# Orthogonal Neural Representations Support Perceptual Judgments of Natural Stimuli

**Authors:** Ramanujan Srinath[1,2], Amy M. Ni[1,2,3], Claire Marucci[3], Marlene R. Cohen[2,4], David H. Brainard[3,4]

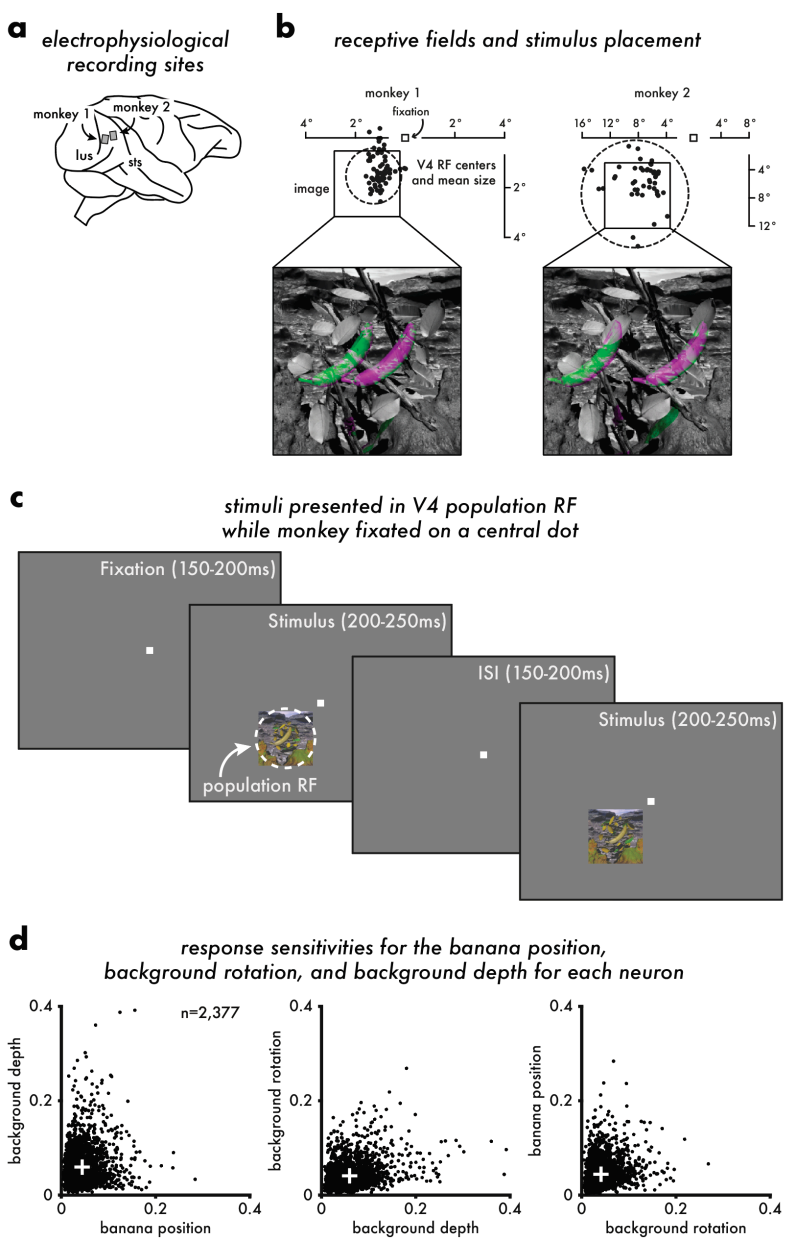
[1] = equal contribution

[2] = Department of Neurobiology and Neuroscience Institute, The University of Chicago, Chicago, IL 60637, USA

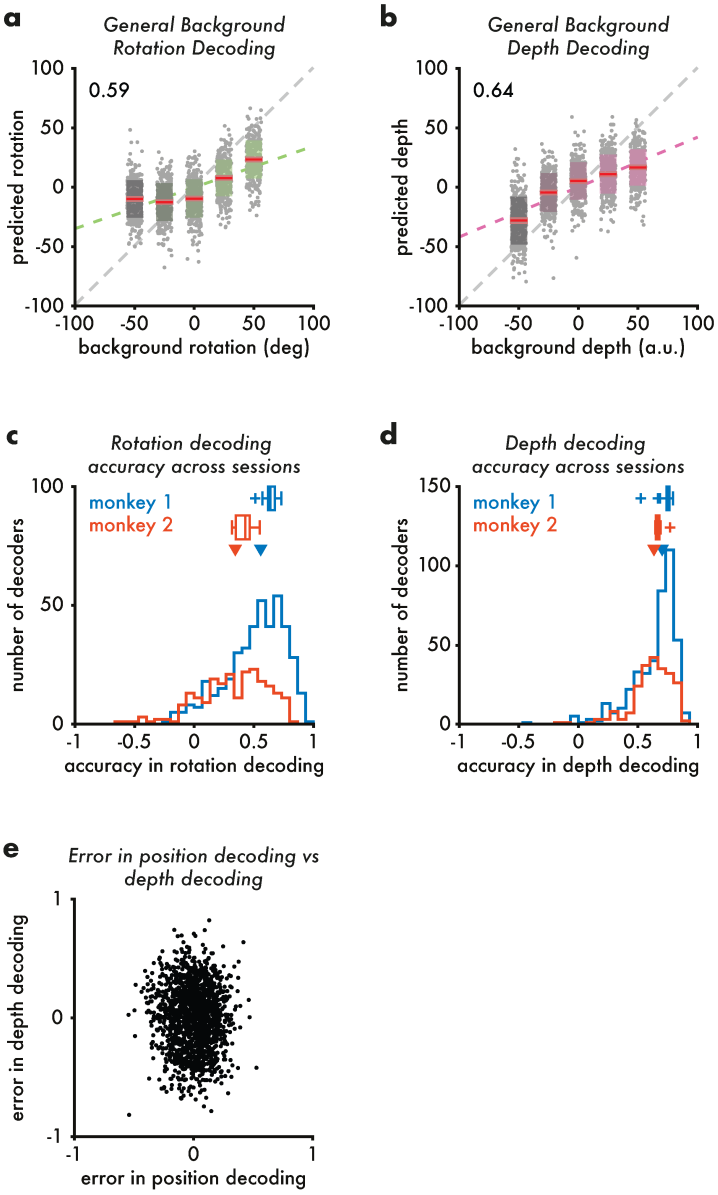
[3] = Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA

[4] = equal contribution

# Supplementary Figures




**Supp. Fig. 1: Recording locations, stimulus positions, behavioral task, and unit responsivity for monkeys.**
**a:** Multielectrode arrays (96 channels each) were chronically implanted in V4 of two monkeys.
**b:** The receptive fields of the recorded multiunits were mapped using Gabor and two-dimensional shape stimuli. The estimated receptive field (RF) center of each visually responsive multiunit is depicted by a black point. The dashed circle depicts the mean size of the RF for each population. The black square indicates the position of the image on the screen. The size and position of the images for each session were chosen so that all the variations in stimulus position and background depth were within the estimated RF of the recorded V4 population. The false color image in each panel indicates the variation in object position. **c:** While monkeys fixated on a central dot, stimuli were flashed on (200-250 ms) and off (150-200 ms) up to eight times before the monkey received a juice reward. Each of the 125 images was repeated between 8-10 times in every session. Electrophysiological recordings were collected using multielectrode arrays implanted in V4. Receptive fields (RFs) were mapped in an independent experiment using Gabor and 2D shape stimuli. Images were placed such that the variation in the central object (the banana) and background overlapped a large majority of RFs. **d:** Comparison of response sensitivity for object position, background depth, and background rotation. Here, sensitivity is defined as the modulation index $(r_{max}-r_{min})/(r_{max}+r_{min})$ where $r_{max}$ and $r_{min}$ are, respectively, the maximum and minimum responses to variation in the corresponding parameter. The white cross represents the mean of the sensitivities across all visually responsive multi-units (n=2377) across all sessions.




**Supp. Fig. 2: Background depth and rotation can also be decoded well from V4 population responses.** Compare with Fig. 2.

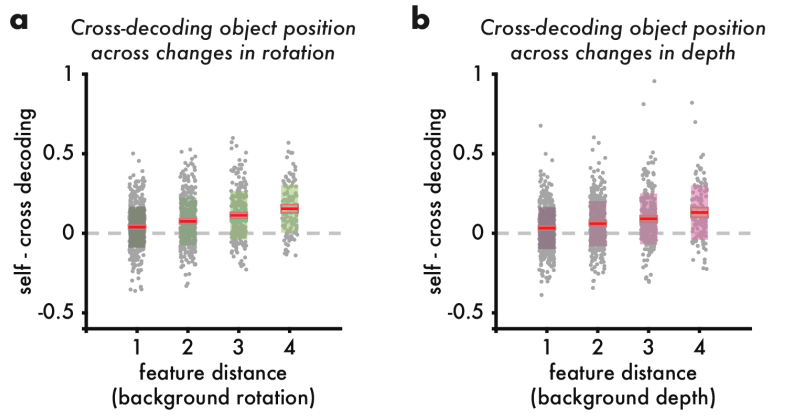
**a:** Same as Fig. 2b but for background rotation. Here, a general linear decoder was used to estimate the background rotation, ignoring variation in background depth and object position (vs constant model $p<0.001$).

**b:** Same as Fig. 2b but for background depth. Here, a general linear decoder was used to estimate the background depth, ignoring variation in background rotation and object position (vs constant model $p<0.001$).

**c:** Same as Fig. 2c for specific background rotation decoding (each session contributes 25 values to the distribution). Here, the background rotation value was decoded while ignoring the other parameter variations across images. Arrows represent median decoding accuracy (0.558 for monkey 1, 0.344 for monkey 2). Box plots above the histograms show the distributions of general decoder performance.

**d:** Same as Fig. 2c for specific background depth decoding (each session contributes 25 values to the distribution). Arrows represent median decoding accuracy (0.703 for monkey 1, 0.634 for monkey 2). Box plots above the histograms show the distributions of general decoder performance.

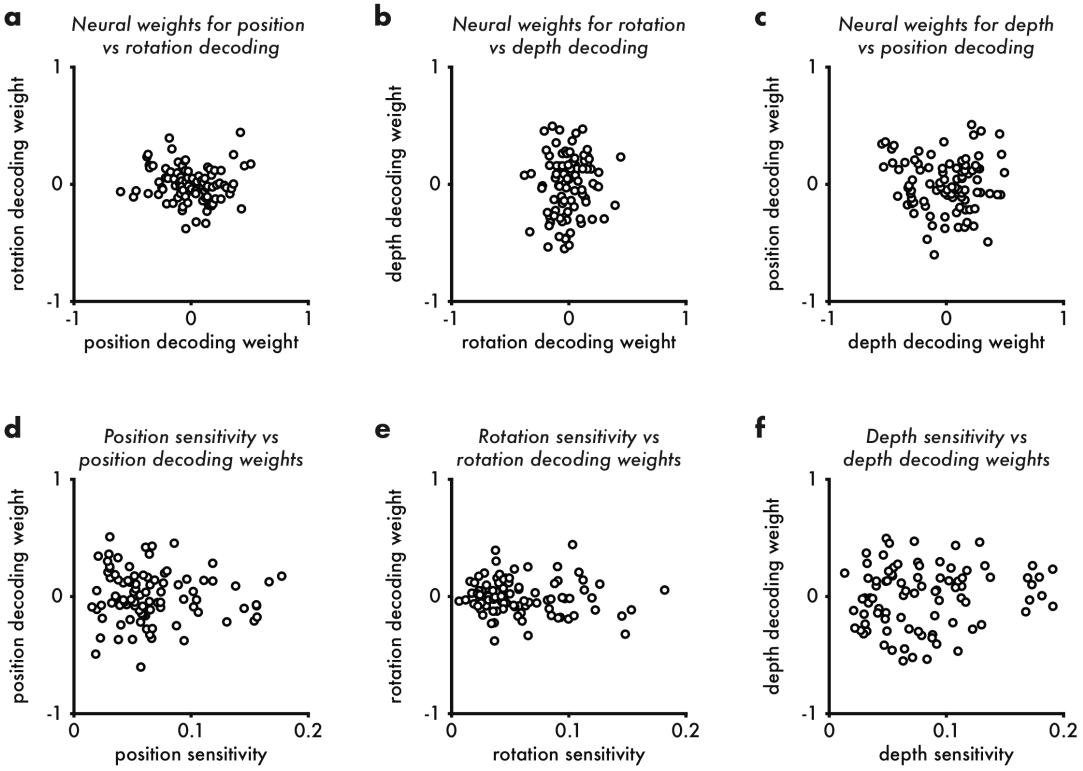
**e:** Comparison of errors for general decoding of object position and general decoding of background depth across trials ($r=-0.03$; $p=0.28$). Compare with Fig. 2d.




**Supp. Fig. 3: Cross-decoding of object position across background conditions reveals slight deviations from orthogonal coding**

**a:** We calculated the performance of object position decoders which were trained on the responses of one background condition (a unique combination of background rotation and depth values) and tested on either held-out trials of the same condition ("self-decoding") or on the responses to images that varied in rotation by 1, 2, 3, or 4 feature value steps in either direction ("cross-decoding"). Each dot represents the difference between the cross-decoding and self-decoding accuracies for a unique pair of training and testing conditions for the example session in Fig. 2 and 3. As the background rotation becomes more different, the cross-condition decoder performs marginally worse (linear fit $r=0.256$, $p<0.0001$). This indicates that across large changes in background rotation, the representations of rotation and object position deviate slightly from pure orthogonality.

**b:** Same as above for decoding object position across changes in background depths. Object position cross-decoding is marginally affected by large changes in background depth values (linear fit $r=0.2$, $p<0.0001$); the median additional cross-decoding error at the largest change in background rotation is 0.12.




**Supp. Fig. 4: No relationship between the linear decoding weights across object position, background rotation, and background depth decoders or with feature sensitivity scores** (Here shown for the example session in Fig. 2 and 3)

**a:** Comparison of object position and background rotation general decoding weights. Feature values were normalized before decoding, and weights were averaged across cross-validation folds. Each dot represents a neuron. No significant relationship was observed (r=-0.07; p=0.5), indicating the independent contribution of each neuron to the two decoders.

**b:** Same as above for comparing background rotation and depth decoding (r=0.09; p=0.389).

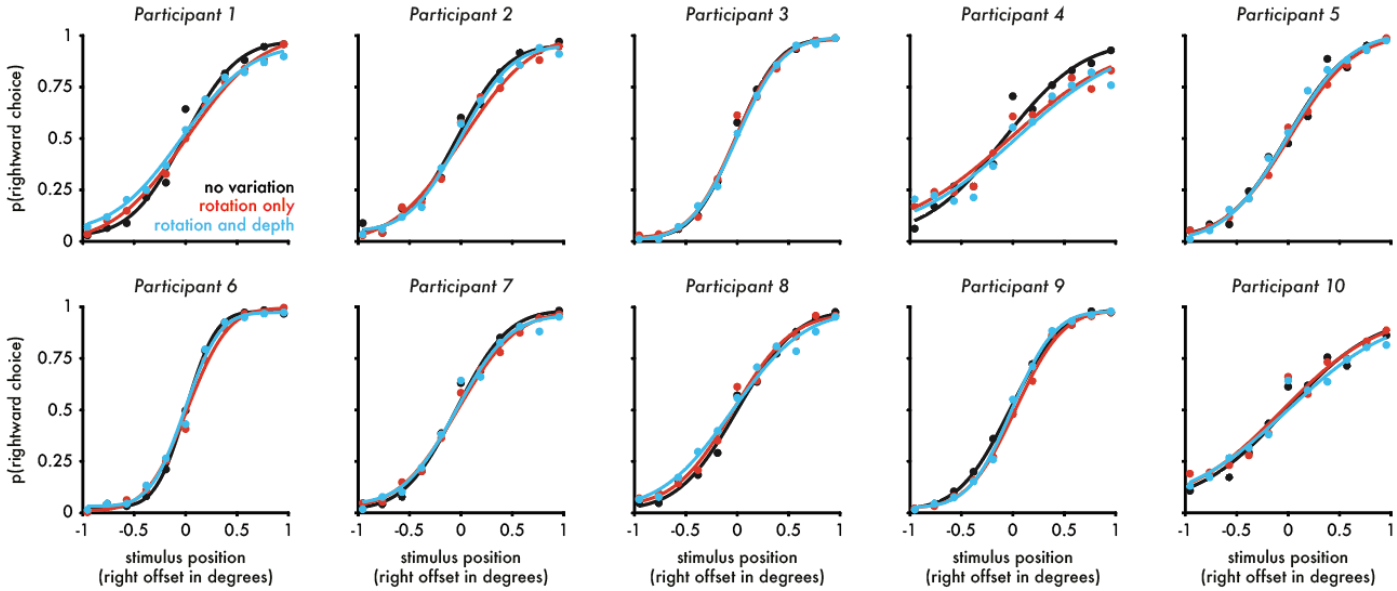
**c:** Same as above for comparing background depth and object position decoding (r=-0.06; p=0.57).

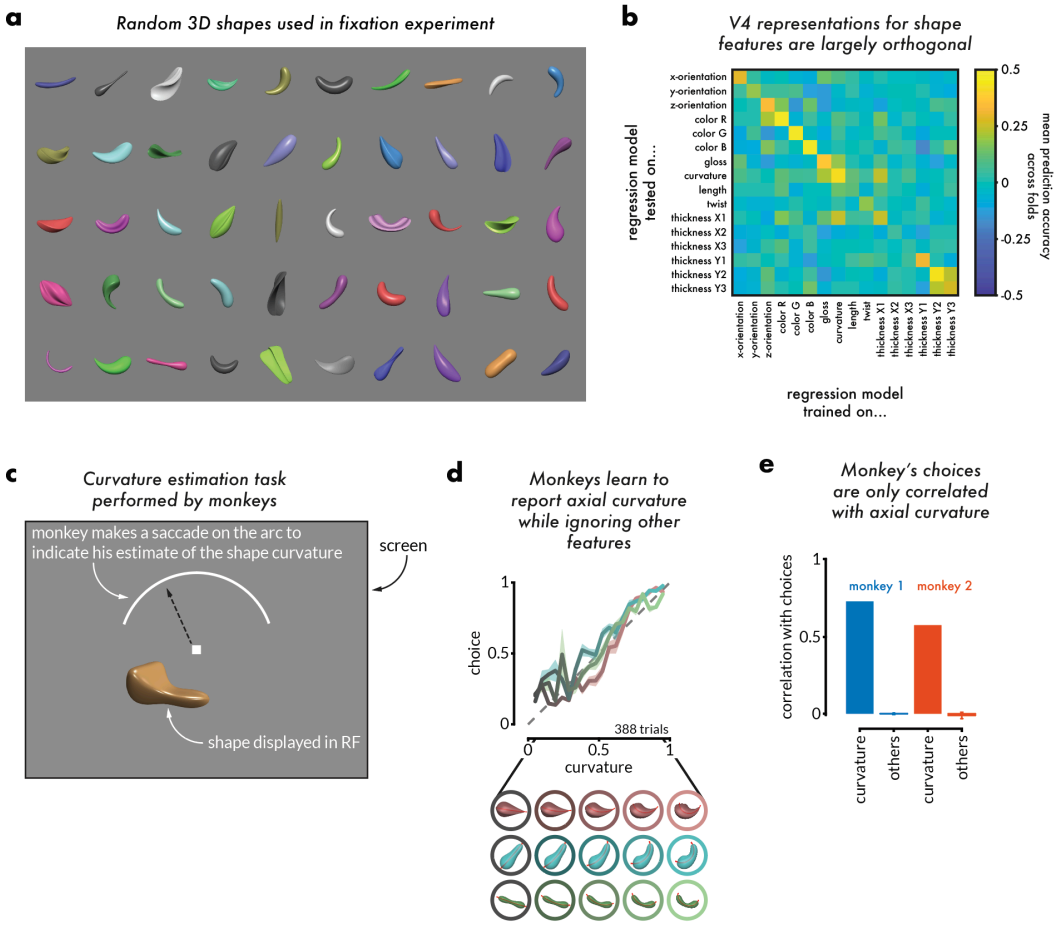
**d:** Comparison of the relationship between the general object position decoding weights of each neuron to its feature sensitivity for object position as calculated in Supp. Fig. 1d. No relationship was observed (r=-0.131; p=0.212), indicating heterogeneous contributions of each neuron to feature decoding.

**e:** Same as d for background rotation. No relationship (r=-0.15; p=0.16).

**f:** Same as d for background depth. No relationship (r=0.01; p=0.92).

**Supp. Fig. 5: Psychometric functions for object position discrimination for each participant.**






**Supp. Fig. 6: Orthogonal representations of 16 object features aid curvature estimation.** In data collected in parallel with the object and background variation experiments, we also displayed 50 randomly generated 3D objects. We generated these objects using 16 parameters like length, 3D orientation, color, gloss, curvature, etc. These parameters were simply randomized, not permuted as in Fig. 5. In an associated behavioral task, we trained the monkey to estimate the axial curvature of the presented object and make a saccade on the displayed target to report the estimate. The stimuli, behavior, and neural recordings are detailed in Srinath et al., 2024[28] and panels c and d have been adapted from Srinath et al., 2024 permission.

**a:** Stimuli that were used in the fixation experiment. Fifty stimuli were generated by randomizing the stimulus generation parameters.

**b:** Similar to Fig. 5, if two features are encoded orthogonally in neural population space, then a decoder trained on one feature should not support decoding of the other feature. All decodable features were represented orthogonal to each other. Minor interactions were observed between gloss and curvature encoding (which is likely due to the randomization of the feature values during stimulus generation), and between two thickness features (which is not unexpected due to the smooth way the thickness of the 3D object varies along the axis.)

**c:** Schematic of the curvature estimation task. Monkeys estimated the axial curvature of the presented 3D shape and reported it on a continuous scale by making a saccade on a curved arc presented in the upper visual field.

**d:** Example choice behavior from a session where curvature variations of three objects were tested. The monkey is able to report the axial curvature successfully without interference from other features.

**e.** Correlation of the monkeys' behavioral response with changes in either axial curvature or other features. We tested the monkeys' estimation behavior for ten unique shapes across 144 sessions (113 for monkey 1, 31 for monkey 2), we calculated the correlation of each feature with the monkeys' choices. Only the task-relevant feature (axial curvature) is correlated with monkeys' choices.