Supplementary Information for:

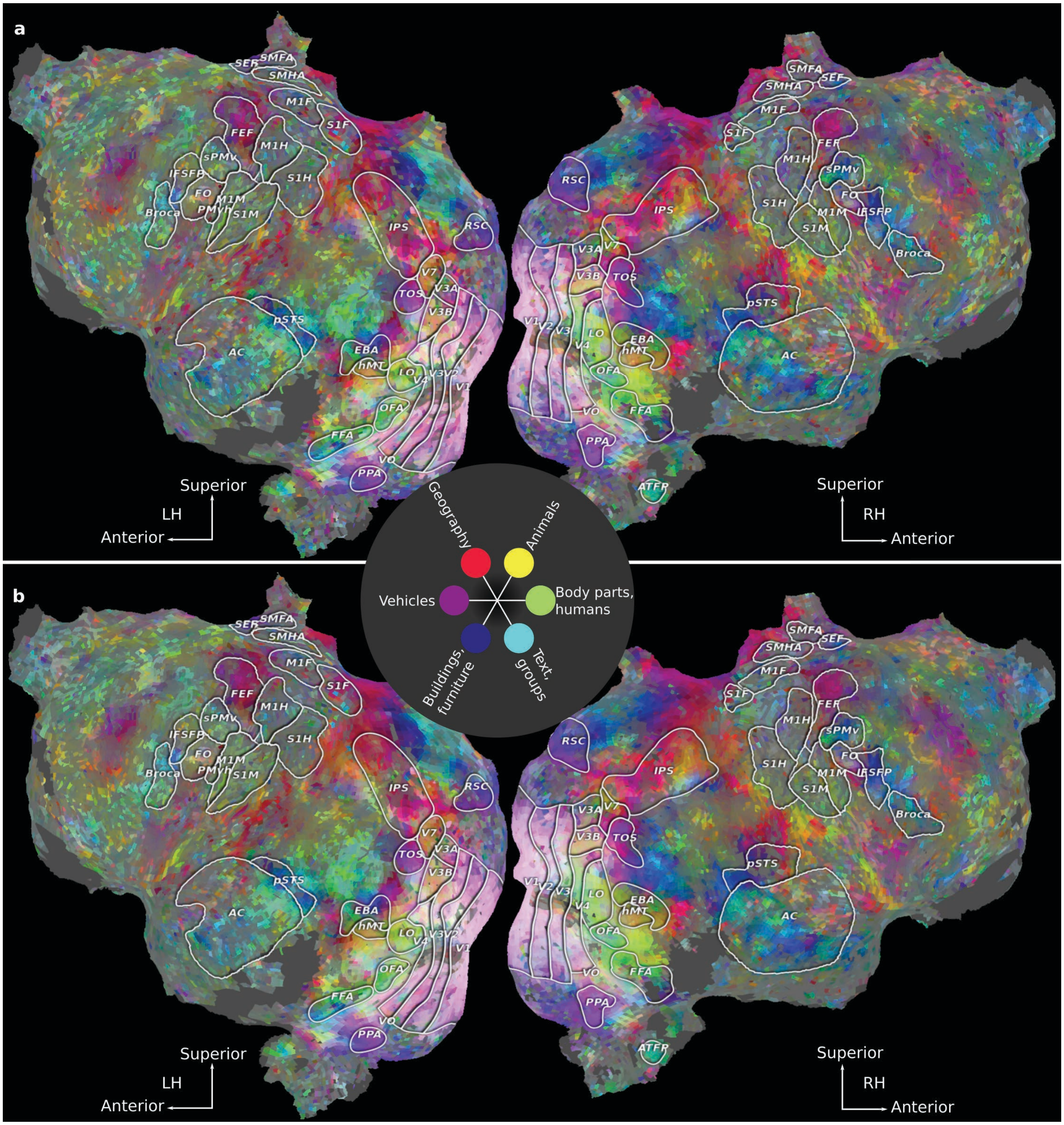
Spatially Informed Voxelwise Modeling for Naturalistic fMRI Experiments

Emin Çelik [a,b], Salman Ul Hassan Dar[b,c], Özgür Yılmaz[b,c], Ümit Keleş[b,d], Tolga Çukur[a,b,c]

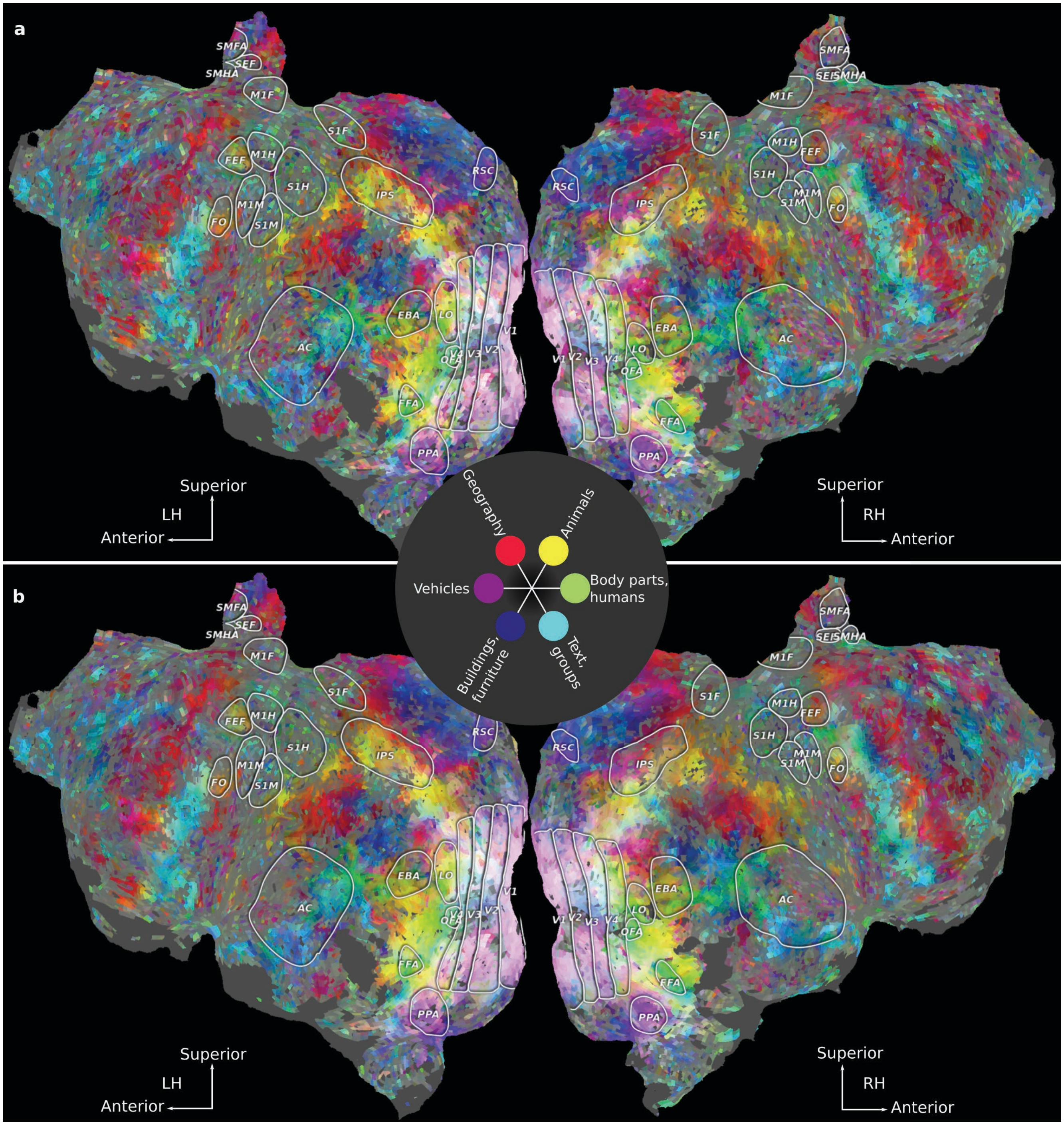
[a]Neuroscience Program, Bilkent University, Ankara, TR-06800, Turkey
[b]National Magnetic Resonance Research Center (UMRAM), Bilkent University, Ankara, TR-06800, Turkey
[c]Department of Electrical and Electronics Engineering, Bilkent University, Ankara, TR-06800, Turkey
[d]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA-91125, USA

Table of Contents

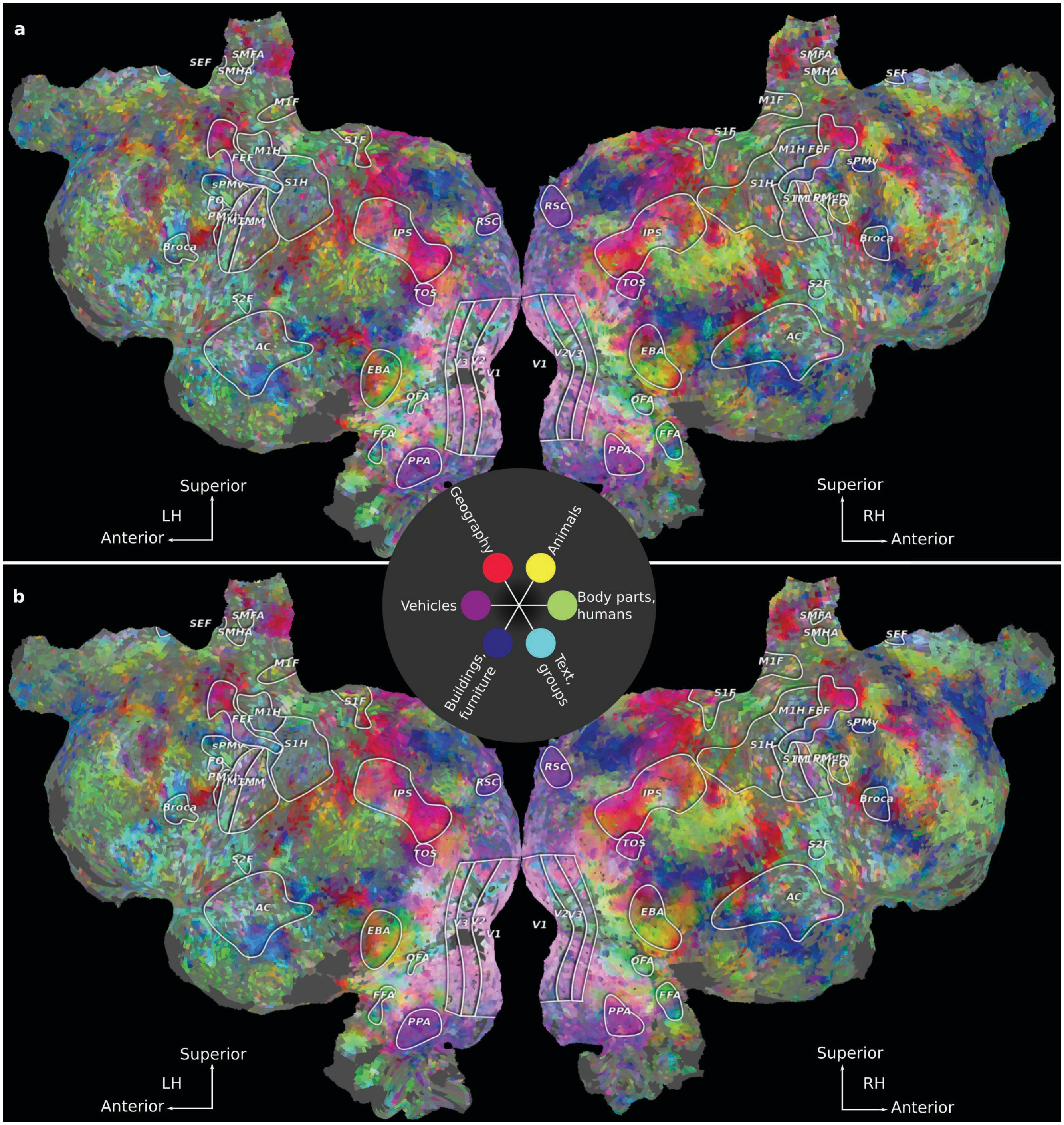
Supplementary Figures



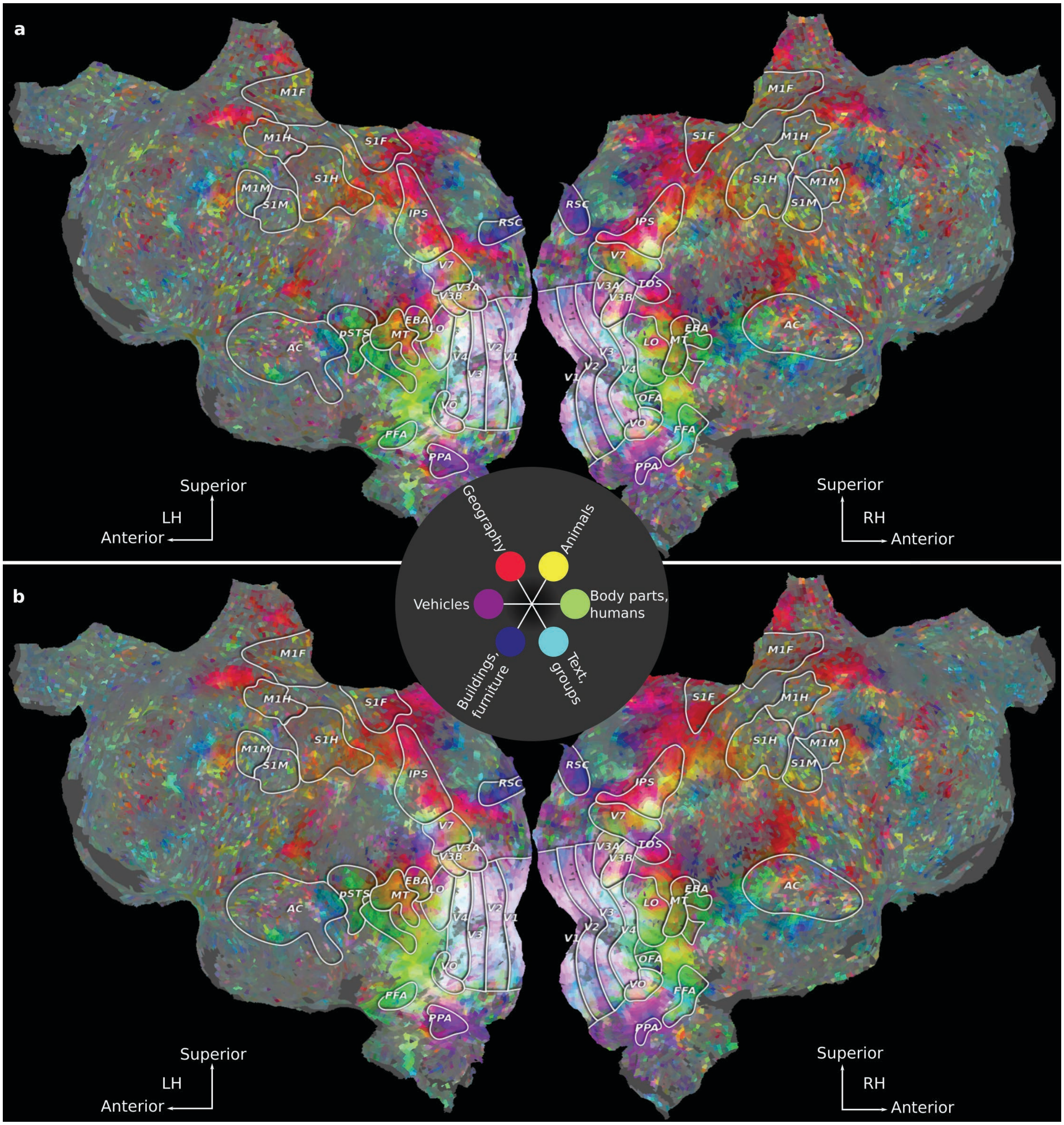
Supplementary Fig. 1. Cortical flatmaps of semantic representation. Cortical flatmaps of semantic representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S1. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Category model weights for each voxel were then projected onto the second, third, and fourth PCs of the group semantic space. Each voxel was assigned a color by representing projections on the second, third, and fourth PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar semantic categories (e.g., dark blue implies selectivity for buildings and furniture, whereas magenta implies selectivity for vehicles). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent semantic maps across many high-level visual and frontal areas.



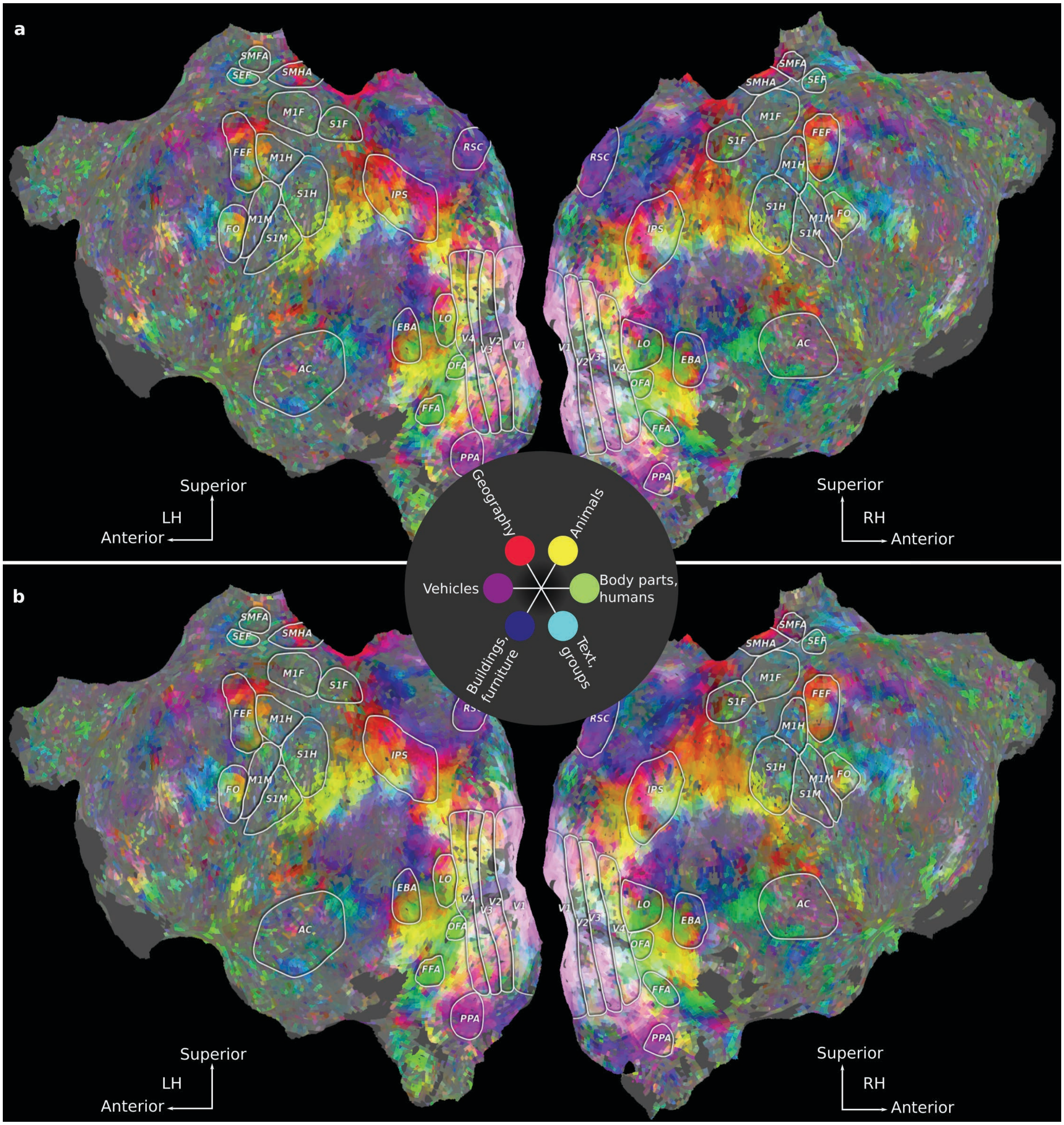
Supplementary Fig. 2. Cortical flatmaps of semantic representation. Cortical flatmaps of semantic representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S2. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Category model weights for each voxel were then projected onto the second, third, and fourth PCs of the group semantic space. Each voxel was assigned a color by representing projections on the second, third, and fourth PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar semantic categories (e.g., dark blue implies selectivity for buildings and furniture, whereas magenta implies selectivity for vehicles). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent semantic maps across many high-level visual and frontal areas.



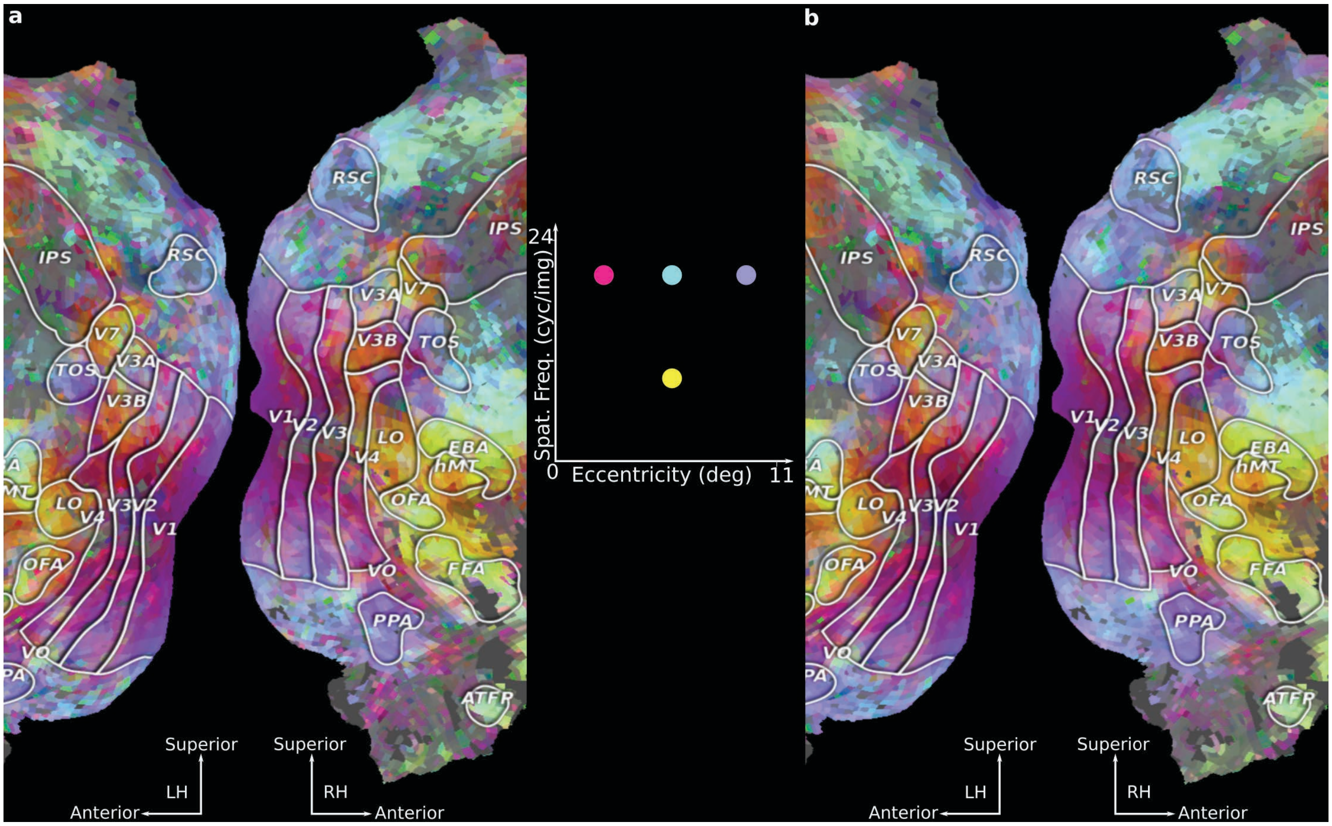
Supplementary Fig. 3. Cortical flatmaps of semantic representation. Cortical flatmaps of semantic representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S3. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Category model weights for each voxel were then projected onto the second, third, and fourth PCs of the group semantic space. Each voxel was assigned a color by representing projections on the second, third, and fourth PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar semantic categories (e.g., dark blue implies selectivity for buildings and furniture, whereas magenta implies selectivity for vehicles). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent semantic maps across many high-level visual and frontal areas.



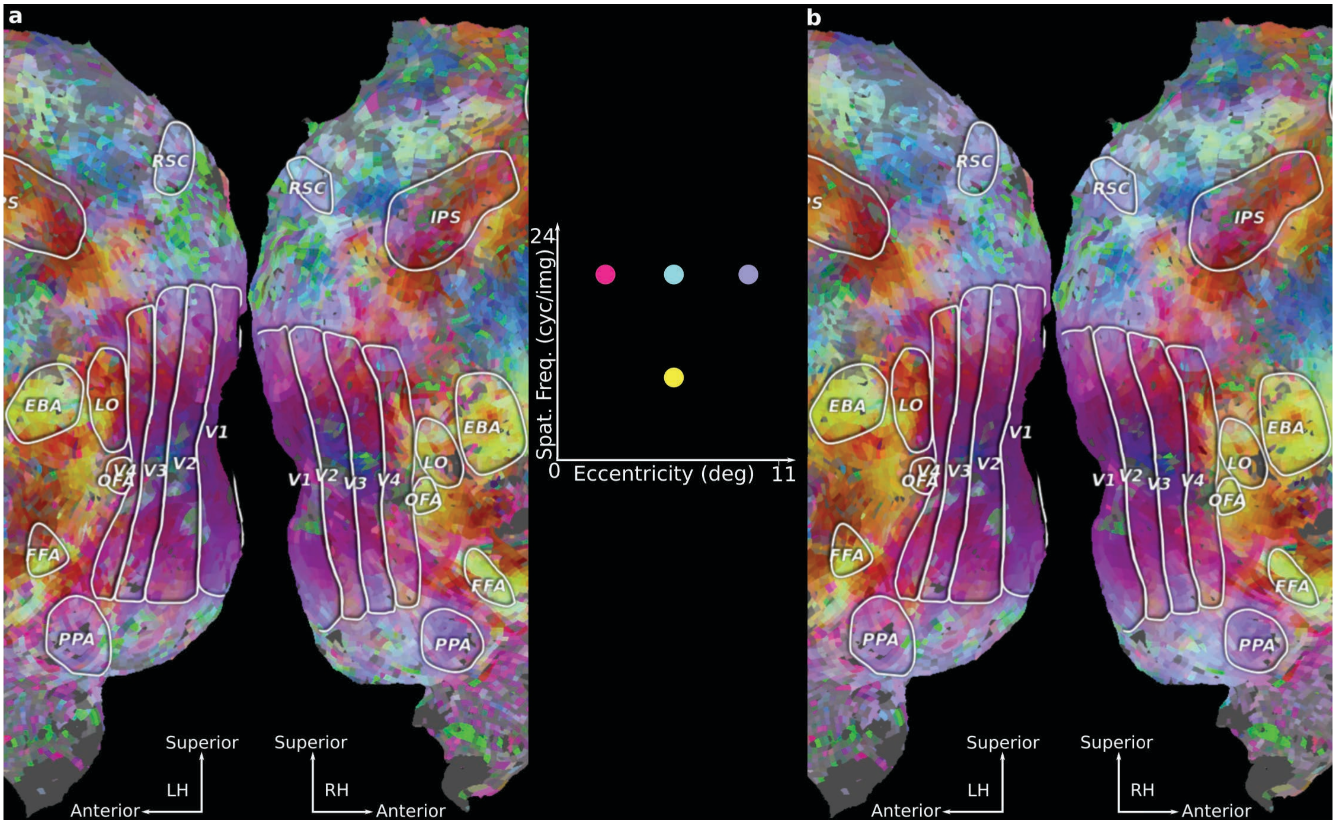
Supplementary Fig. 4. Cortical flatmaps of semantic representation. Cortical flatmaps of semantic representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S4. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Category model weights for each voxel were then projected onto the second, third, and fourth PCs of the group semantic space. Each voxel was assigned a color by representing projections on the second, third, and fourth PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar semantic categories (e.g., dark blue implies selectivity for buildings and furniture, whereas magenta implies selectivity for vehicles). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent semantic maps across many high-level visual and frontal areas.



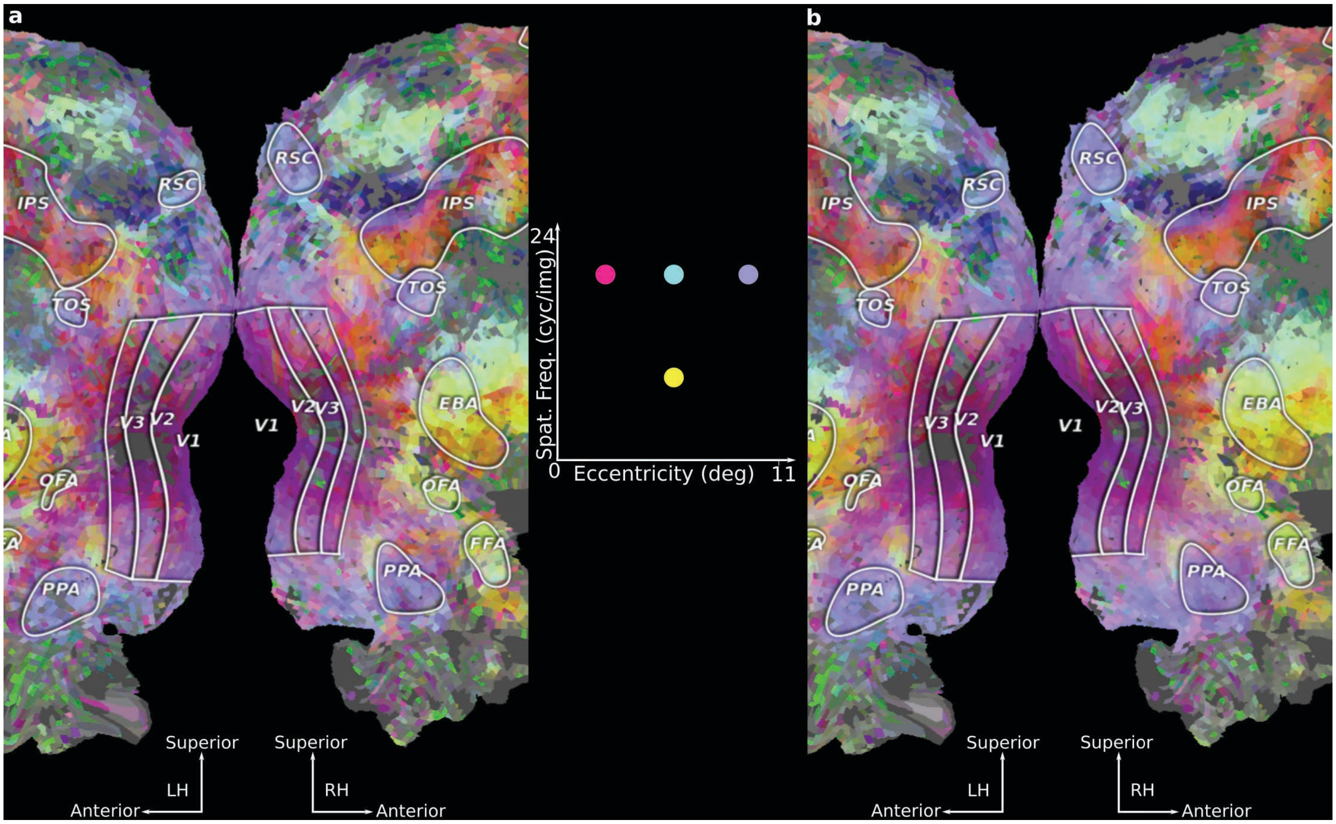
Supplementary Fig. 5. Cortical flatmaps of semantic representation. Cortical flatmaps of semantic representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S5. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Category model weights for each voxel were then projected onto the second, third, and fourth PCs of the group semantic space. Each voxel was assigned a color by representing projections on the second, third, and fourth PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar semantic categories (e.g., dark blue implies selectivity for buildings and furniture, whereas magenta implies selectivity for vehicles). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent semantic maps across many high-level visual and frontal areas.



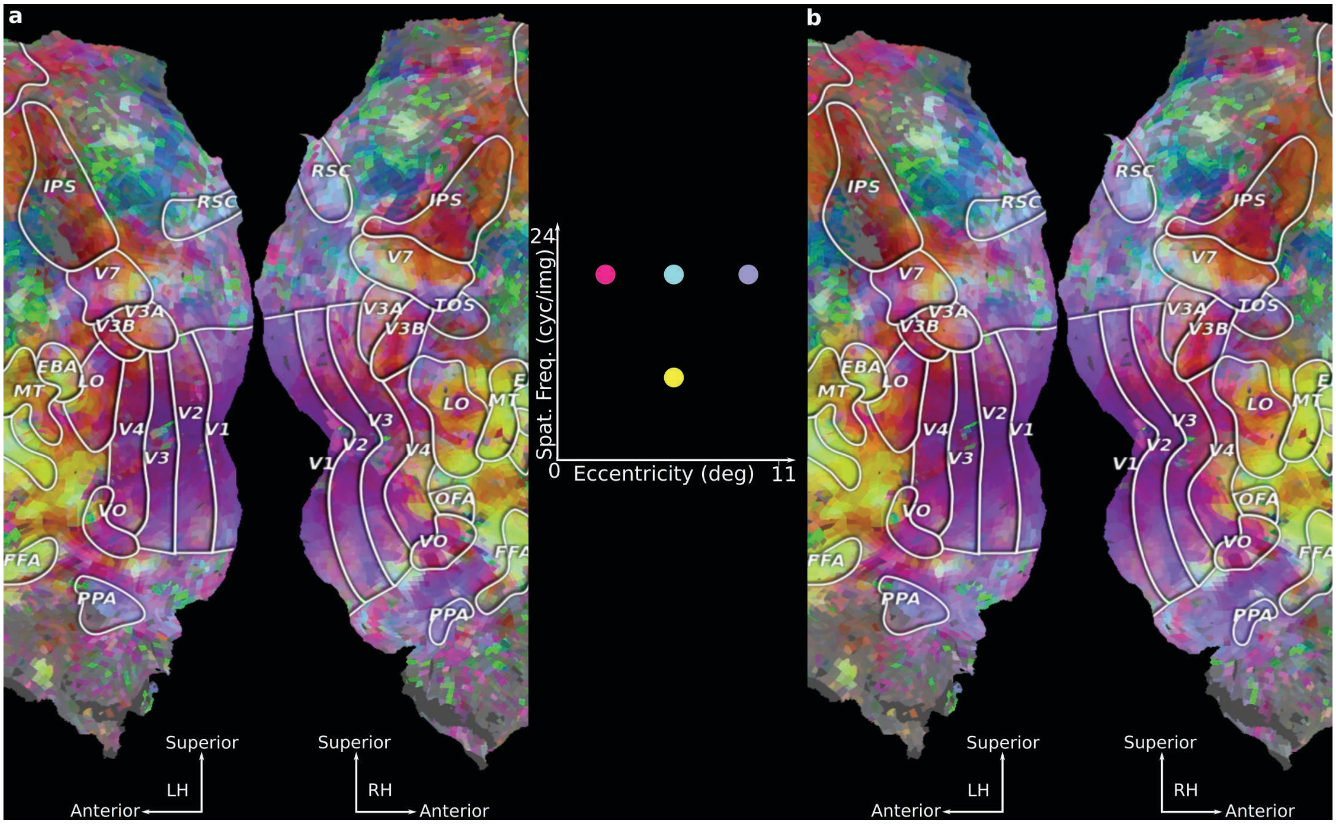
Supplementary Fig. 6. Cortical flatmaps of low-level visual representation. Cortical flatmaps of low-level visual representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S1. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Motion-energy model weights for each voxel were then projected onto the first three PCs of the group Gabor space. Each voxel was assigned a color by representing projections on the first, second, and third PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar low-level features (e.g., yellow signifies medium eccentricity and lower spatial frequency, whereas magenta signifies low eccentricity and higher spatial frequency). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent Gabor maps across early visual areas.



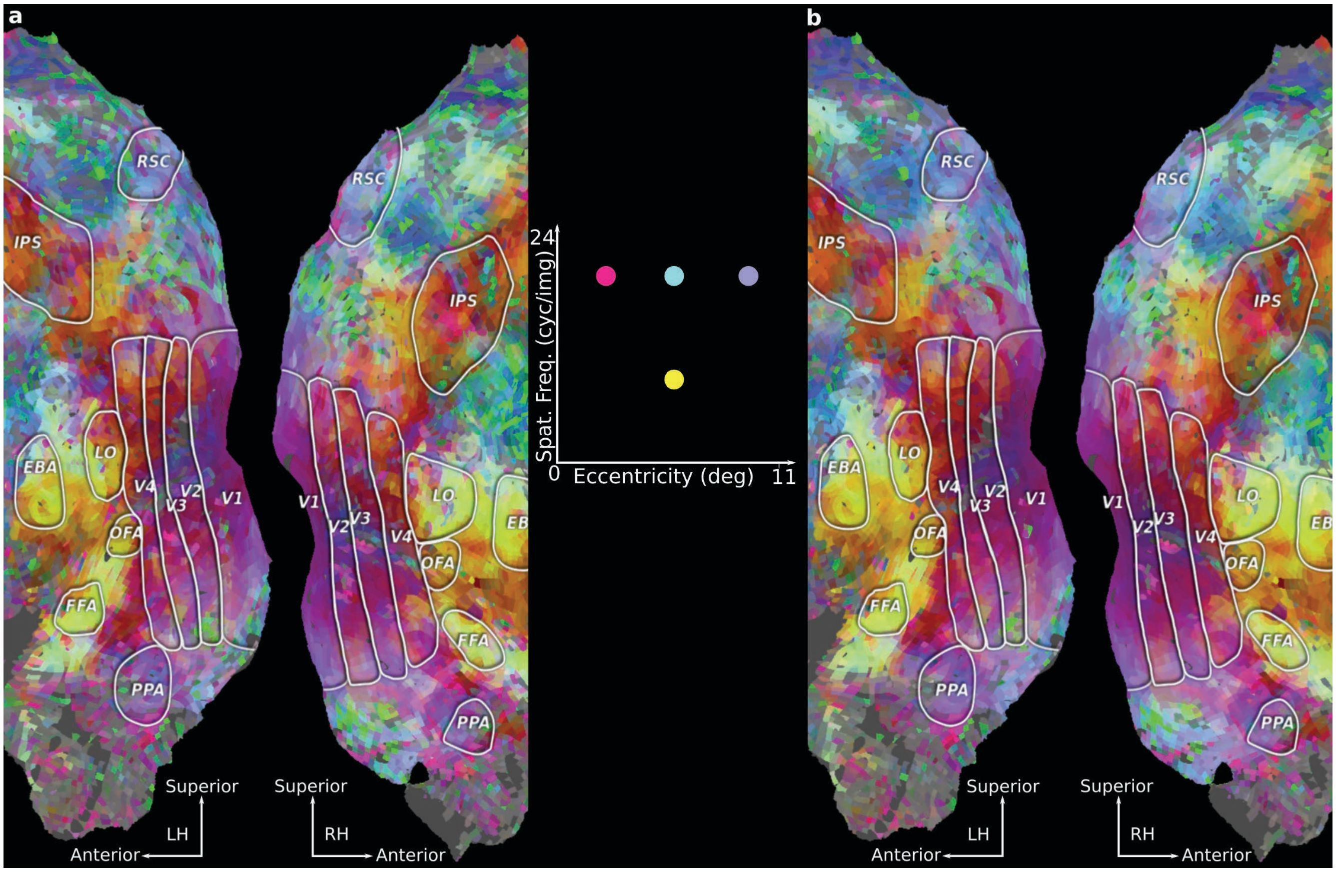
Supplementary Fig. 7. Cortical flatmaps of low-level visual representation. Cortical flatmaps of low-level visual representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S2. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Motion-energy model weights for each voxel were then projected onto the first three PCs of the group Gabor space. Each voxel was assigned a color by representing projections on the first, second, and third PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar low-level features (e.g., yellow signifies medium eccentricity and lower spatial frequency, whereas magenta signifies low eccentricity and higher spatial frequency). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent Gabor maps across early visual areas.



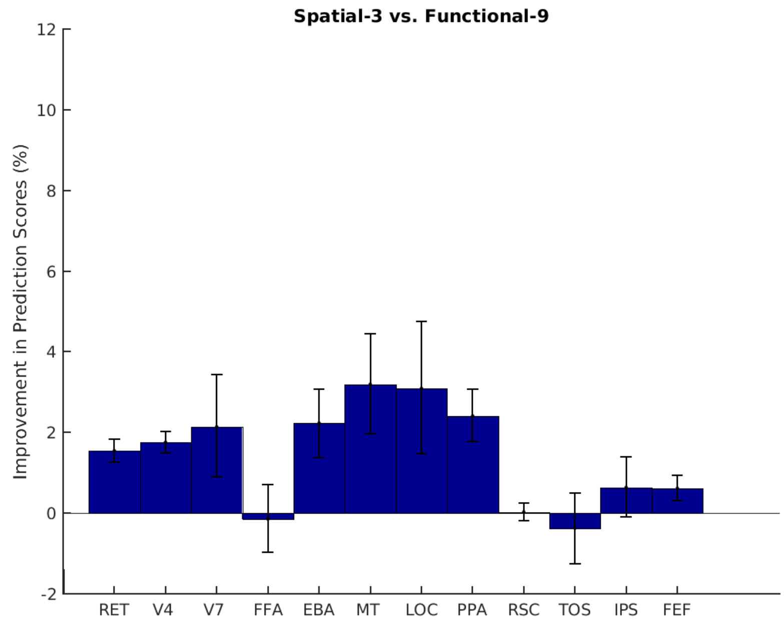
Supplementary Fig. 8. Cortical flatmaps of low-level visual representation. Cortical flatmaps of low-level visual representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S3. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Motion-energy model weights for each voxel were then projected onto the first three PCs of the group Gabor space. Each voxel was assigned a color by representing projections on the first, second, and third PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar low-level features (e.g., yellow signifies medium eccentricity and lower spatial frequency, whereas magenta signifies low eccentricity and higher spatial frequency). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent Gabor maps across early visual areas.



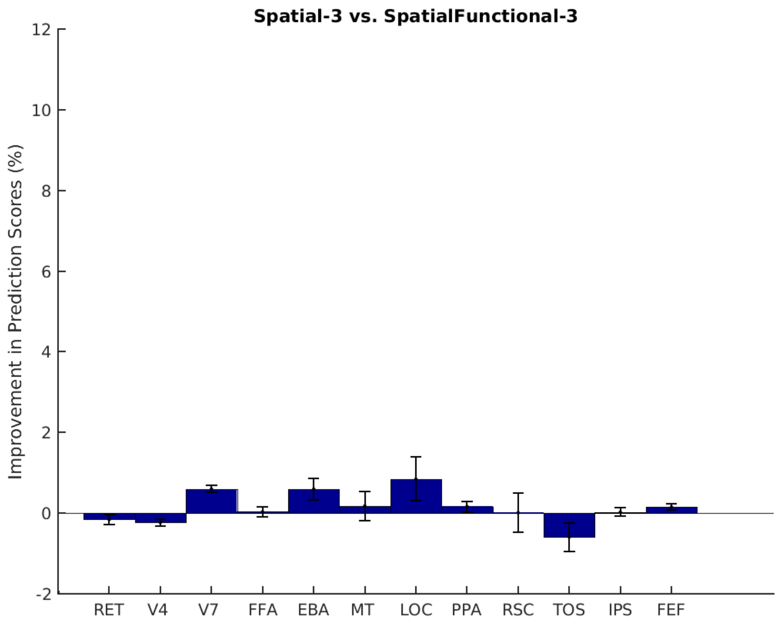
Supplementary Fig. 9. Cortical flatmaps of low-level visual representation. Cortical flatmaps of low-level visual representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S4. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Motion-energy model weights for each voxel were then projected onto the first three PCs of the group Gabor space. Each voxel was assigned a color by representing projections on the first, second, and third PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar low-level features (e.g., yellow signifies medium eccentricity and lower spatial frequency, whereas magenta signifies low eccentricity and higher spatial frequency). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent Gabor maps across early visual areas.



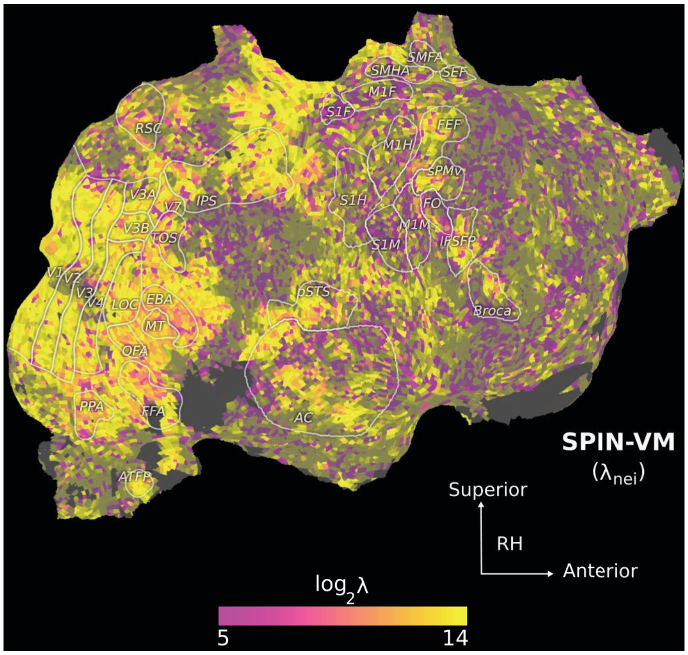
Supplementary Fig. 10. Cortical flatmaps of low-level visual representation. Cortical flatmaps of low-level visual representation as measured by **(a)** VM and **(b)** SPIN-VM for subject S5. To obtain consistent principal components (PCs) across both VM and SPIN-VM models, model weights obtained by both techniques were pooled and PCA was applied. Motion-energy model weights for each voxel were then projected onto the first three PCs of the group Gabor space. Each voxel was assigned a color by representing projections on the first, second, and third PCs with red, green, and blue channels, respectively. Similar colors imply selectivity for similar low-level features (e.g., yellow signifies medium eccentricity and lower spatial frequency, whereas magenta signifies low eccentricity and higher spatial frequency). Compared to VM, estimated selectivities of neighboring voxels are more congruent (i.e., they have more similar colors) for SPIN-VM. Therefore, SPIN-VM produces more coherent Gabor maps across early visual areas.



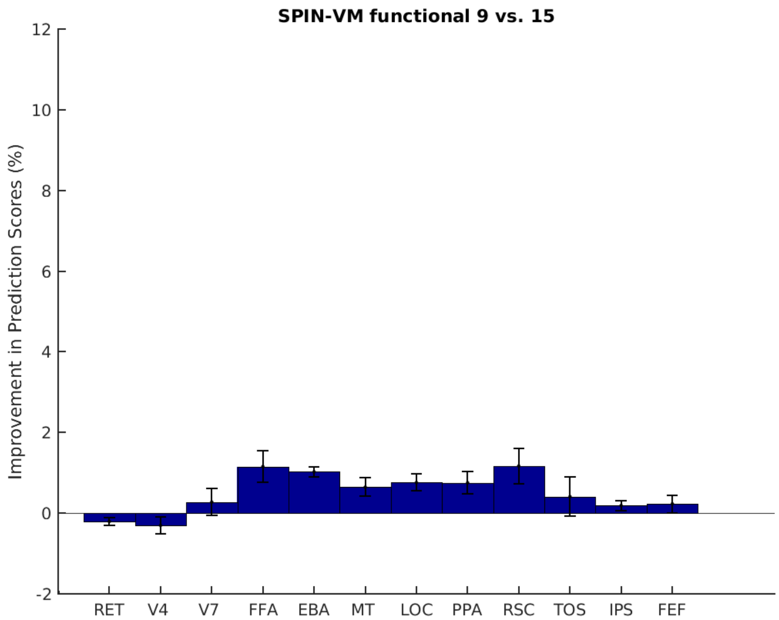
Supplementary Fig. 11. Improvement in prediction scores (Spatial-3 vs. Functional-9). Improvement in prediction scores for the category model with the optimal variant of regular SPIN-VM (Spatial-3; spatial with a window size of 3x3x3) versus the optimal variant of SPIN-VM based on functional correlations (Functional-9; functional with a "window size" of 9x9x9). Positive values indicate higher performance for Spatial-3 compared to Functional-9. Spatial-3 performs better in the majority of the functional ROIs.



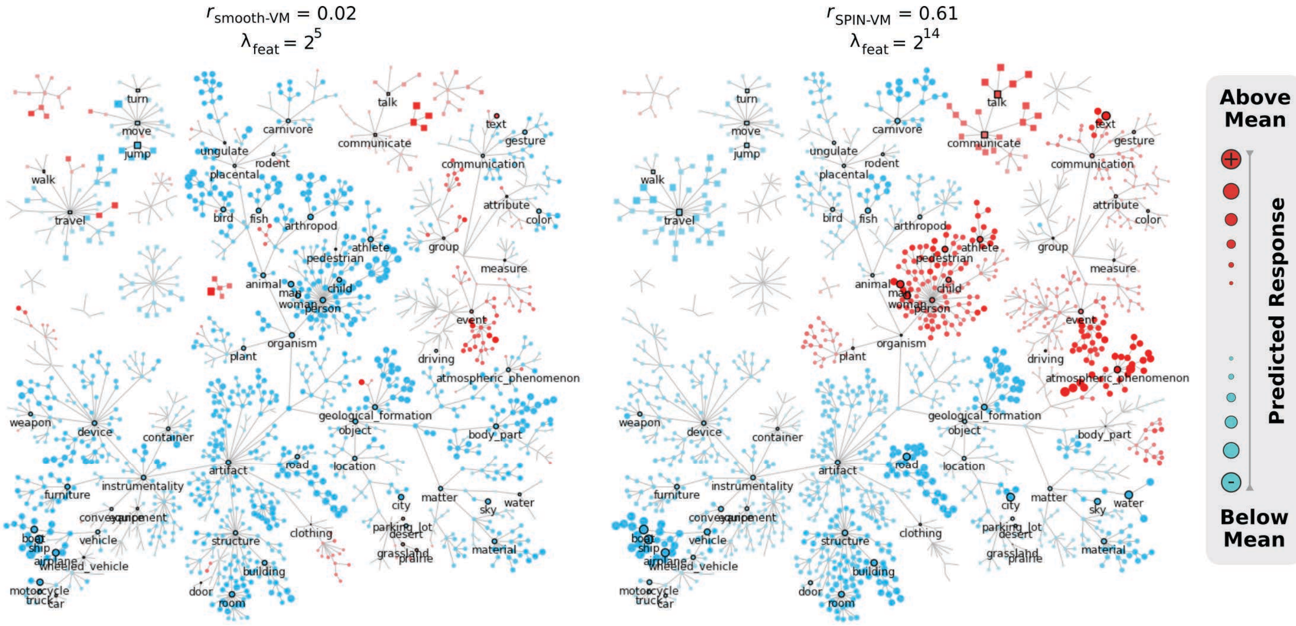
Supplementary Fig. 12. Improvement in prediction scores (Spatial-3 vs. SpatialFunctional-3). Improvement in prediction scores for the category model with the optimal variant of regular SPIN-VM (Spatial-3; spatial with a window size of 3x3x3) versus the optimal variant of SPIN-VM based on the combination of functional correlations and spatial neighborhood (SpatialFunctional-3; spatial + functional with a "window size" of 3x3x3). Spatial-3 performs better in the majority of the ROIs, particularly in the high-level visual areas where the category model works best. Note that the difference is relatively small as weights based on functional correlations closely resemble those based on spatial proximity.



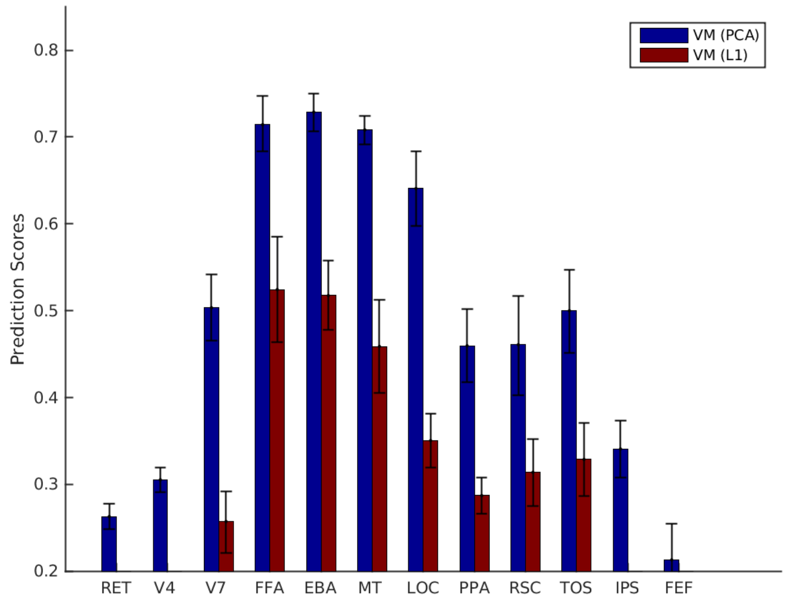
Supplementary Fig. 13. Cortical distribution of regularization parameters. Cortical flatmap of optimal regularization parameters across spatial neighborhoods of voxels (λ_{nei}) displayed in subject S1 for the category model. Optimal λ_{nei} values were determined separately for each voxel during model fitting. Color bar shows the range of λ_{nei} [2^5-2^{14}] in logarithmic scale (pink = low, yellow = high). As expected, we find that optimal λ_{nei} values are relatively higher in both low-level retinotopic and high-level category selective visual areas that are more engaged during viewing of natural movies than non-visual areas in frontotemporal, motor, and somatosensory cortices. These high λ_{nei} values likely compensate for the relatively lower λ_{feat} values in SPIN-VM compared to VM.



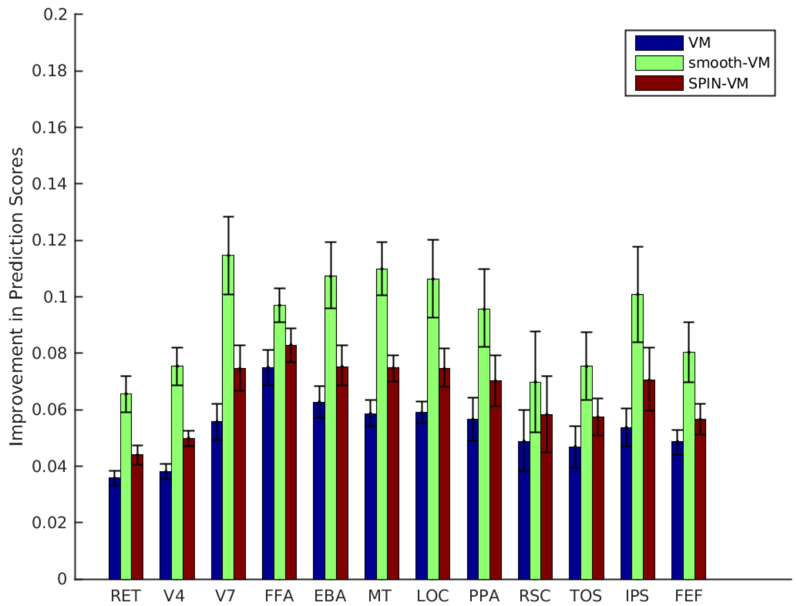
Supplementary Fig. 14. Improvement in prediction scores (Functional-9 vs. Functional-15). Improvement in prediction scores for the category model with the optimal variant of SPIN-VM based on functional correlations (Functional-9; functional with a "window size" of 9x9x9) versus SPIN-VM based on functional correlations with a window size of 15 (Functional-15; functional with a "window size" of 15x15x15). Functional-9 performs better in the majority of the ROIs, particularly in the high-level visual areas where the category model works best.



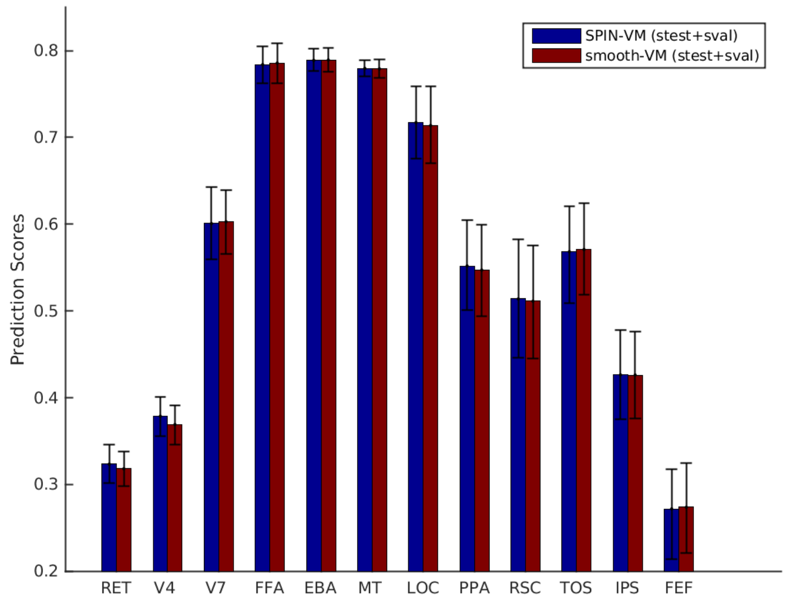
Supplementary Fig. 15. Functional selectivity in a single voxel. Functional selectivity for object and action categories as measured by the category model for a single voxel (voxel #13137) in posterior superior temporal sulcus (pSTS) of subject S1. Functional selectivity obtained by smooth-VM (left) and SPIN-VM (right) is shown. Each node in these graphs represents a distinct object or action organized according to the hierarchical relations in the WordNet lexicon. Some important nodes are labeled to orient the reader. Red nodes correspond to categories that evoke above-mean responses, whereas blue nodes correspond to categories that evoke below-mean responses. The size of each node reflects the magnitude of the category response. The response of voxel #13137 is well-predicted by SPIN-VM ($r = 0.61$), and only poorly-predicted by smooth-VM ($r = 0.02$). pSTS has been implicated in the representation of facial identity and visually-observed social interaction (Srinivasan et al., 2016; Walbrin et al., 2018). While the model obtained via smooth-VM largely fails to capture these representations, SPIN-VM successfully captures selectivity for categories related to individuals such as ‘person’ and ‘man’, as well as categories related to social communication such as ‘talk’ and ‘text’.



Supplementary Fig. 16. Prediction scores with VM using PCA vs. L1-norm. Prediction scores for the category model across functional ROIs via VM based on two distinct regularization methods. The first method first performed PCA on stimulus features for dimensionality reduction, and then used ridge regularization based on L2-norm. The second method only used L1-norm based regularization on the original stimulus features. Color bars show mean prediction scores across five subjects and error bars indicate standard error of the mean (SEM). The conjunction of PCA with L2-norm based regularization yields significantly higher prediction scores than the L1-norm based regularization in all ROIs ($p < 0.05$).



Supplementary Fig. 17. Improvement in prediction scores on smoothed test data. Improvement in prediction scores when model performance was evaluated on smoothed versus raw test data. Results are shown for the category model as measured by VM, smooth-VM, and SPIN-VM. Colored bars show mean prediction score improvements across five subjects and error bars indicate standard error of the mean (SEM). Prediction scores are significantly greater in all twelve ROIs for all three methods when smoothed test data is used ($p < 0.05$). Naturally, smooth-VM benefits relatively more from smoothing of test data.



Supplementary Fig. 18. Prediction scores with SPIN-VM vs. smooth-VM on smoothed validation and data. Mean prediction scores for the category model with SPIN-VM versus smooth-VM when both test and validation data are smoothed. Colored bars show mean prediction scores across five subjects and error bars indicate standard error of the mean (SEM). SPIN-VM and smooth-VM demonstrate almost identical performance on smoothed validation and test data, even though SPIN-VM is trained on unsmoothed data.

Supplementary Tables

	3×3×3	5×5×5	7×7×7	9×9×9	11×11×11	13×13×13	15×15×15
Whole cortex	**.1795 ± .0279***	.1785 ± .0275	.1774 ± .0274	.1764 ± .0272	.1758 ± .0270	.1752 ± .0269	.1746 ± .0267
RET	.2826 ± .0170	.2836 ± .0170	**.2837 ± .0171**	.2830 ± .0168	.2825 ± .0165	.2812 ± .0156	.2805 ± .0156
V4	.3288 ± .0180	**.3301 ± .0176**	.3291 ± .0175	.3277 ± .0171	.3267 ± .0169	.3252 ± .0165	.3243 ± .0164
V7	**.5309 ± .0388***	.5270 ± .0391	.5225 ± .0396	.5176 ± .0392	.5151 ± .0388	.5129 ± .0381	.5102 ± .0384
FFA	**.7227 ± .0301***	.7219 ± .0307	.7205 ± .0308	.7185 ± .0310	.7177 ± .0310	.7171 ± .0315	.7163 ± .0310
EBA	**.7420 ± .0171***	.7412 ± .0172	.7398 ± .0175	.7383 ± .0177	.7373 ± .0181	.7363 ± .0181	.7356 ± .0186
MT	**.7263 ± .0137***	.7246 ± .0141	.7228 ± .0145	.7205 ± .0151	.7186 ± .0152	.7173 ± .0156	.7158 ± .0159
LOC	**.6599 ± .0415***	.6587 ± .0403	.6556 ± .0407	.6528 ± .0407	.6516 ± .0408	.6513 ± .0403	.6497 ± .0405
PPA	**.4839 ± .0458***	.4815 ± .0463	.4772 ± .0464	.4736 ± .0455	.4713 ± .0450	.4704 ± .0445	.4681 ± .0439
RSC	**.4732 ± .0589***	.4707 ± .0596	.4684 ± .0596	.4667 ± .0587	.4664 ± .0586	.4666 ± .0594	.4657 ± .0584
TOS	.5127 ± .0451	**.5136 ± .0445**	.5111 ± .0441	.5098 ± .0435	.5080 ± .0437	.5082 ± .0432	.5067 ± .0435
IPS	**.3567 ± .0355***	.3547 ± .0361	.3517 ± .0363	.3492 ± .0364	.3470 ± .0362	.3458 ± .0359	.3440 ± .0360
FEF	**.2283 ± .0408***	.2260 ± .0411	.2225 ± .0407	.2195 ± .0406	.2171 ± .0400	.2161 ± .0400	.2143 ± .0395

Supplementary Table 1: Prediction scores for the category model estimated with SPIN-VM in functional ROIs. The prediction scores are reported as mean ± SEM across five subjects. Results are given for seven different window sizes. The highest prediction score in each row is annotated in bold font. If the highest prediction score in a row is significantly greater than all the other prediction scores in the same row, it is also marked with a * ($p < 0.05$, Bootstrap test).

	3×3×3	5×5×5	7×7×7	9×9×9	11×11×11	13×13×13	15×15×15
Whole cortex	**.2080 ± .0097***	.2071 ± .0097	.2063 ± .0098	.2055 ± .0098	.2051 ± .0095	.2048 ± .0100	.2040 ± .0096
RET	**.6521 ± .0285***	.6504 ± .0284	.6495 ± .0284	.6487 ± .0282	.6477 ± .0282	.6463 ± .0281	.6458 ± .0283
V4	**.6240 ± .0197***	.6226 ± .0198	.6217 ± .0198	.6202 ± .0198	.6191 ± .0194	.6175 ± .0195	.6163 ± .0196
V7	**.6663 ± .0203***	.6643 ± .0204	.6623 ± .0206	.6609 ± .0203	.6598 ± .0205	.6597 ± .0203	.6583 ± .0203
FFA	**.6528 ± .0354***	.6496 ± .0361	.6454 ± .0363	.6425 ± .0367	.6424 ± .0361	.6416 ± .0362	.6393 ± .0363
EBA	**.6899 ± .0231***	.6886 ± .0227	.6873 ± .0232	.6862 ± .0232	.6847 ± .0233	.6836 ± .0228	.6827 ± .0235
MT	**.6941 ± .0140***	.6932 ± .0135	.6915 ± .0137	.6901 ± .0137	.6889 ± .0135	.6878 ± .0133	.6866 ± .0134
LOC	**.6583 ± .0176**	.6570 ± .0171	.6554 ± .0174	.6539 ± .0178	.6528 ± .0175	.6518 ± .0172	.6513 ± .0179
PPA	**.4642 ± .0309***	.4625 ± .0307	.4615 ± .0309	.4593 ± .0308	.4574 ± .0313	.4559 ± .0310	.4528 ± .0316
RSC	**.3753 ± .0480**	.3748 ± .0472	.3733 ± .0467	.3746 ± .0465	.3722 ± .0466	.3722 ± .0462	.3721 ± .0465
TOS	**.5374 ± .0474**	.5361 ± .0460	.5362 ± .0461	.5357 ± .0474	.5365 ± .0456	.5359 ± .0452	.5355 ± .0463
IPS	**.3952 ± .0410***	.3931 ± .0407	.3906 ± .0406	.3889 ± .0399	.3874 ± .0404	.3867 ± .0399	.3859 ± .0400
FEF	**.2550 ± .0523***	.2536 ± .0519	.2510 ± .0512	.2479 ± .0501	.2477 ± .0503	.2472 ± .0499	.2452 ± .0492

Supplementary Table 2: Prediction scores for the motion-energy model estimated with SPIN-VM in functional ROIs. The prediction scores are reported as mean ± SEM across five subjects. Results are given for seven different window sizes. The highest prediction score in each row is annotated in bold font. If the highest prediction score in a row is significantly greater than all the other prediction scores in the same row, it is also marked with a * ($p < 0.05$, Bootstrap test).

	Gaussian	**Average**	**LoG**
Whole cortex	**.1795 ± .0279**	.1792 ± .0277	.1788 ± .0277
RET	.2826 ± .0170	.2833 ± .0169	**.2844 ± .0171**
V4	.3288 ± .0180	.3300 ± .0175	**.3307 ± .0174**
V7	**.5309 ± .0388**	.5291 ± .0388	.5264 ± .0396
FFA	.7227 ± .0301	**.7232 ± .0303**	.7223 ± .0305
EBA	**.7420 ± .0171**	.7418 ± .0171	.7408 ± .0175
MT	**.7263 ± .0137**	.7258 ± .0139	.7248 ± .0140
LOC	**.6599 ± .0415**	.6594 ± .0405	.6583 ± .0409
PPA	**.4839 ± .0458**	.4837 ± .0466	.4811 ± .0465
RSC	**.4732 ± .0589**	.4727 ± .0597	.4712 ± .0597
TOS	.5127 ± .0451	**.5148 ± .0448**	.5134 ± .0450
IPS	**.3567 ± .0355**	.3564 ± .0358	.3551 ± .0362
FEF	**.2283 ± .0408**	.2282 ± .0411	.2272 ± .0411

Supplementary Table 3: Prediction scores for the category model estimated with SPIN-VM in functional ROIs. The prediction scores are reported as mean ± SEM across five subjects. Results are given for three different types of graph Laplacian filters: Gaussian, average, and LoG. The highest prediction score in each row is annotated in bold font.

	Gaussian	**Average**	**LoG**
Whole cortex	**.2080 ± .0097**	.2078 ± .0097	.2073 ± .0097
RET	**.6521 ± .0285**	.6509 ± .0287	.6513 ± .0284
V4	**.6240 ± .0197**	.6232 ± .0201	.6229 ± .0197
V7	**.6663 ± .0203**	.6656 ± .0202	.6647 ± .0204
FFA	**.6528 ± .0354**	.6505 ± .0361	.6484 ± .0363
EBA	**.6899 ± .0231**	.6893 ± .0230	.6886 ± .0231
MT	**.6941 ± .0140**	.6936 ± .0138	.6932 ± .0139
LOC	**.6583 ± .0176**	.6580 ± .0174	.6566 ± .0174
PPA	**.4642 ± .0309**	.4629 ± .0306	.4621 ± .0305
RSC	.3753 ± .0480	**.3774 ± .0480**	.3749 ± .0471
TOS	**.5374 ± .0474**	.5362 ± .0475	.5371 ± .0467
IPS	**.3952 ± .0410**	.3938 ± .0407	.3931 ± .0408
FEF	**.2550 ± .0523**	.2544 ± .0515	.2544 ± .0519

Supplementary Table 4: Prediction scores for the motion-energy model estimated with SPIN-VM in functional ROIs. The prediction scores are reported as mean ± SEM across five subjects. Results are given for three different types of graph Laplacian filters: Gaussian, average, and LoG. The highest prediction score in each row is annotated in bold font.

	SPIN-VM	VM	smooth-VM
Whole cortex	**.1795 ± .0279***	.1733 ± .0283	.1741 ± .0281
RET	**.2826 ± .0170***	.2626 ± .0140	.2668 ± .0161
V4	**.3288 ± .0180***	.3051 ± .0166	.3077 ± .0168
V7	**.5309 ± .0388***	.5038 ± .0367	.5081 ± .0367
FFA	**.7227 ± .0301***	.7146 ± .0318	.7152 ± .0317
EBA	**.7420 ± .0171***	.7285 ± .0205	.7289 ± .0197
MT	**.7263 ± .0137***	.7083 ± .0177	.7092 ± .0175
LOC	**.6599 ± .0415***	.6412 ± .0414	.6396 ± .0406
PPA	**.4839 ± .0458***	.4604 ± .0421	.4641 ± .0429
RSC	**.4732 ± .0589***	.4614 ± .0590	.4644 ± .0594
TOS	**.5127 ± .0451**	.5006 ± .0448	.5093 ± .0455
IPS	**.3567 ± .0355***	.3410 ± .0325	.3429 ± .0331
FEF	**.2283 ± .0408***	.2133 ± .0386	.2148 ± .0385

Supplementary Table 5: Prediction scores for the category model estimated with SPIN-VM, VM, and smooth-VM, in functional ROIs. The prediction scores are reported as mean ± SEM across five subjects. Bold font is used to annotate the cases in which SPIN-VM achieves significantly higher prediction scores than VM ($p < 0.05$, Bootstrap test). If SPIN-VM also yields significantly higher prediction scores than smooth-VM, it is marked with a * ($p < 0.05$).

	SPIN-VM	**VM**	**smooth-VM**
Whole cortex	**.2080 ± .0097***	.2015 ± .0091	.2024 ± .0095
RET	**.6521 ± .0285***	.6403 ± .0281	.6399 ± .0293
V4	**.6240 ± .0197***	.6070 ± .0192	.6072 ± .0197
V7	**.6663 ± .0203***	.6570 ± .0197	.6596 ± .0194
FFA	**.6528 ± .0354***	.6395 ± .0354	.6408 ± .0348
EBA	**.6899 ± .0231***	.6765 ± .0230	.6775 ± .0236
MT	**.6941 ± .0140***	.6826 ± .0129	.6846 ± .0135
LOC	**.6583 ± .0176***	.6462 ± .0175	.6464 ± .0180
PPA	**.4642 ± .0309***	.4394 ± .0291	.4362 ± .0296
RSC	**.3753 ± .0480***	.3637 ± .0449	.3597 ± .0457
TOS	**.5374 ± .0474***	.5337 ± .0473	.5312 ± .0495
IPS	**.3952 ± .0410***	.3851 ± .0391	.3877 ± .0400
FEF	**.2550 ± .0523***	.2459 ± .0488	.2459 ± .0491

Supplementary Table 6: Prediction scores for the motion-energy model estimated with SPIN-VM, VM, and smooth-VM, in functional ROIs. The prediction scores are reported as mean ± SEM across five subjects. Bold font is used to annotate the cases in which SPIN-VM achieves significantly higher prediction scores than VM ($p < 0.05$, Bootstrap test). If SPIN-VM also yields significantly higher prediction scores than smooth-VM, it is marked with a * ($p < 0.05$).

	SPIN-VM	VM	smooth-VM
Whole cortex	**.7382 ± .0080**	.6549 ± .0074	.7328 ± .0102
RET	.8636 ± .0091	.8359 ± .0120	**.8888 ± .0149***
V4	.8507 ± .0123	.8154 ± .0159	**.8659 ± .0163**
V7	**.8458 ± .0150***	.8063 ± .0228	.8228 ± .0166
FFA	**.8690 ± .0105***	.8395 ± .0144	.8537 ± .0084
EBA	**.8223 ± .0120***	.7793 ± .0130	.7597 ± .0177
MT	**.8247 ± .0162***	.7803 ± .0212	.7673 ± .0198
LOC	**.8645 ± .0190***	.8314 ± .0247	.8201 ± .0253
PPA	**.8696 ± .0074***	.8228 ± .0099	.8445 ± .0084
RSC	**.8889 ± .0119***	.8463 ± .0132	.8713 ± .0144
TOS	**.8544 ± .0103**	.8033 ± .0115	.8369 ± .0242
IPS	**.8293 ± .0099**	.7611 ± .0186	.8125 ± .0209
FEF	.8065 ± .0133	.7240 ± .0252	**.8172 ± .0236**

Supplementary Table 7: Local coherence values for the category model in functional ROIs (mean ± SEM) obtained based on model weights estimated with SPIN-VM, VM, and smooth-VM. The highest local coherence value in each row is annotated in bold font. If this value is significantly higher than both of the alternatives, it is marked with a * ($p < 0.05$, Bootstrap test).

	SPIN-VM	VM	smooth-VM
Whole cortex	**.7848 ± .0051**	.6887 ± .0072	.7721 ± .0134
RET	**.8699 ± .0059***	.8258 ± .0089	.8358 ± .0084
V4	**.8934 ± .0078***	.8491 ± .0119	.8671 ± .0063
V7	**.9028 ± .0072***	.8741 ± .0124	.8858 ± .0095
FFA	**.8681 ± .0103***	.8082 ± .0152	.8253 ± .0088
EBA	**.8772 ± .0061***	.8291 ± .0073	.8180 ± .0125
MT	**.8906 ± .0028***	.8509 ± .0035	.8411 ± .0056
LOC	**.8947 ± .0062***	.8617 ± .0094	.8577 ± .0105
PPA	**.8788 ± .0050***	.7959 ± .0066	.8444 ± .0119
RSC	**.8973 ± .0104***	.8428 ± .0153	.8604 ± .0146
TOS	**.8980 ± .0086***	.8679 ± .0118	.8812 ± .0113
IPS	.8885 ± .0097	.8469 ± .0131	**.9044 ± .0134**
FEF	.8410 ± .0124	.7719 ± .0175	**.8711 ± .0234**

Supplementary Table 8: Local coherence values for the motion-energy model in functional ROIs (mean ± SEM) obtained based on model weights estimated with SPIN-VM, VM, and smooth-VM. The highest local coherence value in each row is annotated in bold font. If this value is significantly higher than both of the alternatives, it is marked with a * ($p < 0.05$, Bootstrap test).

	SPIN-VM	VM	smooth-VM
Whole cortex	.2072 ± .0140	.1949 ± .0136	**.2156 ± .0161***
RET	.3167 ± .0190	.2864 ± .0152	**.3179 ± .0192**
V4	.3677 ± .0202	.3293 ± .0188	**.3688 ± .0219**
V7	.5894 ± .0421	.5449 ± .0410	**.6025 ± .0418***
FFA	.7795 ± .0230	.7662 ± .0264	**.7856 ± .0235***
EBA	.7820 ± .0134	.7637 ± .0168	**.7893 ± .0136***
MT	.7709 ± .0100	.7436 ± .0140	**.7790 ± .0093***
LOC	.7078 ± .0394	.6780 ± .0385	**.7134 ± .0411***
PPA	.5429 ± .0536	.5064 ± .0496	**.5465 ± .0547**
RSC	.5106 ± .0641	.4907 ± .0644	**.5106 ± .0641**
TOS	.5582 ± .0522	.5359 ± .0527	**.5703 ± .0553***
IPS	.4125 ± .0434	.3787 ± .0370	**.4259 ± .0460***
FEF	.2671 ± .0477	.2441 ± .0442	**.2735 ± .0494**

Supplementary Table 9: Prediction scores with VM, smooth-VM, and SPIN-VM for the category model on smoothed test data. The prediction scores are reported as mean ± SEM across five subjects. Bold font is used to annotate the cases in which smooth-VM achieves significantly higher prediction scores than VM ($p < 0.05$, Bootstrap test). If smooth-VM also yields significantly higher prediction scores than SPIN-VM, it is marked with a * ($p < 0.05$).

	SPIN-VM	VM	smooth-VM
Whole cortex	.2375 ± .0117	.2231 ± .0101	**.2464 ± .0126***
RET	.7164 ± .0231	.6943 ± .0229	**.7286 ± .0200***
V4	.6928 ± .0195	.6656 ± .0186	**.7033 ± .0172***
V7	.7229 ± .0187	.7082 ± .0191	**.7361 ± .0191***
FFA	.7280 ± .0288	.7084 ± .0303	**.7380 ± .0279***
EBA	.7401 ± .0168	.7209 ± .0171	**.7442 ± .0164***
MT	.7366 ± .0068	.7189 ± .0070	**.7437 ± .0083***
LOC	.7213 ± .0166	.7034 ± .0154	**.7300 ± .0152***
PPA	.5261 ± .0381	.4890 ± .0352	**.5313 ± .0390***
RSC	.4163 ± .0579	.3940 ± .0510	**.4280 ± .0605***
TOS	.5675 ± .0536	.5566 ± .0537	**.5826 ± .0535***
IPS	.4401 ± .0494	.4185 ± .0455	**.4556 ± .0530***
FEF	.3020 ± .0617	.2820 ± .0559	**.3138 ± .0632***

Supplementary Table 10: Prediction scores with VM, smooth-VM, and SPIN-VM for the motion-energy model on smoothed test data. The prediction scores are reported as mean ± SEM across five subjects. Bold font is used to annotate the cases in which smooth-VM achieves significantly higher prediction scores than VM ($p < 0.05$, Bootstrap test). If smooth-VM also yields significantly higher prediction scores than SPIN-VM, it is marked with a * ($p < 0.05$).

	SPIN-VM	VM	smooth-VM
Whole cortex	.2123 ± .0146	.1967 ± .0139	**.2156 ± .0161**
RET	**.3231 ± .0207**	.2855 ± .0156	.3179 ± .0192
V4	**.3784 ± .0217**	.3263 ± .0190	.3688 ± .0219
V7	.6011 ± .0430	.5437 ± .0413	.6025 ± .0418
FFA	.7840 ± .0222	.7661 ± .0260	.7856 ± .0235
EBA	.7887 ± .0123	.7651 ± .0170	.7893 ± .0136
MT	.7795 ± .0087	.7447 ± .0135	.7790 ± .0093
LOC	.7173 ± .0378	.6784 ± .0393	.7134 ± .0411
PPA	.5525 ± .0558	.5065 ± .0507	.5465 ± .0547
RSC	.5142 ± .0663	.4915 ± .0643	.5106 ± .0641
TOS	.5687 ± .0533	.5391 ± .0535	.5703 ± .0553
IPS	.4267 ± .0483	.3824 ± .0371	.4259 ± .0460
FEF	.2720 ± .0495	.2461 ± .0436	.2735 ± .0494

Supplementary Table 11: Prediction scores with VM, smooth-VM, and SPIN-VM for the category model on smoothed validation and test data. The prediction scores are reported as mean ± SEM across five subjects. Both SPIN-VM and smooth-VM perform significantly better than VM in all twelve ROIs ($p < 0.05$). Bold font is used to annotate the cases in which there is a significant difference between SPIN-VM and smooth-VM ($p < 0.05$, Bootstrap test).

	SPIN-VM	**VM**	**smooth-VM**
Whole cortex	.2464 ± .0129	.2265 ± .0106	.2464 ± .0126
RET	.7287 ± .0217	.6991 ± .0221	.7286 ± .0200
V4	.7053 ± .0178	.6704 ± .0177	.7033 ± .0172
V7	.7315 ± .0199	.7102 ± .0193	**.7361 ± .0191**
FFA	.7310 ± .0295	.7119 ± .0299	**.7380 ± .0279**
EBA	**.7455 ± .0163**	.7225 ± .0173	.7442 ± .0164
MT	**.7451 ± .0083**	.7210 ± .0072	.7437 ± .0083
LOC	.7278 ± .0156	.7062 ± .0152	**.7300 ± .0152**
PPA	**.5385 ± .0402**	.4934 ± .0363	.5313 ± .0390
RSC	.4235 ± .0603	.3963 ± .0528	.4280 ± .0605
TOS	.5783 ± .0520	.5595 ± .0547	**.5826 ± .0535**
IPS	.4502 ± .0520	.4224 ± .0473	**.4556 ± .0530**
FEF	.3109 ± .0641	.2845 ± .0569	.3138 ± .0632

Supplementary Table 12: Prediction scores with VM, smooth-VM, and SPIN-VM for the motion-energy model on smoothed validation and test data. The prediction scores are reported as mean ± SEM across five subjects. Both SPIN-VM and smooth-VM perform significantly better than VM in all twelve ROIs ($p < 0.05$). Bold font is used to annotate the cases in which there is a significant difference between SPIN-VM and smooth-VM ($p < 0.05$, Bootstrap test).

References

Srinivasan, R., Golomb, J.D., Martinez, A.M., 2016. A Neural Basis of Facial Action Recognition in Humans. J. Neurosci. 36, 4434–4442. doi:10.1523/JNEUROSCI.1704-15.2016

Walbrin, J., Downing, P., Koldewyn, K., 2018. Neural responses to visually observed social interactions. Neuropsychologia 112, 31–39. doi:10.1016/j.neuropsychologia.2018.02.023