignment (RHA and CHA, respectively), two complementary methods for aligning data into a topography-independent common space, to functionally align vertex-level BOLD time courses across subjects (Guntupalli et al., 2018; Haxby et al., 2011). We decided to use piecewise hyperalignment, in which vertices are aligned within non-overlapping parcels, instead of searchlight hyperalignment because it has been shown to be both more accurate and more efficient (Bazeille et al., 2021). This approach independently transforms each subject's data within discrete anatomical parcels into the common space, yielding functionally aligned vertex time series that are calculated as weighted linear combinations of the original time series from all other vertices within that same parcel for that subject. We then repeated our BOLD time course and FC profile heritability analyses using these hyperaligned datasets to quantify the extent to which brain response heritability reflects genetic control over

cortical topography. *RHA:* For each Schaefer atlas parcel and day of data collection, we used iterative Procrustes transformations to align vertex-level BOLD time courses to a common model information space. This yielded one invertible transformation matrix per parcel and per subject, which we then used to project data from the other day of data collection (which was not used to generate the transformation matrices) into the common information space. *CHA:* The same iterative Procrustes approach was used for CHA, but here the input data consisted of rest FC profiles for each vertex within a given parcel. Each vertex's functional connectivity profile consisted of the Pearson correlation coefficients between that vertex's time course and the average time courses from all other parcels, the number of which varied with different parcellation resolutions. After training a model whose dimensions correspond to shared rest FC properties (instead of the shared response properties in RHA), we used the corresponding transformation matrices to align movie-watching time series data from the other day of data collection. To quantify the spatial scale at which cortical topographies contribute to brain response heritability, we repeated the RHA and CHA procedures described above for each of the 10 Schaefer atlas resolutions to yield 21 datasets per subject (10 RHA-aligned datasets, 10 CHA-aligned datasets, and the original MSM-aligned dataset).

Relationships between parcel area and heritability

We used power law modeling to characterize the relationship between cortex-level heritability and hyperalignment area. To calculate parcel areas, we first generated vertex-level areas using the *-surface-vertex-areas* function in wb_command and then summed the areas of all vertices included in each parcel. Next, for each hyperalignment method and each day of data collection, we used nonlinear least squares regression to fit a power law model ($y = a \cdot x^b + c$) to the 11 heritability values calculated from the 10 Schaefer atlas resolutions and one from MSM-only aligned data and the 10 average Schaefer atlas parcel areas as well as 0 (corresponding to no hyperalignment in the MSM-only data).

Neural timescale (NT) analyses

To determine the extent to which BOLD time course heritability reflects genetic control over NT, we took an approach used in numerous studies to calculate NT at rest (known as intrinsic neural timescale, or INT; Watanabe et al., 2019; Wengler et al., 2020) and applied it to movie-watching data. NTs were calculated separately for each day of data collection as the sum of the autocorrelation coefficients from the first lag until the first lagged timepoint with a non-positive autocorrelation coefficient (Wengler et al., 2020) for each vertex.

Because time courses that are themselves more temporally autocorrelated will have a higher variance in their correlations with each other (Shinn et al., 2023), and because stimulus-evoked BOLD time courses tend to be positively correlated across subjects, we reasoned that pairs of individuals with longer collective NTs would have more correlated BOLD time courses. To test this directly, we (Spearman) correlated ISC values from one day of data collection with NTs calculated using the other day's data across all possible subject pairs. Because our hypothesis here was unrelated to genetic effects, we tested the significance of the resulting vertex-wise correlation coefficients using a family-compliant quadratic assignment procedure with 1,000 permutations. To account for familial dependencies, we shuffled family units before shuffling individuals within those families. Two-tailed p-values were defined as the proportion of null correlations exceeding the observed coefficient. Finally, we applied FDR correction across all vertices for each day separately, considering results significant only if they were consistent across both independent scan sessions.

We then averaged NT values across all vertices to get a single, cortex-wide NT measure for each subject and each day of data collection. Dyadic NT similarity was quantified as the absolute value of the difference between each subject pair's cortex-wide NT values. This differencing approach yielded several extreme values. We thus used the 1.5×IQR method to identify 4% of dyadic NT values as outliers and excluded them from the group differences analyses. Because these dyads are not independent (as discussed above), we used SOLAR to quantify the heritability of whole-

brain NT (averaged across all vertices) and test for statistical significance, controlling for age, gender, and head motion. We also used the aforementioned multidimensional heritability analysis to quantify the heritability of parcel-level NT topographies (where the phenotypic similarity matrix reflects the Spearman correlation of each pair's 400 × 1 parcel-level NT values).

We also included subject-level NT as a covariate in some multidimensional heritability analyses. When calculating heritability at a given vertex for one day of data collection, the NTs for that vertex from the other day of data collection were used as covariates. To evaluate whether including NT as a covariate decreased BOLD time course heritability, we generated a null distribution of 1,000 heritability values by randomly shuffling vertex-level NT vectors such that the NTs for one subject were paired with the ISC values and covariates from another subject. We then calculated two-sided permutation p-values.

Significance testing for autocorrelated brain maps and FC matrices

We used Spearman correlations to quantify the reliability of ISC and FC heritability maps across days, as well as relationships between these heritability maps and the sensorimotor-association hierarchy ranking from Sydnor et al. (2023). Because cortical spatial maps are significantly autocorrelated, we used the variogram matching approach from BrainSMASH to assess the significance of these correlations. Briefly, this test works by generating autocorrelation-matched surrogates for one of the empirical maps from each correlation, calculating Spearman correlations between these surrogates and the other empirical map, and then comparing these null correlations to the correlation between both empirical maps. We performed independent tests using 1,000 unique surrogates for each hemisphere and averaged the two p-values to get the whole-cortex p-values we report in this manuscript. Because values in FC heritability matrices are similarly not independent of each other, we used Mantel tests with 10,000 permutations to test their reliability.

Results

Similarity in movie-evoked brain activity increases with genetic relatedness

To characterize the heritability of brain responses to complex stimuli, we used 7T fMRI data from 178 HCP Young Adult subjects acquired across two days (using two largely non-overlapping sets of movie stimuli, see Methods) to (1) quantify the heritability of brain activity and connectivity patterns during movie-watching and (2) determine the extent to which the heritability of these dynamic, high-dimensional brain responses is grounded in stable and fundamental aspects of brain function like cortical topographies and neural timescales.

We first aimed to determine how closely movie-evoked brain responses were shared among pairs of individuals, and whether this similarity was influenced by their genetic relationship. Using inter-subject correlation (ISC) of parcellated BOLD time courses to index brain activity similarity, we found that identical (or monozygotic, MZ), fraternal (or dizygotic, DZ) and age- and gender-matched unrelated (UR) dyads differed in their level of brain activity similarity in a manner consistent with their relative degrees of genetic relatedness (although spatial distributions of ISC were consistent across groups, Fig. S1). More specifically, identical twins' BOLD time courses were 59% more similar than those from pairs of unrelated individuals. When comparing identical twins to fraternal twins, the identical twins' brain activity was still more similar, but by a smaller margin of 24%. Finally, fraternal twins had 29% more similar time courses than unrelated pairs. We observed the greatest group differences in parcels with medium levels of ISC (separation between the three traces in Fig. 1B, visualized as scatterplots in Fig. S2), suggesting that floor and ceiling effects may limit the degree to which genetic relatedness impacts brain activity in regions that are not driven by audio-visual stimuli and regions that exhibit highly stereotyped activity across all subjects, respectively.

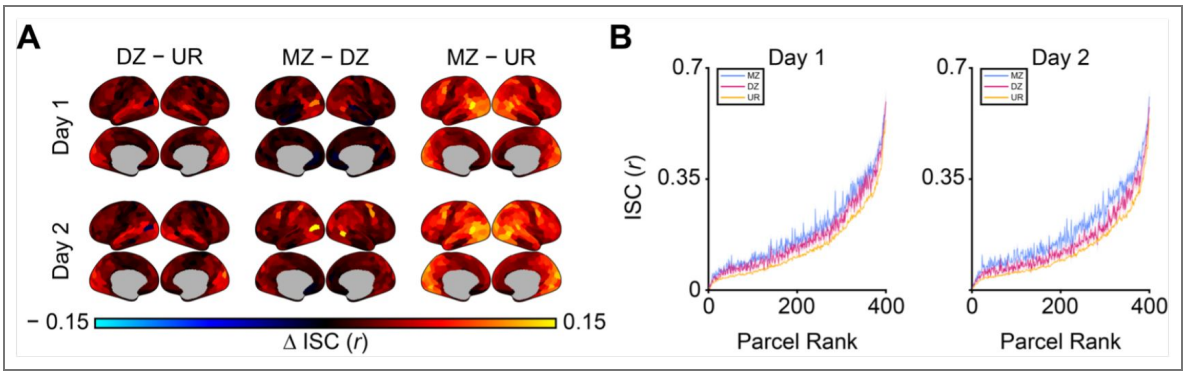


Figure 1. BOLD time course similarity scales with genetic relatedness across the cortex.

(A) Group differences in average BOLD time course similarity (indexed by ISC) show that BOLD time course similarity is greater among dyads who are more genetically related (51 MZ dyads, 34 DZ dyads, 690 UR dyads). Here, top and bottom rows reflect data acquired on different days of data collection while subjects viewed largely non-overlapping sets of movie clips. (B) Group-average ISC values used to create the difference maps in A, plotted in order of average ISC across all subject pairs, show that group differences are most pronounced in parcels with medium to high ISC (shading = SEM).

Patterns of brain activity during movie-watching are heritable

After establishing that more genetically similar individuals share more similar movie-evoked BOLD time courses, we next sought to quantify the heritability of these brain responses. To do this, we leveraged a multidimensional estimator that has been used to assess the heritability of similar brain phenotypes (Anderson et al., 2021; Busch et al., 2023; Ge et al., 2016). Controlling for age, gender, and head motion, we found that movie-evoked BOLD time courses were heritable across almost all of cortex on both days of data collection (Fig. 2A, Day 1 mean h^2 = .064 ± .034 (SD), Day 2 mean h^2 = .068 ± .036, 99% of parcels significant on both days at FDR-corrected P < .05; dorsal/ventral views in Fig. S3A), and the spatial pattern of heritability across the cortex was very consistent across days of data collection (Spearman ρ = .96, $P_{\text{BrainSMASH}}$ < .001). We note that we observed nearly identical results when we repeated this analysis without global signal regression (GSR; Fig. S4). Although heritability studies of one-dimensional traits (e.g., height) tend to require larger samples than the one used here, the multidimensional nature of our analysis affords us considerable power to detect small effects even with our relatively modest sample size (Anderson et al., 2021; Ge et al., 2016). To illustrate this point, we repeated our BOLD time course heritability analysis after excluding up to 90% of families in the present dataset and found that even after excluding half of our subjects, the average difference in h^2 magnitude across parcels and between each subsample and the results reported above was less than .01, and the average spatial correlation (Spearman ρ) between subsample and full sample heritability values was greater than .9 (Fig. S5).

Unsurprisingly, the spatial pattern of BOLD time course heritability was closely related to the spatial pattern of ISC (Day 1: Spearman ρ = .88, $P_{\text{BrainSMASH}}$ < .001, Day 2: Spearman ρ = .86, $P_{\text{BrainSMASH}}$ < .001), reflecting the simple fact that the heritability of movie-evoked BOLD time courses will be lower in parcels with less movie-driven activity to begin with. To characterize the heritability of BOLD time courses relative to the amount of movie-driven activity in each parcel, we regressed parcel-level ISC values (averaged across subject pairs) from heritability values and plotted the residuals in Fig. 2B (dorsal/ventral views in Fig. S3B). Here, negative values in the residual map indicate parcels where heritability is lower than expected based on ISC, while positive values indicate higher-than-expected heritability. Regarding alternative approaches to controlling for ISC, although the heritability model introduced by Ge et al. (2016) allows for the inclusion of covariates defined at the subject level (e.g., age), it does not allow for covariates that are defined at the dyad level (e.g., pairwise ISC). We observed that BOLD time courses were disproportionately more heritable in more associative lateral prefrontal and temporo-parieto-occipital junction parcels, while responses in lower-level auditory areas were less heritable than would be expected given their ISC. This indicates that although these more associative parcels do not encode more stimulus-specific information than unimodal sensory parcels, what information they do encode and/or how they encode it is under increased genetic control compared to auditory parcels.

After establishing that movie-evoked BOLD time courses are heritable, we next sought to determine the extent to which this heritability reflects genetic control over high-vs. low-level sensory processing. To do this, we lever-aged the fact that low-level features of movie stimuli (e.g., visual motion and speech) tend to oscillate on the order of seconds (or faster), whereas higher-level aspects of the stimulus (e.g., social content and narrative structures) are encoded at lower frequencies (Baldassano et al., 2017; Honey et al., 2012; Kauppi et al., 2010). Similar to previous work on frequency-specific ISC (Kauppi et al., 2010), we filtered our data into five non-overlapping frequency bands, each containing an equal proportion of the total spectral power, and generated overall and residualized (with respect to ISC) heritability maps for each band (Fig. 3A and D). We observed that cortex-wide BOLD time course heritability increased monotonically with the period of the frequency band, such that heritability was over 50% higher in the slowest frequency band (0.004–0.02 Hz) compared to the unfiltered data (Fig. 3B and E). This suggests that genetic factors influence the neural processing of complex audio-visual features, and that this influence is greater than for lower-level sensory features. Interestingly, we also observed that both overall and residualized heritability were considerably lower in the one supra-

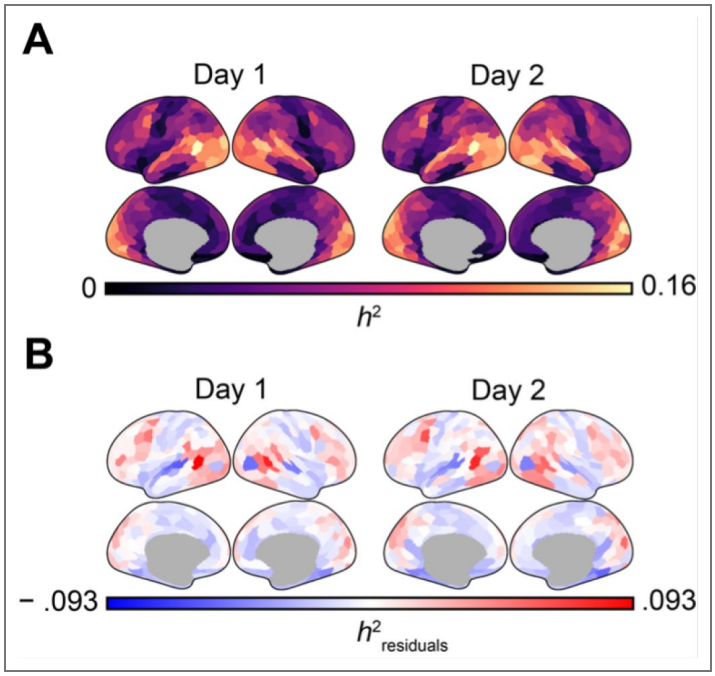


Figure 2. **BOLD time courses are heritable across the cortex.**

(A) Cortical surfaces show heritability of BOLD time courses parcellated using the Schaefer 400 atlas, controlling for age, gender, and head motion (mean h^2 Day 1 / Day 2 = .064 ± .034/ .068 ± .036). (B) Residuals after regressing parcel-level ISC from parcel-level heritability show that BOLD time courses in auditory cortices are less heritable than would be expected based on ISC, whereas the opposite is true for lateral prefrontal and temporo-parieto-occipital junction parcels.

BOLD frequency band (0.14–0.5 Hz, faster than the frequency of BOLD signals resulting from neuronal firing events; Josephs and Henson, 1999, Fig. 3A/D second column from the left) compared to the unfiltered data. This indicates that although there is synchronized high-frequency information in our data (possibly due to aliased cardiovascular and respiratory signals; Pérez et al., 2021), this information is largely not heritable and further supports a BOLD etiology for the heritability results shown above.

Previous studies have shown that during movie-watching, more associative regions process abstract information at longer timescales that range from tens of seconds to minutes, whereas sensory areas encode lower-level features at higher frequencies (Baldassano et al., 2017; Hasson et al., 2008; Honey et al., 2012). As such, we hypothesized that the higher heritability we observed in slower frequency bands was driven by increased heritability in associative (vs. sensory) parcels. To test this hypothesis, we correlated parcel-level differences in heritability between the slowest and fastest BOLD-sensitive frequency bands with sensorimotor-association hierarchy rankings from Sydnor et al. (2023; higher ranking = more associative) and found that heritability increases from the fastest to slowest frequency band were indeed larger for more associative parcels (Day 1: Spearman ρ = .47, $P_{\text{BrainSMASH}}$ < .001, Day 2: Spearman ρ = .35, $P_{\text{BrainSMASH}}$ < .001; Fig. 3C and F). Because removing rest and onset blocks from each clip and concatenating the two movie-watching runs from each day introduced temporal discontinuities that could impact our filtering results, we reran our analyses using the original, uncensored time courses and observed similar results (Fig. S6). We chose to initially analyze BOLD time courses parcellated using the Schaefer 400 atlas because parcellation reduces multiple comparisons, noise, and computational burden. However, we repeated our heritability analyses using data parcellated with 9 other resolutions of the Schaefer atlas and found that heritability reliably increased with average parcel size (Fig. S7). Moreover, the interpretation of parcellated BOLD time course heritability is complicated by the fact that macroscale areal boundaries are known to be heritable (Xu et al., 2016). As such, we use vertex-level data in our subsequent BOLD time course analyses.

Heritable movie-evoked BOLD time courses reflect heritable cortical topographies

Our analyses of data aligned using standard anatomical methods (i.e., MSM) have demonstrated that patterns of movie-evoked brain activity and connectivity are heritable. Importantly, these patterns reflect two distinct and fundamental aspects of brain function: *how* stimuli are processed and *where* stimuli are processed. For example, when analyzing brain responses in a dorsal brain region (as shown in Fig. 4A), two twins (left and center) may appear to process stimulus information more similarly than an unrelated individual (right), based on having higher ISC values for that region. However, this apparent disparity arises purely from spatial differences in where the same information is processed: the unrelated individual in fact exhibits the same functional responses as the twins (i.e., the green and purple time courses), just in different cortical locations. These individual-specific maps of how shared brain functions are spatially distributed are known as cortical topographies (Haxby et al., 2011, 2020), and recent work has shown that cortical topographies defined at rest are influenced by genetic factors (Anderson et al., 2021; Burger et al., 2022; Busch et al., 2023). Therefore, we hypothesized that part of the heritability observed in our previous analyses might reflect genetic control over cortical topography (or “where” information is processed), in addition to genetic influences on information processing itself.

One effective method used to separate the topography-dependent and topography-independent aspects of cortical information processing is known as hyperalignment. Hyperalignment aligns individual brains into a common high-dimensional functional space based on shared functional responses during the same task or stimulus (Haxby et al., 2011). By aligning fMRI data across subjects into a topography-independent functional space, hyperalignment yields datasets that allow for a direct comparison of how information is processed across individuals, independent of individual differences in where that information is processed.

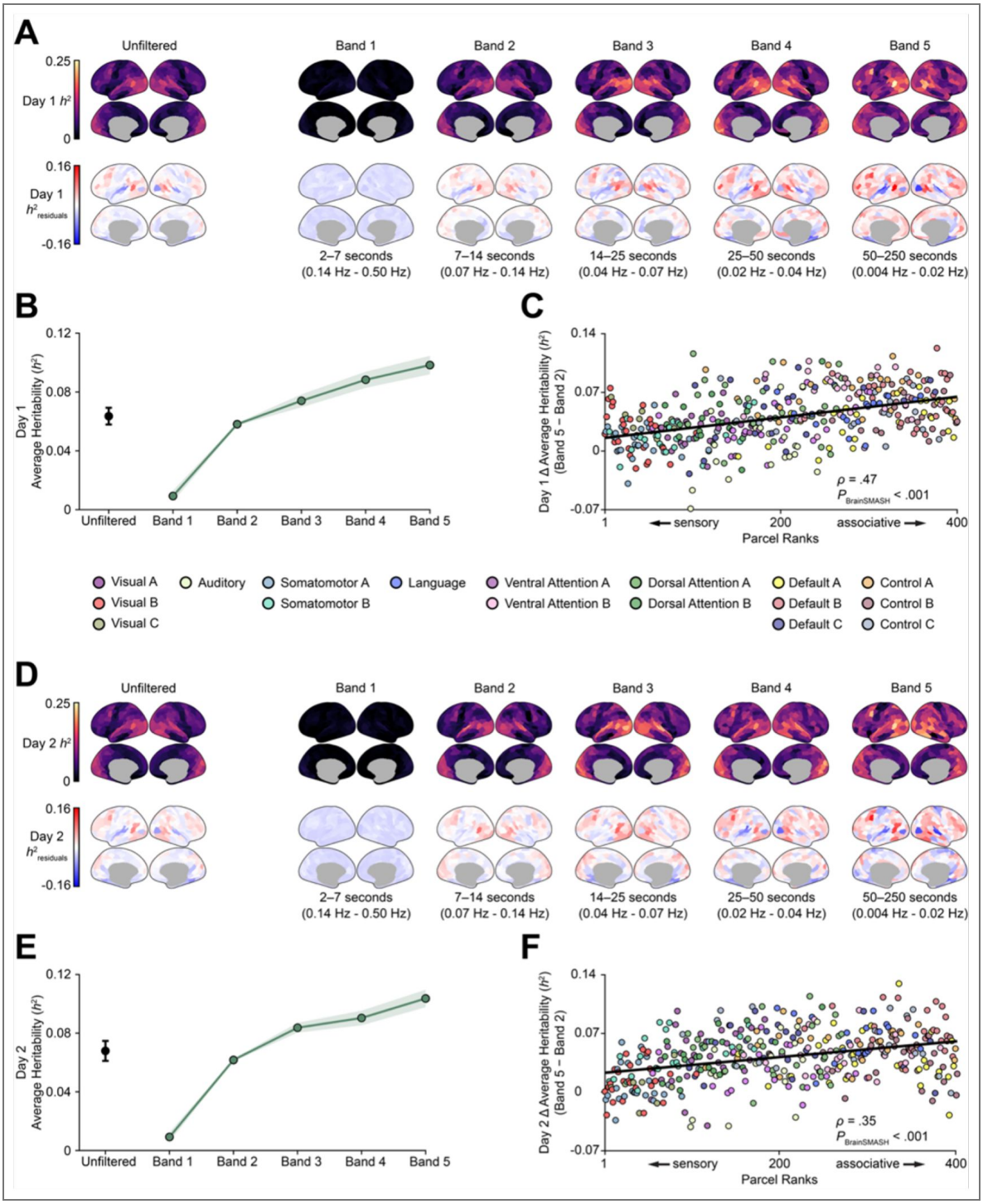


Figure 3. BOLD time course heritability is greater in slower frequency bands, especially for more associative parcels.

(A) Purple/yellow cortical surfaces (upper row) show unfiltered BOLD time course heritability (upper left is identical to Fig. 2A) as well as the heritability of BOLD time courses filtered with five frequency bands, with greater heritability in slower bands for Day 1 data. Red/blue cortical surfaces show BOLD time course heritability residuals after regressing out parcel- and frequency-level differences in ISC (lower left is identical to Fig. 2B), with greater residuals in slower frequencies and more associative parcels. (B) Scatter plot shows heritability averaged across the cortex for each frequency band (i.e., the averages of the upper row of surfaces in A; shading = jackknife SEM). (C) Scatter plot shows the difference in heritability between the slowest and fastest BOLD-sensitive frequency bands for each of the Schaefer 400 parcels plotted against parcel ranks from the Sydnor et al. sensorimotor-association hierarchy (higher = more associative). Least squares lines were added to highlight the positive relationships between average h^2 and parcel ranks but note that these relationships were formally tested with Spearman correlations. (D–F) Same as A–C for Day 2 data.

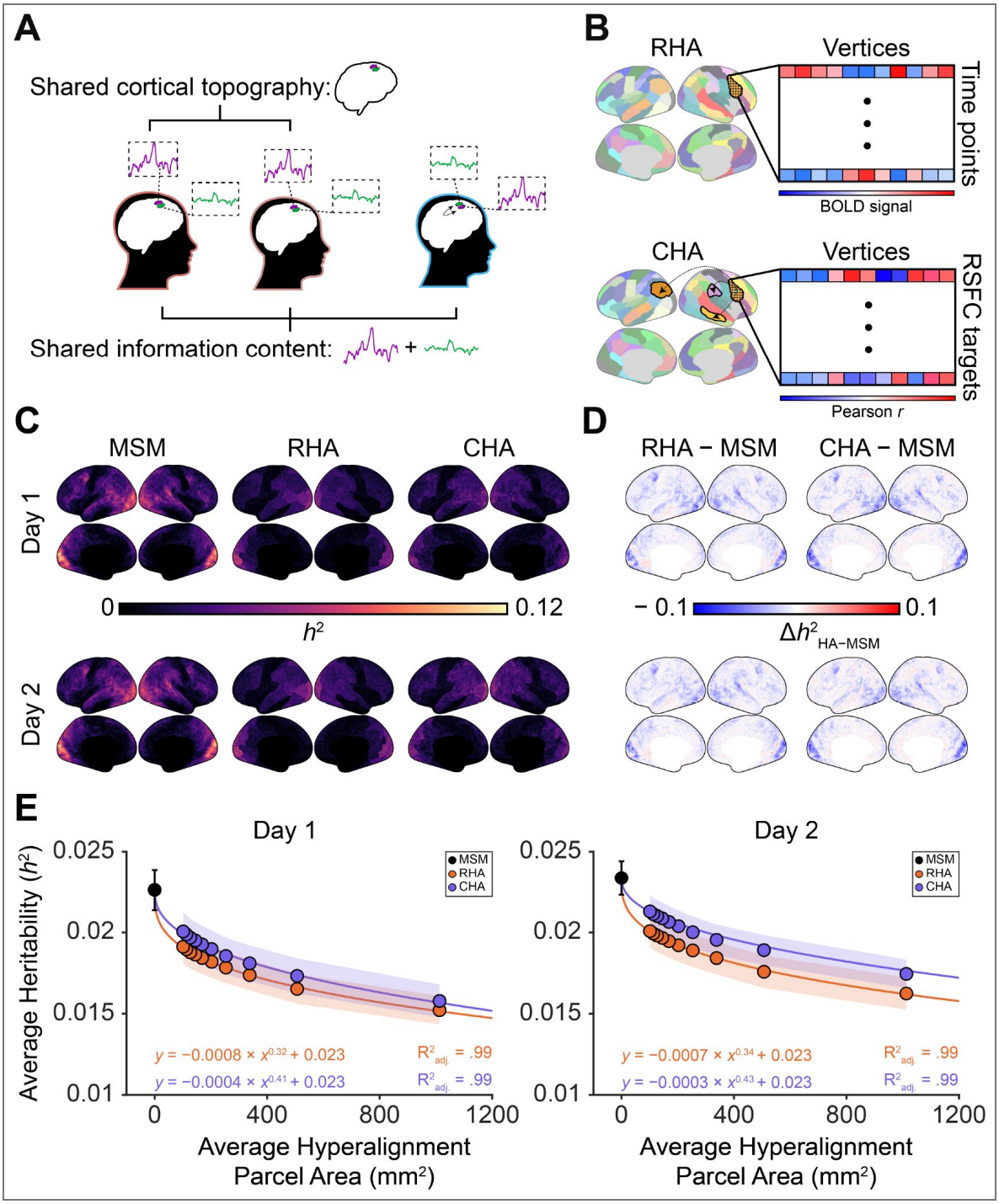


Figure 4. Hyperalignment reduces BOLD time course heritability.

(A) Cartoon illustrates the difference between shared cortical topographies and shared (topography-independent) information content. (B) Diagrams illustrate the inputs to response and connectivity hyperalignment (RHA and CHA, respectively) using the Schaefer 100 atlas. RHA topographies were learned using BOLD time course data from the other day's movie-watching scans, while CHA topographies were learned from vertex-level FC profiles (i.e., correlations between one vertex's BOLD time course and the average time course from each of the 99 other parcels) calculated from the other day's resting state scans. (C) Vertex-level BOLD time course heritability is highest for data aligned via MSM (multimodal surface matching) and lower for data hyperaligned within 100 Schaefer atlas parcels using both response hyperalignment (RHA) and connectivity hyperalignment (CHA). (D) Differences between the MSM-only and hyperaligned heritability maps shown in (C) are distributed across the cortex but are most apparent in visual areas. (E) BOLD time course heritability decreases as a function of hyperalignment parcel size according to a power law (purple and orange lines); each dot corresponds to average cortex-wide heritability for data hyperaligned using one of the 10 Schaefer atlas resolutions (shading = jackknife SEM).

In this study, we used hyperalignment to quantify the extent to which brain response heritability reflects genetic control over *how* vs. *where* information is processed. To hyperalign each subject's movie-watching data to a common functional space for a given day of data collection, we used idiosyncratic transformation matrices that were learned from either the other day's movie activity time courses (response hyperalignment, or RHA) or from rest FC profiles calculated from the other day's scans (connectivity hyperalignment, or CHA). Although RHA and CHA align fMRI data with similar fidelity (Guntupalli et al., 2018; Haxby et al., 2020), using both methods allows us to evaluate whether heritable functional topographies reflect the brain's intrinsic functional architecture or movie watching-specific response functions. We performed both RHA and CHA in a piecewise fashion (Bazeille et al., 2021), aligning vertex-level data within individual Schaefer parcels—in other words, a vertex in one Schaefer parcel would be aligned with other vertices in that parcel and never with vertices from other parcels.

Hyperalignment aligns vertices with similar functional responses across subjects, inherently increasing ISC across subject pairs (Fig. S8). However, to the extent that cortical topographies are under genetic control, twins' brains are intrinsically more aligned than those of unrelated individuals. Therefore, we predicted that hyperalignment would decrease observed heritability across the cortex by eliminating the topography-dependent component of heritability and increasing response similarity more in unrelated dyads than in twin pairs.

Starting with the coarsest Schaefer atlas resolution (100 parcels, mean parcel area = 1,013 mm^2), we found that RHA and CHA significantly decreased BOLD time course heritability to similar degrees across the cortex. Compared to MSM-aligned data (Fig. 4C, left column), hyperalignment reduced BOLD time course heritability across the cortex by 33% on Day 1 (95% CI = [25–40%]) and 31% on Day 2 [22–39%] for RHA (Fig. 4C, middle column), and by 30% [21–39%] and 25% [18–33%] for CHA (Fig. 4C, right column; dorsal/ventral views in Fig. S9). These decreases were most apparent in visual cortex, but were also prominent in associative areas like the right temporoparietal junction and bilateral area 55b (Fig. 4D), and the spatial pattern of this effect was consistent across days (RHA: Spearman ρ = .66, $P_{BrainSMASH}$ < .001, CHA: Spearman ρ = .69, $P_{BrainSMASH}$ < .001) and hyper-alignment methods (Day 1: Spearman ρ = .76, $P_{BrainSMASH}$ < .001, Day 2: Spearman ρ = .74, $P_{BrainSMASH}$ < .001).

To quantify the spatial scale at which cortical topography influences BOLD time course heritability, we then repeated RHA and CHA using the 9 other Schaefer atlas resolutions (200 to 1000 parcels). Because hyperalignment can eliminate more heritable differences in cortical topography when it is performed in larger parcels, we predicted that hyperalignment across larger parcels would decrease BOLD time course heritability to a greater extent. As expected, the magnitude of these reductions decreased as hyperalignment was performed across smaller areas (Fig. 4E). To quantify this relationship, we fit power law models of the form $y = a \cdot x^b + c$ to the 11 hyperalignment resolutions (corresponding to the average parcel areas for the 10 Schaefer atlases as well as 0 for no hyperalignment) and their corresponding average h^2 values. We found that these power law models accurately characterized how heritability scaled with hyperalignment resolution for both RHA and CHA on Day 1 (RHA: $y = -0.0008 \cdot x^{0.32} + 0.023$, $R^2_{adj.}$ = .99, CHA: $y = -0.0004 \cdot x^{0.41} + 0.023$, $R^2_{adj.}$ = .99) and Day 2 (RHA: $y = -0.0007 \cdot x^{0.34} + 0.023$, $R^2_{adj.}$ = .99, CHA: $y = -0.0003 \cdot x^{0.43} + 0.023$, $R^2_{adj.}$ = .99), whereas linear, quadratic, and logarithmic models performed worse (Fig. S10).

Heritability of BOLD time courses is related to neural timescales

In the previous section, we found that individual differences in a stable aspect of brain function (cortical topography) accounted for 30–40% of the heritability of movie-evoked brain responses. Importantly, cortical topography is a largely spatial trait. Although it captures significant inter-individual variability in how brain function varies over space, it is not directly related to how brain responses evolve over time. As such, we reasoned that the heritability of high-dimensional brain responses might also be grounded in the heritability of stable, temporal properties of brain function.

One such property is the neural timescale (NT), which is thought to index the duration of information storage in a given circuit or region. Across the cortex, more associative areas are known to have longer NTs (which is consistent with our frequency-dependent heritability results in Fig. 3), but substantial and behaviorally-relevant variability in NTs also exists across individuals, such that individuals with longer NTs in a given region integrate sensory information across longer periods of time (Wengler et al., 2020). NTs are commonly operationalized as the area under the curve of the autocorrelation function (ACF) until the lag preceding the first negative ACF value (Wengler et al., 2020). Because stimulus-evoked time courses that are more autocorrelated will tend to be more correlated with each other (see Methods), we suspected that NTs could be an important determinant of ISC. To test this, we correlated the sum of each dyad's NTs from one day's movie-watching scan with their ISC from the other movie-watching scan and found that we could explain a considerable portion of the variability in pairwise ISC from NT alone (max/mean correlation Day 1: ρ = 0.56/0.10, Day 2: ρ = 0.65/0.11, 54% of vertices significant at FDR-corrected P_{perm} < .05 on both days, Fig. S11). Given this relationship between NT and ISC as well as the fact that a similar measure of BOLD autocorrelation was recently shown to be heritable (Christova et al., 2022), we hypothesized that some of the BOLD time course heritability not accounted for by topography is underpinned by heritability of NT.

Before testing this hypothesis directly, we first sought to establish that NT magnitude and topography themselves are heritable. After calculating NTs at each vertex from each day's movie-watching data, we averaged these values to generate a single cortex-wide NT (i.e., global NT) for each subject and each day of data collection. We then examined pairwise differences in global NT across the three groups introduced earlier: MZ twins, DZ twins, and age- and gender-matched UR dyads. As expected, we found that differences in MZ twins' global NTs were smaller than in unrelated individuals' NTs on both days (Fig. 5A, Day 1: ΔNT_{MZ} = 0.14 ± .015 (SD), ΔNT_{DZ} = 0.18 ± .022, ΔNT_{UR} = 0.23 ± .007; Day 2: ΔNT_{MZ} = 0.15 ± .013, ΔNT_{DZ} = 0.24 ± .032, ΔNT_{UR} = 0.26 ± .007). We then used SOLAR to formally quantify the heritability of this trait, finding that global NT was heritable on both Day 1 (h^2 = 0.38, P = .003) and Day 2 (h^2 = 0.75, $P = 3.7 \times 10^{-9}$). We note that the discrepancy between h^2 effect sizes on Day 1 and Day 2 reflects decreased power to resolve the heritability of one-dimensional traits, as discussed above. Finally, we used a multidimensional heritability analysis to establish the heritability of parcel-level NT to-pographies, finding that NT topographies (or the scale-invariant pattern of NTs across 400 cortical parcels) was heritable at h^2 = 0.39 on both Day 1 and Day 2 (P_{perm} < .0001 on both days).

To quantify the extent to which BOLD time course heritability reflects genetic control over NT, we repeated the heritability analyses from earlier for each day of data collection, this time including vertex-level NTs calculated from the other day's data as co-variates (in addition to age, gender, and head motion). We observed that controlling for NT significantly reduced BOLD time course heritability on both days in 5.5% of vertices (Fig. 5B upper row, FDR-corrected P_{perm} < .05 separately for each day; raw pre-hyperalignment voxel-level heritability maps available in Fig. S12), most prominently in speech and language areas (e.g., auditory/superior temporal cortices, area 55b) and motion-sensitive visual areas (e.g., medial temporal and medial superior temporal cortices). Although these reductions in heritability were more focal than those observed following hyperalignment, they reached similar strengths, with decreases of 20–30% observed throughout the superior temporal gyri on both days of data collection.

After establishing that cortical topographies and neural timescales both contribute to the heritability of high-dimensional brain responses, we next sought to determine if these contributions are independent of one another. To test this, we recalculated BOLD time course heritability following RHA using the Schaefer 100 parcellation (the most aggressive hyperalignment approach from the previous section), this time controlling for NT calculated from these RHA-aligned data. If cortical topography and neural timescale constitute separable processes through which genetics shapes brain responses, we would expect controlling for NT in the RHA- and MSM-aligned data to reduce brain response heritability to similar degrees. Instead, we observed a mixed result: although controlling for NT after hyperalignment further reduced BOLD time course heritability on both days in 1.5% of vertices (FDR-corrected P_{perm} < .05 separately for

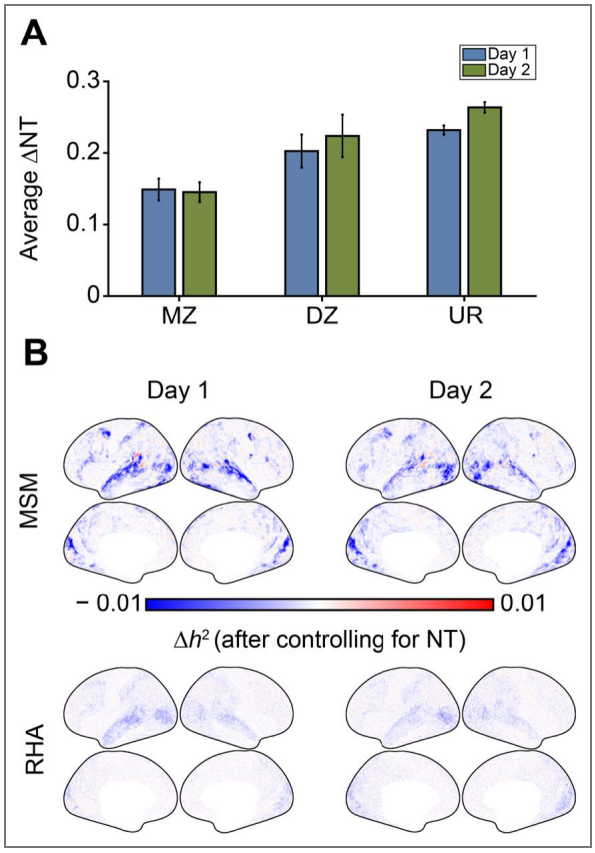


Figure 5. Controlling for neural timescale (NT) reduces heritability of BOLD time courses.

(A) Bar plots show average pairwise differences in cortex-wide NT across MZ, DZ, and UR dyads on both days of data collection. (B) Cortical surfaces show decreases in BOLD time course heritability after NTs calculated from the other day of data collection were included as covariates in the multidimensional heritability analyses for MSM-aligned and RHA-aligned (using the Schaefer 100 parcellation) data, most prominently in mid-level auditory and visual regions. These maps are thresholded at $\Delta h^2 = \pm 0.01$ to aid comparisons of MSM- and RHA-aligned results. The maximum differences in h^2 after controlling for NTs were −0.025 for MSM-aligned data and −0.007 for RHA-aligned data, respectively.

each day, Fig. 5B, lower row), the average magnitude of this decrease across the cortex was 40% smaller than for the MSM-aligned data. This suggests that cortical topography and neural timescale each account for some unique variance in brain response heritability, but their contributions are not entirely independent.

Movie-evoked FC profiles are heritable and reflect heritable cortical topographies

Thus far, we have demonstrated that movie-evoked BOLD time courses are heritable, and that this heritability is related to genetic control over stable spatial and temporal aspects of brain function. However, sensory information is encoded and processed not just in the activities of single regions but also in the functional connectivity (FC) between multiple regions (Chen et al., 2014). To determine if this other kind of movie-evoked brain response is similarly heritable, we repeated our analyses using movie-watching FC (movie FC) profiles. Here, a given individual's movie FC profiles were calculated for each pair of 17 Kong networks and each day of data collection as their set of movie FC values for all connections between parcels in that pair of networks.

Starting again with a dyadic similarity analysis, we observed that for all comparisons, the more genetically related the dyads, the more similar their movie FC patterns (Fig. 6A, MZ>UR / MZ>DZ / DZ>UR: 49% / 26% / 18% greater similarity across network combinations). Compared to group differences in ISC, movie FC profile similarities showed greater separation between groups (Fig. 6B, group average FC profile similarity shown in Fig. S13).

Given the greater between-group separation in FC profile similarity (vs. BOLD time course similarity) in Fig. 6B vs. Fig. 1B, we expected movie FC profiles to be more heritable than movie-evoked BOLD time courses. Applying the same multidimensional heritability analysis to movie FC profiles, we indeed found that FC profiles were around six times as heritable as BOLD time courses (Fig. 7, left column, Day 1 mean h^2 = .36 ± .035, Day 2 mean h^2 = .37 ± .038, all network combinations significant on both days at FDR-corrected P < .05), and the pattern of heritability values across network combinations was very consistent between days (Spearman ρ = .92, P_{perm} < .001). The six-fold higher heritability of multidimensional FC profiles compared to BOLD time courses likely results from each FC profile dimension representing a connectivity strength calculated over many time-points between two cortical areas, whereas each BOLD time course dimension reflects an activity magnitude at a single timepoint in one cortical area. Consequently, FC profile dimensions are less noisy (due to being defined across many timepoints) and capture more individual differences (due to being defined across multiple areas).

Although we are unaware of any previous studies that have investigated FC heritability during movie watching, the heritability of resting state FC (rest FC) measures has been well established (Anderson et al., 2021; Busch et al., 2023; Glahn et al., 2010; Miranda-Dominguez et al., 2018). Compared to rest FC, movie FC profiles serve as better identifiers of individuals (Vanderwal et al., 2017) and better predict individual differences in behavior (Finn & Bandettini, 2021). Given movie FC's increased sensitivity to individual variability, we reasoned that movie FC profiles would be more heritable than rest FC profiles. Using resting state data collected from the same subjects and on the same days of data collection, we calculated rest FC profile heritability using the approach described above and found that heritability was indeed lower than for movie FC profiles (Fig. 7, right column), but this effect was limited to more sensory-oriented networks (Language, Auditory, Somatomotor A and B, Visual A, B, and C, and Dorsal Attention B networks significant at FDR-corrected P_{perm} < .05 on both days). Across these networks, movie (vs. rest) FC profiles were 20% more heritable on Day 1 (min. = 11%, max. = 30%) and 30% more heritable on Day 2 (min. = 19%, max. = 44%). Rest FC profiles were not significantly more heritable than movie FC profiles in any network. Because subjects tend to move more during resting state vs. movie-watching scans, and because this could explain the higher heritability of movie FC profiles, we repeated our analyses after censoring all frames with FD >0.2 mm and found the with- and without-censoring results to be nearly identical (data not shown).

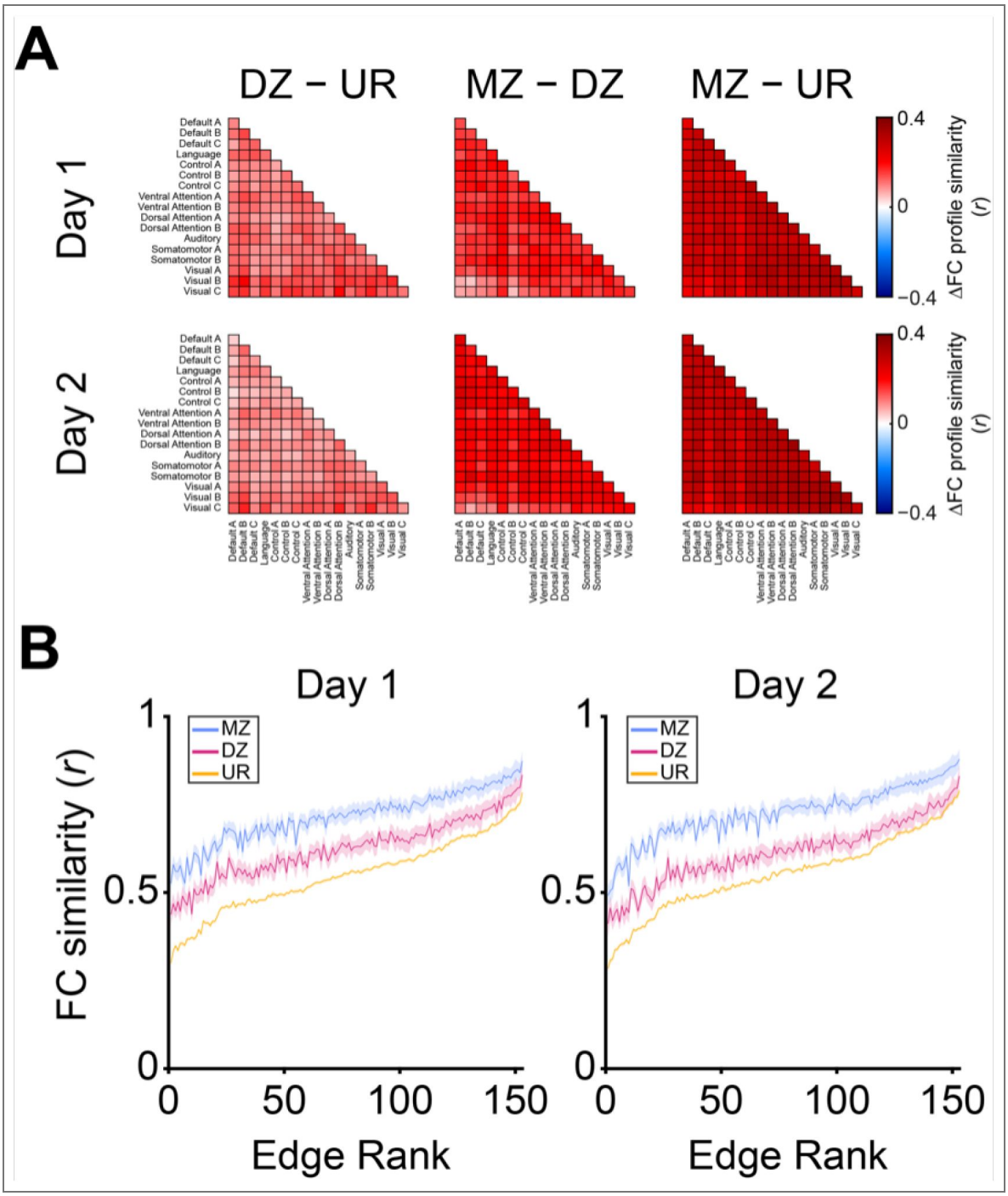


Figure 6. **FC profile similarity scales with genetic relatedness across the cortex.**

(A) Group differences in average FC profile similarity show that FC profiles are more similar for dyads who are more genetically related (51 MZ dyads, 34 DZ dyads, 690 UR dyads). (B) Group-average FC profile similarity values used to create the difference maps in A, plotted in order of average FC profile similarity across all subject pairs (shading = SEM).

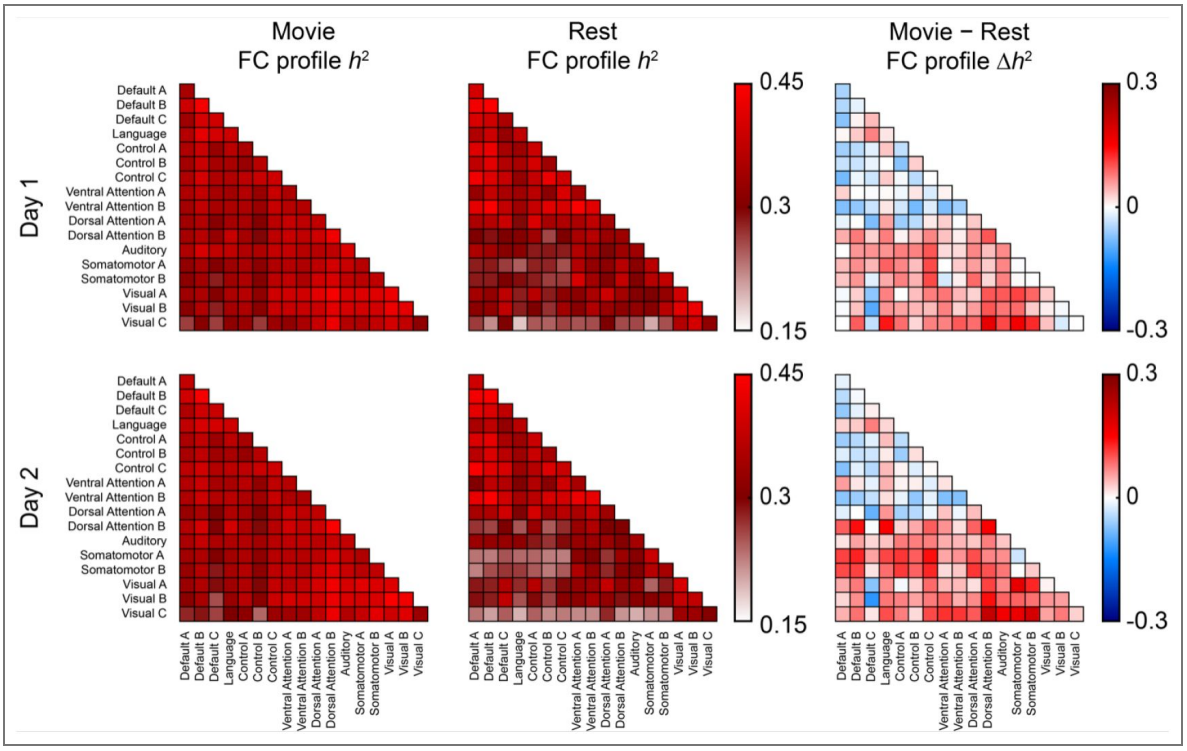


Figure 7. **FC profiles are heritable across network combinations.**

Heatmaps show heritability of FC profiles for all unique within- and between-network combinations of the 17 Kong networks after controlling for age, gender, and head motion. FC profiles during movie-watching (left column) were more heritable than resting state FC profiles (middle column) for more sensory-oriented networks (red rows in the right column).

Our connectivity analyses thus far have focused on FC profiles (i.e., correlations across FC values) instead of FC strengths (i.e., the average of FC values), as the low power afforded by our sample size precludes us from measuring the heritability of one-dimensional phenotypes with high precision (Benson et al., 2022; Busch et al., 2023). However, FC strengths are clinically and behaviorally relevant, and measuring how task conditions and alignment approaches affect FC strength heritability across the cortex requires substantially less power than resolving FC strength heritability values for individual network combinations. With this in mind, we used SOLAR (Almasy & Blangero, 1998) to estimate the heritability of FC strength in this and subsequent analyses and provide the corresponding figures in the Supplementary Materials (Fig. S14).

We next repeated our FC profile (and strength) analyses using the hyperaligned data. Similar to the effects of RHA and CHA on BOLD time course heritability, hyperalignment within the Schaefer 100 parcels significantly reduced FC profile heritability across network combinations by 39% on Day 1 (95% CI = [32–46%]) and 41% on Day 2 [34–48%] for RHA, and by 20% [13–28%] and 18% [12–25%] for CHA (Fig. 8A). However, we did observe some consistent differences between effects of hyperalignment on FC profile vs. BOLD time course heritability. First, RHA decreased FC profile heritability to a significantly greater extent than did CHA, seen in the larger separation between purple and orange traces in Fig. 8B. Second, hyperalignment's effects on FC profile (vs. BOLD time course) heritability were less variable across the different parcellation resolutions, seen in the relatively flat slope between orange/purple dots. This is reflected in the lower power law *b* coefficient values for FC profile (0.08–0.11, Fig. 8B) vs. BOLD time course heritability (0.32–0.43, Fig. 4E). We observed a similar area-independent pattern of results for our FC strength analyses, although here only RHA (and not CHA) significantly decreased FC strength heritability (Fig. S15). Once again, linear, quadratic, and logarithmic models failed to explain the relationships between hyperalignment resolution and FC profile (Fig. S16) and strength (Fig. S17) heritability compared to power law models.

Discussion

In this study, we examined the heritability of movie-evoked BOLD activity and connectivity. First, we showed that BOLD time courses and FC profiles are heritable across the cortex, especially in and between the sensory and associative regions that are most reliably activated by the stimuli. Second, we showed that this heritability is underpinned by genetic control over fundamental spatial and temporal characteristics of brain function that reflect both *where* and *how* individuals process sensory information. More specifically, our findings demonstrate that genetics influences cortical topography as a power law function of cortical area, and that a key property of brain function—the neural timescale—is responsible for an additional portion of BOLD time course heritability, especially in auditory and speech-sensitive areas. Just as importantly, these results suggest a modest ceiling for how much of this stimulus-driven activity and connectivity is under genetic control, leaving the rest to non-genetic individual variation.

Studies using ISC to examine similarity of movie-evoked BOLD activity typically find highly conserved responses in auditory and visual areas. We found that more genetically related individuals exhibited greater ISC not only in these sensory areas, but also across most of cortex. This increased ISC could reflect more similar stimulus processing in a number of ways. For example, high-level attentional effects, such as twins attending to more similar aspects of the stimulus, could account for this increase (Ki et al., 2016; Song et al., 2021). Such an attentional effect would explain our finding that BOLD time courses in auditory cortex (vs. mid-level visual and oculomotor areas) were less (vs. more) heritable than would be expected based on their overall ISC (Fig. 2B), as eye movements gate incoming visual information and associated neural representations (Borovska & de Haas, 2024) but no analogous mechanism exists in the auditory system, and eye movement patterns during complex scene viewing are themselves moderately heritable (Kennedy et al., 2017). Alternatively, low-level stimulus processing effects, such as twins having more similar population tuning than non-twins, could also lead to greater ISC.

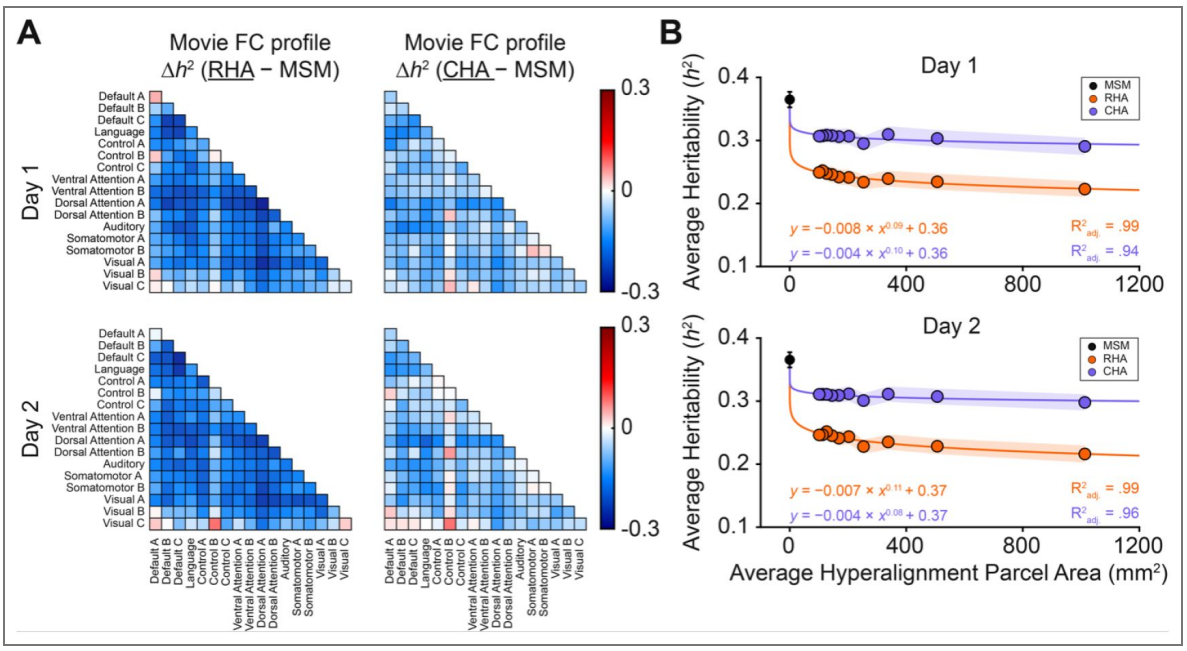


Figure 8. **Hyperalignment reduces FC profile heritability.**

(A) Heatmaps show decreased FC profile heritability for most combinations of 17 Kong networks following RHA (left) and CHA (right) compared to the MSM-only baseline. (B) Scatter plots show that hyperalignment, especially with RHA, decreases FC profile heritability according to a power law function; each dot corresponds to average cortex-wide heritability for data hyperaligned using one of the 10 Schaefer atlas resolutions (or MSM-only alignment, shading = jackknife SEM).

Our frequency-dependent heritability results offer some preliminary insights into the specific aspects of sensory processing that these shared activity patterns represent. Here, we observed that BOLD time course heritability was over 50% greater in the slowest frequency band compared to the unfiltered data, and that this effect was driven by increased low-frequency heritability in more associative parcels. Because these regions and frequency bands encode more abstract stimulus features, this result suggests that the neural processing of high-vs. low-level sensory information is under greater genetic control. Importantly, we note that interpretation of this result is limited by reverse inference, and future studies that directly modulate low- and high-level stimulus information will be necessary to more conclusively answer this question.

Our approach differs from previous studies of stimulus- or task-driven brain activity heritability in that our ISC-based analyses don't require assumptions about the nature of neural responses that, if inaccurate, could decrease the sensitivity of heritability estimates. Furthermore, instead of collapsing brain activity measurements across trials or epochs, our analyses exploit the high-dimensional nature of BOLD time courses by considering the unique information present at each timepoint. This multi-dimensional aspect of our analyses allowed us to leverage the significant amount of data available per subject to detect reliable (spatial $\rho > .9$ for heritability maps across days) effects even at the level of individual vertices. Furthermore, although our BOLD time course heritability effect sizes were modest ($h^2 \leq .25$ for parcels, $h^2 \leq .12$ for individual vertices), we note that these are commensurate with other twin-based heritability estimates of sensory phenotypes measured with fMRI.

Compared to our activity-based analyses, our FC analyses were more in line with previous work. Indeed, the heritability of FC profiles has been investigated on multiple occasions, sometimes with the same multidimensional estimator of heritability used here (Busch et al., 2023; Elliott et al., 2019; Miranda-Dominguez et al., 2018). Still, several aspects of our study allowed us to reveal novel results and add new context to established findings. For example, by analyzing resting state and movie-watching data from the same subjects, we were able to show that FC profiles involving sensory-oriented networks were significantly more heritable during movie-watching than at rest. This finding is consistent with reports that movie FC profiles better identify individuals (Vanderwal et al., 2017) and predict variability in behavioral traits (Finn & Bandettini, 2021) than resting state FC profiles, and that including task data increases estimates of FC profile heritability (Elliott et al., 2019). Relatedly, recent work by Luppi et al. (2025) has shown that FC uniqueness decreases under greater levels of anesthesia, suggesting that resting-state scans, where participants may likewise drift toward reduced arousal states, could mute idiosyncratic (and thus heritable) variation that remains robust when attention is held by an engaging movie stimulus.

Although BOLD time courses and FC profiles often serve as the fundamental brain phenomena to be studied in fMRI experiments, they are complex entities that are underpinned by a variety of lower-level biological processes (Hillman, 2014). As such, the heritability of movie-evoked brain responses established here likely reflects genetic control over more basic aspects of brain function. We demonstrated how two of these aspects, cortical topography and neural timescale, contribute to the heritability of stimulus-driven BOLD activity.

First, we found that controlling for idiosyncratic cortical topographies via response and connectivity hyperalignment (RHA and CHA) decreased activity heritability across the cortex. This decrease was bigger when data were aligned across larger parcels, but the rate of this decrease slowed as a power law function of parcel area, illustrating that genetic control of cortical topography is greatest at the fine scale. In addition to decreasing BOLD time course heritability, we found that RHA and CHA also decreased FC profile heritability, echoing recent work showing that CHA decreases rest FC profile heritability in a developmental population (Busch et al., 2023). Compared to our activity-based analyses, we noticed a far weaker effect of hyperalignment resolution on FC profile heritability, likely because this analysis was performed at the parcel level and across spatially distributed brain networks (thereby reducing the impact of local functional alignment). Although hyperalignment served to reduce heritable individual variability across our analyses, the residual post-hyperalignment heritability might be more behaviorally relevant, as hyperalignment has been shown to increase associations between FC profiles and cognitive test

scores (Feilong et al., 2021). With this in mind, future studies investigating genetic correlations between brain function and behavioral variables may benefit from hyperalignment, as it can factor out individual-specific cortical topography and thus yield more precise estimates of functional heritability.

Just as cortical topographies spatially constrain individual responses to incoming stimuli, neural timescales (NTs) are stable temporal features of brain function that shape high-dimensional activity and connectivity patterns (Shinn et al., 2023; Wengler et al., 2020). More specifically, longer NTs are thought to reflect greater recurrent excitation at the micro-circuit level and yield more stable integration of sensory information (Cavanagh et al., 2020; Watanabe et al., 2019). Here, we found that MZ twins had more similar NTs than unrelated dyads; the fact that NTs measured by fMRI track electrophysiological activity (Watanabe et al., 2019) suggests that this reflects similarities in how movie stimuli were neurally encoded. We next showed that this genetic similarity in NT magnitude contributes to genetic similarity in movie-evoked BOLD time courses, such that controlling for vertex-level NT accounted for up to ~30% of BOLD time course heritability, an effect that was strongest in speech-related areas like the superior temporal gyri. Importantly, controlling for NT had a weaker effect on BOLD time course heritability after hyperalignment. This is evidence that the topographic distribution of NTs, over and above their magnitude, is under genetic control. Beyond genetics, our finding that subjects with longer NTs had more correlated movie-evoked BOLD time courses suggests that decreased ISC in patients with schizophrenia (Patel et al., 2021; Tu et al., 2019), autism (Salmi et al., 2013), and depression (Gruskin et al., 2020) may be underpinned by shorter NTs in these same populations (Watanabe et al., 2019; Wengler et al., 2020; Zheng et al., 2024).

Our work should be considered in light of an important demographic limitation. Almost 90% of subjects in the present sample identified as White, and all subjects were between the ages of 22 and 36. As heritability estimates are known to differ across populations and age groups (Schmitt et al., 2014; Zhang et al., 2023), the generalizability of our findings is limited by the demographic characteristics of the HCP Young Adult sample used here (Ricard et al., 2023). In spite of this limitation, this work constitutes an important first link between the growing fields of neuroimaging genetics and “naturalistic” neuroscience. By considering BOLD time courses and FC profiles alongside cortical topographies and neural timescales derived from independent data, we reveal a multi-layered genetic influence that extends from basic features of brain function to complex, individual-specific sensory processing patterns. This comprehensive approach paves the way for future research to dissect the biological mechanisms that link genetics with sensory processing in both typical and atypical populations. Finally, we note that less than half of the inter-individual variability we observed in movie-evoked BOLD time courses and FC profiles was heritable, leaving the majority of this variability to be explained by gene-environment interactions as well as non-genetic factors such as life experiences and current behavioral state. As such, additional work will be necessary to characterize these non-genetic factors.

Data availability

The raw HCP data used for this project can be downloaded from ConnectomeDB (db.humanconnectome). Code for all analyses will be posted on GitHub upon publication. Analyses were performed in MAT-LAB (R2023b), Python, and R. All cortical surface visualizations were performed with Connectome Workbench (Marcus et al., 2011).

Acknowledgements

We thank Avram Holmes and Erica Busch for helpful conversations regarding this project. Author D.C.G. was supported by an NIH MSTP training grant (T32GM007367). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil, 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at

Washington University. Author G.H.P. receives income and equity from Pfizer, Inc. through family. Authors D.C.G. and D.J.V. have no competing interests to declare. This preprint was created using the LaPreprint template (*https://github.com/roaldarbol/lapreprint*) by Mikkel Roald-Arbøl .

Additional information

Citation diversity statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field (Bertolero et al., 2020; Caplar et al., 2017; Chatterjee & Werner, 2021; Dion et al., 2018; Dworkin et al., 2020; Fulvio et al., 2021; Maliniak et al., 2013; Mitchell et al., 2013; Wang et al., 2021). Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman (Dworkin et al., 2020; Zhou et al., 2020). By this measure and excluding self-citations to the first and last authors of our current paper, our references contain 7.46% woman(first)/woman(last), 5.97% man/woman, 28.36% woman/man, and 58.21% man/man. This method is limited in that a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color (Ambekar et al., 2009; Chintalapati et al., 2023). By this measure (and excluding self-citations), our references contain 13.38% author of color (first)/author of color(last), 15.92% white author/author of color, 22.39% author of color/white author, and 48.32% white author/white author. This method is limited in that a) names and Florida Voter Data to make the predictions may not be indicative of racial/ethnic identity, and b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

Author contributions

David C. Gruskin: Conceptualization, Methodology, Formal Analysis, Visualization, Writing–Original

Draft, Writing–Review & Editing

Daniel J. Vieira: Code review, Writing–Review & Editing

Jessica K. Lee: Supervision, Writing–Review & Editing

Gaurav H. Patel: Conceptualization, Writing–Review & Editing, Supervision

Funding

Funder	Grant reference number	Author
HHS | NIH | National Institute of General Medical Sciences (NIGMS)	T32GM007367	David C Gruskin

Author ORCID iDs

David C Gruskin: https://orcid.org/0000-0001-6504-191X

Additional files

Supplemental Materials

References

Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *American Journal of Human Genetics* **62**:1198-1211 https://doi.org/10.1086/301844 | PubMed

Alvarez I, Finlayson NJ, Ei S, de Haas B, Greenwood JA, Schwarzkopf DS (2021) Heritable functional architecture in human visual cortex. *Neuroimage* **239**:118286 https://doi.org/10.1016/j.neuroimage.2021.118286 | PubMed

Ambekar A, Ward C, Mohammed J, Male S, Skiena S (2009) Name-ethnicity classification from open sources. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 49-58 https://doi.org/10.1145/1557019.1557032

Anderson KM, Ge T, Kong R, Patrick LM, Spreng RN, Sabuncu MR, Yeo BTT, Holmes AJ (2021) Heritability of individualized cortical network topography. *Proceedings of the National Academy of Sciences of the United States of America* **118**:e2016271118 https://doi.org/10.1073/pnas.2016271118 | PubMed

Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA (2017) Discovering event structure in continuous narrative perception and memory. *Neuron* **95**:709-721.e5 https://doi.org/10.1016/j.neuron.2017.06.041 | PubMed

Bazeille T, DuPre E, Richard H, Poline J.-B, Thirion B (2021) An empirical evaluation of functional alignment using inter-subject decoding. *NeuroImage* **245**:118683 https://doi.org/10.1016/j.neuroimage.2021.118683 | PubMed

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* **57**:289-300 https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benson NC, Yoon JMD, Forenzo D, Engel SA, Kay KN, Winawer J (2022) Variability of the surface area of the V1, V2, and V3 maps in a large sample of human observers. *The Journal of Neuroscience* **42**:8629-8646 https://doi.org/10.1523/JNEUROSCI.0690-21.2022 | PubMed

Bertolero MA, Dworkin JD, David SU, Lloreda CL, Srivastava P, Stiso J, Zhou D, Dzirasa K, Fair DA, Kaczkurkin AN, *et al.* (2020) Racial and ethnic imbalance in neuroscience reference lists and intersections with gender. *bioRxiv* https://doi.org/10.1101/2020.10.12.336230

Borovska P, de Haas B (2024) Individual gaze shapes diverging neural representations. *Proceedings of the National Academy of Sciences* **121**:e2405602121 https://doi.org/10.1073/pnas.2405602121 | PubMed

Burger B, Nenning K.-H, Schwartz E, Margulies DS, Goulas A, Liu H, Neubauer S, Dauwels J, Prayer D, Langs G (2022) Disentangling cortical functional connectivity strength and topography reveals divergent roles of genes and environment. *NeuroImage* **247**:118770 https://doi.org/10.1016/j.neuroimage.2021.118770 | PubMed

Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD (2020) Generative modeling of brain maps with spatial autocorrelation. *NeuroImage* **220**:117038 https://doi.org/10.1016/j.neuroimage.2020.117038 | PubMed

Busch EL, Rapuano KM, Anderson KM, Rosenberg MD, Watts R, Casey BJ, Haxby JV, Feilong M (2023) Dissociation of reliability, heritability, and predictivity in coarse- and fine-scale functional connectomes during development. *Journal of Neuroscience* https://doi.org/10.1523/JNEUROSCI.0735-23.2023

Caplar N, Tacchella S, Birrer S (2017) Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy* **1**:0141 https://doi.org/10.1038/s41550-017-0141

Cavanagh SE, Hunt LT, Kennerley SW (2020) A diversity of intrinsic timescales underlie neural computations. *Frontiers in Neural Circuits* **14** https://doi.org/10.3389/fncir.2020.615626 | PubMed

Chatterjee P, Werner RM (2021) Gender disparity in citations in high-impact journal articles. *JAMA Network Open* **4**:e2114509 https://doi.org/10.1001/jamanetworkopen.2021.14509 | PubMed

Chen M, Han J, Hu X, Jiang X, Guo L, Liu T (2014) Survey of encoding and decoding of visual stimulus via FMRI: An image analysis perspective. *Brain imaging and behavior* **8**:7-23 https://doi.org/10.1007/s11682-013-9238-z | PubMed

Chintalapati R, Laohaprapanon S, Sood G (2023) Predicting race and ethnicity from the sequence of characters in a name. *arXiv* https://doi.org/10.48550/arXiv.1805.02109

Christova P, Uğurbil K, Georgopoulos AP (2022) Heritability of brain neurovascular coupling. *Journal of Neurophysiology* **128**:1307-1311 https://doi.org/10.1152/jn.00402.2022 | PubMed

Cole MW, Ito T, Cocuzza C, Sanchez-Romero R (2021) The functional relevance of task-state functional connectivity. *Journal of Neuroscience* **41**:2684-2702 https://doi.org/10.1523/JNEUROSCI.1713-20.2021 | PubMed

Dion ML, Sumner JL, Mitchell SM (2018) Gendered citation patterns across political science and social science methodology fields. *Political Analysis* **26**:312-327 https://doi.org/10.1017/pan.2018.12

Dworetsky A, Seitzman BA, Adeyemo B, Nielsen AN, Hatoum AS, Smith DM, Nichols TE, Neta M, Petersen SE, Gratton C (2024) Two common and distinct forms of variation in human functional brain networks. *Nature Neuroscience* **27**:1187-1198 https://doi.org/10.1038/s41593-024-01618-2 | PubMed

Dworkin JD, Linn KA, Teich EG, Zurn P, Shinohara RT, Bassett DS (2020) The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience* **23**:918-926 https://doi.org/10.1038/s41593-020-0658-y | PubMed

Elliott ML, Knodt AR, Cooke M, Kim MJ, Melzer TR, Keenan R, Ireland D, Ramrakha S, Poulton R, Caspi A, *et al.* (2019) General functional connectivity: Shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks. *NeuroImage* **189**:516-532 https://doi.org/10.1016/j.neuroimage.2019.01.068 | PubMed

Feilong M, Guntupalli JS, Haxby JV (2021) The neural basis of intelligence in fine-grained cortical topographies. *eLife* **10**:e64058 https://doi.org/10.7554/eLife.64058 | PubMed

Finn ES, Bandettini PA (2021) Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage* **235**:117963 https://doi.org/10.1016/j.neuroimage.2021.117963 | PubMed

Finn ES, Corlett PR, Chen G, Bandettini PA, Constable RT (2018) Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. *Nature Communications* **9** https://doi.org/10.1038/s41467-018-04387-2 | PubMed

Fulvio JM, Akinnola I, Postle BR (2021) Gender (im)balance in citation practices in cognitive neuroscience. *Journal of Cognitive Neuroscience* **33**:3-7 https://doi.org/10.1162/jocn_a_01643 | PubMed

Ge T, Reuter M, Winkler AM, Holmes AJ, Lee PH, Tirrell LS, Roffman JL, Buckner RL, Smoller JW, Sabuncu MR (2016) Multidimensional heritability analysis of neuroanatomical shape. *Nature Communications* **7**:13291 https://doi.org/10.1038/ncomms13291 | PubMed

Glahn DC, Winkler AM, Kochunov P, Almasy L, Duggirala R, Carless MA, Curran JC, Olvera RL, Laird AR, Smith SM, *et al.* (2010) Genetic control over the resting brain. *Proceedings of the National Academy of Sciences* **107**:1223-1228 https://doi.org/10.1073/pnas.0909969107 | PubMed

Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, *et al.* (2017) Precision functional mapping of individual human brains. *Neuron* **95**:791-807.e7 https://doi.org/10.1016/j.neuron.2017.07.011 | PubMed

Gruskin DC, Patel GH (2022) Brain connectivity at rest predicts individual differences in normative activity during movie watching. *NeuroImage* **253**:119100 https://doi.org/10.1016/j.neuroimage.2022.119100 | PubMed

Gruskin DC, Rosenberg MD, Holmes AJ (2020) Relationships between depressive symptoms and brain responses during emotional movie viewing emerge in adolescence. *NeuroImage* **216**:116217 https://doi.org/10.1016/j.neuroimage.2019.116217 | PubMed

Guntupalli JS, Feilong M, Haxby JV (2018) A computational model of shared fine-scale structure in the human connectome. *PLOS Computational Biology* **14**:e1006120 https://doi.org/10.1371/journal.pcbi.1006120 | PubMed

Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A hierarchy of temporal receptive windows in human cortex. *The Journal of Neuroscience* **28**:2539-2550 https://doi.org/10.1523/JNEUROSCI.5487-07.2008 | PubMed

Haxby JV, Guntupalli JS, Connolly AC, Halchenko YO, Conroy BR, Gobbini MI, Hanke M, Ramadge PJ (2011) A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* **72**:404-416 https://doi.org/10.1016/j.neuron.2011.08.026 | PubMed

Haxby JV, Guntupalli JS, Nastase SA, Feilong M (2020) Hyperalignment: Modeling shared information encoded in idiosyncratic cortical topographies. *eLife* **9**:e56601 https://doi.org/10.7554/eLife.56601 | PubMed

Hillman EM (2014) Coupling mechanism and significance of the BOLD signal: A status report. *Annual review of neuroscience* **37**:161-181 https://doi.org/10.1146/annurev-neuro-071013-014111 | PubMed

Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**:423-434 https://doi.org/10.1016/j.neuron.2012.08.011 | PubMed

Josephs O, Henson RNA (1999) Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **354**:1215-1228 https://doi.org/10.1098/rstb.1999.0475 | PubMed

Kauppi J.-P, Jääskeläinen IP, Sams M, Tohka J (2010) Inter-subject correlation of brain hemodynamic responses during watching a movie: Localization in space and frequency. *Frontiers in Neuroinformatics* **4**:5 https://doi.org/10.3389/fninf.2010.00005 | PubMed

Kennedy DP, D'Onofrio BM, Quinn PD, Bölte S, Lichtenstein P, Falck-Ytter T (2017) Genetic influence on eye movements to complex scenes at short timescales. *Current biology: CB* **27**:3554-3560.e3 https://doi.org/10.1016/j.cub.2017.10.007 | PubMed

Ki JJ, Kelly SP, Parra LC (2016) Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **36**:3092-3101 https://doi.org/10.1523/JNEUROSCI.2942-15.2016 | PubMed

Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A (2016) Correlations and neuronal population information. *Annual review of neuroscience* **39**:237-256 https://doi.org/10.1146/annurev-neuro-070815-013851 | PubMed

Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, Sun N, Zuo X.-N, Holmes AJ, Eickhoff SB, *et al.* (2019) Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cerebral Cortex* https://doi.org/10.1093/cercor/bhy123 | PubMed

Korgaonkar MS, Ram K, Williams LM, Gatt JM, Grieve SM (2014) Establishing the resting state default mode network derived from functional magnetic resonance imaging tasks as an endophenotype: A twins study. *Human Brain Mapping* **35**:3893-3902 https://doi.org/10.1002/hbm.22446 | PubMed

Li J, Kong R, Liégeois R, Orban C, Tan Y, Sun N, Holmes AJ, Sabuncu MR, Ge T, Yeo BT (2019) Global signal regression strengthens association between resting-state functional connectivity and behavior. *NeuroImage* **196**:126-141 https://doi.org/10.1016/j.neuroimage.2019.04.016 | PubMed

Luppi AI, Golkowski D, Ranft A, Ilg R, Jordan D, Bzdok D, Owen AM, Naci L, Stamatakis EA, Amico E, *et al.* (2025) General anaesthesia decreases the uniqueness of brain functional connectivity across individuals and species. *Nature Human Behaviour* **9**:987-1004 https://doi.org/10.1038/s41562-025-02121-9 | PubMed

Maliniak D, Powers R, Walter BF (2013) The gender citation gap in international relations. *International Organization* **67**:889-922 https://doi.org/10.1017/S0020818313000209

Manyara AM, Purvis A, Ciani O, Collins GS, Taylor RS (2024) Sample size in multistakeholder delphi surveys: At what minimum sample size do replicability of results stabilize?. *Journal of Clinical Epidemiology* **174**:111485 https://doi.org/10.1016/j.jclinepi.2024.111485 | PubMed

Marcus DS, Harwell J, Olsen T, Hodge M, Glasser MF, Prior F, Jenkinson M, Laumann T, Curtiss SW, Van Essen DC (2011) Informatics and data mining tools and strategies for the human connectome project. *Frontiers in Neuroinformatics* **5**:4 https://doi.org/10.3389/fninf.2011.00004 | PubMed

Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, Donohue MR, Foran W, Miller RL, Hendrickson TJ, *et al.* (2022) Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**:654-660 https://doi.org/10.1038/s41586-022-04492-9 | PubMed

Miranda-Dominguez O, Feczko E, Grayson DS, Walum H, Nigg JT, Fair DA (2018) Heritability of the human connectome: A connectotyping study. *Network Neuroscience* **2**:175-199 https://doi.org/10.1162/netn_a_00029 | PubMed

Mitchell SM, Lange S, Brus H (2013) Gendered citation patterns in international relations journals. *International Studies Perspectives* **14**:485-492 https://doi.org/10.1111/insp.12026

Parkes L, Fulcher B, Yücel M, Fornito A (2018) An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage* **171**:415-436 https://doi.org/10.1016/j.neuroimage.2017.12.073 | PubMed

Patel GH, Arkin SC, Ruiz-Betancourt DR, Plaza FI, Mirza SA, Vieira DJ, Strauss NE, Klim CC, Sanchez-Peña JP, Bartel LP (2021) Failure to engage the temporoparietal junction/posterior superior temporal sulcus predicts impaired naturalistic social cognition in schizophrenia. *Brain : a journal of neurology* **144**:1898-1910 https://doi.org/10.1093/brain/awab081 | PubMed

Pérez P, Madsen J, Banellis L, Türker B, Raimondo F, Perlbarg V, Valente M, Niérat M.-C, Puybasset L, Naccache L, *et al.* (2021) Conscious processing of narrative stimuli synchronizes heart rate between individuals. *Cell Reports* **36**:109692 https://doi.org/10.1016/j.celrep.2021.109692 | PubMed

Petersen SE, Seitzman BA, Nelson SM, Wig GS, Gordon EM (2024) Principles of cortical areas and their implications for neuroimaging. *Neuron* **0** https://doi.org/10.1016/j.neuron.2024.05.008 | PubMed

Renvall H, Salmela E, Vihla M, Illman M, Leinonen E, Kere J, Salmelin R (2012) Genomewide linkage analysis of human auditory cortical activation suggests distinct loci on chromosomes 2, 3, and 8. *The Journal of Neuroscience* **32**:14511-14518 https://doi.org/10.1523/JNEUROSCI.1483-12.2012 | PubMed

Ricard JA, Parker TC, Dhamala E, Kwasa J, Allsop A, Holmes AJ (2023) Confronting racially exclusionary practices in the acquisition and analyses of neuroimaging data. *Nature Neuroscience* **26**:4-11 https://doi.org/10.1038/s41593-022-01218-y | PubMed

Robinson EC, Jbabdi S, Glasser MF, Andersson J, Burgess GC, Harms MP, Smith SM, Van Essen DC, Jenkinson M (2014) MSM: A new flexible framework for multimodal surface matching. *NeuroImage* **100**:414-426 https://doi.org/10.1016/j.neuroimage.2014.05.069 | PubMed

Salmi J, Roine U, Glerean E, Lahnakoski J, Nieminen-von Wendt T, Tani P, Leppämäki S, Nummenmaa L, Jääskeläinen I, Carlson S, *et al.* (2013) The brains of high functioning autistic individuals do not synchronize with those of others. *NeuroImage : Clinical* **3**:489-497 https://doi.org/10.1016/j.nicl.2013.10.011 | PubMed

Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X.-N, Holmes AJ, Eickhoff SB, Yeo BT (2018) Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex (New York, N.Y.: 1991)* **28**:3095-3114 https://doi.org/10.1093/cercor/bhx179 | PubMed

Schmitt JE, Neale MC, Fassassi B, Perez J, Lenroot RK, Wells EM, Giedd JN (2014) The dynamic role of genetics on cortical patterning during childhood and adolescence. *Proceedings of the National Academy of Sciences* **111**:6774-6779 https://doi.org/10.1073/pnas.1311630111 | PubMed

Shinn M, Hu A, Turner L, Noble S, Preller KH, Ji JL, Moujaes F, Achard S, Scheinost D, Constable RT, *et al.* (2023) Functional brain networks reflect spatial and temporal autocorrelation. *Nature Neuroscience* **26**:867-878 https://doi.org/10.1038/s41593-023-01299-3 | PubMed

Sinclair B, Hansell NK, Blokland GA, Martin NG, Thompson PM, Breakspear M, de Zubicaray GI, Wright MJ, McMahon KL (2015) Heritability of the network architecture of intrinsic brain functional connectivity. *NeuroImage* **121**:243-252 https://doi.org/10.1016/j.neuroimage.2015.07.048 | PubMed

Song H, Finn ES, Rosenberg MD (2021) Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences* **118**:e2021905118 https://doi.org/10.1073/pnas.2021905118 | PubMed

Sydnor VJ, Larsen B, Seidlitz J, Adebimpe A, Alexander-Bloch AF, Bassett DS, Bertolero MA, Cieslak M, Covitz S, Fan Y, *et al.* (2023) Intrinsic activity development unfolds along a sensorimotor–association cortical axis in youth. *Nature neuroscience* **26**:638 https://doi.org/10.1038/s41593-023-01282-y | PubMed

Tu P.-C, Su T.-P, Lin W.-C, Chang W.-C, Bai Y.-M, Li C.-T, Lin F.-H (2019) Reduced synchronized brain activity in schizophrenia during viewing of comedy movies. *Scientific Reports* **9**:12738 https://doi.org/10.1038/s41598-019-48957-w | PubMed

van Baar JM, Halpern DJ, FeldmanHall O (2021) Intolerance of uncertainty modulates brain-to-brain synchrony during politically polarized perception. *Proceedings of the National Academy of Sciences* **118**:e2022491118 https://doi.org/10.1073/pnas.2022491118 | PubMed

van den Heuvel MP, Hulshoff Pol HE (2010) Exploring the brain network: A review on restingstate fMRI functional connectivity. *European Neuropsychopharmacology* **20**:519-534 https://doi.org/10.1016/j.euroneuro.2010.03.008 | PubMed

Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K (2013) The WUminn human connectome project: An overview. *NeuroImage* **80**:62-79 https://doi.org/10.1016/j.neuroimage.2013.05.041

Vanderwal T, Eilbott J, Finn ES, Craddock RC, Turnbull A, Castellanos FX (2017) Individual differences in functional connectivity during naturalistic viewing conditions. *NeuroImage* **157**:521-530 https://doi.org/10.1016/j.neuroimage.2017.06.027 | PubMed

van Pelt S, Boomsma DI, Fries P (2012) Magnetoencephalography in twins reveals a strong genetic determination of the peak frequency of visually induced gamma-band synchronization. *Journal of Neuroscience* **32**:3388-3392 https://doi.org/10.1523/JNEUROSCI.5592-11.2012 | PubMed

Wang X, Dworkin JD, Zhou D, Stiso J, Falk EB, Bassett DS, Zurn P, Lydon-Staley DM (2021) Gendered citation practices in the field of communication. *Annals of the International Communication Association* **45**:134-153 https://doi.org/10.1080/23808985.2021.1960180 | PubMed

Watanabe T, Rees G, Masuda N (2019) Atypical intrinsic neural timescale in autism. *eLife* **8**:e42256 https://doi.org/10.7554/eLife.42256 | PubMed

Wengler K, Goldberg AT, Chahine G, Horga G (2020) Distinct hierarchical alterations of intrinsic neural timescales account for different manifestations of psychosis. *eLife* **9**:e56151 https://doi.org/10.7554/eLife.56151 | PubMed

Xu T, Opitz A, Craddock RC, Wright MJ, Zuo X.-N, Milham MP (2016) Assessing variations in areal organization for the intrinsic brain: From fingerprints to reliability. *Cerebral Cortex* **26**:4192-4211 https://doi.org/10.1093/cercor/bhw241 | PubMed

Yeshurun Y, Swanson S, Simony E, Chen J, Lazaridi C, Honey CJ, Hasson U (2017) Same story, different story. *Psychological Science* **28**:307-319 https://doi.org/10.1177/0956797616682029 | PubMed

Zhang J, Zhang S, Qiao J, Wang T, Zeng P (2023) Similarity and diversity of genetic architecture for complex traits between East Asian and European populations. *BMC Genomics* **24**:314 https://doi.org/10.1186/s12864-023-09434-x | PubMed

Zheng R, Bu C, Chen Y, Wei Y, Zhou B, Jiang Y, Zhu C, Wang K, Wang C, Li S, *et al.* (2024) Decreased intrinsic neural timescale in treatment-naïve adolescent depression. *Journal of Affective Disorders* **348**:389-397 https://doi.org/10.1016/j.jad.2023.12.048 | PubMed

Zhou D, Cornblath EJ, Stiso J, Teich EG, Dworkin JD, Blevins AS, Bassett DS (2020) Gender diversity statement and code notebook v1.0. *Zenodo* https://doi.org/10.5281/zenodo.3672110

Van Essen et al (2017) Human Connectome Project (HCP) Young Adult 7 Tesla (7T) Dataset. ConnectomeDB. ID HCP_1200 https://db.humanconnectome.org

Peer reviews

Reviewer #1 (Public review):

Summary:

Gruskin and colleagues use twin data from a movie-watching fMRI paradigm to show how genetic control of cortical function intersects with the processing of naturalistic audiovisual stimuli. They use hyperalignment to dissect heritability into the components that can be explained local differences in cortical-functional topography and those that cannot. They show that heritability is strongest at slower-evolving neural time scales, and more evident in functional connectivity estimates than in response time series.

Strengths:

This is a very thorough paper that tackles this question from several different angles. I very much appreciate the use of hyperalignment to factor our topographic differences and found the relationship between heritability and neural time scales very interesting. The writing is clear and the results are compelling. In general, I don't have many complaints after a couple reads through the manuscript; most of my comments below are relatively minor suggestions and points of clarification.

Weaknesses:

The only "weaknesses" I identified were some points where I think the methods, interpretation, or visualization could be clarified:

On page 16, you compare heritability in functional connectivity (FC) and response time series and find that the heritability effect is larger in FC. In general, I agree with your diagnosis that this is in large part due to the fact that FC captures the covariance structure across parcels, whereas response time series only diverge in terms of univariate time-point-by-time-point differences. Another important factor here is that (within-subject) FC can be driven by intrinsic fluctuations that occur with idiosyncratic timing across subjects and are unrelated to the stimulus (whereas time-locked metrics like ISC and time-series differences cannot, by definition). This makes me wonder how this connectivity result would change if you used intersubject functional connectivity (ISFC) analysis to specifically isolate the stimulus-driven components of functional connectivity (Simony et al., 2016). This, to me, would provide a closer comparison to the ISC and response time series results, and could allow the authors to quantify how much of the heritability in FC is intrinsic versus stimulus-driven. I'm not asking that the authors actually perform this analysis, as I don't think it's critical for the message of the manuscript-but it could be an interesting future direction. As the authors discuss on page 17, I also suspect there's something fundamentally shared between response time series and connectivity as they relate to functional topography (Busch et al., 2021) that drives part of the heritability effect.

The observation that regions with intermediate ISC have the largest differences between MZ, DZ, and UR is very interesting, but it's kind of hard to see in Figure 1B. Is there any other way to plot this that might make the effect more obvious? For example, I could imagine three scatter plots where the x- and y-axes are, e.g., MZ ISC and UR ISC, and each data point is a parcel. In this kind of plot, I would expect to see the middle values lifted visibly off the diagonal/unity line toward MZ. You could even color the data points according to networks like in Figure 3C. (You also might not need to scale the ISC axis all the way to r = 1, which would make the differences more visible.)

On page 9, if I understand correctly, you regress the vector of ISC values across parcels out of the vector of heritability values across parcels and then plot the residual heritability values. Do you center the heritability values (or include some kind of intercept) in the process? I'm trying to understand why the heritability values go from all positive (Figure 2A) to roughly balanced between positive and negative (Figure 2B). Important question for me: How should we interpret negative values in this plot? Can you explain this explicitly in the text? (I also wonder if there's a more intuitive way to control for ISC. For example, instead of regressing out ISC at the parcel/map level, could you go into a single parcel and then regress the subject-level pairwise ISC values out when computing the heritability score?)

On page 4 (line 155), you say "we shuffled dyad labels"-is this equivalent to shuffling rows and columns of the pairwise subject-by-subject matrix combined across groups? I'm trying to make sure your approach here is consistent with recommendations by Chen et al., 2016. Is this the same kind of shuffling used for the kinship matrix mentioned at line 189?

I found panel A in Figure 4 to be a little bit misleading because your parcel-wise approach to hyperalignment won't actually resolve topographic idiosyncrasies across a large cortical distance like what's depicted in the illustration (at the scale of the parcels you're performing hyperalignment within). Maybe just move the green and purple brain areas a bit closer to each other so they could feasibly be "aligned" within a large parcel. Worth keeping in mind when writing that hyperalignment is also not actually going to yield a one-to-one mapping of functionally homologous voxels across individuals: it's effectively going to model any given voxel time series as a linear combination of time series across other voxels in the parcel.

References:

Busch, E. L., Slipski, L., Feilong, M., Guntupalli, J. S., di Oleggio Castello, M. V., Huckins, J. F., Nastase, S. A., Gobbini, M. I., Wager, T. D., & Haxby, J. V. (2021). Hybrid hyperalignment: a single high-dimensional model of shared information embedded in cortical patterns of response and functional connectivity. NeuroImage, 233, 117975. *https://doi.org/10.1016/j.neuroimage.2021.117975*

Chen, G., Shin, Y. W., Taylor, P. A., Glen, D. R., Reynolds, R. C., Israel, R. B., & Cox, R. W. (2016). Untangling the relatedness among correlations, part I: nonparametric approaches to inter-subject correlation analysis at the group level. NeuroImage, 142, 248-259. *https://doi.org/10.1016/j.neuroimage.2016.05.023*

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. Nature Communications, 7, 12141. *https://doi.org/10.1038/ncomms12141*

Comments on revised version.

The authors have adequately addressed my previous comments. This is a strong contribution: the methods are sophisticated, the statistical treatment is rigorous, and the results are quite interesting/compelling. I'm happy to endorse the revised manuscript as a finalized version.

Just to confirm: The subjects watched all different movies across the two days, right? For a moment I was wondering "are Day 1 and Day 2 repetitions of the same movies?" Given that Day 1 and Day 2 are an organizational feature of several figures, it might be worth making this very explicit in the Methods and reminding the reader in the Results section.

https://doi.org/10.7554/eLife.106081.2.sa2

Reviewer #3 (Public review):

Strengths:

It's sort of novel to study the heritability of movie-watching fMRI data. The methodology the authors used in the paper is also supportive of their findings. Figures are nicely organized and plotted. They finally found that sensory processing in the human brain is under genetic control over stable aspects of brain function (here referring to neural timescale and resting state connectivity).

Weaknesses:

What I am worried about most is the sample size and interpretation of heritability.

(1) Figure 1. I assumed that the authors just calculated the ISC within each group (MZ, DZ, and UR). Of course, you can get different variations between each group. Therefore, there is heritability. Why not calculate ISC across the whole sample, then separate MZ, DZ, and UR?

(2) Heritability scores in the paper are sort of small. If the sample size is small, please consider p-values, which will tell more about the trustworthiness of your heritability.

(3) I don't understand the high-frequency signals in fMRI data. It's always regarded as noise, the band 1 here in particular.

(4) The statement "we show that the heritability of brain activity patterns can be partially explained by the heritability of the neural timescale" should come from Figure 5. However, after controlling for NT, the heritability decreased max. 0.025 in temporal areas. I am not sure this change supports the statement. If the visual cortex is outlined, and combining ISC changes in the visual cortex, I think this would somehow be answered. Instead of delta h2, adding a new model h2 would be obvious to the readers.

(5) Figures 7 and 8, when getting the difference of heritability, please also consider the standard errors of the heritability estimates. Then you can compare across networks/regions.

(6) I think movie VS resting state is a really important result in this paper. However, there is almost no discussion. Discussing this part would be more beneficial for understanding the genetic control over the neuron arousal and excitation circuits.

Comments on revised version.

The whole manuscript has been improved a lot, and the concerns have been clarified.

https://doi.org/10.7554/eLife.106081.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *Gruskin and colleagues use twin data from a movie-watching fMRI paradigm to show how genetic control of cortical function intersects with the processing of naturalistic audiovisual stimuli. They use hyperalignment to dissect heritability into the components that can be explained by local differences in cortical-functional topography and those that cannot. They show that heritability is strongest at slower-evolving neural time scales and is more evident in functional connectivity estimates than in response time series.*
>
> *Strengths:*

> *This is a very thorough paper that tackles this question from several different angles. I very much appreciate the use of hyperalignment to factor out topographic differences, and I found the relationship between heritability and neural time scales very interesting. The writing is clear, and the results are compelling.*

We thank Reviewer 1 for their kind words and enthusiastic support of our manuscript.

> *Weaknesses:*
>
> *The only "weaknesses" I identified were some points where I think the methods, interpretation, or visualization could be clarified.*
>
> *(1) On page 16, the authors compare heritability in functional connectivity (FC) and response time series, and find that the heritability effect is larger in FC. In general, I agree with your diagnosis that this is in large part due to the fact that FC captures the covariance structure across parcels, whereas response time series only diverge in terms of univariate time-point-by-time-point differences. Another important factor here is that (within-subject) FC can be driven by intrinsic fluctuations that occur with idiosyncratic timing across subjects and are unrelated to the stimulus (whereas time-locked metrics like ISC and timeseries differences cannot, by definition). This makes me wonder how this connectivity result would change if the authors used inter-subject functional connectivity (ISFC) analysis to specifically isolate the stimulus-driven components of functional connectivity (Simony et al., 2016). This, to me, would provide a closer comparison to the ISC and response time series results, and could allow the authors to quantify how much of the heritability in FC is intrinsic versus stimulus-driven. I'm not asking that the authors actually perform this analysis, as I don't think it's critical for the message of the manuscript, but it could be an interesting future direction. As the authors discuss on page 17, I also suspect there's something fundamentally shared between response time series and connectivity as they relate to functional topography (Busch et al., 2021) that drives part of the heritability effect.*

We agree that investigating the heritability of ISFC (or stimulus-driven functional connectivity) would make for a very interesting future direction. Ultimately, we chose to analyze FC (vs. ISFC) profiles to allow for direct comparison with the sizable existing literature on the heritability of FC (such as in our Movie vs. Rest FC analysis) and decided to refrain from analyzing ISFC data in order to keep the present manuscript focused. ISFC analysis of this dataset will be a focus of future work.

> *(2) The observation that regions with intermediate ISC have the largest differences between MZ, DZ, and UR is very interesting, but it's kind of hard to see in Figure 1B. Is there any other way to plot this that might make the effect more obvious? For example, I could imagine three scatter plots where the x- and y-axes are, e.g., MZ ISC and UR ISC, and each data point is a parcel. In this kind of plot, I would expect to see the middle values lifted visibly off the diagonal/unity line toward MZ. The authors could even color the data points according to networks, like in Figure 3C. (They also might not need to scale the ISC axis all the way to r = 1, which would make the differences more visible.)*

We thank R1 for this helpful suggestion- we originally set the y-axis limits to r = 1 in order to facilitate comparison between ISC (Fig. 1B) and FC profile (Fig. 6B) similarity, but we agree that this renders the group differences harder to discern and have updated the plot accordingly (along with thicker lines to enhance readability). We prefer to keep the line plots in the main body as they allow for direct comparison of all three groups on the same plot, but we have included the scatter plot version in Fig. S2 for those who are interested.

> *(3) On page 9, if I understand correctly, the authors regress the vector of ISC values across parcels out of the vector of heritability values across parcels, and then plot the*

> *residual heritability values. Do they center the heritability values (or include some kind of intercept) in the process? I'm trying to understand why the heritability values go from all positive (Figure 2A) to roughly balanced between positive and negative (Figure 2B). Important question for me: How should we interpret negative values in this plot? Can the authors explain this explicitly in the text? (I also wonder if there's a more intuitive way to control for ISC. For example, instead of regressing out ISC at the parcel/map level, could they go into a single parcel and then regress the subject-level pairwise ISC values out when computing the heritability score?).*

We indeed included an intercept in this model using MATLAB's fitlm function. This means that the model estimates the best-fitting line of the following form: $\text{heritability}_i=\beta0+\beta1\text{ISC}_i+\varepsilon_i$. We agree that the interpretation of these ε_i values and alternative approaches to controlling for ISC should be clarified. As such, we have added the following passages to the text:

Methods: "Because the heritability of ISC is constrained by the degree of synchronization in a given area, we also sought to identify areas in which BOLD time courses were more/less heritable than would be expected based on ISC alone by fitting a linear model of the form $\text{heritability}_i=\beta0+\beta1\text{ISC}_i+\varepsilon_i$ and plotting the residuals. Regarding alternative approaches to controlling for ISC, although the heritability model introduced by Ge et al. allows for the inclusion of covariates defined at the subject level (e.g., age), it does not allow for covariates that are defined at the dyad level (e.g., pairwise ISC)."

Results: "Here, negative values in the residual map indicate parcels where heritability is lower than expected based on ISC, while positive values indicate higher-than expected heritability."

> *(4) On page 4 (line 155), the authors say "we shuffled dyad labels"- is this equivalent to shuffling rows and columns of the pairwise subject-by-subject matrix combined across groups? I'm trying to make sure their approach here is consistent with recommendations by Chen et al., 2016. Is this the same kind of shuffling used for the kinship matrix mentioned in line 189?*

Briefly, shuffling the kinship matrix involved permuting the rows and columns of the matrix in the same manner (also known as the quadratic assignment procedure), whereas shuffling the dyad labels involved random permutations of the three group labels (MZ, DZ, unrelated), which could not be done through matrix operations as the age- and gender matching precluded the use of a complete similarity matrix. However, given concerns raised by Reviewer 2, we have removed our significance claims from this (and similar) sections, which we discuss in more detail in response to Reviewer 2's weakness A.

> *(5) I found panel A in Figure 4 to be a little bit misleading because their parcel-wise approach to hyperalignment won't actually resolve topographic idiosyncrasies across a large cortical distance like what's depicted in the illustration (at the scale of the parcels they are performing hyperalignment within). Maybe just move the green and purple brain areas a bit closer to each other so they could feasibly be "aligned" within a large parcel. Worth keeping in mind when writing that hyperalignment is also not actually going to yield a one-to-one mapping of functionally homologous voxels across individuals: it's effectively going to model any given voxel time series as a linear combination of time series across other voxels in the parcel.*

We agree that our efforts to present a simplified depiction of hyperalignment may mislead less familiar readers and have amended Fig. 4A according to this suggestion. We have also added text to the methods section (below) to clarify that the outputs of hyperalignment are time series that reflect linear combinations of other voxels' time series from that parcel.

"This approach independently transforms each subject's data within discrete anatomical

parcels into the common space, yielding functionally aligned vertex time series that are calculated as weighted linear combinations of the original time series from all other vertices within that same parcel for that subject."

> *(6) I believe the subjects watched all different movies across the two days, however, for a moment I was wondering "are Day 1 and Day 2 repetitions of the same movies?" Given that Day 1 and Day 2 are an organizational feature of several figures, it might be worth making this very explicit in the Methods and reminding the reader in the Results section.*

We agree that this would be helpful and have added the following text to the relevant sections:

"All clips were only viewed once by each subject, with the exception of the brief montage which was included at the end of each of the four runs for test-retest purposes."

"To characterize the heritability of brain responses to complex stimuli, we used 7T fMRI data from 178 HCP Young Adult subjects acquired across two days (using two largely non-overlapping sets of movie stimuli, see Methods)…"

References:

Busch, E. L., Slipski, L., Feilong, M., Guntupalli, J. S., di Oleggio Castello, M. V., Huckins, J. F., Nastase, S. A., Gobbini, M. I., Wager, T. D., & Haxby, J. V. (2021). Hybrid hyperalignment: a single high-dimensional model of shared information embedded in cortical patterns of response and functional connectivity. NeuroImage, 233, 117975. *https://doi.org/10.1016/j.neuroimage.2021.117975*

Chen, G., Shin, Y. W., Taylor, P. A., Glen, D. R., Reynolds, R. C., Israel, R. B., & Cox, R. W. (2016). Untangling the relatedness among correlations, part I: nonparametric approaches to inter-subject correlation analysis at the group level. NeuroImage, 142, 248259. *https://doi.org/10.1016/j.neuroimage.2016.05.023*

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. Nature Communications, 7, 12141. *https://doi.org/10.1038/ncomms12141*

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *The authors attempt to estimate the heritability of brain activity evoked from a naturalistic fMRI paradigm. No new data were collected; the authors analyzed the publicly available and well-known data from the Human Connectome Project. The paper has 3 main pieces, as described in the Abstract:*
>
> *(1) Heritability of movie-evoked brain activity and connectivity patterns across the cortex.*
>
> *(2) Decomposition of this heritability into genetic similarity in "where" vs. "how" sensory information is processed.*
>
> *(3) Heritability of brain activity patterns, as partially explained by the heritability of neural timescales.*
>
> *Strengths:*
>
> *The authors investigate a very relevant topic that concerns how heritable patterns of brain activity among individuals subjected to the same kind of naturalistic stimulation are. Notably, the authors complement their analysis of movie-watching data with resting-state data.*

Weaknesses:

The paper has numerous problems, most of which stem from the statistical analyses. I also note the lack of mapping between the subsections within the Methods section and the subsections within the Results section. We can only assess results after understanding and confirming the methods are valid; here, however, Methods and Results, as written, are not aligned, so we can't always be sure which results are coming from which analysis.

(A) Intersubject correlation (ISC) (section that starts from line 143): "We used nonparametric permutation testing to quantify average differences in ISC for each parcel in the Schaefer 400 atlas for each day of data collection across three groups: MZ dyads, DZ dyads, and unrelated (UR) dyads, where all UR dyads were matched for gender and age in years." ... "some participants contributed to ISC values for multiple dyads (thus violating independence assumptions)"

This is an indirect attempt to demonstrate heritability. And it's also incorrect since, as the authors themselves point out, some subjects contribute to more than one dyad.

Permutation tests don't quantify "average differences", they provide a measure of evidence about whether differences observed are sufficient to reject a hypothesis of no difference.

Matching subjects is also incorrect as it artificially alters the sample; covarying for age and sex, as done in standard analyses of heritability, would have been appropriate.

It isn't clear why the authors went through the trouble of implementing their own nonparametric test if HCP recommends using PALM, which already contains the validated and documented methods for permutation tests developed precisely for HCP data.

The results from this analysis, in their current form, are likely incorrect.

We appreciate that permutation tests do not quantify average differences and intended to write “We used non-parametric permutation testing to quantify [the significance of] average differences…”. Our intention with this analysis was not to demonstrate heritability, but rather to quantify group differences in ISC in a manner that is interpretable for readers who are unfamiliar with h^2 (e.g., “identical twins’ BOLD time courses were 59% more similar than those from pairs of unrelated individuals”) and motivate the formal heritability analysis used later in the paper. Indeed, all of the heritability analyses in this paper leveraged a validated multidimensional heritability method first introduced by Ge et al. (2016) and used by many other investigators since then. Furthermore, we covaried for age and sex at the subject level in all our heritability analyses, and always tested the significance of these heritability values using a validated permutation procedure (the quadratic assignment procedure; Hubert & Schultz, 1976) that respects the non-independence of dyadic data.

Regarding the shuffling procedure used for Figure 1, while PALM is the standard for univariate, subject-level GLMs in the HCP pipeline and can accommodate nested designs (i.e., subjects within families), it is not designed to handle the unique relational dependencies of dyadic ISC analysis (i.e., the same subject contributing to multiple dyads). Although the element-wise resampling approach was the most appropriate approach available, it is known to inflate the false positive rate (Chen et al., 2016; doi:10.1016/j.neuroimage.2016.05.023); given that this analysis was simply meant to motivate our later hypothesis testing heritability analyses, we have removed significance claims from this section of the manuscript. Still, we emphasize that this has no bearing on the validity of our conclusions which were supported

by our formal heritability analyses; throughout our paper we have correctly used the appropriate methods to back the stated claims.

> *(B) Functional connectivity (FC) (section that starts from line 159): Here the authors compute two 400x400 FC matrix for each subject, one for rest, one for movie-watching, then correlate the correlations within each dyad, then compared the average correlation of correlations for MZ, DZ, and UR. In addition to the same problems as the previous analysis, here it is not clear what is meant by "averaging correlations [...] within a network combination". What is a "network combination"? Further, to average correlations, they need to be r-to-z transformed first. As with the above, the results from this analysis in its current form are likely incorrect.*

We regret that R2 had difficulty understanding our analysis and have added the following text to the relevant Methods section to clarify our approach:

“For example, there are 16 parcels in the Kong et al. Auditory network and 17 parcels in the Language network, so the FC profile for a given subject’s Auditory-Language network combination consists of the (16 * 17 =) 272 correlation coefficients between all unique pairs of one parcel from each network.”

As we stated in the previous Methods paragraph, “All Pearson *r* values in this and all other analyses were Fisher *z*-transformed before averaging (and converted back to Pearson *r* for visualization)”. Thus, contrary to the reviewer’s assertion, these analyses were performed correctly. Once again, we emphasize that this analysis was not intended to demonstrate heritability, but rather to describe group differences in FC in familiar units.

> *(C) ISC and FC profile heritability analyses (section that starts from line 175): Here, the authors use first a valid method remarkably similar to the old Haseman-Elston approach to compute heritability, complemented by a permutation test. That is fine. But then they proceed with two novel, ill-described, and likely invalid methods to (1) "compare the heritability of movie and rest FC profiles" and (2) to "determine the sample size necessary for stable multidimensional heritability results". For (1), they permute, seemingly under the alternative, rest and movie-watching timeseries, and (2), by dropping subjects and estimating changes in the distribution.*
>
> *The (1) might be correct, but there are items that are not clearly described, so the reader cannot be sure of what was done. What are the "153 unique network combinations"? Why do the authors separate by day here, whereas the previous analyses concatenated both days? Were the correlations r-to-z transformed before averaging?*
>
> *The (2) is also not well described, and in any case, power can be computed analytically; it isn't clear why the authors needed to resort to this ad hoc approach, the validity of which is unknown. If the issue is the possibility that the multidimensional phenotypic correlation matrix is rank-deficient, it suffices that there are more independent measurements per subject than the number of subjects.*

Regarding (1), we have clarified in section 2.6 that the 153 unique network combinations reflect each unique pair of 17 Kong networks. All of our analyses, including this one, were performed separately for each day of data collection, as we state throughout the paper and visualize in our figures (although we acknowledge that, on some occasions, we [conservatively] performed FDR-correction on a combined set of p-values, as discussed in our response to K). Given that the null hypothesis for this analysis is that rest FC and movie FC are equally heritable, we are not sure why permuting rest and movie FC matrices would be invalid. All Pearson r values were z-transformed before averaging, as we stated in our paper.

Regarding (2), we included this analysis in response to editorial concerns that our heritability analyses were not sufficiently powered, and we chose this approach because it serves as a

simple way to demonstrate the stability of our results at various sample sizes whose validity is self-evident. Furthermore, this sort of subsampling approach has been used many times before in our field (e.g., Marek et al., 2022) and others (e.g., Manyara et al., 2024) to demonstrate the sample-size dependence and stability of statistical effects. We have added text explaining this to the relevant Methods section (2.6).

> *(D) Frequency-dependent ISC heritability analysis (from line 216): Here, the authors decompose the timeseries into frequency bands, then repeat earlier analyses, thus bringing here the same earlier problems and questions of non-exchangability in the permutations given the dyads pattern, r-z transforms, and sex/age covariates.*

We did not use dyadic permutation testing for any of the frequency-dependent ISC analyses; rather, we used the jackknife SEMs to compare heritability across frequency bands and have added an explicit description of this to section 2.7. We have addressed the r-z transform and covariate concerns in previous comments.

> *(E) FC strength heritability analysis (from line 236): Here, the authors use the univariate FC to compute heritability using valid and well-established methods as implemented in SOLAR. There is no "linkage" being done here (thus, the statement in line 238 is incorrect in this application. SOLAR already produces SEs, so it's unclear why the authors went out of their way to obtain jackknife estimates. If the issue is non-normality, I note that the assumption of normality is present already at the stage in which parameters themselves are estimated, not just the standard errors; for non-normal data, a rank-based inversenormal transformation could have been used. Moreover, typically, r-to-z transformed values tend to be fairly normally distributed. So, while the heritabilities might be correct, the standard errors may not be (the authors don't demonstrate that their jackknife SE estimator is valid). The comparison of h2 between dyads raises the same questions about permutations, age/sex covariates, and r-z transforms as above.*

We used jackknife SEs for these analyses to maintain consistency with the multidimensional heritability package used here, which only outputs jackknife SEs. We note that this jackknife approach (and the corresponding multidimensional heritability analysis) was detailed in prior work (Anderson et al., 2021), and that the leave-one-family-out jackknife has a long history of being used to estimate SEs in heritability studies, especially when working with smaller samples (Knapp et al., 1989). We are also not sure what “the comparison of h2 between dyads” means- heritability cannot be compared “between” dyads; rather, it is defined across dyads.

> *(F) Hyperalignment (from line 245): It isn't clear at this point in the manuscript in what way hyperalignment would help to decompose heritability in "where vs. how" (from the Abstract). That information and references are only described much later, from around line 459. The description itself provides no references, and one cannot even try to reproduce what is described here in the Methods section. Regardless, it isn't entirely clear why this analysis was done: by matching functional areas, all heritabilities are going to be reduced because there will be less variance between subjects. Perhaps studying the parameters that drive the alignment (akin to what is done in tensor-based and deformation-based morphometry) could have been more informative. Plus, the alignment process itself may introduce errors, which could also reduce heritability. This could be an alternative explanation for the reduced heritability after hyperalignment and should be discussed. An investigation of hyperaligment parameters, their heritability, and their co-heritability with the BOLD-phenotypes can inform on this.*

To help set up our hyperalignment analyses, we have added text to the introduction explaining how hyperalignment would help to decompose heritability. The description in the Methods section included a reference to Bazeille et al., 2021, in which the hyperalignment method used here is discussed in detail. Still, we have added citations to additional papers

(also cited in the Bazeille et al. paper, and elsewhere in our paper) in case that might be helpful. We note that it is not the case that all heritabilities were reduced by hyperalignment- as can be seen in Figs. 4D, 8A, and S15, hyperalignment did increase heritability in some voxels and network combinations. This would be expected under the alternative (albeit unlikely) hypothesis that functional topographies are not heritable, such that topographic variation between related individuals would obscure similarities in their (heritable) topography-independent brain responses. Recognizing that this alternative is unlikely, we believe the main novelty of this analysis comes from the magnitude of the hyperalignment effect (up to 40% of brain-wide heritability) and its spatial pattern (e.g., larger heritability decreases in visual vs. auditory cortex, the opposite of our NT result).

We agree that we would see lower post-hyperalignment heritability if the alignment process itself introduced errors/noise, but this would be deeply surprising as hyperalignment increases ISC by design (and errors/noise could only decrease ISC). To demonstrate this, we have added Figure S7 which shows that (as expected) ISC across all voxels and subject pairs increases after hyperalignment (and that this increase is larger when hyperalignment is performed in larger parcels). Given that hyperalignment increased ISC, and that it is blind to twin status, we are unsure how it could have introduced errors that would have confounded this result.

> *(G) Relationships between parcel area and heritability (from line 270): As under F), how much the results are distorted likely depends on the accuracy of the alignment, and the error variance (vs heritable variance) introduced by this.*

We agree that alignment accuracy could potentially impact parcel-level differences in how much heritability changes following hyperalignment, and we included the frequency dependent $h^2_{residuals}$ (controlling for differences in ISC) in Fig. 3 for this reason, as more accurate hyperalignment should result in greater increases in ISC, raising the heritability ceiling. We note that we observe similar relationships between parcel rank and frequency dependent changes in these residualized maps, suggesting that our parcel-level differences are not simply the result of better alignment in more sensory parcels.

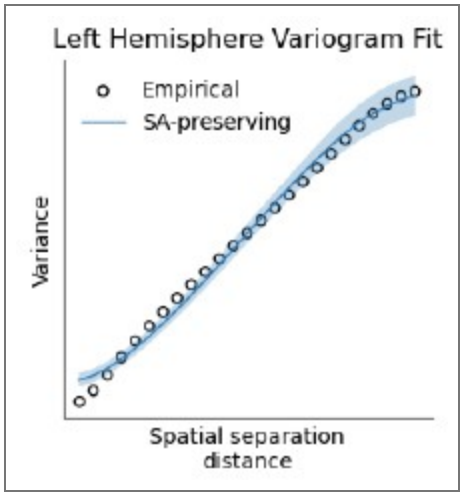
> *(H) Neural timescale analyses (from line 280): Here, a valid phenotype (NT) is assessed with statistical methods with the same limitations as those previously (exchangability of dyads, age/sex covariates, and r-z transforms). NT values are combined across space and used as covariates in "some multivariate analyses". As a reader, I really wanted to see the results related to NT, something as simple as its heritability, but these aren't clearly shown, only differences between types of dyads.*

We have addressed the exchangeability, covariates, and r-z transform comments above (in A). As we explained for our FC strength analyses, we are underpowered to evaluate the heritability of unidimensional traits (like the heritability of NT magnitude), and the heritability of a closely-related measure (BOLD turnover magnitude) has already been established in a larger sample of HCP subjects (*https://doi.org/10.1152/jn.00402.2022*). Still, we agree that more results related to the heritability of NTs would be of interest to our readers. As such, we have added an analysis in section 3.4 quantifying the heritability of multivariate NT topographies and used SOLAR to quantify the heritability of NT magnitudes, with the disclaimer that this and similar analyses are underpowered (hence the large difference in day 1 and day 2 heritability effect sizes). We also removed significance claims for the dyadic NT similarity analysis.

> *(I) Significance testing for autocorrelated brain maps and FC matrices (from line 310): Here, the authors suddenly bring up something entirely different: reliability of heritability maps, and then never return to the topic of reliability again. As a reader, I find this confusing. In any case, analyses with BrainSMASH with well-behaved, normally distributed data are ok. Whether their data is well behaved or whether they ensured that*

> *the data would be well behaved so that BrainSMASH is valid is not described. As to why Spearman correlations are needed here, Mantel tests, or whether the 1000 "surrogate" maps are valid realizations of the data under the null, remains undemonstrated.*

We brought up reliability in this section because we show the reliability of our results across the two days of data collection several times in the paper. R2 is correct to point out that BrainSMASH was validated using normally distributed brain maps, and although some of our brain maps contain normally distributed values, others are right skewed (due largely to the fact that many voxels/parcels exhibit low ISC while visual/auditory areas have very high ISC). In preparing our original manuscript, we visualized BrainSMASH's variogram outputs for one of the most skewed inputs (vertex-wise BOLD time course heritability) and found that the autocorrelation structures of the empirical and null maps were well-matched. We did not include this in the original manuscript as it is not commonplace in the field to report the variograms, see Author response image 1. Furthermore, our use of Spearman (vs. Pearson) correlations renders these distributional differences less relevant, as the Spearman correlation transforms all inputs to a uniform distribution. To empirically check that these distributional differences do not bias our results, we retested the significance of all brain map associations using the spin test (10.1016/j.neuroimage.2018.05.070), an alternative method that does not assume normally distributed inputs, and obtained identical p-values for all analyses (P<.001 in all cases).



Author response image 1.

> *(J) Global signal was removed, and the authors do not acknowledge that this could be a limitation in their analyses, nor offer a side analysis in which the global signal is preserved.*

Although we agree that GSR is a contentious preprocessing step for certain analyses, it has explicitly been shown to increase ISC signal-to-noise without compromising FC fingerprints (Graff et al., 10.1016/j.dcn.2022.101087), and it is uncommon to perform ISC analyses with and without GSR. Still, we have added additional text to our Methods section explaining our rationale for using GSR and that this could affect our results. We also re-ran our main analysis (BOLD time course heritability) with and without GSR and found that GSR had little impact on our results; we have included this in our manuscript as Fig. S4.

Specifically, we see that GSR resulted in a slight increase in heritability (average Day 1 h^2 with/without GSR = .064/.060; Day 2: .068/.061) and almost no effect on the spatial pattern of our results (With GSR/without GSR Spearman ρ = .99, $P_{\text{brainSMASH}} < .001$ on both Day 1 and Day 2).

> *(K) FDR is used to control the error rate, but in many cases, as it's applied to multiple sets of p-values, the amount of false discoveries is only controlled across all tests, but not within each set. The number of errors within any set remains unknown.*

We agree that the FDR usage in our original manuscript was inconsistent, in that for two analyses we FDR-corrected p-values from the two days of data collection together (instead of correcting p-values from each day separately and reporting voxels/parcels/etc. that were significant at q<.05 on both days, as in the rest of our analyses). We note that both approaches are more conservative than reporting significant results at q<.05 separately; regardless, to maintain consistency we have updated all analyses such that FDR correction is always performed separately for each day of data collection.

> *(L) Generally, when studying the heritability of a trait, the trait must be defined first. Here, multiple traits are investigated, but are never rigorously defined. Worse, the trait being analyzed changes at every turn.*

Here, we analyze the heritability of movie-evoked BOLD time courses (Figures 1-5) as well as FC profiles (Figures 6-8). We defined FC profiles in our Introduction as an individual's pattern of pairwise FC strengths (and further detailed how we quantified FC profiles in the relevant Methods section), and believe that "BOLD time course" is a well understood phrase in the field and does not need to be further defined. We also used hyperalignment to decompose the heritability of these traits into topography-dependent and independent portions, and (new to this version) also explicitly quantify the heritability of neural timescales, which we defined as the AUC of the ACF until the first negative ACF value in both the relevant Results and Methods sections.

To make this clearer, we have modified the last paragraph of our Introduction to begin with:

In the present work, we address these questions by analyzing 7T fMRI recordings of a twin sample acquired by the Human Connectome Project (Van Essen et al., 2013) to quantify the heritability of two distinct high-dimensional traits—stimulus-evoked BOLD time courses and functional connectivity profiles—across the cortex.

> ***Reviewer #3 (Public review):***
>
> *Strengths:*
>
> *It's sort of novel to study the heritability of movie-watching fMRI data. The methodology the authors used in the paper is also supportive of their findings. Figures are nicely organized and plotted. They finally found that sensory processing in the human brain is under genetic control over stable aspects of brain function (here referring to neural timescale and resting state connectivity).*
>
> *Weaknesses:*
>
> *What I am worried about most is the sample size and interpretation of heritability.*
>
> *(1) Figure 1. I assumed that the authors just calculated the ISC within each group (MZ, DZ, and UR). Of course, you can get different variations between each group. Therefore, there is heritability. Why not calculate ISC across the whole sample, then separate MZ, DZ, and UR?*

We believe that this question is getting at the difference between pairwise ISC (i.e., correlating one BOLD time course from one subject with that from another subject) and leave-one-subject-out ISC (i.e., correlating one BOLD time course from one subject with the corresponding average time course across all other subjects). We chose to use the pairwise ISC method because it allows us to capitalize on the information contained in the n^2 pairwise

ISC matrix (whereas the other approach averages out meaningful information to yield a n^1 ISC matrix) and leverage a more sophisticated multidimensional heritability approach. Also, the leave-one-subject-out approach introduces additional issues re: handling family-level data (e.g., should we include a subject's twin in the leave-one-subject-out average? If so, how should we handle subjects who don't have a twin in the dataset, as averaging data from different numbers of subjects will lead to different ISC magnitudes? etc.).

> *(2) Heritability scores in the paper are sort of small. If the sample size is small, please consider p-values, which will tell more about the trustworthiness of your heritability.*

We report p-values for heritability throughout our paper (e.g., stating that BOLD time courses are significantly heritable in 99% of parcels in Figure 2), and we believe that the reliability of our spatial maps across days of data collection (also quantified with p-values) further demonstrates the trustworthiness of our results. Finally, as we demonstrate in Figure S5, our sample size is more than sufficient to reliably detect small effects.

> *(3) I don't understand the high-frequency signals in fMRI data. It's always regarded as noise, the band 1 here in particular.*

In addition to driving shared neuronal responses (which are captured in BOLD signal oscillations <.1 Hz or so), movies also elicit shared cardiac, respiratory, and motion responses across participants at higher frequencies. Although we used a relatively conservative denoising approach here, we believe some of these non-neuronal signals are still present in our data; alternatively, it is also possible that these signals reflect "fast" BOLD responses at >.15 Hz (as discussed in 10.1016/j.neuroimage.2021.118658). In any case, the fact that information in this frequency band is considerably less heritable than information in slower frequency bands supports the idea that this band is noisier and suggests that our heritability results are driven by canonical neuronal activity-related BOLD signals.

> *(4) The statement "we show that the heritability of brain activity patterns can be partially explained by the heritability of the neural timescale" should come from Figure 5. However, after controlling for NT, the heritability decreased max. 0.025 in temporal areas. I am not sure this change supports the statement. If the visual cortex is outlined, and combining ISC changes in the visual cortex, I think this would somehow be answered. Instead of delta h2, adding a new model h2 would be obvious to the readers.*

Although the decrease of 0.025 is small, we note that this constitutes around ~50% of BOLD time course heritability in some voxels (seen in comparison to Fig. 4C), and the spatial pattern of this result is quite consistent across days of data collection, indicating its reliability. Furthermore, the whole-brain distributions of results shown in Fig. 5B are clearly skewed towards negative values, indicating that controlling for NT partially reduces (or "explains") BOLD time course heritability. Still, we agree that showing raw h^2 values in addition to the difference maps would be helpful for some readers and have added a corresponding supplementary figure (S12) which shows these.

> *(5) Figures 7 and 8, when getting the difference of heritability, please also consider the standard errors of the heritability estimates. Then you can compare across networks/regions.*

We did consider adding standard errors for these heritability estimates, but found that visualizing standard errors for each of the 153 unique network combinations in our heatmaps rendered the visualizations difficult to parse, and given that our hypotheses concerned global (e.g., hyperaligned vs. MSM-aligned) or network-level (e.g., sensory vs. associative) patterns, we focused on calculating standard errors/p-values for these analyses (although we note that dyad-level standard errors can be found in Fig. 6B, where they are clearly marginal compared to the group effects).

> *(6) I think movie VS resting state is a really important result in this paper. However, there is almost no discussion. Discussing this part would be more beneficial for understanding the genetic control over the neuron arousal and excitation circuits.*

We agree that this result was relatively under-explored in our Discussion section and have added additional text (lines 851-855) to connect this result to recent work on arousal-dependent uniqueness of FC.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) Do the authors have any ideas why we see this hotspot of heritability in pMTG/LOTC? It really jumps out in Figure 1A and Figure 2. The more posterior sensory MT+ area seems to drop when regressing out ISC in Figure 2B, but this pMTG area stays hot. Is there anything special about this kind of multimodal biological motion/action observation / social perception area (Pitcher & Ungerleider, 2021)? I don't think this is necessary to discuss in the manuscript, but I'm curious if the authors have any speculation.*

We are not certain as to why BOLD time courses in this parcel are particularly heritable-although this area is associated with biological motion, that particular function tends to be more right lateralized, and here we see nominally higher heritability in the left hemisphere. Per a Neurosynth review (and consistent with the left lateralization), we believe this may have more to do with speech processing, but a more definitive answer will require further investigation.

> *(2) Page 3, line 127: "More information on these clips"-it might be worth saying a little bit more here just to make sure people understand that these are audiovisual clips, they include language, they're long enough to convey meaningful social and narrative information, etc.*

We agree and have added additional details on the clip composition to the relevant methods paragraph.

> *(3) Figure 1 caption: can you add a sentence reminding readers what's going on with Day 1 and Day 2?*

We thank R1 for this suggestion and have added a sentence to this effect at this location.

> *(4) Page 9, line 379: "although these more associative parcels do not encode a substantial amount of stimulus-specific information"-is this really true? I suspect these association areas still have decent ISCs, even if there are many processing stages downstream of the raw stimulus.*

Although these parcels are not the most synchronized by the stimulus, we agree that it is unfair (and vague) to say that they do not encode a substantial amount of stimulus-specific information. We have edited this sentence to make a more specific claim and highlight the relatively lower ISC in these parcels vs. more unimodal sensory areas.

> *(5) Page 9, line 417: Can you unpack a bit more what you mean by "supra-BOLD frequency band"?*

Here, we refer to the fact that BOLD signals resulting from neuronal firing events have frequencies below ~.15 Hz (Josephs and Henson, 1999). We have added additional text and the Josephs and Henson citation to this line to further unpack this point.

> *(6) Page 18, line 695: This discussion of how attention and gaze might partly shape response time series reminded me of recent work by Borovska & de Haas (2024)-might be worth citing.*

We are grateful to R1 for alerting us to this very relevant work and have included a reference to it in our discussion.

> *(7) Page 19, line 755: I'm not sure I'd describe the hyperalignment results here as a "deleterious effects [on] heritability"-my reading was that hyperalignment allows you to say something more specific about heritability of function by allowing you to effectively factor out heritability effects that reduce to individual differences cortical topography; this seems like a good thing!*

We agree that "deleterious" was a poor word choice given its negative connotation, and have edited this sentence to read:

"With this in mind, future studies investigating genetic correlations between brain function and behavioral variables may benefit from hyperalignment, as it can factor out individual-specific cortical topography and thus yield more precise estimates of functional heritability."

> *(8) I would love to see a ventral view in some of these plots! Not asking you to recreate the figures, but the ventral temporal cortex is an area of interest for many folks in the movie fMRI space (e.g., Haxby et al., 2011).*

We agree that ventral views would be of interest to some readers and have added the corresponding maps for our main results in supplementary figures S3 and S9.

References:

Borovska, P., & de Haas, B. (2024). Individual gaze shapes diverging neural representations. Proceedings of the National Academy of Sciences, 121(36), e2405602121. *https://doi.org/10.1073/pnas.2405602121*

Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., Hanke, M., & Ramadge, P. J. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron, 72(2), 404416. *https://doi.org/10.1016/j.neuron.2011.08.026*

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. Trends in Cognitive Sciences, 25(2), 100-110. *https://doi.org/10.1016/j.tics.2020.11.006*

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) To address the common core analytical problems listed under A), B), C), D), E), and basically throughout the methods:*
>
> *(a) Conduct permutations with exchangability restrictions to account for the pattern of dyad-relationships as e.g. implemented in PALM.*
>
> *(b) Control for age and sex covariates as covariates (e.g. as in SOLAR), rather than by matching.*
>
> *(c) Perform r-to-z transforms when conducting further analyses on correlations that assume normality.*
>
> *(d) For all analyses that assume normal distributions, e.g. in SOLAR and BrainSMASH, check that this is the case.*

We have explained how PALM is not suited for the study of effects that are defined at the dyad level (A), that we controlled for age and sex covariates in all our formal heritability analyses in our original submission (B), that we always performed r-to-z transforms when indicated in our original submission (C), and that our spatial permutation results don't hinge on distributional differences (D).

> *(2) Replace SEs derived from kacknife approach with those from SOLAR, or provide a comparison and motivation and/or demonstrate that SEs are correct.*

A more thorough explanation of the block jackknife procedure can be found in prior work introducing the multidimensional heritability method used here (Anderson et al., 2021).

> *(3) Given problem (F & G):*
>
> *(a) Consider studying the parameters that drive the hyperalignment. They can be included as covariates in heritability analyses, and/or their heritability is of interest to understand the reasons for the heritability reduction post-hyperaligment.*

We agree that this would be interesting but the specific parameters that drive hyperalignment are beyond the scope of this study.

> *(b) Include the alternative explanation of hyperalignment-induced noise in the discussion.*

We have added a figure showing that hyperalignment does not increase noise in ISC and explained here why "hyperalignment-induced noise" does not constitute a reasonable alternative explanation for our results.

> *(4) Add heritability results for NT phenotypes.*

We have added heritability analyses for NT topography and (global) NT magnitude, as detailed above.

> *(5) Motivate global signal removal, and acknowledge this process typically alters results substantially.*

We have added an explanation of our rationale for using GSR and shown in this response that it does not in fact substantially alter the results.

> *(6) Rephrase and/or clarify the following:*
>
> *(a) "permutations quantify average differences" (under A).*
>
> *(b) "network combinations" and related analyses (under B & C).*
>
> *(c) why some analyses are separated per visit/day and others not (C).*
>
> *(d) methods and reasons for sample size estimation (C).*

We have rephrased or clarified all of the above.

> ***Reviewer #3 (Recommendations for the authors):***
>
> *(1) Participants should be recleared. I know HCP 7T data has 184 subjects. How can the authors have 176 twins and 690 unrelated subjects?*

As we reported in our Methods section, 178 subjects had complete movie-watching datasets, and 176 subjects had complete movie-watching and resting-state datasets. Of the 178 subjects with complete movie-watching data, we identified 690 age- and sex-matched dyads.

> *(2) Figure 1. I don't find Figure S1A in Figure S1.*

We thank R3 for catching this error- we have amended this reference to read Fig. S1.

> *(3) I could also suggest putting Figure 1 and Figure 2 together.*

We thank R3 for this suggestion- ultimately, we prefer to keep these figures separate to reinforce the difference between our dyadic similarity and formal heritability analyses.

https://doi.org/10.7554/eLife.106081.2.sa0