Spontaneously emerging cortical representations of visual objects and features in human visual cortex

DoHyun Kim[1], Tomer Livne[2], Nicholas V. Metcalf [3], Maurizio Corbetta[3,4,5,6], Gordon L. Shulman[3,4]

Departments of [1]Biomedical Engineering, [3]Neurology, [4]Radiology, and [5]Anatomy & Neurobiology at Washington University School of Medicine, St. Louis, MO 63110, USA
Department of [2]Neurobiology, Weizmann Institution of Science, Rehovot 76100, Israel.
Department of [6]Neuroscience and [7]Padova Neuroscience Center (PNC), University of Padova, Padova, 35122, Italy

Correspondence: Gordon L. Shulman

Washington University Medical School, St. Louis, MO 63110, USA

Department of Neurology, Campus Box 8225

4525 Scott Avenue, Rm 2121

St. Louis, MO. 63110

Phone: (314) 362-7666

Fax: (314) 362-6911

E-mail: gshulman@wustl.edu

Keywords: BOLD, fMRI, Spontaneous activity, multivariate pattern analysis (MVPA), Resting-state, functional connectivity.

Abstract

The function of spontaneous brain activity is unknown. Here we test the novel hypothesis that patterns of spontaneous activity reflect not only synaptic homeostasis or synchronization of neuronal populations, but code representational patterns evoked by stimuli and tasks. We compared in visual cortex the spatial patterns of spontaneous activity to the patterns evoked by ecological visual stimuli (faces, bodies, scenes) and low-level visual features (e.g. phase-scrambled faces). We identified regions that preferred particular stimulus categories during localizer scans, measured multivariate spatial patterns for each category during task scans, and then spatially correlated these stimulus-evoked patterns to the pattern measured in each frame of resting-state scans. The mean correlation coefficient was essentially zero for all regions/stimulus categories, indicating that resting activity patterns were not biased toward particular stimulus categories. However, the spread of correlation coefficients, i.e. both positive and negative, was significantly greater for a stimulus category over the ROIs preferring that category (e.g. the body category in body-preferring ROIs). Therefore, the putative representational content of spontaneous activity was related to stimulus-evoked spatial activity patterns. This content also governed the temporal correlation or functional connectivity of spatial patterns of spontaneous activity between individual regions. Resting spatial activity patterns related to an object category (e.g. bodies) fluctuated preferentially between ROIs preferring the same category. Moreover, activity patterns related to different categories fluctuated independently within their respective preferred ROIs. These results support the general proposal that spontaneous multi-voxel activity patterns are linked to stimulus-evoked patterns, consistent with a representational function for spontaneous activity.

Significance Statement

Neurons throughout the brain are spontaneously active. Although this activity was once thought to reflect only noise, the remarkable spatiotemporal regularities of spontaneous activity have motivated functional explanations. Here we address the hypothesis that spontaneous activity codes the representational structure evoked by stimuli. We show that the spatial pattern of stimulus-evoked activity across high-level areas of visual cortex that prefer visual categories such as bodies or scenes is related to the spatial pattern of spontaneous activity across the same areas. This spatial structure partly governs the spontaneous spatiotemporal interactions between regions known as functional connectivity, resulting in correlated fluctuations of spatial activity patterns that are specific to particular stimulus representations. These results support a representational framework for understanding spontaneous activity.

Introduction

Spontaneous neural activity is observed throughout the brain, yet its function remains mysterious. An important clue, however, comes from work that has uncovered striking similarities between spontaneous activity and the activity evoked by a task (1-12). For example, the temporal correlation of spontaneous activity between brain regions (functional connectivity, FC) closely resembles the spatial topography of task-evoked activity (13-18), links distributed brain regions into functional networks, and can be used to predict task activation (5, 6).

The remarkable spatiotemporal regularities of spontaneous activity and widespread findings that abnormalities in inter-regional correlations of spontaneous activity in humans are associated with neurological and psychiatric disorders (e.g. (19, 20)) have motivated a search for functional explanations. One hypothesis is that spontaneous activity has a role in the synaptic homeostasis of structural connections (21). Another idea is that fluctuations of spontaneous activity between regions constitute a spatiotemporal prior that facilitates the recruitment of task circuitries during behavior (22, 23). Both hypotheses have been mainly concerned with explaining the FC between regions.

A novel hypothesis is that spontaneous activity and inter-regional FC has a role in representing behaviorally relevant information. Genetically determined circuitries generate spontaneous activity that is shaped in the course of development by experience through Hebbian statistical learning (1). Conversely, spatial and temporal patterns of spontaneous activity constrain task-evoked patterns As a result of this cyclic process, both spontaneous and task-evoked activity code similar representations of internal and external states (24, 25). The same process determines the spontaneous interactions between regions, which reflect connectivity patterns that are coded as synaptic efficacies in cortical networks (25, 26). **Figure 1** illustrates schematically the representation hypothesis of spontaneous brain activity.

The '*representation* hypothesis' is supported by human and animal work. In animals, imaging of neural activity at a scale that extends across many cortical columns

has shown that the macro-scale spatial pattern of spontaneous neural activity within a sensory area in an anesthetized animal mirrors the pattern of activity evoked by stimulation of a specific visual feature (7, 8). In humans, fMRI studies of early visual cortex have shown that FC between individual voxels respects the stimulus-evoked selectivity of voxels for polar angle, eccentricity, and low-level stimulus features (9-12). Recent work has also shown that voxel-wise resting FC in visual cortex is better approximated by the FC evoked by movies than by more artificial stimuli such as rotating checkerboards or static pictures of stimuli (26, 27).

However, these studies have not considered the spatial patterns of activity within a region, or the temporal correlation of these spatial patterns across cortical regions. Yet to-date, the best evidence in humans that evoked activity in cortex codes for behaviorally relevant information such as stimulus categories, retrieved memories, or cognitive processes (e.g. attention) has come from multivariate spatial pattern analyses of fMRI signals in visual and associative cortex (28-35). Critical tests of the 'representation' hypothesis of spontaneous interactions therefore include predictions about the spatial pattern of activity within a region and the interaction of that spatial pattern with the patterns from other regions.

First, if spontaneous activity carries information about stimulus categories such as faces, bodies, or scenes, then spontaneous spatial activity patterns, i.e. patterns of activity observed at rest in the absence of any stimulation, in functionally specialized occipital regions such as the extrastriate body area (EBA, (36)) should be more related to the spatial patterns of activity evoked by the preferred stimulus category (e.g. bodies) than by other categories. This predicted relationship between spontaneous and task-evoked spatial activity patterns putatively reflects the entrainment of task patterns into spontaneous activity in the course of development and experience. Second, if the FC between regions at least partly reflects correlated fluctuations of the spontaneous representational content of those regions, then the resting inter-regional temporal correlation of patterns of spatial activity should systematically depend not only on whether the regions prefer the same visual stimulus category, but also on whether the correlated spatial activity patterns code the preferred visual category.

Results

The first goal of the experiment was to compare multi-vertex activity patterns measured in the resting state with fMRI to stimulus-evoked activity patterns reflecting a variety of stimulus categories, including those that are more or less ecological (e.g. photographs of faces, tools, and scenes vs. phase-scrambled images of those stimuli). This comparison was conducted in regions of higher-order visual cortex that activated more strongly to stimuli from a particular category (e.g. bodies) relative to other categories (e.g. chairs and tools), and was also conducted in early visual cortical regions (37). To measure spontaneous activity, we ran a set of resting-state scans in which human observers fixated a central point on a blank screen. This activity was measured first to prevent possible learning effects from the other conditions (**Fig. S1**).

Localization of regions with visual category preferences

To identify category-specific visual regions, we ran a set of localizer scans in which multiple stimuli belonging to one of five stimulus categories (faces, bodies, scenes, man-made objects (chairs and tools), and phase-scrambled versions of these stimuli) were presented in a blocked design (**Figs. S1** and **S2**). We used standard contrasts (as in (38)) to identify category-preferring regions. For instance, activity evoked by body stimuli was subtracted from activity evoked by man-made objects (tools and chairs) to localize body-preferring regions such as EBA (see **Figs. S3A and S3B, and Table S1** for all category-preferring regions). Separate contrasts identified regions more active for whole-objects (face+scene+bodies+(tools+chairs)) than for low-level visual features (phase-scrambled objects). Phase-scrambled objects activated more strongly in regions of early visual cortex (V1-V3 based on the maps of (37)), while whole objects activated more strongly lateral and ventral occipital cortex, including some category-preferring regions (**Figs. S4B** and **S4B, Table S1**).

Representational similarity analysis of task-evoked patterns

During task scans (**Fig. S1**), we randomly presented individual stimuli belonging to each category to extract the stimulus-evoked spatial pattern in a particular ROI for each stimulus and corresponding category. Two general linear models (GLMs) were conducted to estimate the stimulus-evoked patterns. One model used stimulus-specific β weights to estimate the spatial activity pattern evoked by each individual stimulus, and the other used category-specific β weights to estimate the stimulus-evoked spatial activity pattern associated with each category.

To show that our stimuli and procedure generated spatial patterns that were consistent with the literature, we conducted a representational similarity analysis (RSA). Importantly, the representational similarity analysis was conducted using the task scans, which were completely independent of the localizer scans used to determine category-preferring ROIs. **Figure 2A** shows the spatial similarity of multi-vertex patterns evoked by individual stimuli within several classical category-preferring ROIs. In left EBA the highest representational similarity was found between human bodies, and the next highest between pictures of mammals, which included their bodies. In the right fusiform face area (FFA (39)), faces and other animate stimuli (bodies, mammals) generated more similar patterns than stimuli from the inanimate categories chair, tool, and scene, with the most consistent representational similarity found between face exemplars (32, 40). In the scene-preferring region right parahippocampal place area (PPA (41)), the activity patterns evoked by different scenes were well correlated, with low correlations between and within all other categories.

We conducted a second representational similarity analysis using the pattern evoked by a stimulus category, as estimated by the category regressor in a GLM, rather than using the patterns evoked by individual stimuli. Instead of conducting this analysis separately within each localizer-defined ROI, we grouped each set of category-preferential ROIs for an individual into a single joint-ROI. For instance, the body joint-ROI included left and right EBA, left and right fusiform body area (FBA), and so forth, and the scene joint-ROI included constituent regions such as PPA, the transverse occipital sulcus (TOS), and retrosplenial cortex (RSC) (see **Table S1** for a complete listing). The results of this analysis (**Fig. 2B)** were consistent with the literature. Both in

body and face joint-ROIs, the highest representational similarity was found between animate categories (face, bodies, mammals) as compared to other categories (32, 40). For instance, the similarity of body- and face-evoked spatial activity patterns in the face joint-ROI was ρ=0.73, while the similarity of face- and scene-evoked patterns in the face joint-ROI was ρ=0.41. **Table S2** indicates the representational similarity between the task-evoked spatial patterns corresponding to the face, body, and scene categories within each joint-ROI.

Task-rest pattern similarity analysis in category-preferring regions

We next tested the first prediction of the representation hypothesis, namely that spatial patterns of spontaneous activity at rest in category-preferring regions should be more related to the task-evoked spatial pattern for preferred than non-preferred categories. For each category, the category-evoked spatial pattern was spatially correlated on each resting frame with the spatial activity pattern in a joint-ROI to determine a resting timeseries of correlation coefficients and a corresponding frequency distribution of coefficient values. The upper 90% value (U90 value) of the distribution was used as a summary measure of the relationship between the stimulus-evoked and resting spatial activity patterns.

Figure 3A illustrates this procedure in a single subject using a region that prefers scenes (PPA) and the corresponding category-evoked spatial pattern for scenes. The scene-evoked pattern in PPA is a multi-vertex set of normalized activation values (**Fig. 3A**, left). This evoked pattern is spatially correlated (ρ) with the spontaneous patterns of activity on each frame of the resting state scans. A timeseries of ρ-values is generated (**Fig. 3A,** middle), as well as a corresponding frequency distribution of ρ-values (**Fig. 3A**, the histogram in blue). A U90 value for the joint-ROI and category is then determined from the distribution. The insets in the middle panel show resting frames in which the spontaneous activity (real data) was not correlated (ρ=0.003, outlined in green), positively correlated (ρ=0.81, outlined in magenta), or negatively correlated (ρ=-0.74, outlined in cyan) with the scene-evoked spatial activity pattern. The same procedure

was used to generate a U90 value for each of eight categories (faces, bodies, mammals, chairs, tools, scenes, grid-scrambled, phase-scrambled) in each joint-ROI (body, face, scene).

Figure 3B shows the distributions of correlation coefficients across all subjects for each preferred category within its corresponding joint ROI (green hue). For example, the leftmost graph shows the distribution using the body-evoked activity pattern within the body joint-ROI. A second distribution, generated using the spatial activity pattern evoked by phase-scrambled objects within the same body joint-ROI, has been superimposed (black hue). Theoretically, the two distributions might differ in the mean, variance, skewness, or some other parameter. For each joint-ROI, the distribution of correlation coefficients for both the preferred stimulus category and the phase-scrambled category were symmetric and centered on zero. However, the spread of the distribution was higher for the preferred category-evoked spatial activity pattern, meaning that larger correlation coefficients, both positive and negative, were observed for the preferred than phase-scrambled category (red arrows in **Fig. 3B**). Therefore, a larger U90 value indicates the presence of larger positive matches and negative matches of the resting spatial pattern to the category-evoked spatial pattern. Similar findings were obtained for face (middle panel) and scene activity patterns (rightmost panel) in the corresponding joint-ROIs, as compared to phase-scrambled patterns.

The categorical specificity of spontaneous activity patterns in each joint-ROI was tested by comparing U90 values for different categories. **Figure 3C** shows mean U90 values for a joint-ROI's preferred category, defined from its localizer contrast (green symbol; e.g. body in the Body joint-ROI), "non-preferred" categories (red symbols), grid-scrambled category (blue symbol) and phase-scrambled category (black symbol) averaged across subjects. We first conducted an overall repeated measures analysis of variance ANOVA on U90 values with joint-ROI (body, face, and scene) and Category (8 levels) as factors. The main effects of joint-ROI ($F(2, 30)=50.4$, $p<.0001$) and Category ($F(7, 105)=3.46$, $p=.002$), and the interaction of joint-ROI by Category ($F(14, 210)=4.37$, $p<.0001$) were all significant. The interaction indicated that the variation of U90 values across categories depended on the joint-ROI. Separate repeated measures ANOVAs

for each joint-ROI with Category (8 levels) as a factor indicated a highly significant main effect of Category in each joint-ROI (Body: $F(7,105)=7.25$, $p<.0001$; Face: $F(7,105)=3.25$, $p=.004$; Scene: $F(7,105)=3.93$, $p=.0008$). Therefore, for each joint-ROI, the spread of stimulus-evoked-to-rest spatial similarity values significantly depended on the category of the stimulus-evoked spatial pattern.

To compare the U90 value for the joint-ROI's preferred category vs. each other category, we conducted paired t-tests with a Holm-Bonferroni correction for multiple comparisons. Significant, multiple-corrected comparisons are indicated in the figure by plus signs. In the Body joint-ROI, the U90 value for bodies was significantly larger than for chairs, scenes, and phase-scrambled stimuli. In the Face joint-ROI, the U90 value for faces was significantly larger than for scenes. Conversely, in the Scene joint-ROI, the U90 value for scenes was significantly larger than all other categories.

Therefore, the 'animate' Body and Face joint-ROIs and the 'inanimate' Scene joint-ROI showed a significant double dissociation involving the corresponding categories, with U90 values in the Face and Body joint-ROIs significantly greater for the face and body categories, respectively, than for scene categories, and the U90 value in the inanimate Scene joint-ROI significantly greater for scenes than for either faces or bodies. The U90 values for face and body categories within each joint ROI were similar, reflecting the fact that both are animate categories and have greater cross-category representational similarity with each other than with scenes (**Table S2**). These results provide some support for the first prediction of the representation hypothesis, namely that spontaneous activity patterns in category-preferring regions are more related to the patterns for some categories than for others. However, 'more related' means a greater spread of extreme similarity values, both positive and negative, rather than a shift in the mean to more positive similarity values.

Task-rest pattern similarity analysis in regions preferring whole vs. phase-scrambled objects

The above results showed that the spatial pattern of spontaneous activity in regions of high-level visual cortex that respond preferentially to ecological visual categories was more related to the spatial pattern evoked by one category than another (e.g. bodies vs. scenes). We next asked whether a similar result would be found in regions that show stimulus preferences for low-level features as compared to more ecological categories such as face or body. This result would support a general conclusion that the stimulus preferences of a region largely drive the spatial pattern of spontaneous activity. We used the localizer scans to identify ROIs in which stimulus-evoked responses were stronger or weaker for phase-scrambled objects than for the union of the whole-object categories (face, body, mammal, chair, tool, and scene). The resulting 'Phase-scrambled objects' joint-ROI was located in medial posterior visual regions in early visual cortex (V1-V3 according to the Wang template (37) while the 'Whole-objects' joint-ROI was located in lateral and ventral visual cortex, in association visual cortex (**Fig. S4B**).

A representational similarity analysis in the Phase-scrambled objects joint-ROI showed high similarity between phase-scrambled, grid-scrambled, and scene stimuli, while the Whole-objects joint-ROI showed low similarity between those categories (**Fig 4A)**. **Figure 4B** shows the results of a task-rest pattern similarity analysis based on U90 values in each joint-ROI, which support the general conclusion that task-rest pattern similarities are not necessarily stronger for more ecological stimuli. Instead, task-rest correspondences reflect stimulus preferences, which are different in high- and low-level visual cortical regions (**Fig. 4B**). An ANOVA with ROI-type (Whole-objects, Phase-scrambled objects) and Stimulus-type (whole-objects, grid-scrambled, phase-scrambled objects) as factors indicated that the critical interaction of ROI-type by Category ($F(2,30)=14.2$, $p<.0001$) was significant. A significant interaction was also found for a 2 x 2 sub-ANOVA restricted to the categories whole objects and phase-scrambled objects ($F(1,15)=23.5$, $p<.0001$).

Within each joint-ROI, we compared the U90 value for the preferred category vs. the two "non-preferred" categories using paired t-tests with a Holm-Bonferroni correction for the four comparisons over the two joint-ROIs. In the Phase-scrambled joint-ROI, U90

values were significantly higher for both scrambled stimulus categories than for the whole-objects category, and in the Whole-objects joint-ROI, the U90 value for the whole-object category was significantly greater than for the phase-scrambled object category. The grid-scrambled pattern, which contains both high-level and low-level features (e.g. a high density of contour terminators), showed U90 values both in early visual and higher-order visual cortex that were not distinguishable from the regions' preferred stimulus category.

These results demonstrate a second double dissociation relating the dependence of U90 values on both the category-evoked spatial activity pattern and the joint-ROI in which similarity of the evoked pattern to spontaneous patterns was evaluated. They are consistent with the interpretation that spontaneous activity patterns in visual cortex are strongly affected by the stimulus preferences of the region, irrespective of whether those preferences favor more or less ecological categories.

U90 values correlate with activation strength

Another prediction of the representation hypothesis is that task-evoked patterns will entrain spontaneous activity patterns in the course of development and experience. Therefore, one expects a positive relationship between the magnitude of the stimulus specific response and the strength of the relationship between category specific patterns and spontaneous activity patterns (U90 values). **Figure 5** shows the correlation across category between task activation magnitude and U90 values of task-rest pattern similarity for each joint-ROI. **Figure S5** shows the mean activation strengths for different categories during the Task scans. The magnitude of the stimulus-evoked response in a joint-ROI was generally strongest for the preferred category (**Fig. S5**). Since joint-ROIs were defined from localizer scans that were independent of the task scans, this result indicates the stability of the ROI assignments.

There was a positive and significant correlation between task activation values and U90 values in each joint-ROI (**Fig. 5**). The greater the activation strength of a category, the greater the U90 value, a relationship that held for joint-ROIs irrespective of

their stimulus preferences. This relationship also was significant when the correlation coefficient between activation strength and U90 value was computed separately for each participant, and a group 1-sample t-test was conducted. Correlation coefficients were highly significant in all joint-ROIs (Body, p=.0004, Face, p=.0047, Scene, p=.0014; Whole-object, p=.0002; Phase-scrambled, p<.0001).

Pattern-based functional connectivity at rest

FC analyses typically evaluate the correlation between the timeseries of activity for single voxels or between voxel-averaged timeseries. However, recent task-based studies have also measured the inter-regional temporal correlation of spatial activity patterns (42-44). We used a similar approach to determine whether resting fluctuations of the multi-vertex spatial pattern for a category in each constituent ROI of a joint-ROI fluctuated synchronously or independently across the constituent ROIs. For instance, we determined whether spatial patterns for scenes in regions such as PPA, TOS, and RSC, which were previously combined to form the Scene joint-ROI (**Table S1**), fluctuated synchronously at rest. Synchronous fluctuations would indicate temporal variations of an inter-regional brain state specific for a particular category. We computed the temporal correlation of spatial activity patterns across the constituent ROIs of the Body and Scene joint-ROIs. The Face joint-ROI was not included in these analyses, since that ROI only included two regions and one of them overlapped with a constituent body ROI. In contrast, the Scene and Body joint-ROIs contained multiple ROIs that were all disjoint.

For each body- and scene-preferring constituent ROI, separate body and scene "pattern-to-rest" correlation timeseries were computed based, respectively, on the spatial correlation of the activity pattern on each resting frame with the body-evoked activity pattern and the scene-evoked activity pattern. This procedure is illustrated in **Figure S6** using data from segments of resting scans in one subject. Pattern-to-rest-correlation timeseries are shown for two scene-preferring regions (right PPA and right TOS) and two body-preferring regions (right EBA and right FBA). Each pattern-to-rest-

correlation timeseries shows the similarity values over time of resting spatial patterns to a category-evoked pattern in a single ROI. For example, the left lower dark blue timeseries in **Figure S6** shows the similarity of the resting pattern on each frame to the body-evoked pattern in the body constituent region EBA. The left lower graph shows that the pattern-to-rest correlation timeseries for body-preferring regions (right EBA and right FBA) that were computed using body-evoked activity patterns are positively correlated. In contrast, the left upper graph shows that the pattern-to-rest correlation timeseries for scene-preferring regions (right PPA and right TOS) that were computed using body-evoked activity patterns are uncorrelated. Conversely, when pattern-to-rest correlation timeseries were computed using the scene-evoked activity pattern, the opposite results are found. Now the pattern-to-rest correlation timeseries for scene-preferring regions show positively correlated fluctuations (right upper graph), while the timeseries for body-preferring regions are weakly correlated (right lower graph).

The data from all resting scans of all subjects were analyzed and the results are summarized in **Figures 6A** and **6B**. **Figure 6A** shows three resting 'pattern-based FC' matrices. A pattern-based 'body' FC matrix (leftmost matrix) was constructed by computing all pairwise inter-regional correlations between the pattern-to-rest correlation timeseries computed from body-evoked spatial patterns. Similarly, a pattern-based 'scene' FC matrix (middle matrix) was computed using the pattern-to-rest correlation timeseries computed using scene-evoked patterns. Qualitatively, body-preferring ROIs showed stronger positively correlated spontaneous fluctuations for body-evoked than scene-evoked patterns, and scene-preferring ROIs showed stronger positively correlated spontaneous fluctuations for scene-evoked than body-evoked patterns.

The graphs in **Figure 6B** summarize the pattern-based Body and Scene FC matrices by averaging the inter-regional correlations for body-preferring regions (the lower left block of each matrix **in Fig. 6A** outlined in blue) and scene-preferring regions (the upper right block of each matrix outlined in red). The left and middle graphs show respectively the results when pattern-to-rest timeseries were computed using body-evoked and scene-evoked spatial patterns. A 2-factor ANOVA on the mean pairwise pattern-based FC values with ROI-type (body, scene) and Category-evoked-pattern

(body, scene) as factors yielded a main effect of ROI-type ($F(1,15)=5.41$, $p=.035$), reflecting the larger FC values in body-preferring regions and, critically, a significant interaction of ROI-type by Category-evoked pattern ($F(1,15)=8.96$, $p=.009$). This effect was further supported by paired t-tests of specific contrasts. When pattern-to-rest timeseries were computed using a body-evoked spatial pattern, inter-regional correlations were significantly higher in body- than scene-preferring ROIs. Conversely, when pattern-to-rest timeseries were computed using a scene-evoked spatial pattern, inter-regional correlations were significantly higher in scene- than body-preferring ROIs.

Therefore, spontaneous fluctuations of spatial patterns of activity were more strongly correlated for the spatial patterns corresponding to the regions' more preferred stimulus. This result indicates that resting-state FC is modulated by the representational content of spontaneous activity. In addition, a paired t-test indicated that in the pattern-based Body FC matrix, the average FC was less in the scene-body block than in the body-body block of the matrix ($p<.001$; **Fig. 6A**, leftmost matrix, orange vs blue outlined blocks). Similarly, in the pattern-based Scene FC matrix, the average FC was less in the scene-body block than in the scene-scene block of the matrix ($p=.035$; **Fig. 6A**, middle matrix, gray vs red outlined blocks). Therefore, pattern-based FC was greater between regions preferring the same category than between regions preferring different categories.

A related question was whether putative body and scene representations fluctuated independently at rest. The rightmost matrix in **Figure 6A** shows a Preferred-Category pattern-based FC matrix in which the pattern-to-rest correlation timeseries in body-preferring regions were computed using body-evoked spatial patterns and the pattern-to-rest correlation timeseries in scene-preferring regions were computed using scene-evoked spatial patterns. Accordingly, the lower left and upper right blocks of the Preferred-Category matrix match, respectively, the lower left block of the Body pattern-based FC matrix and the upper right block of the Scene pattern-based FC matrix.

The 'scene-body' block of the Preferred Category matrix outlined in green is of primary interest. The correlation between scene and body regions was uniformly low under conditions in which the inter-regional correlation involved timeseries from scene

and body regions that respectively indicated the fluctuations of scene- and body-evoked spatial patterns (see rightmost graph, **Fig. 6B,** for average correlation values for scene-body blocks. Therefore, periods in which a body-evoked pattern was maximally present in body-preferring ROIs were largely independent of periods in which a scene-evoked pattern was maximally present in scene-preferring ROIs. Paired t-tests indicated that correlations in scene-body blocks from the Preferred-Category matrix were significantly lower than the correlations from scene-scene ($p=.009$) and body-body region blocks ($p < .0001$).

Finally, a standard FC matrix (**Fig. S7**, left panel) was constructed by computing vertex-averaged resting timeseries for each region, followed by pairwise correlation of the regional timeseries. As in previous work (45-48), vertex-averaged FC was category specific, with stronger FC between body-preferring regions and between scene-preferring regions, than between body- and scene-preferring regions. Pattern-based FC matrices were moderately-to-strongly correlated with the vertex-averaged FC matrix. The largest correlation was with the preferred-category matrix rather than the matrices generated using a single category (body-evoked pattern, $r=0.57$; scene-evoked pattern, $r=0.48$; preferred-category, $r=0.65$).

Category selectivity of U90 values in constituent ROIs

Supplementary Figure 8 shows the category selectivity of U90 values for the individual constituent ROIs within a joint-ROI. For each joint-ROI, we conducted a two-factor ANOVA with Category and Constituent-ROI as factors and U90 value as the dependent measure. A main effect of Category with no interaction between Category and Constituent-ROI was observed for both the body joint-ROI ($F(7,63)=2.45$, $p=.028$) and the scene joint-ROI ($F(7,63)=2.41$, $p=.03$), indicating a consistent profile of U90 values over categories across the constituent ROIs of each joint-ROI. Nevertheless, variability in the category profiles over the constituent ROIs is evident. Since there were many fewer vertices in each constituent ROI than in the associated joint-ROI, some variability in category selectivity over constituent ROIs is expected due to noise.

Additionally, we argue in the discussion that the category selectivity of U90 values for a joint-ROI is aided by the category selectivity of the pattern-based FC between its constituent ROIs.

Discussion

The goal of the experiment was to test representational theories of spontaneous activity by determining whether in regions of human visual cortex there is a link between spatial patterns of spontaneous activity, measured using resting-state fMRI, and the spatial patterns evoked by more or less ecological visual stimuli such as bodies or phase-scrambled bodies.

We obtained two main results. First, resting spatial activity patterns in regions of visual cortex were more closely related to the activity patterns evoked by the regions' more preferred stimulus categories. This relationship did not reflect a greater average similarity of resting patterns to the patterns for more preferred categories, but instead a greater spread of resting similarity values for more preferred categories, i.e. both larger positive and larger negative similarity values. This result was demonstrated statistically by two significant double dissociations. Body- and face-preferring regions showed larger U90 values, indexing the spread of similarity values, for faces and bodies than for scenes, while scene-preferring regions showed larger U90 values for scenes than for faces and bodies (**Fig. 3**). Regions preferring whole objects vs. phase-scrambled objects showed a similar significant double dissociation (**Fig. 4**). This result was further strengthened by the positive correlation between U90 values and stimulus specific activation magnitudes. The more strongly a stimulus activated a region, the higher the spread of spatial similarity values between the stimulus-evoked spatial pattern and the spontaneous activity pattern (**Fig.5**). The latter result is consistent with the notion that task-evoked patterns entrain spontaneous activity patterns in the course of development and experience, and therefore serve as a prior for task activation (**Fig. 1**)

The second main result was that spatial activity patterns coding for a category fluctuated more synchronously at rest between cortical regions preferring that category. For example, in the resting-state the spatial pattern evoked by bodies was more

positively correlated between body-preferring ROIs than between scene-preferring ROIs. The spatial pattern coding for scenes showed the opposite result (**Fig. 6**). Finally, the spatial patterns coding for scenes and bodies fluctuated largely independently within the preferred regions for those categories.

To our knowledge, this is the first demonstration that spatial patterns of spontaneous activity within regions of human cortex, and fluctuations of those patterns between regions, code for stimulus and category specific information. In this respect, our work is more related to seminal work conducted in cats (7) and monkeys (8, 49) than to previous work in humans, which has focused on measurements of voxelwise functional connectivity (26, 27, 45, 46, 50-52).

Spontaneous activity patterns for objects and features in visual cortex

Spontaneous activity patterns were not more similar on average to preferred stimulus evoked-patterns, but showed the greatest variation with respect to those patterns (**Fig. 3**). Interestingly, animal studies thus far have found the same result with some caveats. For instance, Kenet et al. (7) recorded voltage sensitive dye imaging in anesthetized cat visual cortex and found a significant spatial correlation (r=0.6) between spontaneous activity patterns and orientation selective stimulus evoked patterns. Positive and negative values of the correlation distribution were higher (as in our experiment), rather than the mean, as compared to a control distribution obtained by flipping the orientation selective map. This finding was replicated recently in anesthetized monkey visual cortex (8). In auditory monkey cortex, spontaneous spatial covariations of gamma activity recorded with cortical grids from auditory cortex resemble tonotopic maps derived with auditory stimuli. Also in this case, the correlation involves both positive and negative high correlation values as compared to a control distribution obtained by shuffling the tonotopic map (49).

In our experiment, the control distribution was not spatially shuffled because this control might not preserve the local structure of the vascular architecture that is the anatomical basis of the measured BOLD signal. Therefore, we instead compared task-

rest correlation distributions for two different stimuli (e.g. bodies vs. scenes). Although the animal and human experiments differed in many ways, in all experiments the reported match between spontaneous and task-evoked activity was not a shift in the mean, but rather a higher frequency of more extreme matches/mismatches of the spatial patterns (e.g. Compare our **Fig.6A** to (8) **Figs.1-2**).

An issue for future work is whether spontaneous activity patterns more resemble the average pattern evoked by a category or the patterns evoked by individual exemplars from the category. In rat hippocampus, spontaneous replay in anesthetized and awake animals of sequences of activity during navigation, a phenomenon qualitatively similar to what reported here (e.g. (53); see also (54)), seem to reflect individual experiences rather than averages. Hippocampal replay sequences have been mainly conceptualized as reflecting a mechanism for consolidating information in long-term memory.

We interpret our findings on spontaneous activity as constituting a prior for task processing, e.g. object recognition. An important rationale for postulating a representational function of resting activity is that limits on the information processing capacity of the brain may be mitigated by the incorporation of useful prior information. Appropriate priors will generally depend on context and therefore will change dynamically. The perceptual priors appropriate to walking alone through a forest vs. eating a family meal at the dinner table are quite different. These putative dynamic changes are thought to reflect generative models of the expected input via top-down pathways (55).

Resting scans are usually conducted under conditions in which subjects lie in a dark tube while fixating a cross in an otherwise blank display, which would not seem a fertile context for a perceptual prior. However, some aspects of an appropriate prior may not heavily depend on context. For example, recent work in monkey inferotemporal cortex indicates that individual faces can be coded by face cell assemblies whose firing rate is distributed along a small number of orthogonal dimensions (56). Therefore one possibility is that the spontaneous activity patterns found in the present work reflect fluctuations along canonical low dimensional configurations.

Synchronous fluctuations of representational content

The results for pattern-based FC show that the putative representational content of spontaneous activity, as indexed by multivertex spatial patterns, fluctuates more synchronously for the preferred category of ROIs that are linked by a common stimulus preference. Pattern-based FC between body-preferring regions was significantly larger when computed using a body- than scene-evoked pattern, while pattern-based FC between scene-preferring regions was significantly larger when computed using a scene-evoked than body-evoked activity pattern (**Fig. 6**).

An interesting, novel result from the pattern-based FC analysis was that at rest different putative representational states, as indexed by spatial activity patterns, fluctuated largely independently. Coherent fluctuations associated with body-evoked patterns in body-preferring ROIs occurred independently of fluctuations associated with scene-evoked patterns in scene-preferring ROIs, as shown by the very low correlations between scene and body regions in the analysis of the Preferred Category FC matrix (**Fig. 6**). Therefore, resting activity across category-selective regions of visual cortex cannot be described in terms of a single representational state. To our knowledge, the current results on resting pattern-based FC have not previously been reported and provide new constraints on theories of the function of FC.

Although the putative representational content of resting FC is often unspecified, an important exception comes from studies of early visual cortex, which have shown that resting FC respects the tuning of single voxels for polar angle, eccentricity, and low-level stimulus features (9-12). Most task-based studies of representation in higher-order visual and associative regions, however, have not involved measurements of voxelwise tuning functions but instead have identified task-evoked representations through measurements of regional spatial patterns. Therefore, pattern-based FC (42-44) could provide insights into the putative representational FC of spontaneous activity in high-level brain regions that are complementary to those provided by approaches based on the tuning properties of single voxels.

We suggest two additional ways in which pattern-based FC might inform studies of resting-state organization. First, pattern-based FC may help fractionate existing resting-state networks and identify the functional factors associated with that fractionation. For example, pattern-based resting FC between regions that prefer a particular category might depend on selectivity for features within the category, such as gender for face-preferring regions.

Second, pattern-based FC might uncover resting FC organizations that differ substantially from the normative whole-brain structure that has been described over the past decade (6, 15, 16, 57), although this structure does vary over individuals (58-60). In the current work, pattern-based FC was measured within category-preferring regions. Because regions that co-activate tend to show greater resting FC (61), and because previous studies have shown that regions preferring the same category show preferential FC (45, 46, 50-52), novel FC organizations were not expected. Accordingly, pattern-based FC matrices were moderately-to-strongly correlated with inter-regional vertex-averaged FC matrices. However, divergent FC organizations may be more likely in studies that use task-evoked activity patterns based on frequently occurring processes that combine different domains: for example, activity patterns based on integration of voice and face information during person-to-person interactions, visuomotor coordination during object manipulation, or biologically significant stimulus-reward or response-reward contingencies. Cross-domain pattern-based FC that cuts across standard networks might reflect synergies (62-64) or routines for implementing frequently occurring processes.

Pattern-based FC and correspondence of resting and evoked activity patterns

We suggest that the synchronous fluctuations of representational content evidenced by pattern-based FC (**Fig. 6**) is partly responsible for the correspondence of stimulus-evoked activity patterns and spontaneous activity patterns that was observed in joint-ROIs (**Figs. 3 and 4**). In the spatial analysis, the largest positive or negative similarity values for a resting frame in a joint-ROI will occur when the similarity values in

the constituent ROIs on that frame are simultaneously large and have the same sign. Otherwise, across constituent ROIs the similarity values will tend to cancel or average to a lower value. Therefore, a larger spread of similarity values in a joint-ROI is more likely to be observed if the similarity values in the constituent ROIs fluctuate in a temporally correlated or coherent fashion.

This mechanism explains the pattern of U90 values across the different joint-ROIs shown in **Figure 3C**. U90 values averaged across categories were highest in the face joint-ROI, intermediate for the body joint-ROI, and lowest for the scene joint-ROI, i.e. were inversely related to the number of constituent regions in each joint-ROI (2 for face, 5 for body, and 9 for scenes). The greater the number of constituent ROIs in a joint-ROI, the more the overall U90 value for the joint-ROI was decreased by sub-optimal coherence. As noted in the results section, in a two-factor ANOVA on U90 values with joint-ROI (body, face, scene) and Category (8 levels) as factors, the main effect of joint-ROI was significant. We additionally computed a single U90 value for each joint-ROI for each participant by averaging over categories. Paired t-tests on these averaged U90 values, with a Bonferroni-Holm correction for multiple comparisons (3 tests), indicated significantly larger U90 values within the Face than Body joint-ROIs ($p<.0001$), Face than Scene joint-ROIs ($p<.0001$), and Body than Scene joint-ROIs ($p=.01$).

In addition, since the spatial activity patterns for a non-preferred category do not fluctuate as coherently across constituent ROIs as those for a preferred category (**Fig. 6**), the resulting similarity values across frames in the joint-ROI for that non-preferred category will show less variation from zero, resulting in smaller U90 values. Therefore, within a joint-ROI, differences in U90 values between categories should be more reliable for joint-ROIs comprised of more constituent ROIs. This suggestion is also consistent with the results in **Figure 3C**, with the fewest significant differences found for the face joint-ROI and the most for the scene joint-ROI. Although other factors besides the number of constituent ROIs are clearly important in determining the category selectivity of U90 values in a joint-ROI, the larger point is that the temporal coherence of

preferred categories increases the incidence of extreme spatial matches and mismatches between evoked and spontaneous patterns in a category specific fashion.

Low- and high-level visual correspondences at rest

The spread of spatial similarity values between resting activity patterns and stimulus-evoked patterns was determined by how well a stimulus activated the region, irrespective of whether the stimulus was more or less ecological. In many higher-level visual ROIs, stimulus preferences favored a particular whole-stimulus category (e.g. bodies) over another whole-stimulus category (e.g. scenes) or over the phase-scrambled category. Conversely, in early visual cortex, preferences favored stimuli that weighted low-level features, resulting in larger U90 values for scrambled than whole-stimulus categories. The larger U90 values for scrambled stimuli in early visual cortex do not contradict an overall framework in which resting activity patterns reflect the statistical distribution of features in the environment. Rather, this result suggests that resting activity patterns in regions that primarily extract low-level visual features are relatively independent of the patterns associated with higher-order features/statistics that define categories of more ecological stimuli.

Grid-scrambled objects showed greater U90 values than whole-objects in the Phase-scrambled joint-ROI and equivalent U90 values to whole-objects in the Whole-object joint-ROI (**Fig. 4)**. This latter equivalence may have reflected the fact that the union of different category-preferential regions in the Whole-object joint-ROI eliminated or reduced the importance of features selective for a particular ecological category. Instead, the U90 value reflected features common to different ecological categories that were also present in grid-scrambled objects.

Therefore, the distribution of spatial matches between resting and evoked activity patterns can be driven by a variety of stimulus features that reflect local (e.g. contour-related features) or global (e.g. faces) stimulus characteristics depending on the tested regions.

Limitations

Stimuli were not controlled for low-level variables that might have differentially activated visual regions. As noted, grid-scrambled stimuli may have included contour terminators to a greater extent than many whole-object stimuli, increasing the activation of early visual cortex. However, this factor was not explicitly controlled or manipulated. Also, stimuli were presented in a non-naturalistic context. Wilf et al. (27) have shown that in early visual cortex, resting FC patterns are better accounted for by movies than by standard retinotopic stimuli, while Strappini et al. (26) have shown that in higher-level visual cortex, resting FC patterns are better accounted for by movies than by static pictures of stimuli similar to those used here. Therefore, the present results may have underestimated the spatial correspondences between resting and evoked activity patterns.

Materials and Methods

More detailed information concerning ***Participants***, ***Stimuli***, ***Scanning Procedure***, ***Imaging Parameters, fMRI pre-processing, and Definition of ROIs*** Is included in SI: Materials and Methods.

Participants. The study included 16 healthy young adult volunteers and was approved by the Institutional Review Board (IRB) of Washington University in St. Louis School of Medicine.

Stimuli. Images from seven 'whole-object' categories (human faces, human bodies, mammals, chairs, tools, scenes, and words), phase-scrambled images, and grid-scrambled images were presented on task scans. Word stimuli were included for exploratory analyses and results are not considered here. Phase- and grid-scrambled stimuli categories were constructed from the whole-object stimuli excluding words. Phase-scrambled stimuli preserved the spatial frequency amplitude spectrum of the whole-objects stimuli, and grid-scrambled stimuli included basic visual properties of the whole-objects images such as line segments and connectors. Localizer scans included

human faces, human bodies, objects (chairs and tools), scenes, words, false font character strings and phase-scrambled images. The categories for localizer and task scans differed slightly since the former was only used to define regions of interest (ROIs).

Scanning Procedure. Subjects participated in two sessions on separate days (**Fig. S1**). Session one included 3 resting scans, 2 localizer scans, and 8 task scans. Session two include 2 resting scans, 2 localizer scans, 8 task scans, and 2 post-task resting scans. During resting scans participants maintained fixation on a centrally presented cross. Localizer scans contained blocks consisting of 20 presentations (300 ms duration, 700 ms interstimulus interval, ISI) of exemplars from a single category. In the task scans, exemplars from all categories were presented in random order (duration=300 ms, jittered ISI of 3.7 to 8.7 sec). In both localizer and task scans, subjects performed a minimal cognitively engaging task by pressing a button if the presented image changed its size or position.

Imaging Parameters and fMRI pre-processing. Structural and fMRI images were obtained from a Siemens 3T Prisma MRI scanner. FMRI scans involved a gradient echo-planar sequence sensitive to BOLD contrast (TE = 26.6 ms, flip angle = 58°, 2.4 x 2.4 x 2.4 mm voxels, 48 contiguous slices, TR = 1.0 s, and multiband factor of 4). FMRI data were pre-processed as described in (65).

Defining ROIs from localizer activation contrasts. ROIs were defined for each subject from univariate vertex-wise statistical contrasts of the localizer conditions. One set of contrasts isolated ROIs that preferred a particular category (face, body, or scene) relative to the object category (chairs + tools). Vertices from all ROIs that preferred a particular category (e.g. bodies) were grouped into a single 'joint-ROI', excluding all vertices located in early visual areas (V1 to V3) (26), as estimated from (37). A second set of contrasts identified ROIs that preferred whole-objects relative to phase-scrambled objects (face + body + scene + object > phase-scrambled) or the reverse (phase-scrambled > face + body + scene + object)(see **Table S1** for ROI descriptive statistics, **Fig. S3B** and **Fig. S4B** for group-mean locations of ROIs from set1 and set2 contrasts)**.**

Task scans: multi-voxel activation patterns

For each joint-ROI from each subject, the multi-vertex activation pattern for each stimulus category (except the word category) in the task scans was estimated via a GLM that included a category regressor for all stimulus presentations involving the category. In addition, the GLM included a target regressor for trials in which a stimulus was perturbed in size or position, and baseline and linear trend regressors for each scan. The category and target regressors were each convolved with an assumed hemodynamic response function, yielding a stimulus-evoked BOLD multi-voxel pattern for each category (e.g. the pattern outlined by the red square in **Fig. S3A)** and for target trials.

For each joint-ROI from each subject, a subject's representational similarity matrix (RSM) was computed by spatially correlating the obtained categorical β weights (e.g. the average scene-evoked spatial pattern outlined by the red square in **Fig. S3A**) across categories. A group-averaged categorical RSM in a joint-ROI was computed by averaging all 16 subjects' RSM within a joint-ROI with Fisher-Z transformations and reverse Fisher-Z transformations (**Fig. 2B**). Additionally, for each constituent ROI in each joint-ROI from each subject, the multi-vertex activation pattern for each of the stimulus exemplars (24 exemplars for each of 6 object categories) in the task scans was estimated via an exemplar-specific GLM similar to the above GLM. A group-averaged RSM for the stimulus exemplars was then computed for each constituent ROI in each joint-ROI (**Fig. 2A**). Finally, in order to determine the task-evoked magnitude for each stimulus category, a β weight matrix was separately computed using spatially non-normalized BOLD timeseries from the task scans.

Determining similarity of resting multi-vertex patterns and stimulus-evoked patterns

For each participant's individual joint-ROI and the associated constituent ROIs, we determined the degree to which the multi-vertex pattern for a stimulus-evoked activity pattern for a category matched the multi-vertex pattern on each resting frame. The procedure is illustrated in **Figure 3A** for a single subject using real data. In the first

step, as described above, the multi-vertex pattern evoked by a category in a region was determined (e.g. the 'Scene' activity pattern outlined by the red square in **Fig. 3A, 'Task BOLD'**). Then, framewise intrinsic activity patterns were obtained from resting-state scans and the average stimulus-evoked pattern for a category was spatially correlated with the resting activity pattern on a frame (**Fig. 3A, 'Resting-state BOLD')**. A high positive correlation coefficient indicates that the multi-vertex resting activity pattern on a given frame was very similar to the pattern evoked by the category (e.g. the resting frame with a 'Scene'-like resting activity pattern outlined by the magenta square in **Fig. 3A**). A near zero correlation coefficient indicates that the multi-vertex resting activity pattern on a given frame was not similar to the pattern evoked by the category (e.g. the resting frame with a not-'Scene'-like resting activity pattern outlined by the green square in **Fig. 3A**). Finally, a high negative correlation coefficient indicates that the multi-vertex resting activity pattern on a given frame was very similar to the inverse of the pattern evoked by the category (e.g. the resting frame with a 'Scene'-inverted resting activity pattern outlined by the cyan square in **Fig. 3A**). This procedure was repeated across all resting frames, resulting in a 'pattern-to-rest' correlation timeseries (one correlation coefficient per resting frame) for a particular category in a particular ROI, as shown by the timeseries in **Figure 3A**.

From each timeseries, we constructed a corresponding distribution of correlation coefficients (**Fig. 3A,** distribution shown in blue). The upper 90% value of each distribution, hereafter termed the **U90-value**, was then determined. The U90 value computed for a category and ROI served as a measure of the relationship between resting activity patterns and the patterns evoked by a category mean. The U90-value was used as an alternative measure of variance since the U90-value refers to a spatial correlation coefficient value, which indicates the degree of pattern similarity between the task-evoked and resting state activity pattern. Using the U90-value instead of the variance of the distribution as a summary measure does not change the results of the analysis. For analyses that involved the Whole-Objects joint-ROI rather than joint-ROIs that preferred a particular object category such as faces, U90-values for the six whole-object categories (face, body, mammal, chair, tool, scene) were averaged together to form a whole-objects U90 value.

Statistical analysis of U90 values

U90 values were analyzed via repeated measures ANOVAs and post-hoc paired t-tests. For example, the statistical significance of an overall dependence of U90 values for a joint-ROI on the stimulus category was determined by conducting repeated-measures ANOVAs with Category-Type as factors. Paired t-tests were then conducted to test specific contrasts, with a Bonferroni-Holm correction for multiple comparisons.

Pattern-based resting functional connectivity

For the following pattern-based FC analysis, we used constituent ROIs from the joint-ROI that preferred a particular category (face, body, or scene) relative to the object category (chairs + tools). Since only two face constituent ROIs were found and one of those ROIs largely overlapped with a Body constituent ROI in ventral temporal cortex, Face constituent ROIs were not included in the FC analysis. Therefore, pattern-based FC was computed over 14 ROIs: 5 from the Body joint-ROI and 9 from the Scene joint-ROI.

Three pattern-based FC matrices were computed. **Figure S6** illustrates the procedure for computing the cells of an FC matrix using the pattern-to-rest correlation timeseries from two scene regions (TOS, PPA) and two body regions (EBA, FBA). **Figure 6A** shows the resulting matrices. First, for each participant, pattern-to-rest correlation timeseries for all 14 ROIs were generated using only the body-evoked pattern for each ROI (i.e. the spatial activity pattern evoked by bodies in that ROI during the Task scans). The correlation between the timeseries for all pairings of the 14 ROIs was then computed (i.e. body-ROI-to-body-ROI, scene-ROI-to-scene-ROI, and body-ROI-to-scene-ROI pairings). For example, the leftmost graphs in **Figure S6** show the correlation between TOS and PPA (top graph) and the correlation between EBA and FBA (bottom graph) using the pattern-to-rest correlation timeseries generated in each ROI from the body-evoked pattern. The correlation coefficients were then entered into the corresponding cells of the pattern-based FC matrix, “Using Body pattern-to-Rest correlation timeseries only”, shown in **Figure 6A.**

A similar procedure was used to generate a pattern-based FC matrix using only the scene-evoked pattern for each ROI. Pattern-to-rest correlation timeseries for all 14 ROIs were generated using only the scene-evoked pattern for each ROI (i.e. the pattern evoked by scenes in that ROI during the Task scans). Then the correlation between the pattern-to-rest correlation timeseries for all pairings of the 14 ROIs was computed. For example, the middle graphs in **Figure S6** show the correlation between TOS and PPA, and between EBA and FBA using the pattern-to-rest correlation timeseries generated in each ROI using the scene-evoked pattern. Finally, the correlation coefficients were entered into the corresponding cells of the pattern-based FC matrix, "Using Scene pattern-to-rest correlation timeseries only", shown in **Figure 6A.**

To generate the third pattern-based FC matrix in **Figure 6A** ("using Preferred Pattern-to-Rest correlation timeseries), the pattern-to-rest correlation timeseries in a Body ROI was generated using the body-evoked pattern for that ROI and the pattern-to-rest correlation timeseries in a Scene ROI was generated using the scene-evoked pattern for that ROI. Then the correlation between the correlation timeseries for all pairings of the 14 ROIs was computed and entered into the appropriate cells of the pattern-based FC matrix.

Finally, a vertex-averaged FC matrix was computed by first averaging the resting BOLD timeseries across all vertices of an ROI to generate a vertex-averaged timeseries, and then temporally correlating these averaged timeseries for all pairs of ROIs. Vertex-averaged FC matrices, which correspond to the standard regional FC matrices found in the literature, eliminate any information carried by the spatial pattern of BOLD activity within ROIs.

Pattern-based FC values were analyzed via repeated measures ANOVAs and paired t-tests. For example, we statistically evaluated whether the magnitude of pattern-based FC depended on both the category of the stimulus-evoked spatial activity pattern and the preferred category of the ROIs by conducting a repeated-measures ANOVA with the Category-evoked pattern (body, scene) and ROI-Type (body, scene) as factors. Paired t-tests were conducted to test differences between specific evoked patterns/ROI

combinations. For example, pattern-based FC values for body ROIs vs. scene ROIs were compared for correlation timeseries generated using body-evoked spatial patterns.

Acknowledgements

This work was supported by the National Institutes of Health RO1 MH096482 to MC.

References

1. Berkes P, Orban G, Lengyel M, & Fiser J (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331(6013):83-87.
2. Fiser J, Chiu C, & Weliky M (2004) Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature* 431(7008):573-578.
3. Tsodyks M, Kenet T, Grinvald A, & Arieli A (1999) Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science* 286(5446):1943-1946.
4. Fox MD, Snyder AZ, Vincent JL, & Raichle ME (2007) Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* 56(1):171-184.
5. Tavor I*, et al.* (2016) Task-free MRI predicts individual differences in brain activity during task performance. *Science* 352(6282):216-220.
6. Cole MW, Ito T, Bassett DS, & Schultz DH (2016) Activity flow over resting-state networks shapes cognitive task activations. *Nature neuroscience* 19(12):1718-1726.
7. Kenet T, Bibitchkov D, Tsodyks M, Grinvald A, & Arieli A (2003) Spontaneously emerging cortical representations of visual attributes. *Nature.* 425(6961):954-956.
8. Omer DB, Fekete T, Ulchin Y, Hildesheim R, & Grinvald A (2018) Dynamic Patterns of Spontaneous Ongoing Activity in the Visual Cortex of Anesthetized and Awake Monkeys are Different. *Cereb Cortex*.
9. Heinzle J, Kahnt T, & Haynes JD (2011) Topographically specific functional connectivity between visual field maps in the human brain. *Neuroimage* 56(3):1426-1436.
10. Raemaekers M*, et al.* (2014) Patterns of resting state connectivity in human primary visual cortical areas: A 7T fMRI study. *Neuroimage* 84:911-921.
11. Ryu J & Lee SH (2018) Stimulus-Tuned Structure of Correlated fMRI Activity in Human Visual Cortex. *Cereb Cortex* 28(2):693-712.
12. Arcaro MJ, Honey CJ, Mruczek RE, Kastner S, & Hasson U (2015) Widespread correlation patterns of fMRI signal across visual cortex reflect eccentricity organization. *eLife* 4.
13. Biswal B, Yetkin FZ, Haughton VM, & Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 34(4):537-541.
14. Greicius MD, Krasnow B, Reiss AL, & Menon V (2003) Functional connectivity in the resting brain: a nework analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A* 100(1):253-258.
15. Yeo BT*, et al.* (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106(3):1125-1165.

16. Power JD*, et al.* (2011) Functional network organization of the human brain. *Neuron* 72(4):665-678.
17. de Pasquale F*, et al.* (2010) Temporal dynamics of spontaneous MEG activity in brain networks. *Proc Natl Acad Sci U S A* 107(13):6040-6045.
18. He BJ, Snyder AZ, Zempel JM, Smyth MD, & Raichle ME (2008) Electrophysiological correlates of the brain's intrinsic large-scale functional architecture. *Proc Natl Acad Sci U S A* 105(41):16039-16044.
19. Grefkes C & Fink GR (2014) Connectivity-based approaches in stroke and recovery of function. *Lancet Neurol* 13(2):206-216.
20. Northoff G & Duncan NW (2016) How do abnormalities in the brain's spontaneous activity translate into symptoms in schizophrenia? From an overview of resting state activity findings to a proposed spatiotemporal psychopathology. *Progress in neurobiology* 145-146:26-45.
21. Deco G*, et al.* (2013) Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J Neurosci* 33(27):11239-11252.
22. Raichle ME (2011) The restless brain. *Brain Connect* 1(1):3-12.
23. Petersen SE & Sporns O (2015) Brain Networks and Cognitive Architectures. *Neuron* 88(1):207-219.
24. Fiser J, Berkes P, Orban G, & Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci* 14(3):119-130.
25. Harmelech T & Malach R (2013) Neurocognitive biases and the patterns of spontaneous correlations in the human cortex. *Trends Cogn Sci* 17(12):606-615.
26. Strappini F*, et al.* (2018) Resting-State Activity in High-Order Visual Areas as a Window into Natural Human Brain Activations. *Cereb Cortex*.
27. Wilf M*, et al.* (2017) Spontaneously Emerging Patterns in Human Visual Cortex Reflect Responses to Naturalistic Sensory Stimuli. *Cereb Cortex* 27(1):750-763.
28. Haxby JV*, et al.* (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293(5539):2425-2430.
29. Kamitani Y & Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nature neuroscience* 8(5):679-685.
30. Haynes JD & Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature neuroscience* 8(5):686-691.
31. Kriegeskorte N, Goebel R, & Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103(10):3863-3868.
32. Kriegeskorte N*, et al.* (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60(6):1126-1141.
33. Kuhl BA & Chun MM (2014) Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J Neurosci* 34(23):8051-8060.
34. Serences JT & Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55(2):301-312.
35. Greenberg AS, Esterman M, Wilson D, Serences JT, & Yantis S (2010) Control of spatial and feature-based attention in frontoparietal cortex. *J Neurosci* 30(43):14330-14339.
36. Downing PE, Jiang Y, Shuman M, & Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293(5539):2470-2473.
37. Wang L, Mruczek RE, Arcaro MJ, & Kastner S (2015) Probabilistic Maps of Visual Topography in Human Cortex. *Cereb Cortex* 25(10):3911-3931.
38. Bracci S & Op de Beeck H (2016) Dissociations and Associations between Shape and Category Representations in the Two Visual Pathways. *J Neurosci* 36(2):432-444.

39. Kanwisher N, McDermott J, & Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17(11):4302-4311.
40. Grill-Spector K & Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nature reviews. Neuroscience* 15(8):536-548.
41. Epstein R & Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392(6676):598-601.
42. Coutanche MN & Thompson-Schill SL (2013) Informational connectivity: identifying synchronized discriminability of multi-voxel patterns across the brain. *Front Hum Neurosci* 7:15.
43. Anzellotti S & Coutanche MN (2018) Beyond Functional Connectivity: Investigating Networks of Multivariate Representations. *Trends Cogn Sci* 22(3):258-269.
44. Chen RH, Ito T, Kulkarni KR, & Cole MW (2018) The Human Brain Traverses a Common Activation-Pattern State Space Across Task and Rest. *Brain Connect*.
45. Hutchison RM, Culham JC, Everling S, Flanagan JR, & Gallivan JP (2014) Distinct and distributed functional connectivity patterns across cortex reflect the domain-specific constraints of object, face, scene, body, and tool category-selective modules in the ventral visual pathway. *Neuroimage* 96:216-236.
46. Zhu Q, Zhang J, Luo YL, Dilks DD, & Liu J (Resting-State Neural Activity across Face-Selective Cortical Regions Is Behaviorally Relevant. *J Neurosci* 31(28):10323-10330.
47. Konkle T & Caramazza A (2017) The Large-Scale Organization of Object-Responsive Cortex Is Reflected in Resting-State Network Architecture. *Cereb Cortex* 27(10):4933-4945.
48. Wang X*, et al.* (2016) The Hierarchical Structure of the Face Network Revealed by Its Functional Connectivity Pattern. *J Neurosci* 36(3):890-900.
49. Fukushima M, Saunders RC, Leopold DA, Mishkin M, & Averbeck BB (2012) Spontaneous high-gamma band activity reflects functional organization of auditory cortex in the awake macaque. *Neuron* 74(5):899-910.
50. Zhang H, Tian J, Liu J, Li J, & Lee K (2009) Intrinsically organized network for face perception during the resting state. *Neuroscience letters* 454(1):1-5.
51. Turk-Browne NB, Norman-Haignere SV, & McCarthy G (2010) Face-Specific Resting Functional Connectivity between the Fusiform Gyrus and Posterior Superior Temporal Sulcus. *Front Hum Neurosci* 4:176.
52. Stevens WD, Kravitz DJ, Peng CS, Tessler MH, & Martin A (2017) Privileged Functional Connectivity between the Visual Word Form Area and the Language System. 37(21):5288-5297.
53. Karlsson MP & Frank LM (2009) Awake replay of remote experiences in the hippocampus. *Nature neuroscience* 12(7):913-918.
54. Buzsaki G & Draguhn A (2004) Neuronal Oscillations in Cortical Networks. *Science* 304(5679):1926-1929.
55. Mumford D (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological cybernetics* 66(3):241-251.
56. Chang L & Tsao DY (2017) The Code for Facial Identity in the Primate Brain. *Cell* 169(6):1013-1028.e1014.
57. Gordon EM*, et al.* (2016) Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb Cortex* 26(1):288-303.
58. Gordon EM*, et al.* (2017) Precision Functional Mapping of Individual Human Brains. *Neuron* 95(4):791-807.e797.
59. Gratton C*, et al.* (2018) Functional Brain Networks Are Dominated by Stable Group and Individual Factors, Not Cognitive or Daily Variation. *Neuron* 98(2):439-452.e435.
60. Laumann TO*, et al.* (2015) Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron* 87(3):657-670.

61. Smith SM*, et al.* (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A* 106(31):13040-13045.
62. Schieber MH & Santello M (2004) Hand function: peripheral and central constraints on performance. *Journal of applied physiology (Bethesda, Md. : 1985)* 96(6):2293-2300.
63. Santello M, Baud-Bovy G, & Jorntell H (2013) Neural bases of hand synergies. *Front Comput Neurosci* 7:23.
64. Leo A*, et al.* (2016) A synergy-based hand control is encoded in human motor cortical areas. *eLife* 5.
65. Siegel JS*, et al.* (2016) Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. *Proc Natl Acad Sci U S A*.

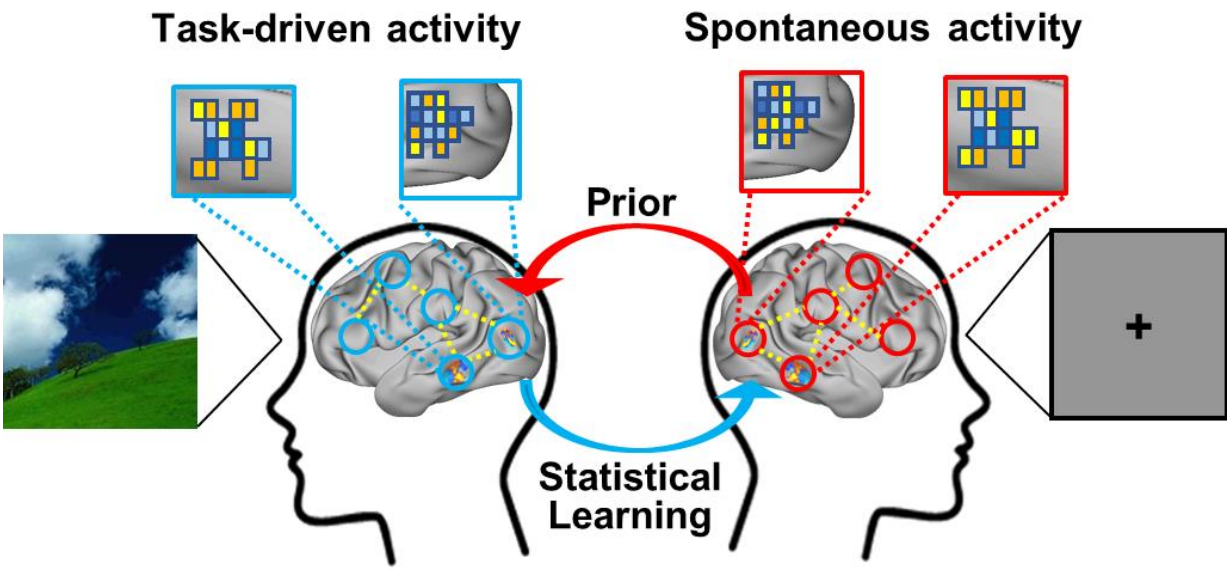


Figure 1. The putative cyclic interplay between brain activity evoked by real-world experiences and resting-state activity.

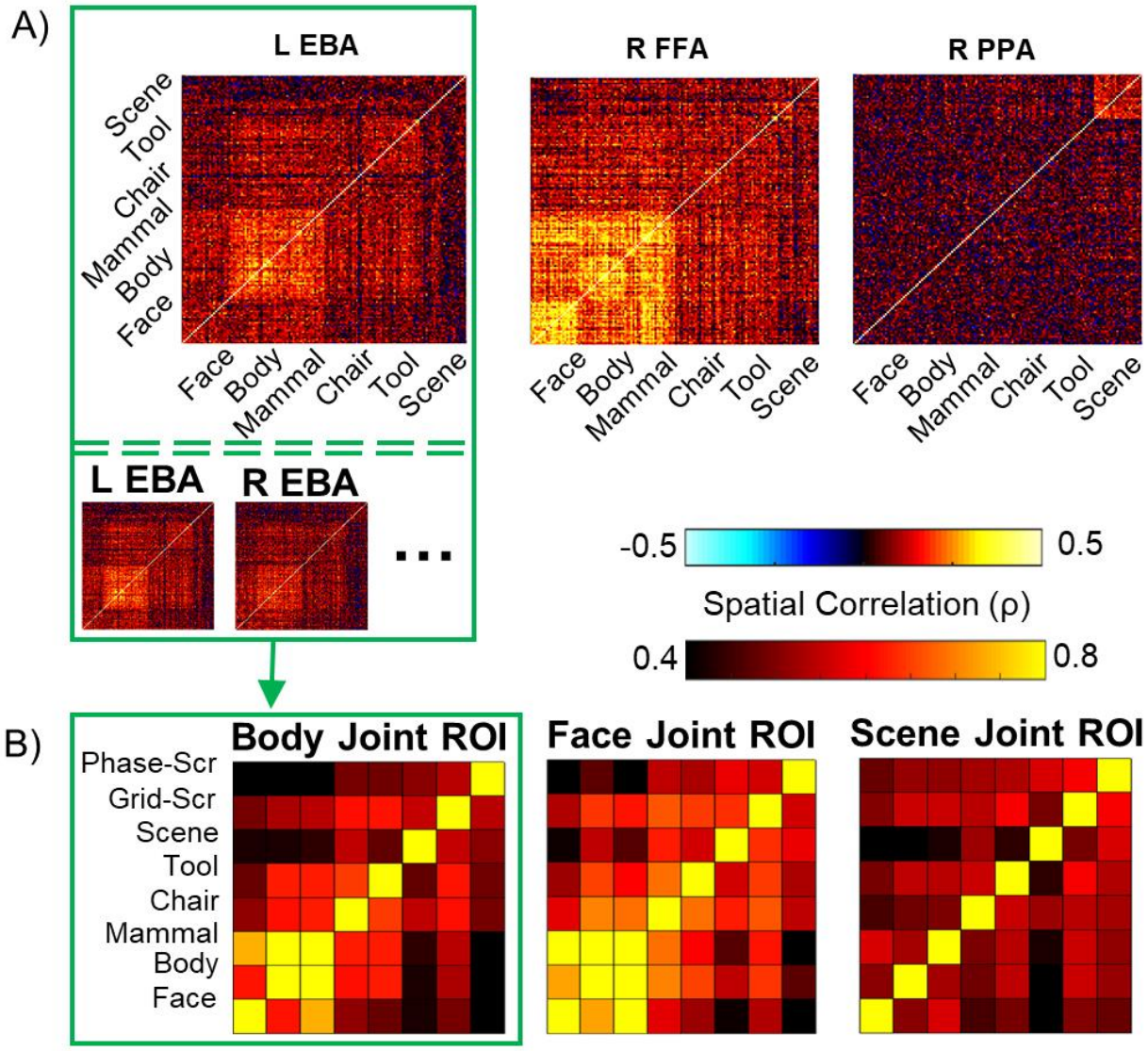


Figure 2. Representational similarity analysis (RSA) of individual stimuli and categories. **(A)** RSA for individual exemplars for each 'whole-object' category in 3 classical category-preferential areas. **(B)** RSA based on the average spatial pattern evoked for each category.

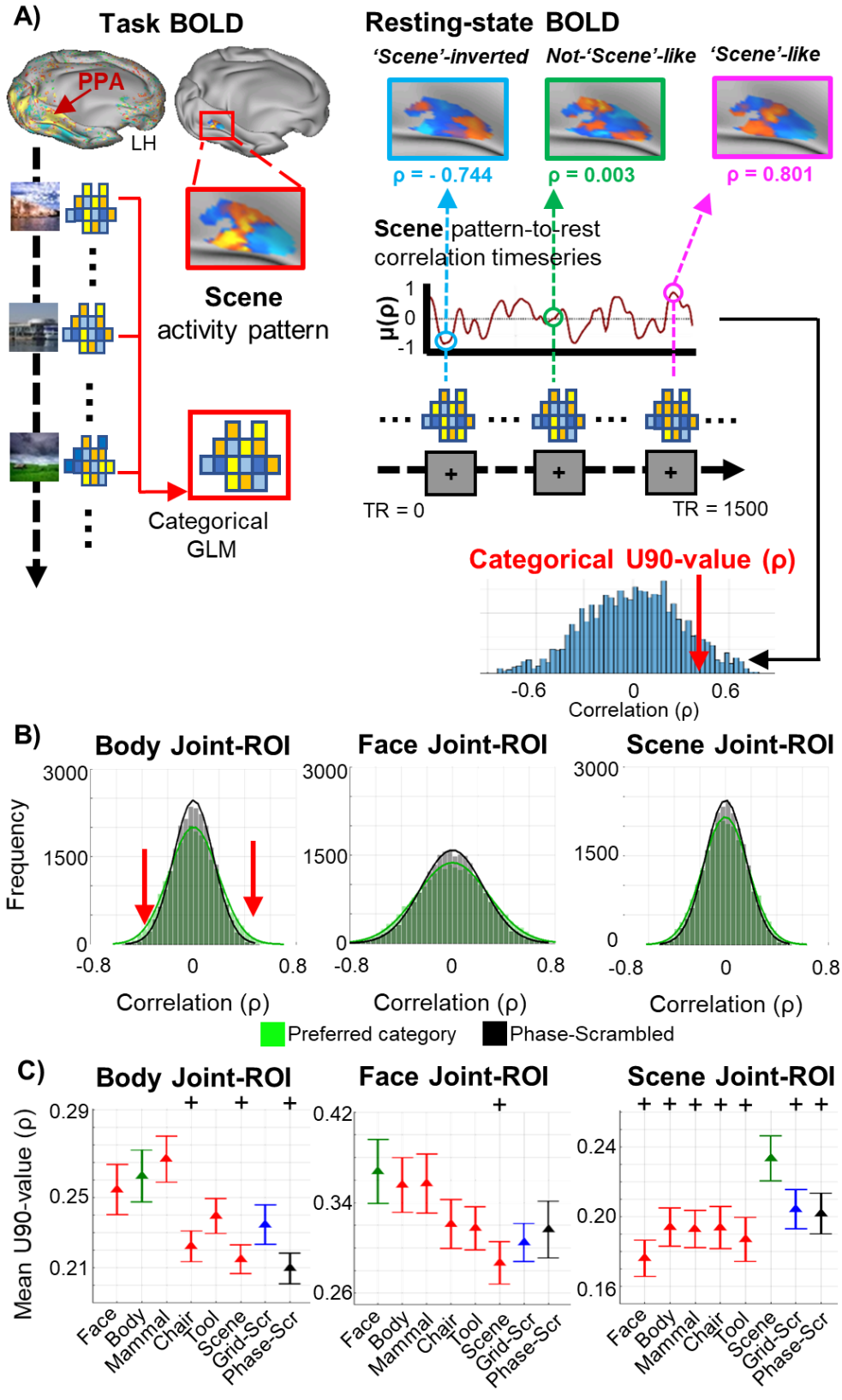


Figure 3. Task-rest pattern similarity analysis in visual regions preferring specific

categories. **(A)** Procedure for computing U90-values: determining the category-evoked spatial pattern on task scans, spatially correlating that pattern with the activity pattern on each resting frame, computing a U90 value from the resulting distribution of correlation coefficients. **(B)** Superimposed distributions of correlation coefficients for a joint-ROI's preferred stimulus category, which was defined by the corresponding localizer contrast (light green; e.g. Body in the Body-preferred joint-ROI) and the phase-scrambled category (gray). **(C)** Group-averaged U90 values for the joint-ROI's preferred category (green symbol), other whole-object categories (red symbols), grid-scrambled category (blue symbol), and phase-scrambled category (gray symbol). Black symbols indicate significant paired t-test between the joint-ROI's preferred category and indicated category (++ = Bonferroni-Holm corrected p-val ≤ 0.005). Error bars indicate ±SEM.

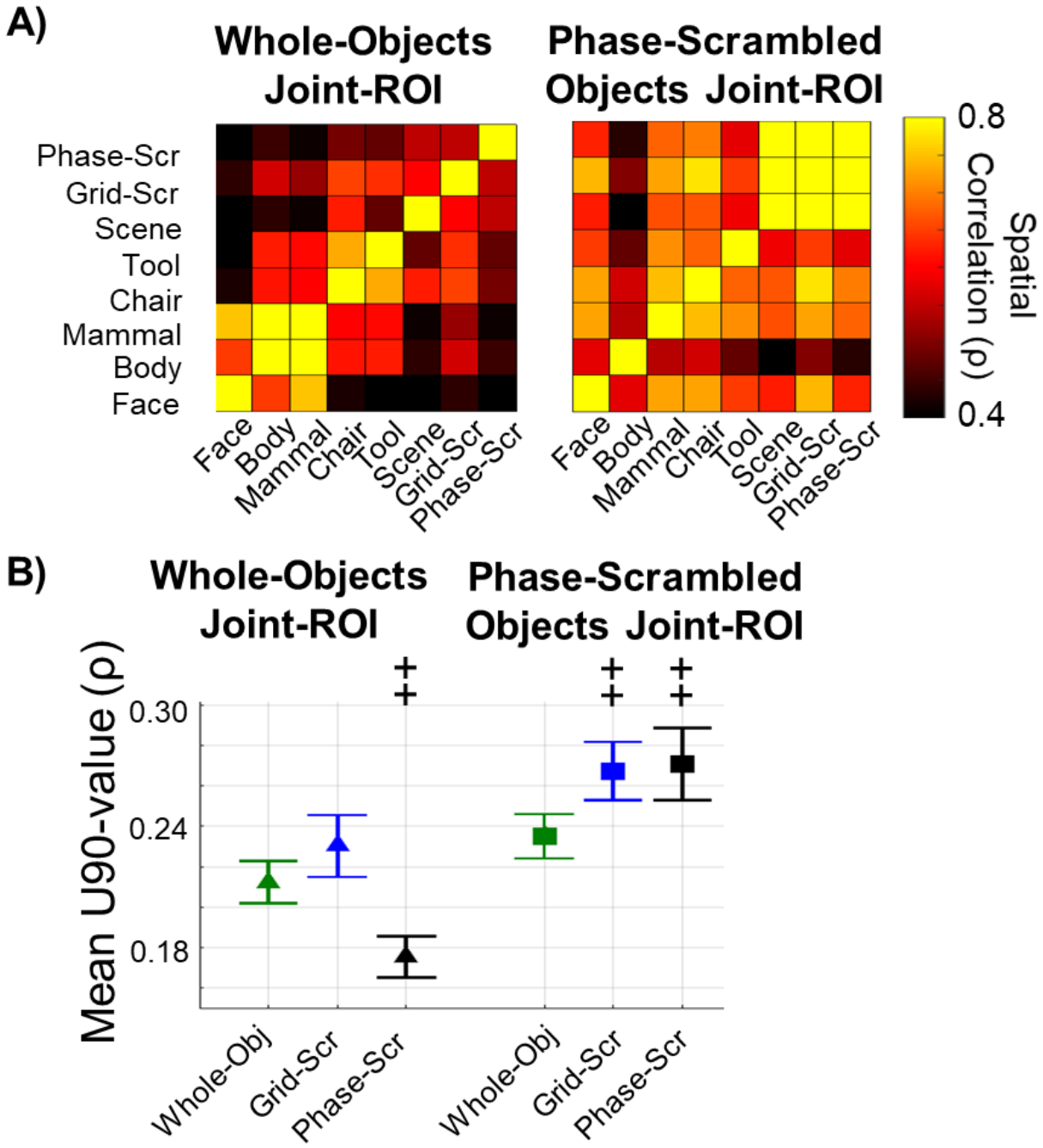


Figure 4. Task-rest pattern similarity analysis in regions preferring phase-scrambled objects or whole-objects. **(A)** RSAs for Whole-Objects and Phase-Scrambled Objects joint-ROIs based on the average spatial pattern for each category. **(B)** Group-averaged U90 values. Black symbols indicate significant paired t-test between the whole-object category and scrambled category (++ = Bonferroni-Holm corrected p-val ≤ 0.005). Error bars indicate ±SEM.

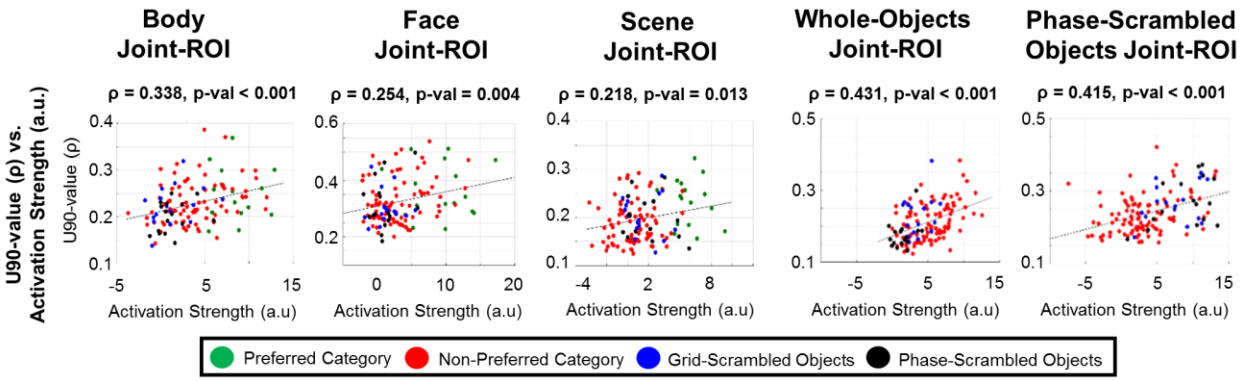


Figure 5. Correlation between activation strengths from task scans and U90 values across all categories and participants.

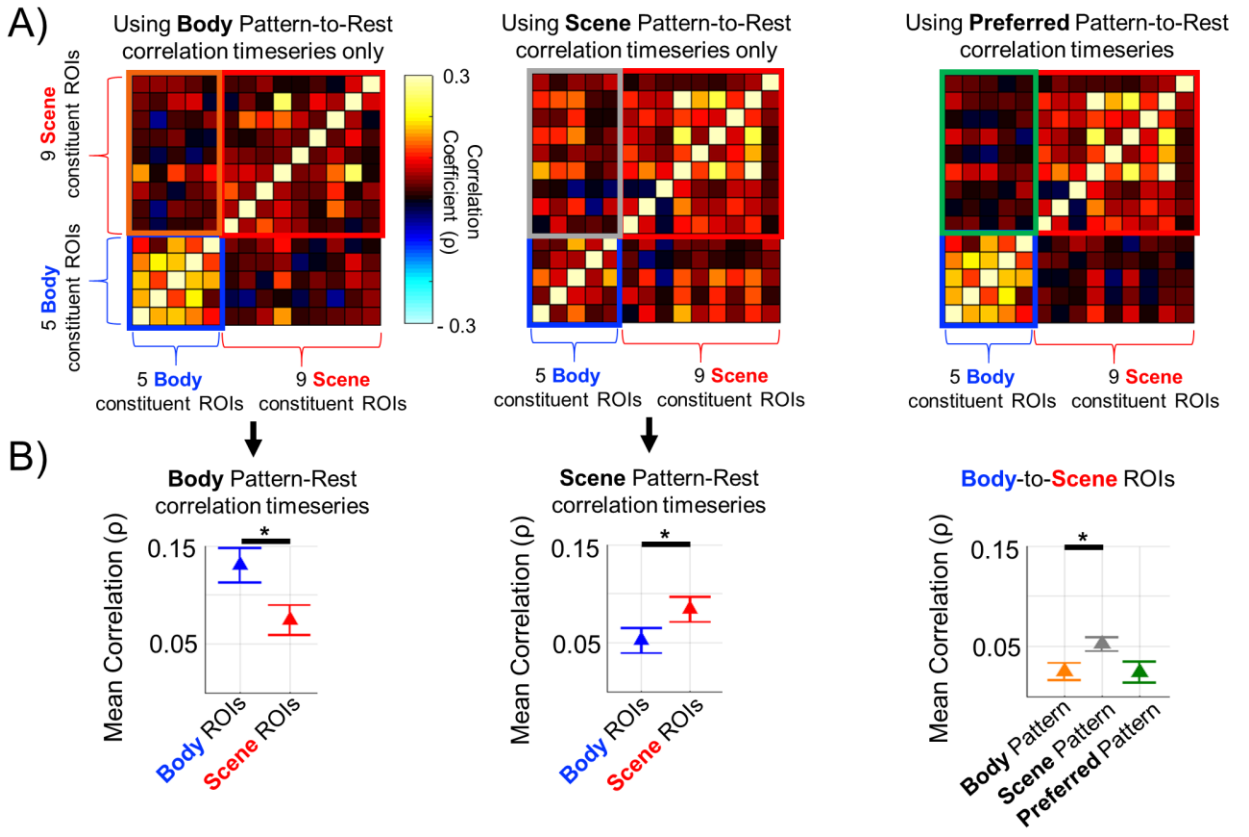


Figure 6. Pattern-based FC. **(A)** Pattern-based FC matrices computed using body-evoked, scene-evoked, or preferred category-evoked spatial patterns (see text for details). **(B) Left, Middle**: Group-averaged pattern-based FC between body-preferring regions and between scene-preferring regions computed using body-evoked patterns (left) or scene-evoked patterns (middle). **Right**: Group-averaged pattern-based FC between body and scene regions computed using body-evoked, scene-evoked, or preferred-category-evoked activity patterns. Black symbols indicate significant group paired t-test comparing correlation (ρ) values (* = p-val ≤ 0.05). Error bars indicate ±SEM.

Supplementary Information Text

Materials and Methods

Participants

The study included 16 healthy young adult volunteers (10 female; age 21 – 35 years-old) with no prior history of neurological or psychiatric disorders. All participants were right-handed native English speakers with normal or corrected-to-normal vision. All participants gave informed consent to take part in the experiment, and the study was approved by the Institutional Review Board (IRB) of Washington University in St. Louis School of Medicine.

Stimuli

Nine categories of color images subtending 8^o x 8^o of visual angle were included in event-related 'task' fMRI scans. Seven 'whole-object' categories consisted of images that are encountered in real-world environments: human faces, human bodies, mammals, chairs, tools, scenes, and words. Stimuli, excluding the word category, were obtained from Downing et al., 2006 (1). Faces, bodies and mammals served as animate categories, and chairs, tools and scenes as inanimate categories (2). Word stimuli were included for exploratory analyses and results for those stimuli will not be considered in this paper.

Two control stimulus categories were constructed from the above stimuli excluding the word stimuli. A low-level control consisted of phase-scrambled stimuli that preserved the spatial frequency amplitude spectrum of the whole-objects images. An intermediate-level control consisted of grid-scrambled stimuli that included basic visual properties of the whole-objects images such as line segments and connectors. For the low-level control condition, 2D phase-scrambled images of the exemplars from the 6 categories were generated by applying the same set of random phases to each 2-dimensional frequency component of the original image while keeping the magnitude constant (3). Exemplars from all six whole-objects categories except real word stimuli were 2D phase-scrambled, yielding a total of 144 2D Phase-scrambled stimuli. For the intermediate-level control condition, grid-scrambled images of exemplars from the six whole-objects categories were generated by sub-dividing each image into a 10 x 10 grid (each grid is

0.8^{o} x 0.8^{o}) and randomly rearranging the individual grid segments. **Supplementary Figure 2** shows exemplar stimuli from the 8 categories used.

Color images of exemplars from seven categories were included in blocked-design localizer scans: human faces, human bodies, objects (chairs and tools), scenes, words, false font character strings and phase-scrambled images. The categories for the localizer scans differed slightly from the categories for the task scans since the former was only used to define the regions of interest (ROIs). ROIs related to the false font and word stimuli will not be considered in this paper.

Stimuli were presented using the Psychophysics Toolbox package (4) in MATLAB (The MathWorks). Stimulus images were projected onto a screen and were viewed through a mirror mounted on the head coil. All stimuli were presented centrally on a gray background.

Scanning Procedure

The study consisted of two separate sessions, each conducted on a separate day (**Fig. S1**). In session one, subjects received 3 resting state runs, 2 localizer runs, and 8 task runs. In session 2, subjects received 2 resting state runs, 2 localizer runs, 8 task runs, and 2 post-task resting state runs. One subject had a total of 13 task runs over the two sessions instead of 16.

Resting state runs. Participants received a total of 7 resting state scans, each lasting 5 min (300 TRs). During a scan the participant was asked to maintain fixation on a cross that was displayed at the center of the screen during the entire run. Five resting scans (3 for first session and 2 for second session) were conducted before any localizer or task scans to collect stimulus-free intrinsic activities. For the second session only, two additional 5 min resting state scans were conducted after the task scans to investigate potential post stimuli-driven effects on intrinsic activity. The results from the post-task resting scans will not be discussed here.

Localizer runs. Each session included 2 localizer runs (4 in total), each lasting 5 min and 40s (340 TRs), and each localizer scan was presented in a blocked fMRI design. Each block of a localizer run contained 20 images of a single category, and those images were different from the images used in the task scans. A fully randomized sequence of eight blocks, consisting of the 7 stimulus categories and a fixation block, was repeated twice within each run. At the beginning and the end of each run, an additional fixation block was presented for 4s and 16s.

Within each category block, images were presented for 300ms with an inter-stimulus interval (ISI) of 700ms. A fixation cross was continuously present at the center of the screen during the ISI and during fixation blocks. During category blocks, participants performed a minimally cognitively engaging task by pressing a button if the initially presented image was changed in size or position during the 300ms presentation.

Task runs. Each session included 8 task runs (16 in total), each lasting 5 min and 15s (315 TRs). For each subject and for each run, stimulus presentation order and inter-stimulus intervals were fully randomized using Optseq2 (5). Each stimulus presentation lasted for 300ms and the interval between stimuli was jittered between 3.7s and 8.7s. A fixation cross was continuously present at the center of the screen during the ISI. In each whole-objects category, there were 24 separate exemplars (e.g. 24 different faces) and each exemplar was repeated 4 times. In each scrambled category, there were 96 exemplars, each presented once. Participants performed a minimally cognitively engaging task by pressing a button if the presented image changed its size or position during a 300ms presentation, the same task as that performed during the localizer scans.

Imaging parameters

Data were obtained from a Siemens 3T Prisma MRI scanner. Structural images for atlas transformation and lesion segmentation were acquired using T1-weighted magnetization prepared-rapid gradient echo (MP-RAGE) (1 x 1 x 1 mm voxels; echo time [TE] = 2.36 ms, repetition time [TR] = 1700 ms, TI=1000 ms, flip angle = 8°) and T2-weighted fast spin echo sequences (1 x 1 x 1 mm voxels; TE = 564 ms, TR = 3200 ms). FMRI scans were collected using a gradient echo-planar sequence sensitive to BOLD contrast (TE = 26.6 ms, flip angle = 58°, 2.4 x 2.4 x 2.4 mm voxels, 48 contiguous slices, TR = 1.0 s, and multiband factor of 4).

fMRI pre-processing

fMRI data underwent pre-processing as previously described (6). This included: (1) compensation for asynchronous slice acquisition using sinc interpolation; (2) elimination of odd/even slice intensity differences resulting from interleaved acquisition; (3) whole brain intensity normalization to achieve a mode value of 1000; (4) spatial realignment within and

across fMRI runs; and (5) resampling to 2.4 mm cubic voxels in atlas space, including realignment and atlas transformation in one resampling step. Cross-modal (e.g. T2-weighted to T1-weighted) image registration was accomplished by aligning image gradients.

Surface generation and processing of functional data followed procedures similar to Glasser et al (7). First, anatomical surfaces were generated for each subject's T1 MRI using FreeSurfer automated segmentation (8). This step included brain extraction, segmentation, generation of white matter and pial surface, inflation of the surfaces to a sphere, and surface shape-based spherical registration to the subjects' "native" surface to the fs_average surface. The left and right hemispheres were then resampled to 164,000 vertices and registered to each other (9).

Data were passed through several additional preprocessing steps: (i) removal by regression of the following sources of spurious variance: (a) six parameters obtained by rigid body correction of head motion, (b) the signal averaged over the whole brain (global signal regression), (c) signal from ventricles and CSF, and (d) signal from white matter; (ii) temporal filtering retaining frequencies in the 0.009–0.08-Hz band; and (iii) frame censoring (framewise displacement (FD) ≥ 0.5mm). The first four frames of each BOLD run were excluded.

To account for magnitude variability between different task and resting state runs, the BOLD timeseries for each vertex were Z-normalized across time within the task and the resting state runs. This Z-normalization was not applied to the localizer scans. Also, it was not applied to the Task scans for a separate analysis described below in which task-evoked activation magnitudes were determined (see below, ***Task scans: multi-voxel activation patterns).***

Defining ROIs from localizer activation contrasts

ROIs were defined from univariate vertex-wise statistical contrasts on the localizer activation magnitudes for different categories. For example, face-selective areas were defined from the vertices for the contrast of faces minus objects, where objects consisted of chairs and tools. First, for each participant a general linear model (GLM) was applied to their functional localizer scans. The GLM consisted of separate regressors for each stimulus category (e.g. faces) using an assumed hemodynamic response function from the Statistical Parametric Mapping (SPM12), a baseline term, and a linear trend term. Condition contrasts were formed to identify vertices showing a preference for each category, using a scheme similar to that of Bracci and Op de Beeck (2016) (10): body-preference (body > objects, i.e. chairs and tools),

face-preference (face > objects), scene-preference (scene > objects), whole-objects-preference (face + body + scene + object (chair+tool) > phase-scrambled), and phase-scrambled-objects-preference (phase-scrambled > face + body + scene + object (chair+tool)).

A group random-effect statistical Z-map for each contrast was then computed from the single-subject localizer GLMs (see **Fig. S3A** for the group z-statistic maps for body-, face-, and scene-preferences). The Z-values obtained were sorted in magnitude. From the highest Z-values from the map, the group peak with the next highest Z-value was generated until the Z-value was <= 2.0. Group peaks had to be separated by at least 38.4mm (9.6 mm x 4) in the sphere mesh to prevent a vertex being assigned to multiple ROIs in a subject. ROIs were then defined separately for each participant based on the individual's univariate statistical maps (11, 12). From each group peak defined above, the corresponding peak for an individual subject peak was defined as the vertex with the highest Z-value within a sphere of 9.6 mm radius centered around the group peak in each subject's sphere mesh. The single-subject ROI was formed from the vertices exceeding Z= 2.0 in a sphere of 9.6 mm radius centered around the peak in the subject's mesh. All ROIs used in the following analysis contained at least 175 vertices in at least 14 subjects. ROIs in individual subjects with less than 175 vertices were discarded.

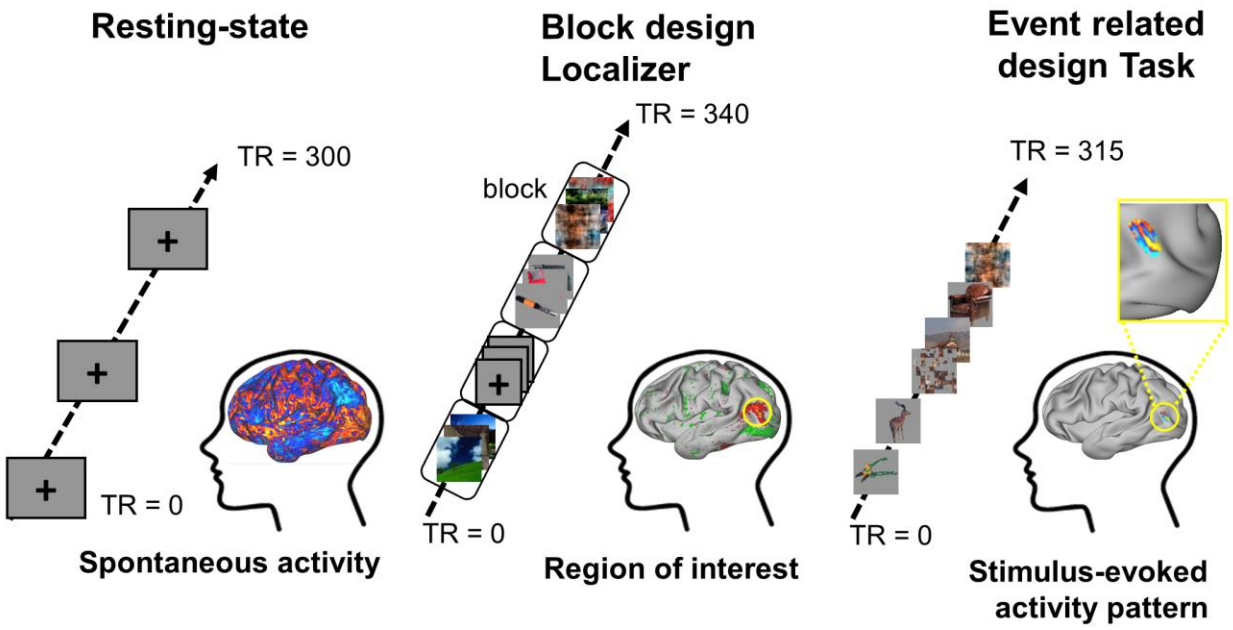
To remove differences in BOLD magnitude across MR frames, for each ROI a z-normalization was applied across the vertices of each frame of the resting and task scans. This within-frame Z-normalization was not applied to the localizer scans. Also, it was not applied to the Task scans for a separate analysis described below in which task-evoked activation magnitudes were determined (see below, ***Task scans: multi-voxel activation patterns).***

Two sets of ROIs were created for use in different analyses. The first set was created from the localizer-defined ROIs that preferred a particular target category (face, body, or scene) relative to the object category (chairs + tools). We grouped the vertices from the target-preferred ROIs into a single target-preferred 'joint-ROI'. For each joint-ROI, all vertices located in early visual areas (V1 to V3) were excluded (13), as estimated from surface topology using the template created by Wang et al. (14). A second set of ROIs consisted of Whole-Object ROIs (face + body + scene + object > phase-scrambled) and Phase-Scrambled Object ROIs (phase-scrambled > face + body + scene + object). Whole-Object and Phase-Scrambled Object ROIs were grouped, respectively, into a Whole-Objects joint-ROI and a Phase-Scrambled Objects joint-ROI. **Supplementary Table 1** summarizes the mean MNI coordinate, mean Z-score for the

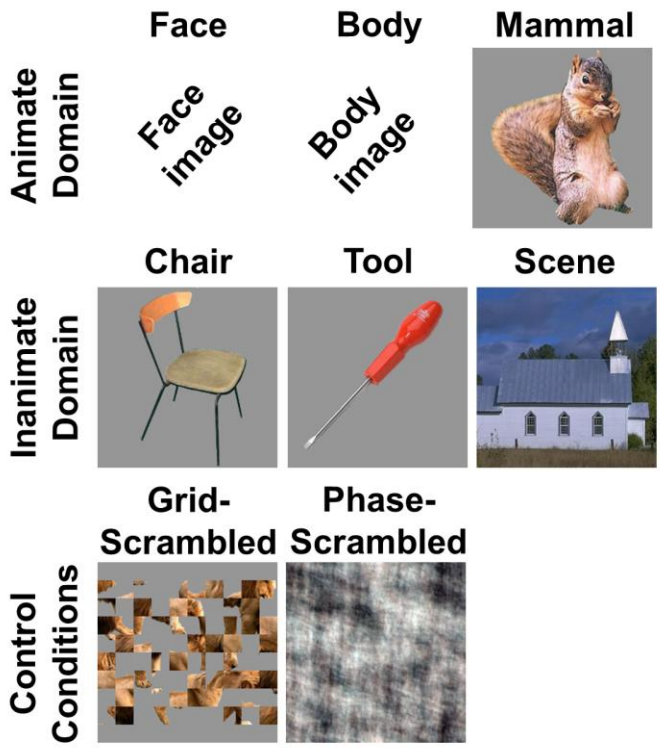
obtained group peak, and mean number of vertices for all constituent ROIs in each joint-ROI. **Supplementary Figure 3B** schematically indicates the position of all constituent ROIs in the Face-, the Body- and Scene-joint-ROIs based on their group peak locations. **Supplementary Figure 4B** shows the location of all constituent ROIs in the Whole-Objects joint-ROI and the Phase-Scrambled Objects joint-ROI superimposed on a surface map of V1-V3 using the template from Wang et al. (14).

Supplementary References

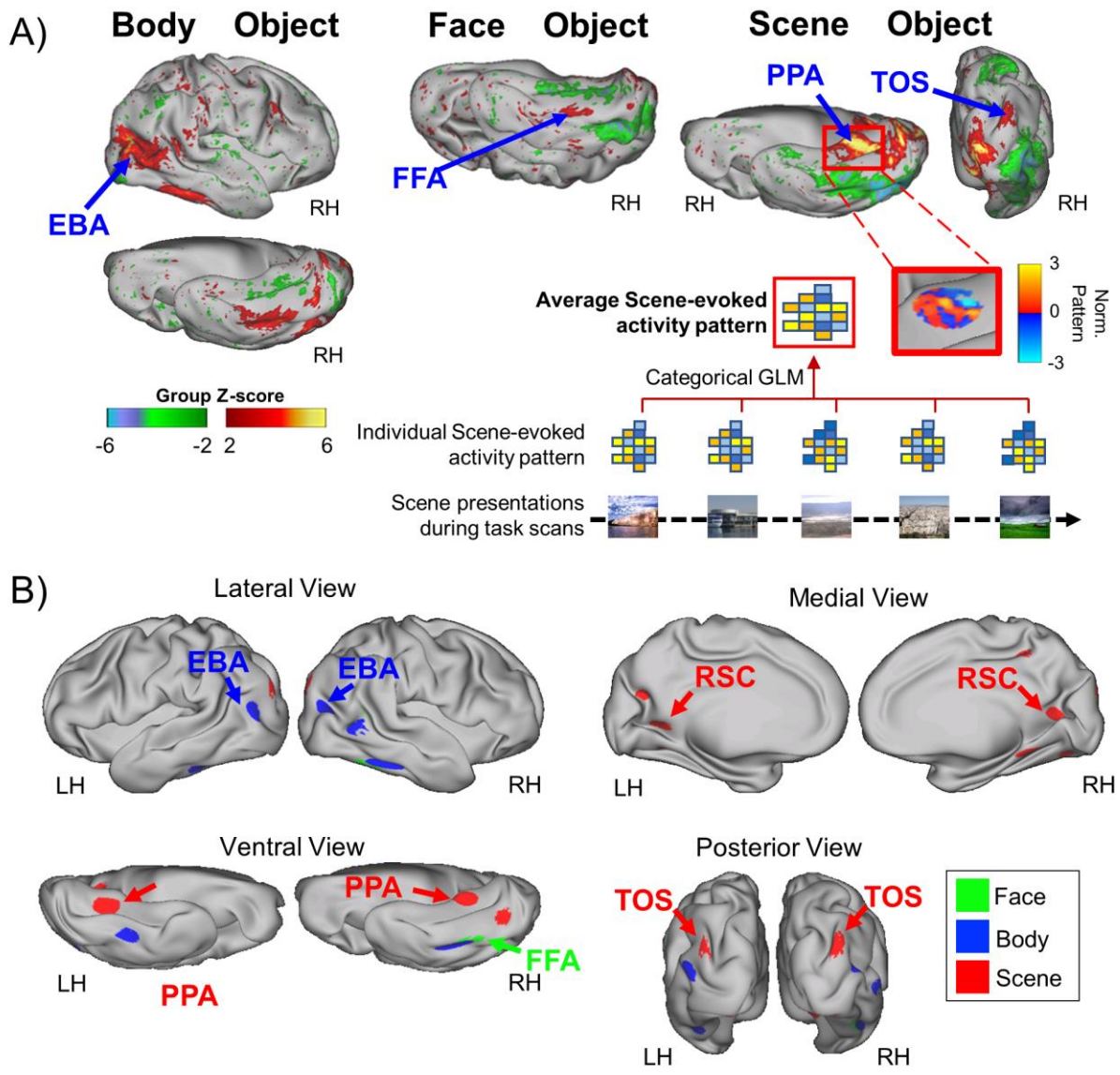
1. Downing PE, Chan AW, Peelen MV, Dodds CM, & Kanwisher N (2006) Domain specificity in visual cortex. *Cereb Cortex* 16(10):1453-1461.
2. Kriegeskorte N*, et al.* (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60(6):1126-1141.
3. Watson DM, Hymers M, Hartley T, & Andrews TJ (2016) Patterns of neural response in scene-selective regions of the human brain are affected by low-level manipulations of spatial frequency. *Neuroimage* 124(Pt A):107-117.
4. Brainard DH (1997) The Psychophysics Toolbox. *Spatial vision.* 10(4).
5. Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8(2-3):109-114.
6. Siegel JS*, et al.* (2016) Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. *Proc Natl Acad Sci U S A*.
7. Glasser MF*, et al.* (2013) The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80:105-124.
8. Fischl B, Sereno MI, Tootell RB, & Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8(4):272-284.
9. Van Essen DC*, et al.* (2001) An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform Assoc* 8(5):443-459.
10. Bracci S & Op de Beeck H (2016) Dissociations and Associations between Shape and Category Representations in the Two Visual Pathways. *J Neurosci* 36(2):432-444.
11. Oosterhof NN, Tipper SP, & Downing PE (2012) Viewpoint (in)dependence of action representations: an MVPA study. *J Cogn Neurosci* 24(4):975-989.
12. Wurm MF, Ariani G, Greenlee MW, & Lingnau A (2016) Decoding Concrete and Abstract Action Representations During Explicit and Implicit Conceptual Processing. *Cereb Cortex* 26(8):3390-3401.
13. Strappini F*, et al.* (2018) Resting-State Activity in High-Order Visual Areas as a Window into Natural Human Brain Activations. *Cereb Cortex*.
14. Wang L, Mruczek RE, Arcaro MJ, & Kastner S (2015) Probabilistic Maps of Visual Topography in Human Cortex. *Cereb Cortex* 25(10):3911-3931.



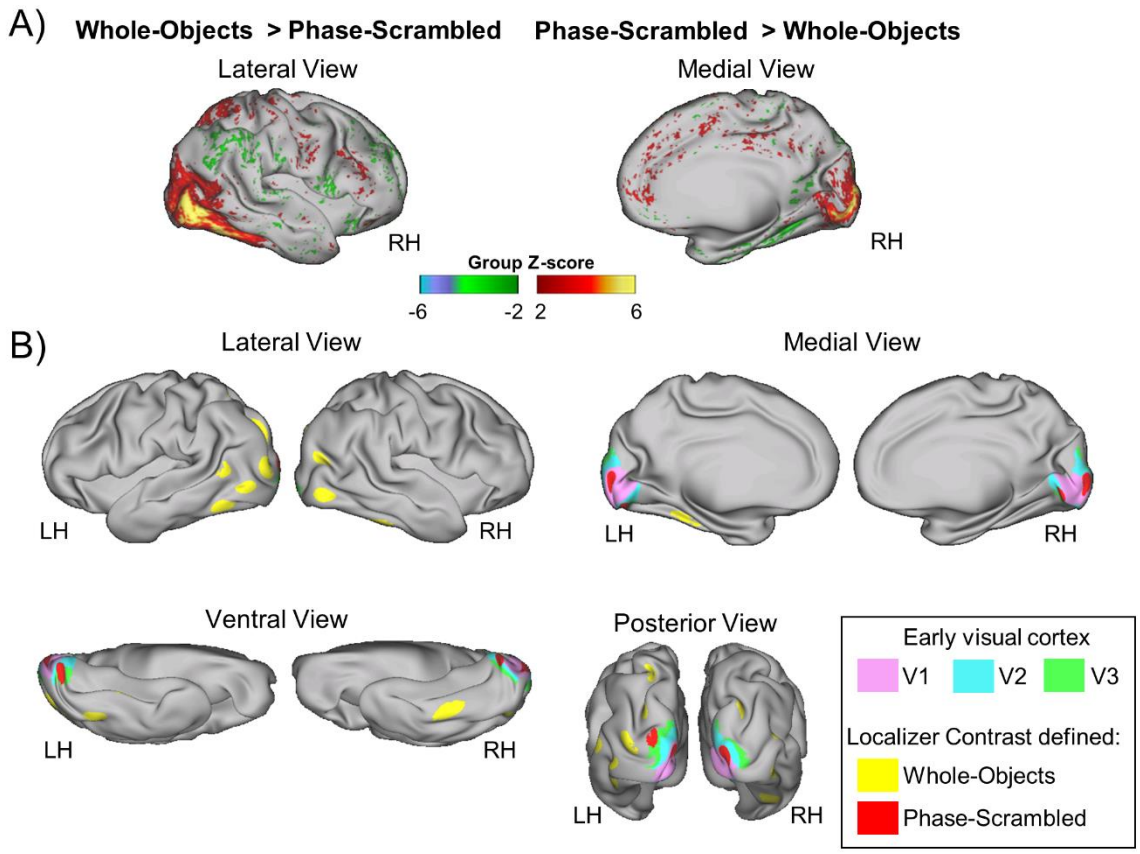
Supplementary Figure 1. Experimental design, with separate resting scans, blocked-design localizer scans, and event-related task scans.



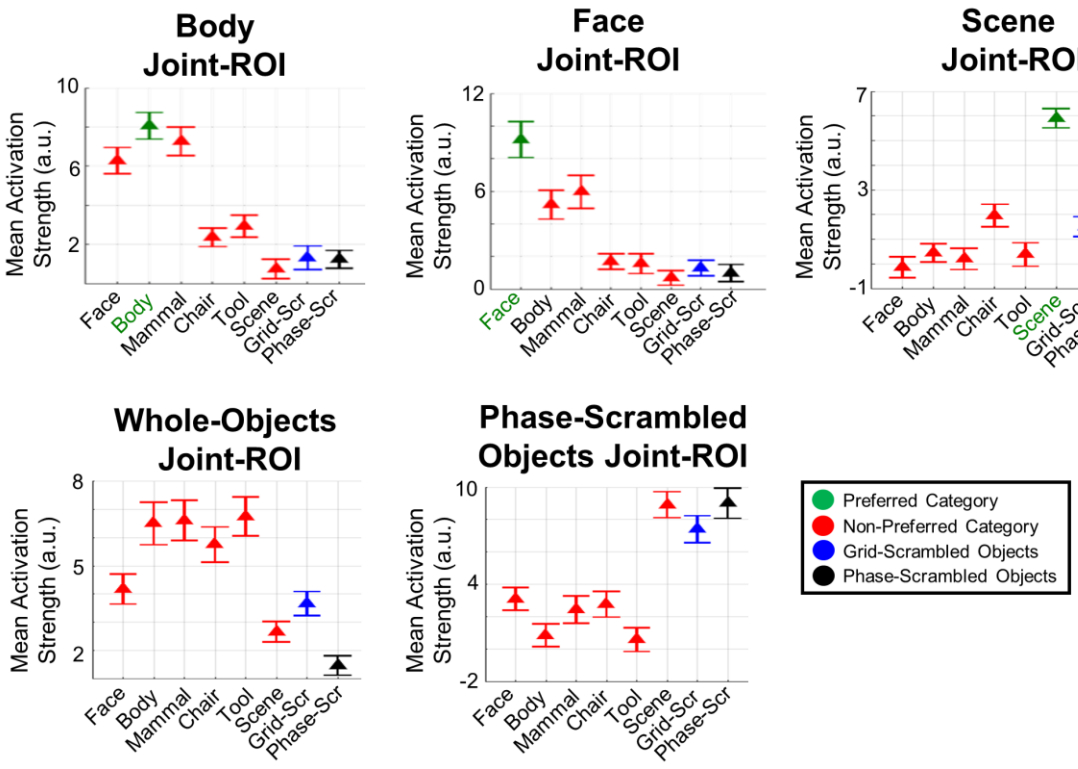
Supplementary Figure 2. Stimulus categories used in the experiment.



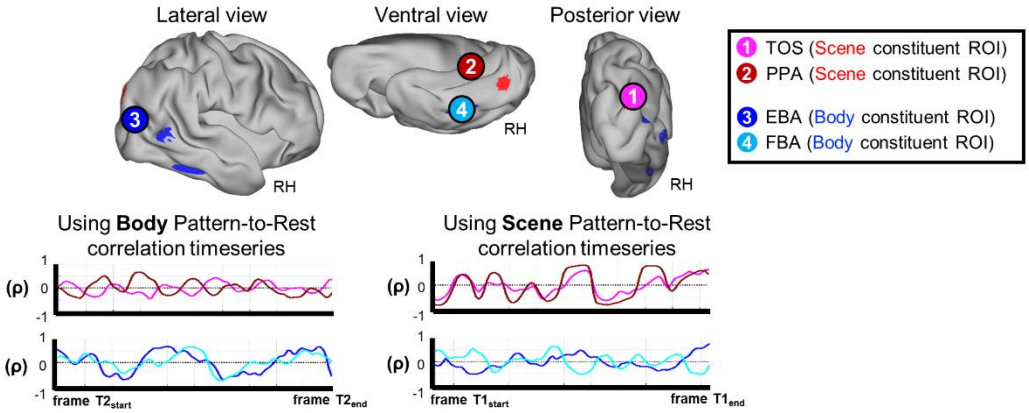
Supplementary Figure 3. Body, Face, and Scene ROIs (regions of interest). **(A)** Group z-statistic Localizer maps for category-preferring visual regions. ROIs were separately defined for each individual from their localizer maps using the group foci as a constraint. **(B)** Schematic rendering of three sets of category-preferential ROIs for faces, bodies, and scenes using the object category (tools and chairs) as the baseline.



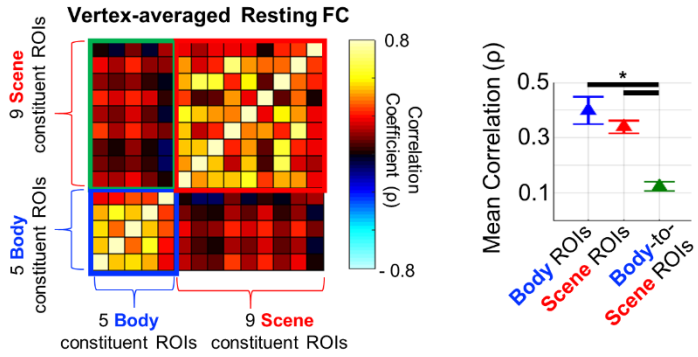
Supplementary Figure 4. Whole-object and Phase-scrambled ROIs. **(A)** Group z-statistic Localizer maps of visual regions that prefer whole-objects or phase-scrambled objects. **(B)** Schematic rendering of Whole-Objects and Phase-Scrambled ROIs. ROIs were separately defined for each individual from their localizer maps using the group foci as a constraint. Surface renderings of V1, V2, and V3 from the Wang et al. template (14) are superimposed.



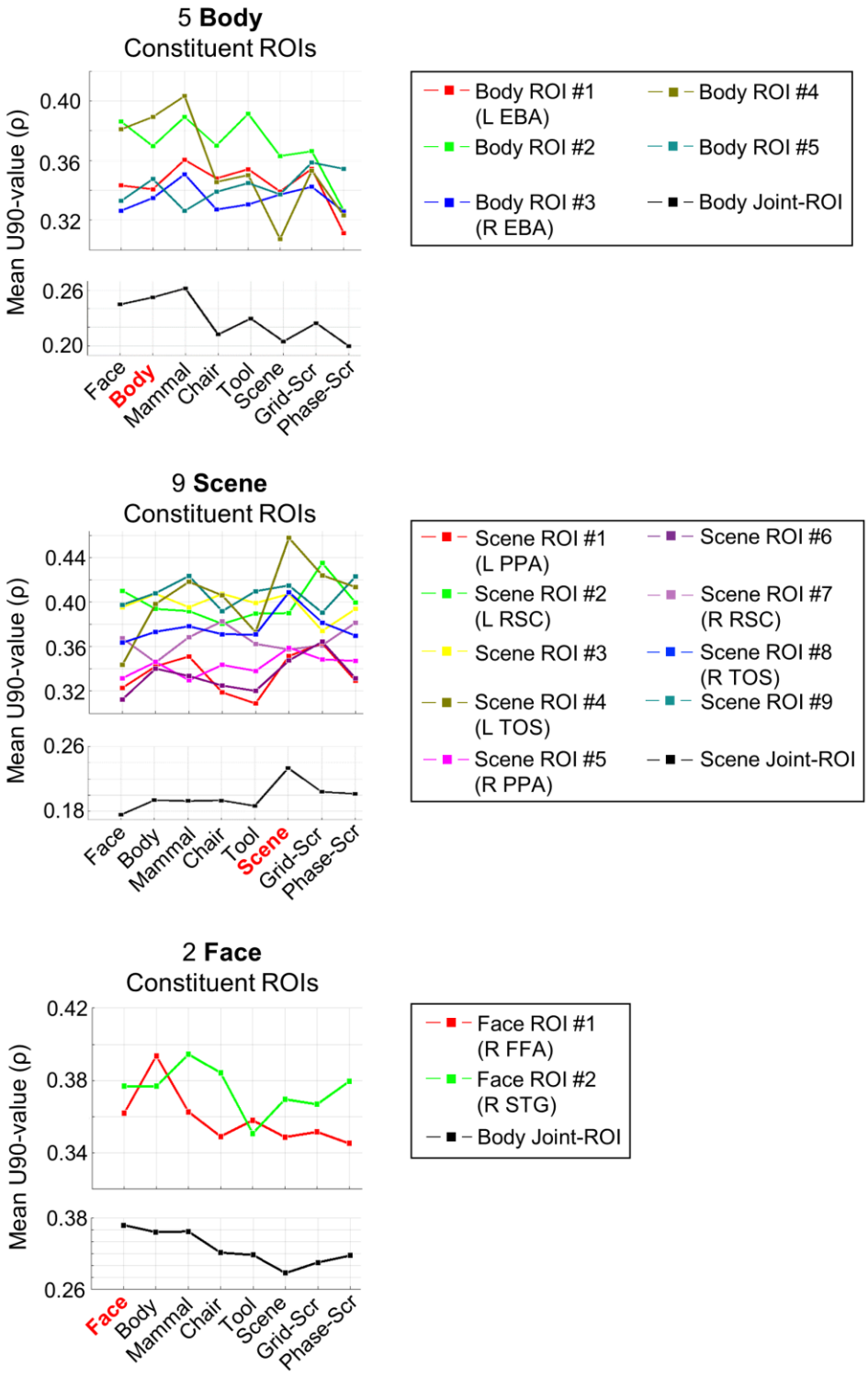
Supplementary Figure 5. Group mean activation strength from task scans for all stimulus categories in each joint-ROI. Error bars indicate ±SEM.



Supplementary Figure 6. Pattern-to-rest correlation timeseries computed using body-evoked and scene-evoked activity patterns in two scene-preferring and two body-preferring ROIs.



Supplementary Figure 7. A standard non-pattern-based FC matrix was computed by first averaging the vertexwise timeseries of the BOLD signal within each region, and then correlating the resulting timeseries across all pairs of regions (**Left**). All cells involving FC between body regions, between scene regions, and between body and scene regions are averaged and plotted **(Right)**.



Supplementary Figure 8. Graphs of the profile of group-averaged U90 values across stimulus categories for each constituent body region, scene region, and face region.

		Mean MNI across subjects					# Vertices		
Contrast	Hemisphere	x	y	z	Group Peak Z-value	# Subjects	Mean	STD	ROI Name
Face > Objects	RH	41	-50	-16	3.76	14	262	70	R FFA
		54	-49	12	2.71	15	232	74	R STG
Body > Objects	LH	-41	-73	12	6.61	16	289	45	L EBA
		-37	-44	-17	4.42	14	253	73	
	RH	42	-65	15	6.28	15	312	52	R EBA
		41	-44	-13	4.62	14	280	69	
		56	-52	4	3.91	14	276	73	
Scene > Objects	LH	-21	-47	-4	8.14	16	327	37	L PPA
		-14	-51	7	5.88	16	294	56	L RSC
		-17	-59	21	3.95	14	236	42	
		-30	-80	23	3.19	14	211	57	L TOS
	RH	20	-41	-10	8.35	16	279	45	R PPA
		27	-68	-10	6.79	16	254	43	
		21	-54	10	6.02	14	285	76	R RSC
		35	-77	29	4.55	15	273	64	R TOS
		13	-36	41	4.01	14	237	65	
Whole-Objects > Phase-Scrambled Objects	LH	-41	-73	-1	8.46	16	374	31	L LO
		-35	-66	-12	7.50	16	334	36	
		-20	-72	30	5.92	15	252	42	
		-30	-83	11	5.71	16	280	63	
		-26	-39	-14	5.57	15	243	48	
		-22	-56	45	4.12	14	245	70	
		-51	-59	8	3.10	14	231	49	
	RH	44	-72	-7	8.18	16	322	38	R LO
		36	-44	-20	6.47	16	295	43	
		26	-74	30	5.03	15	237	60	
		40	-64	16	4.74	14	263	82	
Phase-Scrambled Objects > Whole-Objects	LH	-5	-89	3	8.25	16	276	34	
		-8	-80	-11	5.84	16	267	40	
		-19	-89	16	3.71	15	282	54	
	RH	10	-93	4	7.41	16	354	37	
		9	-77	-4	7.4	16	247	35	

Supplementary Table 1. ROI Summary

	Similarity (ρ) between Face-evoked activity pattern and Body-evoked activity pattern	Similarity (ρ) between Face-evoked activity pattern and Scene-evoked activity pattern	Similarity (ρ) between Body-evoked activity pattern and Scene-evoked activity pattern
Face Joint-ROI	0.729	0.415	0.550
Body Joint-ROI	0.618	0.421	0.423
Scene Joint-ROI	0.506	0.356	0.379

Supplementary Table 2. Spatial correlation between face-evoked, body-evoked, and scene-evoked activity patterns in three joint-ROIs.