Representational Tuning Models: Parametric Encoding Within Arbitrary Stimulus Domains

Ben Harvey

b.m.harvey@uu.nl

Utrecht University https://orcid.org/0000-0003-2694-618X

Stan Bergey

Tilburg University

Luis Martín-Roldán Cervantes

Technical University Eindhoven

Jorge Almeida

University of Coimbra https://orcid.org/0000-0002-6302-7564

Technical Report

Keywords: Representational Tuning, Encoding Models, Representational Similarity Analysis (RSA), Functional Connectivity, Population Receptive Field (pRF)

Posted Date: June 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9940836/v1

Additional Declarations: There is **NO** Competing Interest.

Representational Tuning Models: Parametric Encoding Within Arbitrary Stimulus Domains

Stan Bergey[1], Luis Martín-Roldán Cervantes[2], Jorge Almeida[3], Ben M. Harvey[4]

[1]Department of Computational Cognitive Science, Tilburg University, Netherlands

[2]EngD Program Software Technology, Technical University Eindhoven, Netherlands

[3]PROACTION lab, Faculty of Psychology and Educational Sciences, University of Coimbra, Portugal

[4]Department of Experimental Psychology, Utrecht University, Netherlands

Abstract

Encoding models characterize parametric response functions for individual recordings but require predefined stimulus dimensions. Conversely, Representational Similarity Analysis (RSA) accommodates arbitrary stimuli but relies on population-level patterns, obscuring individual-site tuning. To bridge this gap, we introduce Representational Tuning Models (RTMs). RTMs utilize RSA and multidimensional scaling to project arbitrary stimuli into continuous, latent representational spaces, within which parametric tuning functions are fit to predict individual recording responses. Validating this framework with 7T fMRI responses to natural scenes, we show that RTMs capture voxel-specific tuning in high-level (but not early) visual maps with performance approaching the noise ceiling. Our results reveal systematic increases in selectivity across the visual hierarchy and show that latent representational space dimensions largely reflect cortical topography. Finally, we quantify representational sampling and functional connectivity by modeling voxels as functions of distant representational spaces. RTMs provide a unified, modality-agnostic tool for mapping neural tuning and information transformations.

Keywords: Representational Tuning, Encoding Models, Representational Similarity Analysis (RSA), Functional Connectivity, Population Receptive Field (pRF).

Highlights:

-RTM unifies RSA and encoding models for parametric tuning of individual recordings.

-Enables site-specific parametric modeling for arbitrary or high-dimensional stimuli.

-Reveals that latent representational dimensions reflect cortical topography.

-Uses Between-ROI models to quantify representational sampling connectivity.

-Identifies increased image selectivity across the visual hierarchy.

Significance Statement:

Understanding how the brain encodes information typically requires a trade-off between the stimulus flexibility of population-level pattern analyses and the individual-site specificity of encoding models. We introduce Representational Tuning Models (RTMs), a framework that bridges this gap by deriving continuous tuning parameters directly from latent representational spaces.

By applying RTMs to 7T fMRI data, we demonstrate that this approach recovers voxel-specific tuning for high-dimensional natural images and reveals a systematic increase in tuning selectivity across the visual hierarchy. Furthermore, we show that Between-ROI models can characterize neural responses as a sampling of representational spaces from distant regions, providing a novel metric for functional connectivity. RTMs offer a unified, data-driven tool for mapping neural tuning and information transformation across diverse stimulus domains, recording modalities and brain regions.

Introduction

Parametric encoding models, such as population receptive field (pRF) models [1], are widely used to characterize individual recording sites by mapping responses as functions of to continuous stimulus dimensions. This approach has successfully quantified neural tuning across visual, auditory, and sensorimotor domains[1–3], revealing how response functions change across cortical maps[4–7] and are affected by attention, development, aging, and clinical conditions[6,8–10]. Beyond early sensory maps, encoding models have also elucidated higher-level processing for numerosity, object properties, and event timing[11–19]. However, many naturalistic stimuli vary in complex, high-dimensional ways, making it difficult to hypothesize the low-dimensional parameters necessary for traditional encoding models.

When stimulus parameters are unknown, researchers often turn to Representational Similarity Analysis (RSA)[20]. By calculating a Representational Dissimilarity Matrix (RDM)[21] (Fig. 1A), RSA identifies "fingerprints" of regional representational content without requiring a priori stimulus parameterization. RSA is highly flexible, allowing comparisons across species or with artificial vision models[22]. Nevertheless, RSA's pattern-level focus obscures the tuning of individual recording sites and precludes the powerful parametric analyses enabled by encoding models.

To bridge these methodologies, we introduce **Representational Tuning Models (RTMs)** (See Fig. 1, Methods at a Glance box, Methods section and Algorithmic Appendix). Using multidimensional scaling to project high-dimensional response pattern similarity from arbitrary stimulus sets into a low-dimensional representational space[23] (Fig. 1B), we define a parametric representational tuning function (here a circular Gaussian) for each recording site within this representational space. Specifically, we fit this function to the individual recording site's response amplitudes to all conditions (Fig. 1C), where the positions of the conditions are taken from the two-dimensional representational space of a larger region of interest (ROI) (Fig. 1B). This allows us to quantify which conditions a site responds to using a parametric function (Fig. 1D), and so reveal how this function's parameters change between recording sites using the same interpretive framework as low-dimensional encoding models. We first test if RTMs can predict responses to independent measurements and assess if their parameters vary systematically within and between visual field maps. Furthermore, by fitting tuning functions for each recording site in the representational spaces of other distant visual field maps, we reveal how the responses of recording sites in one brain area are functionally connected to responses throughout another area, thereby extend the concept of connective field modeling[24] into representational spaces.

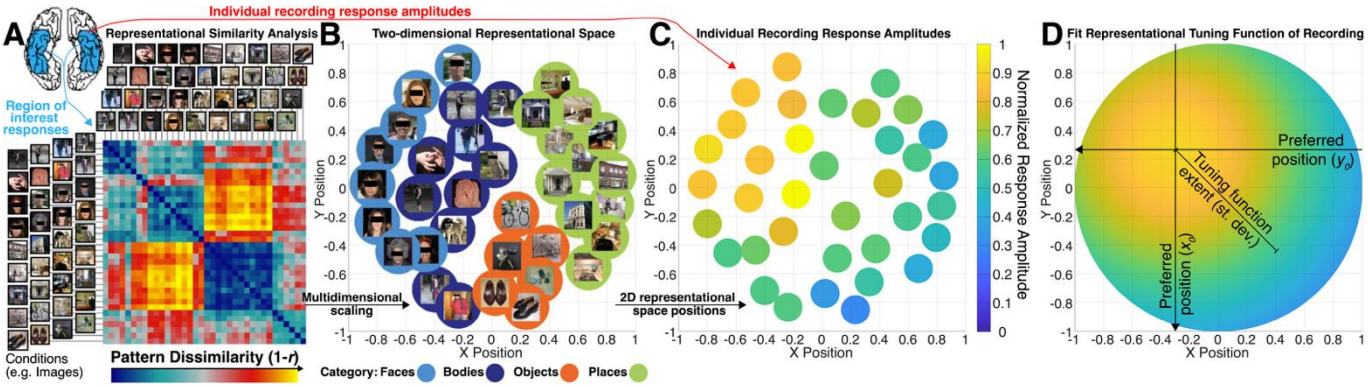

Figure 1: Schematic of the Representational Tuning Modeling (RTM) procedure. (**A**) RSA computes pairwise dissimilarities (1-correlation) between the response pattern across a large set of recordings for all conditions (images) to generate a Representational Dissimilarity Matrix (RDM). (**B**) Multidimensional scaling (MDS) projects these dissimilarities into positions in a two-dimensional representational space, where distances between conditions reflect pattern dissimilarities. (**C**) For each individual recording, an RTM uses the positions from this two-dimensional representational space and the individual recordings' response amplitudes for the same conditions. (**D**) The RTM then fits a parametric response function to these response amplitudes. Here we use a circular Gaussian parameterized by two preferred position (x_0, y_0) parameters and an extent (standard deviation, st. dev). While the representational space (B) is shared across the ROI, the tuning function is specific to each recording's unique response amplitudes profile, enabling individual-site characterization within a common representational framework. Images (A) and (B) adapted from Charest et al. [25].

We validated this approach using the Natural Scene Dataset (NSD)[26]. This 7T fMRI dataset provides small-scale recording sites (voxels), complex naturalistic stimuli that defy simple parameterization, and independent visual field mapping for ROI definition (Fig. 2). Critically, the repetition of thousands of stimuli allows for robust cross-validation of the RTM framework across an exceptionally comprehensive representational space.

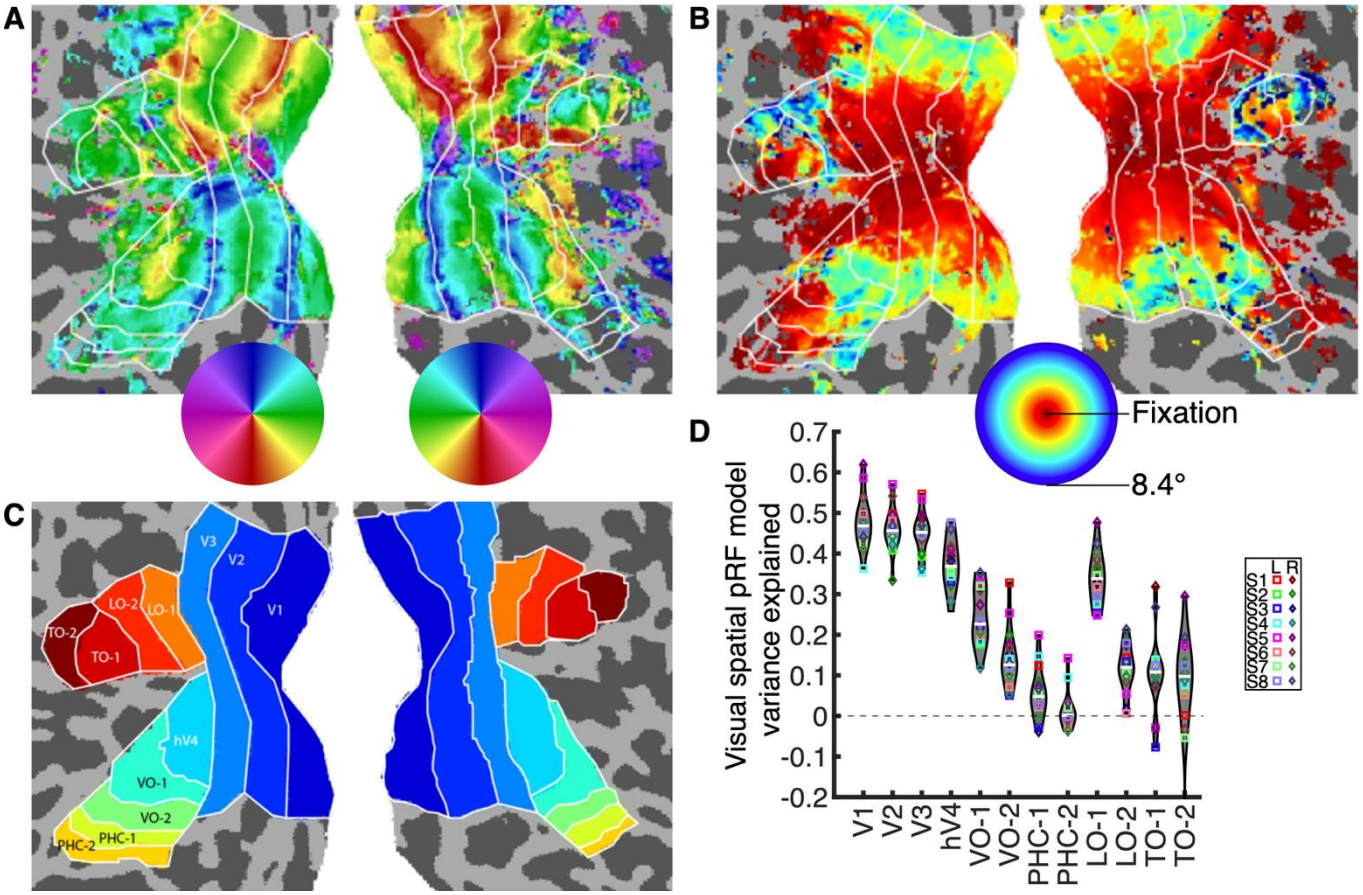

Figure 2: Visual field mapping data and ROI definitions. (**A**) Preferred visual field position polar angle and (**B**) eccentricity (colors) of position-selective voxels in a flat map of the early visual cortex of subject 1, which were used to identify visual field map borders (lines). (**C**) The resulting visual field map ROI definitions for this subject. (**D**) Variance explained of the visual field mapping pRF model in each visual field map and hemisphere (colored markers, white line shows the mean).

Methods at a Glance: Representational Tuning Modeling (RTM)

The RTM framework bridges the gap between population-level pattern analysis and individual-site encoding. It is a four-stage pipeline designed for any recording modality (fMRI, ECoG, single-unit, neural network) and arbitrary stimulus sets. For details see the Methods and Algorithmic Appendix.

1. Constructing the Representational Space

- **Input:** A matrix of neural responses (Recording sites × Conditions) from a defined source region of interest.
- **RSA:** Compute a Representational Dissimilarity Matrix (RDM) using (1 – Pearson's *r* or Euclidean distance) for all condition pairs (Fig. 1A).
- **Dimensionality Reduction:** Apply Multidimensional Scaling (MDS) to project the RDM into a low-dimensional (e.g., 2D) coordinate system. Each stimulus now has a fixed position (*x,y*) in this "latent" representational space (Fig. 1B).

2. Fitting the Representational Tuning Function

- **Recording Site Selection:** Identify a target response recording (within the source ROI for *Within-ROI* models, or elsewhere for *Between-ROI* models).
- **Parameterization:** Fit a parametric function (e.g., a circular Gaussian) to the target recoding's response amplitudes across all conditions (Fig. 1C-D).
- **Model Equation:** The predicted response for a condition at position (*x,y*) is defined by:

$$g(v) = A \cdot e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}} + k$$

Where (x_0, y_0) is the preferred representational position and σ is the tuning function extent.

3. Model Validation & Performance

- **Cross-Validation:** Fit parameters on a training set and evaluate on an independent test set.
- **Variance explained:** Calculate the cross-validated variance explained in the test set.
- **Model Performance:** Normalize variance explained by the Noise Ceiling (the maximum explainable variance given the data's repeatability) to determine the proportion of biologically available signal captured by the RTM.

4. Between-ROI Connectivity & Mapping

- **Source ROI Selection:** For any target region, iteratively test different source ROIs. The source that yields the highest training-set variance explained (across all examples of the target ROI) is identified as the primary representational source.
- **Whole-Brain Mapping:** Apply the selected Between-ROI model to the entire cortical surface to visualize locations responding to specific parts of the source ROI's representational space.

Results

Model Goodness of Fit and Performance

We first evaluated the RTM's ability to explain neural responses across visual field maps (Fig. 3A). The Within-ROI model explained variance significantly above zero (under cross-validation) in V1 and high-level maps (LO-2, PHC-2, TO-1; Table S1). However, this model explained a relatively small proportion of the response variance in our test set, which consists of a single presentation of each condition. Because test-set repeatability (noise ceiling) was generally low (<0.15) and differed between visual field maps, we calculated Model Performance, the proportion of the noise ceiling captured by the model, to normalize for varying data quality (Fig. 3B). This revealed that while V1 had significant variance explained due to high repeatability, its Model Performance was low. In contrast, higher-level maps such as PHC-2, LO-2, and TO-1 showed Model Performance significantly above zero (Fig. 3B, Table S2) and often exceeding 50%.

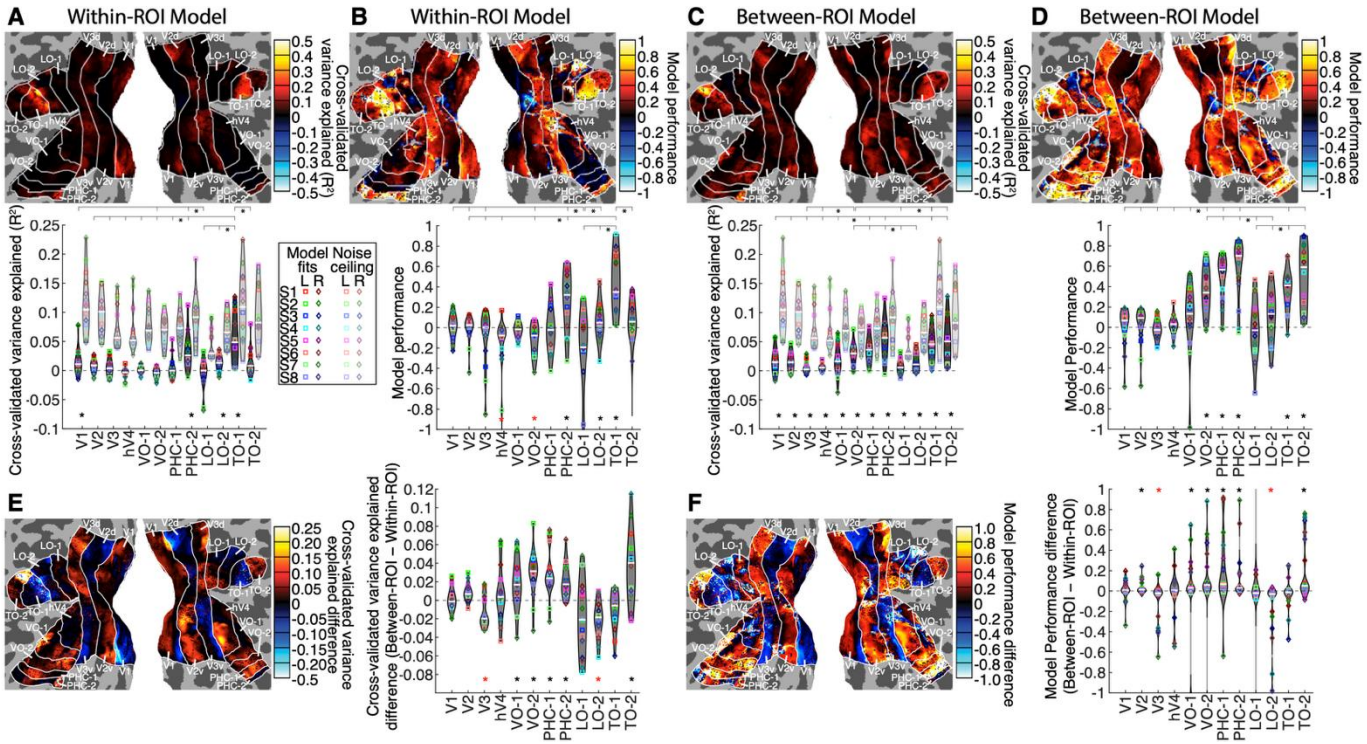

Figure 3: Goodness-of-fit of representational tuning models across visual field maps. (**A**) Cross-validated variance explained for the Within-ROI model in an example subject (top) and averaged surface vertices in each visual field map in each hemisphere (colored markers, bottom). Light violins/markers indicate the average noise ceiling in each hemisphere. (**B**) Model Performance (variance explained normalized by noise ceiling) shows increased explanatory power in higher-level maps. (**C-D**) Variance explained and Model Performance for the Between-ROI model, following the same format as (A) and (B). Variance explained and Model performance progressively increased through the ventral (hV4, VO-1/2, PHC-1/2) and lateral (LO-1/2, TO-1/2) streams. Here, variance explained was significantly positive in all visual field maps, but Model Performance was only significant in higher-level visual field maps (VO-2, PHC-1/2, TO-1/2). (**E-F**) Direct comparisons revealed that Between-ROI models significantly outperformed Within-ROI models especially in these high-level visual field maps. Black and red stars indicate median values significantly above or below zero, respectively. Brackets denote significant differences between map distributions with markers to the left and right of the star. White lines represent the mean across hemispheres.

Two-factor ANOVAs (visual field map, subject) confirmed significant differences between maps for both variance explained ($F_{11,173}$ = 15.05, p = 3×10^{-20}) and Model Performance ($F_{11,173}$ = 23.59, p = 5×10^{-20}). Notably, performance increased from early to late visual maps, unlike traditional spatial pRF models in the same subjects ($F_{11,173}$ = 105, $p = 2\times10^{-70}$), which perform best in early visual areas (Fig 2D). This dissociation suggests RTM fits are driven by representational content rather than simple spatial contrast selectivity.

To investigate whether voxels are better described by the representations of other regions that may provide their functional inputs, we developed a Between-ROI model. For each target map, we identified the "source" ROI whose representational space best predicted target voxel responses in the training set when averaged across all hemispheres (Table S3). Generally, V1–V3 were best predicted by each other, while TO-1 served as the optimal source for many higher-level maps. However, as differences between multiple candidate source visual field maps were often non-significant (Table S4), we interpret these as primary rather than exclusive sampling relationships.

The Between-ROI model explained test set variance significantly above zero in all tested visual field maps (Fig. 3C, Table S5). Normalizing for repeatability, Model Performance was significantly above zero in higher-level maps (VO-2, PHC-1/2, TO-1/2; Fig. 3D, Table S6), sometimes exceeding 60% of the explainable response. Consistent with the Within-ROI results, a two-factor ANOVA confirmed significant differences between maps for both variance explained ($F_{11,173}$ = 15.44, $p = 10^{-20}$) and Model Performance ($F_{11,173}$ = 5.8, $p = 6\times10^{-08}$), with increases from earlier to later visual field maps.

Critically, Wilcoxon signed-rank tests demonstrated that the Between-ROI approach provided significantly better fits than the Within-ROI model in almost all higher-level visual field maps where Model Performance was significantly above zero (VO-1/2, PHC-1/2, TO-2; Fig. 3E, F, Table S7-S8).

Representational Tuning Function Extents

We next assessed whether visual field maps differed in tuning function extent: the range of the specific subsets of the conditions in the representational space to which voxels respond. The extent is defined as the standard deviation (σ) of the fit Gaussian; smaller values indicate more selective responses to specific conditions, or a larger range of response amplitudes between conditions. Given that the representational space has a radius of one (maximum distance of two), larger values imply that while response amplitudes may vary, all conditions elicited relatively strong responses.

For both Within-ROI (Fig. 4A) and Between-ROI (Fig. 4B) response models, tuning function extents were largest in early visual field maps (V1–V3: σ ≈ 6-8) and significantly smaller in later maps (PHC-2, TO-1/2: σ ≈ 2-5). Two-factor ANOVAs confirmed a significant effect of visual field map on tuning function extent for both Within-ROI ($F_{11,173}$ = 25.2, p = 2×10^{-30}) and Between-ROI models ($F_{11,173}$ = 48.1, p = 8×10^{-47}). While Wilcoxon signed-rank tests showed isolated differences between the two models in a few maps (V3, PHC-1, LO-2), these lacked a consistent direction. Thus, while tuning function extent varied systematically across the visual hierarchy, it remained consistent across both modeling frameworks.

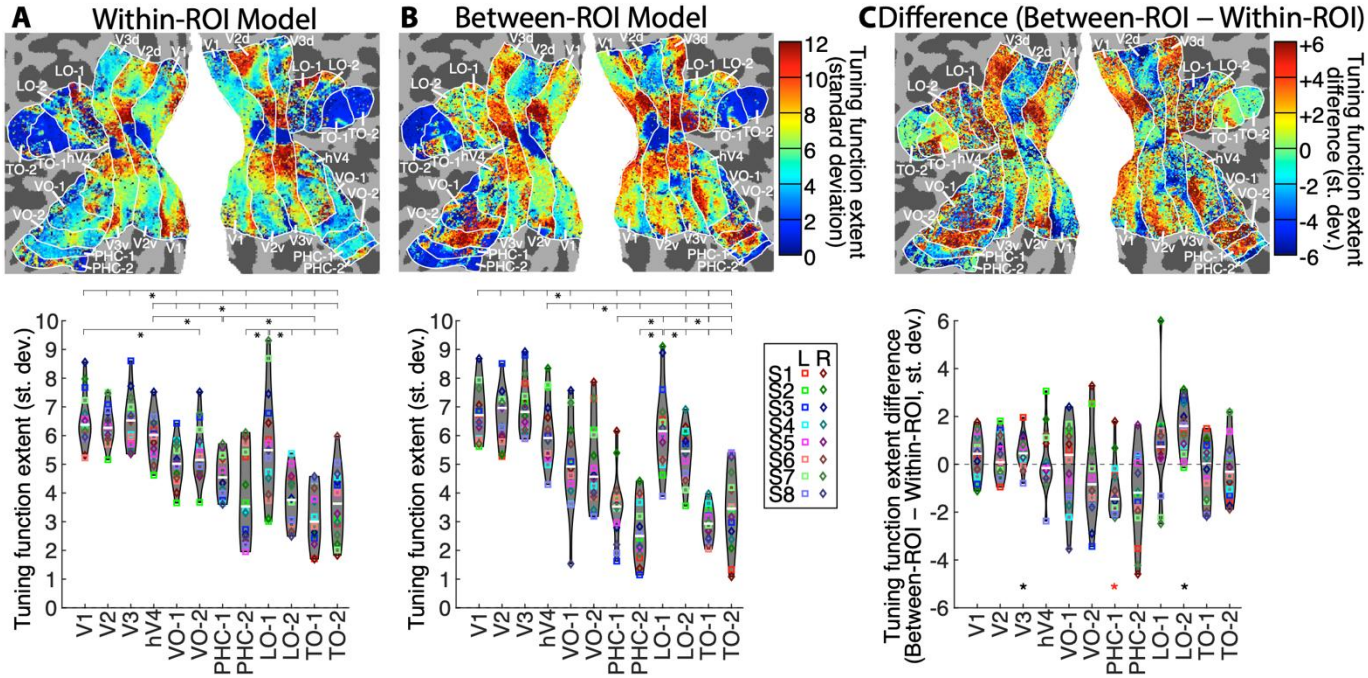

Figure 4: Representational tuning function extents across visual field maps. (**A**) Extent (standard deviation, st. dev.) of the Within-ROI representational tuning function in an example subject (top) and averaged across surface vertices (bottom) for each hemisphere (colored markers). Extents decreased through the visual hierarchy. (**B**) Extent of the Between-ROI representational tuning function show a similar hierarchical decrease. Brackets denote significant differences between map distributions with markers to the left and right of the star. (**C**) Differences in tuning function extent between models show large local variations (top) but showed no systematic difference considering all hemispheres. Black and red stars indicate median differences significantly above or below zero, respectively. White lines represent the mean across hemispheres.

Spatial Organization of Preferred Representational Positions

We next examined whether voxels' preferred representational positions (x_0 and y_0) were systematically organized on the cortical surface (Fig. 5A–B). Because MDS-derived representational spaces are rotation-invariant, we first identified a global alignment axis to visualize these preferences. For each voxel pair, we calculated the representational distance and direction vector divided by their cortical surface distance. Bayesian analysis across all maps and hemispheres favored

a von Mises (circular Gaussian) distribution over a uniform distribution (Tables S9-S10), confirming a consistent direction of maximal change. We used the circular mean of these directions across all visual field maps and hemispheres (86.59° Within-ROI; 88.14° Between-ROI) to rotate all preferred positions around the origin. This transformation provided a consistent orientation for visualization (Fig. 5A-B) without affecting quantitative distance-based analyses.

To quantify the relationship between representational and cortical locations, we calculated the Spearman correlation (ρ) between pairwise cortical surface distances and pairwise distances in representational space. In every visual field map and both models, median correlations were significantly above zero (Wilcoxon signed-rank tests; Fig. 5C & E, Tables S11-12). Two-factor ANOVAs revealed significant differences in these correlations between maps (Within-ROI: $F_{13,203} = 5.98$, $p < 10^{-70}$. Between-ROI model: $F_{13,203} = 5.02$, $p < 10^{-70}$). Specifically, the coupling between representational and cortical distance was strongest in early visual field maps (V2, V3), lowest in lateral maps (LO-1/2, TO-1/2), and intermediate in ventral maps (VO-2, PHC-1/2). These results indicate that the spatial structure of responses across the cortex partially reflects the principal components of the representational space and vice-versa.

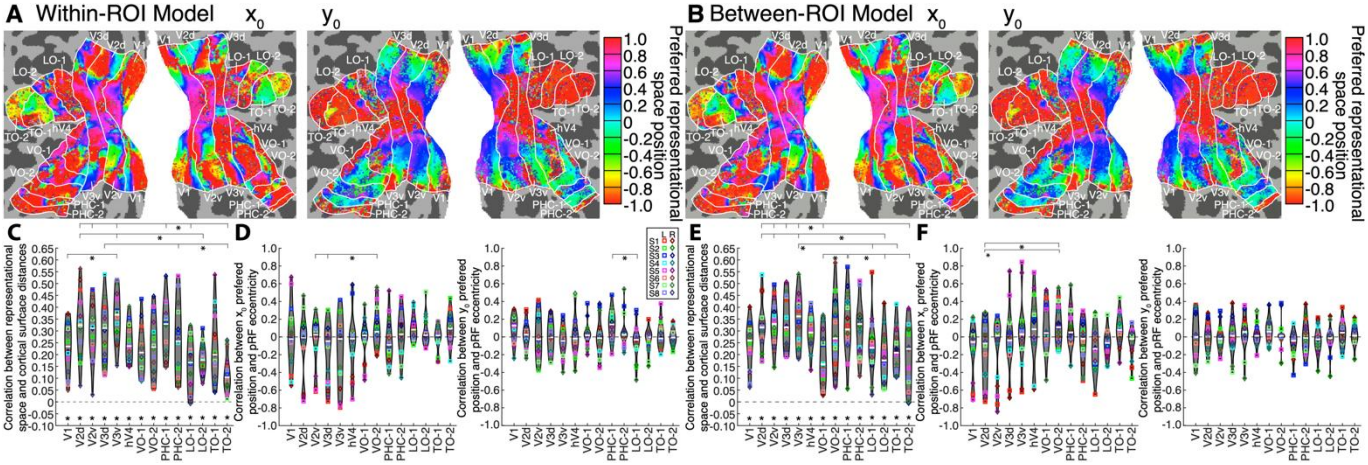

Figure 5: Spatial organization of preferred representational positions. (**A–B**) Preferred positions (x_0, y_0) of representational tuning functions progressed gradually across the cortical surface in many regions. (**C & E**) For both models, every visual field map exhibited a significant positive Spearman correlation between pairwise cortical surface distances and pairwise representational distances. This spatial coupling was strongest in early visual field maps and generally decreased through the hierarchy. (**D & F**) Despite global alignment of representational spaces, no visual field map showed a significant correlation between preferred representational positions and visual field eccentricity preferences. Black stars indicate median correlation coefficients significantly above zero; brackets denote significant differences between map distributions with markers to the left and right of the stars. White lines represent the mean across hemispheres.

Visual field maps also exhibit gradual shifts in spatial position preferences across the cortical surface. We therefore tested whether the progression of preferred representational space positions followed the progression of visual field eccentricity preferences. Despite aligning representational dimensions across all maps and participants, no visual field map showed a consistent correlation between preferred eccentricity and representational position in either dimension (Fig. 5D & F, Tables S13–S16).

Whole-Brain Application of Between-ROI Models

While our previous analyses focused on visual field map ROIs, the Between-ROI framework allows for characterizing recording sites anywhere in the brain by sampling from a distant source representation. As an exploratory analysis, we evaluated RTM performance across the entire cortical surface using visual field map TO-1 as the source ROI, as it proved to be the best fitting source for multiple higher-level visual field maps.

This whole-brain approach revealed that RTMs based on TO-1's two-dimensional representational space predicted responses across an extensive band of the temporal and parietal lobes (Fig. 6B), largely anterior to the traditionally defined

visual field maps (Fig. 6A). Good fits were also observed in the superior temporal sulcus, medial frontal cortex, and the lateral sulcus, including Heschl's gyri. In these regions, Model Performance remained high (Fig. 6C), reaching levels comparable to those found in higher-level visual field maps. Tuning function extents in these areas were consistently small (Fig. 6D), and preferred representational tuning function positions (particularly the y-coordinate) exhibited gradual spatial variations across the cortical surface (Fig. 6E-F), mirroring the results observed in the visual hierarchy. These results suggest that Between-ROI RTMs can capture advanced stages of visual processing by modeling neural responses as functions of high-level representations rather than low-level image properties.

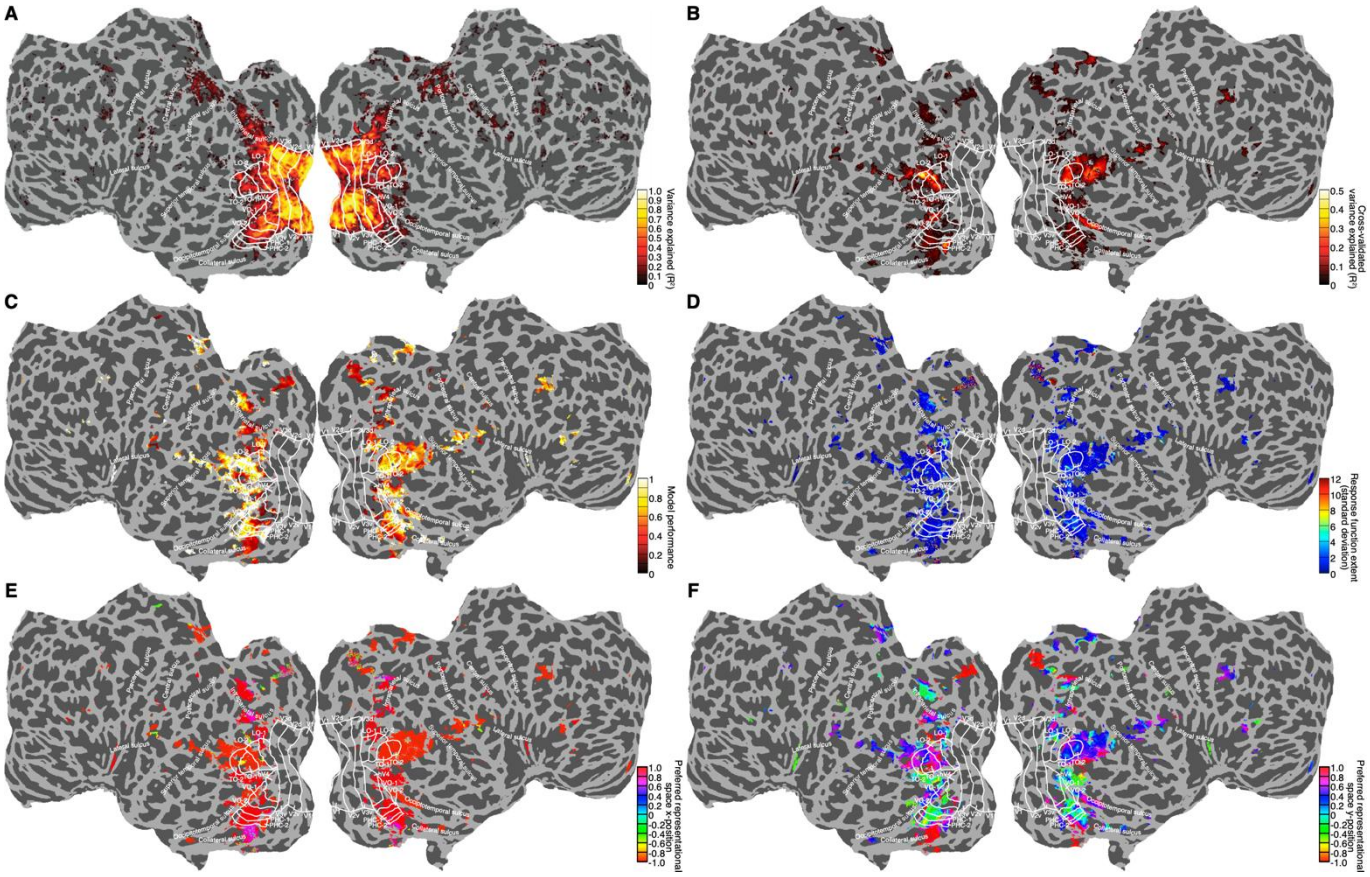

Figure 6: Between-ROI RTMs captured responses beyond visual field maps. (**A**) Visual spatial pRF models captured responses only in the visual field maps of the occipital and superior parietal lobes. (**B**) In contrast, an RTM using TO-1 as the source ROI captured responses in an extensive band through the parietal and temporal lobes, the superior temporal sulcus, medial frontal lobe, and lateral sulcus. **(C)** Model performance remained high across these regions. **(D)** Tuning function extents were consistently small in these areas. **(E-F)** Preferred representational positions, particularly in the y-dimension, exhibited gradual progressions across the cortical surface of responsive regions.

Discussion

In this study, we introduced and evaluated Representational Tuning Models (RTMs), a novel framework that bridges population-level similarity analysis and individual-site encoding models. By combining RSA-based dimensionality reduction with parametric encoding models, RTMs derive continuous tuning functions directly from high-dimensional or arbitrary stimuli, eliminating the need for predefined stimulus parameters. Our results demonstrate that RTMs, particularly the Between-ROI variant, significantly predict independent test-set responses in mid-level visual maps, revealing that neural populations here preferentially respond to a sub-set of images in a continuous region of the representational space. RTMs reveal systematic spatial organization of preferred representational positions and increasing selectivity through the visual hierarchy. The superior performance of Between-ROI models suggests that neural populations preferentially sample representations from other regions rather than representing part of the representational space of their local area, thereby transforming representations between regions. Critically, we show RTMs transcend anatomical boundaries and visual field maps, successfully modeling advanced image processing stages throughout the brain. This generalizability

positions RTMs as a scalable, powerful tool for mapping neural tuning and functional connectivity across diverse stimulus domains and modalities, offering a unified path to understanding how the brain transforms information.

Relationships to Visual Position Tuning

RTMs are designed for arbitrary stimuli, but our use of natural images, which contain specific spatial contrasts, and visual field map ROIs might imply that RTMs simply reflect visual field position tuning. Several lines of evidence suggest otherwise. First, unlike Gabor wavelet models that predict responses based on local contrasts throughout the image[27], RTMs fit all voxels within a representational space where each image occupies only one position. Second, while spatial pRF models perform best in early visual areas, RTMs show no significant fits in early visual areas, with superior fits in later visual field maps (Fig. 3) and even areas beyond the visual field maps, where spatial models fail (Fig. 6). Third, while pRF sizes increase through the visual hierarchy, RTM tuning function extents decrease, signaling increased selectivity for specific image subsets. Finally, we found no consistent relationship between representational and eccentricity preferences (Supplementary Tables 8–11). Although early visual cortex shows the strongest correlations between representational distance and cortical surface distance (Fig. 5), the poor cross-validated RTM fit in these areas suggests that RSA here primarily reflects cortical anatomy rather than repeatable, structured responses to stimuli. In contrast, the high performance and lower anatomical correlation in later maps indicate that RTMs capture intrinsic geometry of stimulus representations here.

Applicability to Different Dimensionalities, Response Functions and Modalities

We employed two-dimensional MDS and circular Gaussian functions due to their common use, interpretability and alignment with cortical surface topology[1,23]. However, the RTM framework is dimension-agnostic. Scaling to a single dimension would characterize the first principal component of a representation, while higher-dimensional scaling could capture more complex tuning profiles.

Similarly, the choice of a symmetrical Gaussian is a starting point; the framework can accommodate more complex functions, such as anisotropic elliptical[28,29], power-law exponents[30], difference-of-Gaussian functions with inhibitory surrounds[31] or divisive normalization[32]. Comparing the predictive accuracy of these varied functional forms, particularly in Between-ROI models, may reveal the specific computations applied as information is transformed between brain regions[33].

The RTM framework is compatible with any recording modality providing both population-level response patterns and spatially specific individual sites, including fMRI, ECoG, high-density electrophysiology (single- or multi-unit), and artificial neural networks. In datasets featuring multiple brain regions or network layers, Between-ROI models can characterize representational transformations across the hierarchy.

Crucially, RTMs facilitate cross-modal integration: representational spaces derived from one modality (e.g., neural network activations) can be used to predict individual responses in another (e.g., fMRI voxels). While RTMs are widely applicable to existing high-resolution datasets, they are not suited for modalities that lack fine-grained individual recording sites, such as EEG or behavioral dissimilarity matrices, where traditional RSA remains effective.

Properties of Between-ROI Models:

Between-ROI models characterize a target voxel's response as a sampling of a distant source ROI's representational space. This has three major implications. First, unlike Within-ROI models, Between-ROI RTMs can be applied brain-wide. We demonstrated that using TO-1 as a source ROI captures responses in parietal and temporal regions that are invisible to spatial pRF models (Fig. 6), providing a method to characterize high-level visual transformations. Second, the superior performance of Between-ROI models over Within-ROI models in later visual field maps suggests that voxel responses are better understood as samplings of the representational space that provides the voxel's inputs rather than as a sub-set of the target ROI's representational space. Third, while we selected a single best-fitting source ROI, cortical

activity at any location inherently integrates inputs from many source areas[34]. Future iterations could employ weighted sums of multiple source spaces, potentially integrated via Bayesian network models[35] to map complex functional connectivity.

Interpretation of Representational Tuning Parameters

Because RTMs derive dimensions from neural patterns rather than physical stimulus properties (e.g., visual position, frequency or numerosity), the resulting dimensions lack immediate semantic labels. While RSA on large ROIs can differentiate broad categories like faces, places or animacy[36] (Fig. 1A-B), our use of smaller visual field maps and complex natural images makes such interpretations difficult. However, RTMs provide a quantitative way to interpret these dimensions by testing for correlations with independent parameters like cortical distance or eccentricity. Methods to interpret dimensions of representational spaces[37] provide the same interpretation to RTM dimensions.

A representation tuning function's extent quantifies response variation across the stimulus set. In our data, standard deviations were typically above two, implying that all images elicited some response as expected for visually responsive areas. In regions with higher categorical selectivity (where many stimuli elicit no response), we predict extents below one. We observed an inverse relationship between tuning function extent and model performance: smaller tuning functions predict greater variance in response amplitudes, allowing the model to explain more variance above the noise floor.

While RSA identifies the geometry of a representation, the underlying neural encoding (whether it arises from tuned response functions or other mechanisms) remains ambiguous[38]. RTMs resolve this by explicitly testing for tuned responses within the representational space. Similarly, although representational similarity can emerge from gradual topographic shifts in tuning across the cortical surface, RSA cannot isolate this spatial component from other response structure. RTMs provide a direct test for topographic representational organization. Our findings indicate that representational similarity heavily reflects the spatial structure of neural responses, particularly in brain regions or conditions lacking stable, stimulus-driven response patterns.

Conclusions

Representational tuning models unify encoding models and RSA into a general method for characterizing individual recording sites within arbitrary representational spaces. This approach parameterizes neural responses to naturalistic, high-dimensional stimuli that defy traditional encoding models. Beyond fMRI, RTMs are applicable to any measure where individual recordings are analyzed as parts of larger patterns. By extending parametric modeling into higher-level processing through Between-ROI sampling, RTMs provide a powerful tool for understanding how the brain transforms information across the cortical hierarchy, even beyond the reach of stimulus-referred encoding models.

Methods

Dataset, Participants and Preprocessing

Analyses were performed using the Natural Scenes Dataset (NSD), an open-access 7T fMRI dataset[26]. Eight participants (two males and six females; age range, 19–32 years) are included in that data set. All participants had normal or corrected-to-normal visual acuity, with no known cognitive deficits or color blindness, and were right handed. The data set used no statistical methods to pre-determine the sample size: the experimental approach emphasizes collecting extensive data from each participant, which enables the demonstration and replication of effects in individual participants. The data set provides additional participant information including height, weight, handedness and visual acuity, which we did not use.

Informed written consent was obtained from all participants, and the experimental protocol (including public data sharing) was approved by the University of Minnesota institutional review board.

Each participant viewed between 9,000 and (up to) 10,000 natural scenes (COCO dataset[39]) across 30–40 scanning sessions, with most images presented three times in an event-related design. A common set of 1,000 images was viewed by all participants. Images were presented using a 3-s ON/1-s OFF trial structure. Participants were instructed to fixate the central dot and press a button with their right index finger if the presented image had never been presented before, and a different button using their right middle finger if the image had even been presented before. Analyses of behavioral responses are reported in the NSD publication[26].

Functional data collection used gradient-echo EPI at 1.8-mm isotropic resolution with whole-brain coverage. 84 axial slices were collected, each with a matrix size of 120 × 120, giving field-of-view 216 mm (FE) × 216 mm (PE). The phase encode direction was anterior-to-posterior. TR was 1,600 ms, TE was 22.0 ms, flip angle was 62°. Echo spacing was 0.66 ms, bandwidth was 1,736 Hz per pixel, partial Fourier 7/8, iPAT 2, multi-band slice acceleration factor 3. The NSD dataset used custom preprocessing methods that avoided spatial or temporal smoothing[26]. Data analyses were on an individual participant basis maintain the spatial structure of high-resolution activation patterns.

Single-trial response amplitudes (betas) were estimated via general linear modeling[40]. We utilized the *nativesurface* beta version, resampled to each participant's 1 mm resolution cortical surface model and averaged across cortical depth onto the cortical surface ribbon. Betas were z-scored within each session to minimize inter-session variance. Data from subjects 6 and 8 contained missing values (*NaN*), leading to the exclusion of affected voxels (2.9% and 0.22% of total voxels, respectively).

Cross-Validation Framework

We implemented a leave-one-out cross-validation scheme. Images with only a single presentation were discarded. For images with two or three presentations, the beta from the final presentation was assigned to the test set. The training set consisted of either the first presentation beta (for images with two presentations) or the mean of the first two presentations (for images with three presentations), resulting in up to 10,000 training conditions and 10,000 test conditions per participants.

ROI Definition

Using the NSD's visual field mapping data and pRF models, we extended the provided V1, V2, and V3 definitions to include hV4, VO-1/2, PHC-1/2, LO-1/2, and TO-1/2. ROI boundaries were identified based on reversals in cortical polar angle progressions.

Construction of 2D Representational Spaces

To generate the latent space for representational tuning, we first used the training set to calculate a Representational Dissimilarity Matrix (RDM) for each ROI in each participant. Dissimilarity between condition pairs (i, j) was defined as 1-r, where r is the Pearson correlation between multi-voxel activity patterns.

The resulting RDM underwent classical Multidimensional Scaling (MDS) (i.e. principle coordinate analysis or Torgerson scaling) using De Leeuw's algorithm[41]. This projected the dissimilarities into a two-dimensional representational space where Euclidean distances between points approximated the original pattern dissimilarities. This space was centered at (0,0) with a radius of 1 (maximum pairwise distance of 2). While we utilized 2D spaces for interpretability and

visualization, the framework is extensible to fewer or more dimensions.

Global Alignment of Representational Spaces

Because MDS-derived spaces are rotation-invariant, their orientations are initially arbitrary, complicating inter-regional and inter-subject comparisons. We aligned all spaces to a common reference, Subject 1's VO-1 space, using a multi-step rotation procedure:

1. Intra-subject Alignment: For each participant, all ROI-specific spaces were rotated to match their local VO-1 space using the Kabsch optimal rotation algorithm[42] to minimize distance between shared condition positions.
2. Inter-subject Alignment: Each participant's VO-1 space was then rotated to match Subject 1's VO-1 space using the positions of the 1,000 shared images.
3. Transformation Application: These calculated rotations were applied to all ROIs within the corresponding participant.

This ensured that representational tuning functions from all visual field maps and all subjects were positioned within a consistent, shared coordinate system. The arbitrary orientation of this common reference was addressed at a later stage.

Fitting Representational Tuning Functions

We modeled individual voxel responses as a function of condition positions within the 2D representational space. We compared two approaches:

1. Within-ROI: Voxels were fit using the representational space derived from their own ROI.
2. Between-ROI: Voxels were fit using the representational space of a different "source" ROI to characterize how populations sample representations from distant regions.

In both models, for each voxel we fit the response amplitudes for the different conditions a circular symmetric Gaussian (Equation 1) defined by three parameters: preferred position (x_0, y_0) and tuning function extent (σ). The slope (A) and intercept (k) of each tuning function were linearly solved. We used scikit-learn's curve_fit to minimize the sum of squared residuals between predicted amplitudes and measured training-set betas.

$$g(v) = A \cdot e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}} + k$$

(Equation 1)

In principle, any form of parametric representational tuning function could be fit to these positions and amplitudes, in a representational space with any number of dimensions. If the neural population in each voxel responds with different amplitudes to different conditions, and conditions yielding similar response amplitudes are closer in the representational space, the distribution of response amplitudes across conditions may be summarized by a continuous parametric function in this representational space.

Constraints and Initialization

To ensure biological interpretability and prevent overfitting, we applied specific bounds (Table 1):

- Position: Centers were restricted to [-1.05 1.05] to keep preferred positions within or near the unit-radius representational space. Centers far outside the representational space predict monotonic (rather than tuned) changes in response amplitude across the representational space, and these predictions differ little from representational tuning functions with centers just outside the representational space.
- Extent: was lower-bounded at 0.01 to prevent single-condition overfitting. While we used no upper bound on extent,

very large extents predict little variance across the representational tuning function so fit poorly.

- Amplitude: the tuning function was restricted to positive slope values to ensure interpretability.

Parameters were randomly initialized within narrower ranges (Table 2) to begin the search near the space's center and using amplitudes consistent with typical ranges of z-scored fMRI response amplitudes.

	Lower Bound	Upper Bound
x_0	-1.05	1.05
y_0	-1.05	1.05
σ	0.01	inf
A	10^{-08}	inf
k	— inf	inf

Table 1: Lower and upper bounds of our Gaussian function parameters.

	Lower Bound	Upper Bound
x_0	-0.4	0.4
y_0	-0.4	0.4
σ	0.01	2
A	0.1	10
k	-2	2

Table 2: Lower and upper bounds of the starting values of our parameters, which were randomly initialized withing these ranges

Evaluation of Model Fits

Model accuracy was quantified using cross-validated variance explained. We used the parameters optimized on the training set to give predicted response amplitudes (\widehat{y_i}) which were compared to measured response amplitudes in the test set (y_i) following Equation 2:

$$\text{Variance Explained} = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

(Equation 2)

To account for low fMRI SNR and single-trial test-set variability, we normalized variance explained by the noise ceiling—a model-free measure of data repeatability. First, to account for response amplitude differences between scanning sessions[43], we linearly rescaled training-set betas to best fit test-set betas as follows:

$$\hat{y}_{train,i} = a \cdot \beta_{train,i} + b$$

(Equation 3)

Where (\beta_{train, i}) is the original training set response amplitude for condition *i*, and *a* and *b* are constants found via linear regression that minimize the difference between the training and test set amplitudes.

Once rescaled, the training data ({\hat{y}}_{train, i}) is used as the "perfect" model to predict the test data (y_i), giving the noise ceiling as follows:

$$\text{Noise Ceiling} = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_{\text{train},i}\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

(Equation 4)

We calculated Model Performance by normalizing variance explained by this noise ceiling as follows:

$$\text{Model Performance} = \frac{\text{Variance Explained}}{\text{Noise Ceiling}}$$

(Equation 5)

Source ROI Selection for Between-ROI Models

For the Between-ROI analysis, we first fit representational tuning functions for every voxel using each of the other 11 visual field maps as potential source ROIs. To allow for a unified comparison across the hierarchy, we then selected a single "best" source ROI for each target region. V1 was the only exception, where it could serve as its own source to account for feedforward inputs from the thalamus not being modelled.

For all other visual field maps, we identified the optimal source by finding the visual field map whose representational space yielded the highest mean variance explained (R^2) within the training set, averaged across all voxels and hemispheres. By selecting the source ROI based solely on training data, we ensured that the subsequent comparison of cross-validated variance explained between Within-ROI and Between-ROI models remained statistically independent and non-circular.

Optimal Cortical Alignment

Since two-dimensional representational spaces are rotation-invariant, we identified a global rotation angle to align their x-axis with the predominant direction of cortical change. In each ROI in each hemsiphere, we calculated the vector difference between representational positions for all voxel pairs and divided by their cortical surface distance. We ran a Bayesian circular analysis[44], with a conjugate prior of $p(\mu, \kappa) \propto I_0(\kappa)^{-1}$ where $\mu \in [0, 2\pi)$ is the mean direction, $\kappa \in R^+$ is a concentration parameter and $I_0(\cdot)$ is the modified Bessel finciton of the first kind and order zero. This confirmed these directions followed a non-uniform von Mises distribution. All preferred positions were then rotated by the circular mean across all ROIs and hemispheres (Within-ROI: 86.59°. Between-ROI 88.14°) for standardized visualization.

Statistics and Comparisons

Goodness of Fit: We used Wilcoxon signed-rank tests (FDR-corrected) to compare mean variance explained and Model Performance against zero and to compare Within-ROI vs. Between-ROI performance. Two-factor ANOVAs (ROI, Participant) tested for differences across visual field maps.

Tuning Extents: Mean standard deviation values were compared across ROIs using two-factor ANOVAs. Differences between models were assessed via Wilcoxon signed-rank tests (FDR-corrected) and compared across ROIs using two-factor ANOVAs.

Topographic Organization: We calculated Spearman correlations (ρ) between pairwise representational space distances and cortical surface distances. However, there were far more pairwise distances than voxels, because each voxel is paired with each

other voxel and the voxels in the cortical surface model (1x1x1 mm) were smaller than the acquired voxels (1.8x1.8x1.8 mm). To correct the number of independent measures in our correlation, we used a downscaling factor: the square root of the number of pairs, divided by 1.8^3. The sample size for our Spearman correlation becomes an approximation of the original number of recorded voxels based on that scaling factor. Then, for 1000 permutations in each ROI, we randomly selected this downsampled number of voxel pairs, computed the Spearman correlation between their representational space and cortical surface distances, and took the median correlation coefficient and p values across these permutations. We did this in the two hemispheres separately to avoid considering cortical surface distances between voxels in the left and right hemispheres. Moreover, in V2 and V3 we separated dorsal and ventral quadrant maps (V2d, V2v, V3d and V3v) to prevent arbitrarily large distances resulting from the anatomical separation of these regions.

Visual Position Relationships: We correlated preferred representational space positions (x_0, y_0) with pRF position eccentricity using Spearman's correlations, with significance assessed by Wilcoxon tests on the distribution of correlation coefficients (ρ) across hemispheres.

Algorithmic Appendix:

To facilitate the reproducibility and platform-independent adoption of Representational Tuning Models (RTMs), we provide a standalone software toolbox: https://github.com/StanBgy/representational_tuning. However, the RTM architecture is modular and language-agnostic. The core analytical pipeline can be implemented independently of our specific codebase by executing the data transformations, optimization constraints, and cross-validation loops detailed in the algorithmic blueprint below.

Input:

- Beta matrices B_s for subjects s ∈ {1...8} and ROIs r ∈ {1...12}
- Test/Train condition splitting indices
- Target voxels V_target

Output:

- Within-ROI and Between-ROI Model Variance Explained and Performance (Normalized variance explained)
- Voxel Tuning Parameters (x0, y0, σ)
- Cortical Topography Alignment Coordinates

STAGE 1: PREPROCESSING & CROSS-VALIDATION SPLITTING

For each subject s:

  For each session:

    Z-score betas across condition presentations to minimize session variance

  Identify condition presentation counts:

    Discard single-presentation images

    For images with 2 presentations:

      Assign presentation 1 to Training Set (B_train)

      Assign presentation 2 to Test Set (B_test)

    For images with 3 presentations:

      Assign Mean(presentation 1, presentation 2) to Training Set (B_train)

      Assign presentation 3 to Test Set (B_test)

  Exclude any voxels containing NaN values (e.g., in Subjects 6 and 8)

STAGE 2: LATENT REPRESENTATIONAL SPACE CONSTRUCTION

For each ROI r in each subject s:
  Compute Representational Dissimilarity Matrix: RDM_r = 1 - Pearson_r(B_train)
  Apply Classical Multidimensional Scaling (MDS) using De Leeuw's algorithm:
    Project RDM_r → 2D coordinate space X_r, Y_r (radius = 1, center = 0,0)

Align all representational spaces globally to a common reference:
  For each subject s:
    Rotate all ROI spaces to match local VO-1 space via Kabsch algorithm
  Rotate each subject's VO-1 space to match Subject 1's VO-1 space via shared images
  Apply calculated rotation matrices across all internal ROI spaces

STAGE 3: TUNING FUNCTION FITTING (Within-ROI & Between-ROI)
Define Target Function (here, Circular Symmetric Gaussian):
  g(x, y) = m * exp(-((x - x0)^2 + (y - y0)^2) / (2 * σ^2)) + k

For each Target Voxel v in V_target:
  If Model Type == Within-ROI:
    Use source coordinates (X_source, Y_source) from voxel's own ROI space
  If Model Type == Between-ROI:
    Select Source ROI maximizing mean training variance explained across hemispheres
    Use source coordinates (X_source, Y_source) from selected extrinsic ROI space

  Initialize parameters randomly within Table 2 bounds
  Fit parameters (x0, y0, σ, m, k) using non-linear least squares optimization:
    Minimize: Σ (B_train_v - g(X_source, Y_source))^2
    Subject to constraints: x0,y0 ∈ [-1.05, 1.05], σ ≥ 0.01, m > 0

STAGE 4: EVALUATION AND PERFORMANCE NORMALIZATION
For each Target Voxel v:
  Generate predictions using optimized parameters: y_pred = g(X_test, Y_test)
  Calculate Cross-Validated Variance Explained (CVVE):
    CVVE = 1 - [Σ(B_test_v - y_pred)^2 / Σ(B_test_v - Mean(B_test_v))^2]

  Calculate Noise Ceiling:
    Fit Ordinary Least Squares linear regression: B_train_v → B_test_v
    Rescale training data: B_train_rescaled = a * B_train_v + b
    Noise_Ceiling = 1 - [Σ(B_test_v - B_train_rescaled)^2 / Variance(B_test_v)]

  Calculate Model Performance:
    Model_Performance = CVVE / Noise_Ceiling

STAGE 5: SPATIAL COUPLING AND CORTICAL PROJECTON
For each ROI r and each hemisphere:
  Identify all unique voxel pairs (i, j)
  Calculate D_cortex(i, j) = Geodesic distance on the cortical surface ribbon

Calculate D_RTM(i, j) = Euclidean distance between fit centers (x0_i, y0_i) and (x0_j, y0_j)

Apply spatial autocorrelation downscaling factor:

N_samples = Sqrt(Total_Pairs) / 1.8

Execute 1,000 permutations:

Randomly sub-sample N_samples from the unique voxel pairs

Compute Spearman_rho between D_cortex and D_RTM for the sub-sample

Store median Spearman_rho across permutations as the ROI's topographic coupling strength

Acknowledgements:

This work is supported by the European Research Council under the European Union's Horizon 2020 Research and Innovation Program (Starting Grant number 802553, "ContentMAP" to JA), and the European Research Executive Agency Widening Programme under the European Union's Horizon Europe Grant (101087584, "CogBooster" to JA), by Netherlands Organization for Scientific Research (#452.17.012 to BH), by the Portuguese Foundation for Science and Technology (#IF/01405/2014 to BH).

Data and code availability:

The original natural scenes dataset is available for free upon request from: https://forms.gle/eT4jHxaWwYUDEf2i9.

The visual training set and test set response amplitudes within the visual field maps, two-dimension representational coordinates, the model outputs and the files needed for the cortical projections are publicly available at: https://doi.org/10.6084/m9.figshare.32125993

All code is publicly available at: https://github.com/StanBgy/representational_tuning

Author Contributions:

S.B., L.M.-R.C. and B.M.H. contributed to the **Methodology** of the study, conducted the **Investigation** and were responsible for **Formal Analysis**. S.B. and L.M.-R.C. were responsible for **Data Curation**, developed specialized **Software** and generated the data **Visualization**. B.M.H. and J.A. secured the **Funding Acquisition**. B.M.H. provided **Supervision** and managed **Project Administration**. S.B. performed the **Validation** of the experimental results. S.B. and B.M.H. drafted the manuscript (**Writing – Original Draft**), and all authors contributed to the **Conceptualization** and **Writing – Review & Editing** and approved the final version.

References

1. Dumoulin, S. O. & Wandell, B. A. Population receptive field estimates in human visual cortex. *NeuroImage* **39**, 647–660 (2008).
2. Thomas, J. M. *et al.* Population receptive field estimates of human auditory cortex. *NeuroImage* **105**, 428–439 (2015).
3. Schellekens, W. *et al.* A touch of hierarchy: population receptive fields reveal fingertip integration in Brodmann areas in human primary somatosensory cortex. *Brain Struct. Funct.* **226**, 2099–2112 (2021).
4. Himmelberg, M. M., Winawer, J. & Carrasco, M. Linking individual differences in human primary visual cortex to contrast sensitivity around the visual field. *Nat. Commun.* **13**, 3309 (2022).
5. Harvey, B. M. & Dumoulin, S. O. The Relationship between Cortical Magnification Factor and Population Receptive Field Size in Human Visual Cortex: Constancies in Cortical Architecture. *J. Neurosci.* **31**, 13604–13612 (2011).
6. Silva, M. F. *et al.* Radial asymmetries in population receptive field size and cortical magnification factor in early visual

cortex. *NeuroImage* **167**, 41–52 (2018).
7. Puckett, A. M., Bollmann, S., Junday, K., Barth, M. & Cunnington, R. Bayesian population receptive field modeling in human somatosensory cortex. *NeuroImage* **208**, 116465 (2020).
8. Klein, B. P., Harvey, B. M. & Dumoulin, S. O. Attraction of Position Preference by Spatial Attention throughout Human Visual Cortex. *Neuron* **84**, 227–237 (2014).
9. Gomez, J., Natu, V., Jeska, B., Barnett, M. & Grill-Spector, K. Development differentially sculpts receptive fields across early and high-level human visual cortex. *Nat. Commun.* **9**, 788 (2018).
10. Elul, D. & Levin, N. The Role of Population Receptive Field Sizes in Higher-Order Visual Dysfunction. *Curr. Neurol. Neurosci. Rep.* **24**, 611–620 (2024).
11. Harvey, B. M., Klein, B. P., Petridou, N. & Dumoulin, S. O. Topographic Representation of Numerosity in the Human Parietal Cortex. *Science* **341**, 1123–1126 (2013).
12. Harvey, B. M., Fracasso, A., Petridou, N. & Dumoulin, S. O. Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 13525–13530 (2015).
13. Harvey, B. M., Dumoulin, S. O., Fracasso, A. & Paul, J. M. A Network of Topographic Maps in Human Association Cortex Hierarchically Transforms Visual Timing-Selective Responses. *Curr. Biol. CB* **30**, 1424-1434.e6 (2020).
14. Cai, Y. *et al.* Topographic numerosity maps cover subitizing and estimation ranges. *Nat. Commun.* **12**, 3374 (2021).
15. Cai, Y., Hofstetter, S., Harvey, B. M. & Dumoulin, S. O. Attention drives human numerosity-selective responses. *Cell Rep.* **39**, 111005 (2022).
16. Hendrikx, E., Paul, J. M., van Ackooij, M., van der Stoep, N. & Harvey, B. M. Visual timing-tuned responses in human association cortices and response dynamics in early visual cortex. *Nat. Commun.* **13**, 3952 (2022).
17. Hofstetter, S., Cai, Y., Harvey, B. M. & Dumoulin, S. O. Topographic maps representing haptic numerosity reveals distinct sensory representations in supramodal networks. *Nat. Commun.* **12**, 221 (2021).
18. Paul, J. M., van Ackooij, M., ten Cate, T. C. & Harvey, B. M. Numerosity tuning in human association cortices and local image contrast representations in early visual cortex. *Nat. Commun.* **13**, 1340 (2022).
19. Tsouli, A. *et al.* Adaptation to visual numerosity changes neural numerosity selectivity. *NeuroImage* **229**, 117794 (2021).
20. Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, (2008).
21. Diedrichsen, J. & Kriegeskorte, N. Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis. *PLOS Comput. Biol.* **13**, e1005508 (2017).
22. Khaligh-Razavi, S.-M. & Kriegeskorte, N. Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLOS Comput. Biol.* **10**, e1003915 (2014).
23. Douglas Carroll, J. & Arabie, P. Chapter 3 - Multidimensional Scaling. in *Measurement, Judgment and Decision Making* (ed. Birnbaum, M. H.) 179–250 (Academic Press, San Diego, 1998). doi:10.1016/B978-012099975-0.50005-1.
24. Haak, K. V. *et al.* Connective field modeling. *NeuroImage* **66**, 376–384 (2013).
25. Charest, I., Kievit, R. A., Schmitz, T. W., Deca, D. & Kriegeskorte, N. Unique semantic space in the brain of each beholder predicts perceived similarity. *Proc. Natl. Acad. Sci.* **111**, 14565–14570 (2014).
26. Allen, E. J. *et al.* A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat. Neurosci.* **25**, 116–126 (2022).
27. Kay, K. N., Naselaris, T., Prenger, R. J. & Gallant, J. L. Identifying natural images from human brain activity. *Nature* **452**, 352–355 (2008).
28. Lerma-Usabiaga, G., Winawer, J. & Wandell, B. A. Population Receptive Field Shapes in Early Visual Cortex Are Nearly Circular. *J. Neurosci.* **41**, 2420–2427 (2021).
29. Silson, E. H., Reynolds, R. C., Kravitz, D. J. & Baker, C. I. Differential Sampling of Visual Space in Ventral and Dorsal Early Visual Cortex. *J. Neurosci.* **38**, 2294–2303 (2018).
30. Kay, K. N., Winawer, J., Mezer, A. & Wandell, B. A. Compressive spatial summation in human visual cortex. *J. Neurophysiol.* **110**, 481–494 (2013).
31. Zuiderbaan, W., Harvey, B. M. & Dumoulin, S. O. Modeling center-surround configurations in population receptive fields using fMRI. *J. Vis.* **12**, 10 (2012).
32. Aqil, M., Knapen, T. & Dumoulin, S. O. Divisive normalization unifies disparate response signatures throughout the human visual hierarchy. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2108713118 (2021).
33. Invernizzi, A., Haak, K. V., Carvalho, J. C., Renken, R. J. & Cornelissen, F. W. Bayesian connective field modeling using a Markov Chain Monte Carlo approach. *NeuroImage* **264**, 119688 (2022).
34. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex N. Y. N 1991* **1**, 1–47 (1991).
35. Rajapakse, J. C. & Zhou, J. Learning effective brain connectivity with dynamic Bayesian networks. *NeuroImage* **37**, 749–

760 (2007).
36. Mur, M. *et al.* Categorical, Yet Graded – Single-Image Activation Profiles of Human Category-Selective Cortical Regions. *J. Neurosci.* **32**, 8649–8662 (2012).
37. Almeida, J. *et al.* Neural and behavioral signatures of the multidimensionality of manipulable object processing. *Commun. Biol.* **6**, 940 (2023).
38. Kriegeskorte, N. & Wei, X.-X. Neural tuning and representational geometry. *Nat. Rev. Neurosci.* **22**, 703–718 (2021).
39. Lin, T.-Y. *et al.* Microsoft COCO: Common Objects in Context. in *Computer Vision – ECCV 2014* (eds Fleet, D., Pajdla, T., Schiele, B. & Tuytelaars, T.) 740–755 (Springer International Publishing, Cham, 2014). doi:10.1007/978-3-319-10602-1_48.
40. Prince, J. S. *et al.* Improving the accuracy of single-trial fMRI response estimates using GLMsingle. *eLife* **11**, e77599 (2022).
41. de Leeuw, J. Applications of Convex Analysis to Multidimensional Scaling. https://escholarship.org/uc/item/7wg0k7xq (2005).
42. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A* **32**, 922–923 (1976).
43. Roth, Z. N. & Merriam, E. P. Representations in human primary visual cortex drift over time. *Nat. Commun.* **14**, 4422 (2023).
44. Mulder, K. T. & Klugkist, I. Bayesian tests for circular uniformity. *J. Stat. Plan. Inference* **211**, 315–325 (2021).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.pdf