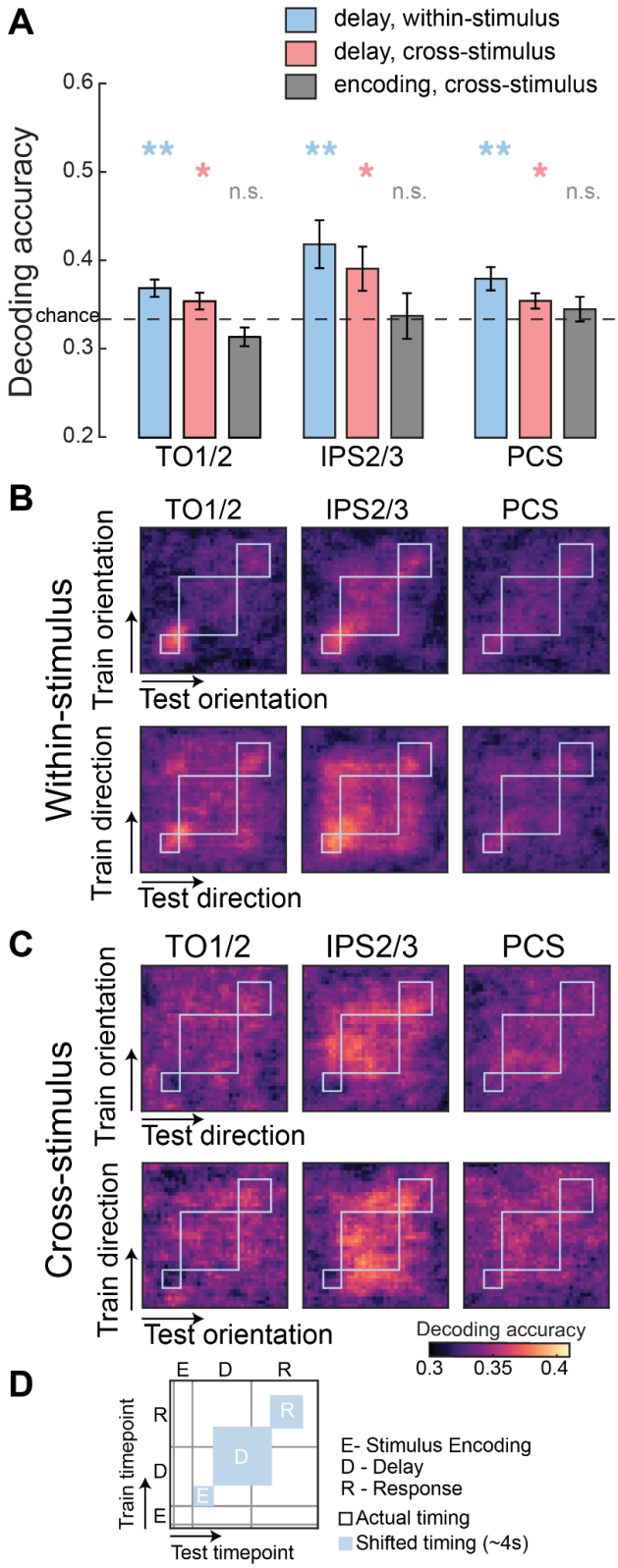

Figure S1. Working memory representations for orientation and motion direction share a common format, related to Figure 2. Decoding analysis was performed for other ROIs. **(A)** In WM, there are robust representations of orientation and motion direction information (blue bars), and these representations share a common format (red bars). The shared representational format is not observed during stimulus encoding (gray bars). **(B-C)** The temporal generalization matrix was generated by training and testing on each time point. **(D)** Schematics of the matrix plots. Gray lines denote the actual timing of events, and blue boxes show each of these events shifted by ~4s assuming hemodynamic lag. *p < 0.05, **p < 0.01, n.s. not significant, corrected (p values in Table S1). Error bars represent ±1 SEM.

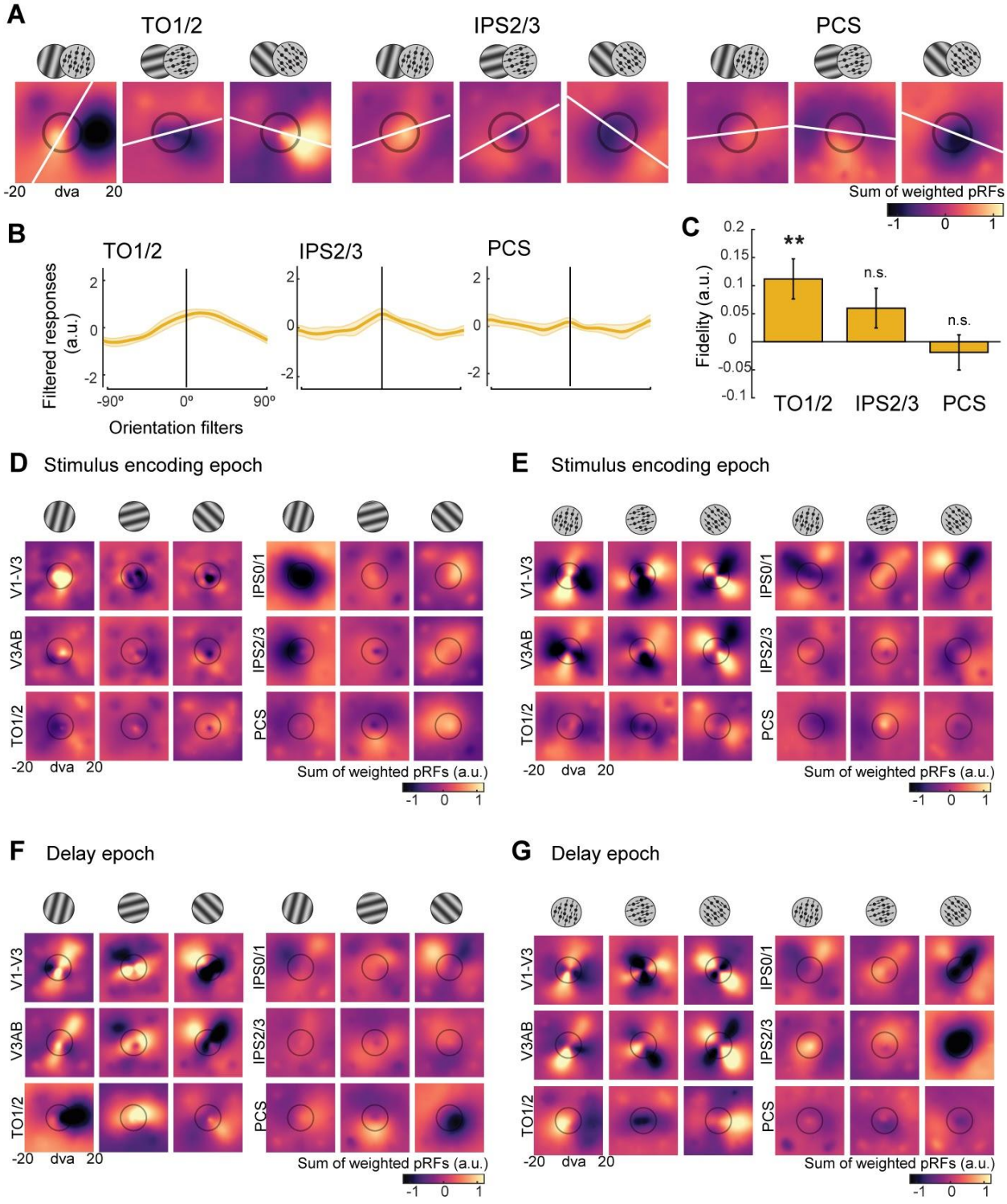

Figure S2. Unveiling the re-coded formats of working memory representations for orientation and motion direction, related to Figure 3. The spatial reconstruction analysis was performed for other ROIs.

(A) Population reconstruction maps for orientation and motion direction combined. Best fitting lines of the reconstruction are plotted in white. The size of each map is from -20° to 20° of visual angle (dva), and the stimulus size is shown for reference in black circles.
(B) Filtered responses from reconstruction maps in (A).
(C) Reconstruction fidelity calculated from filtered responses in (B). $^{**}p < 0.01$, n.s. not significant, corrected for multiple comparisons (*p* values in Table S1). Error bars represent ±1 SEM. For more information on filtered responses and fidelity, see STAR Methods.
(D-E) Reconstruction maps during the stimulus encoding epoch, separately for orientation and motion direction. For grating orientation, no line format is observed, likely due to a combination of orthogonal drifting motion and aperture bias during perception, alongside the emergence of orientation representation to be maintained in memory. For motion direction, we do observe a line-like format in some areas, mostly due to the emergence of motion direction representation to be maintained during the delay period.
(F-G) Reconstruction maps during the delay epoch, separately for orientation and motion direction. Both stimulus types were represented as a line format in WM across a wide range of brain regions.
For all figures, the size of each map is from -20° to 20° of visual angle (dva), and the stimulus size is shown for reference in black circles.

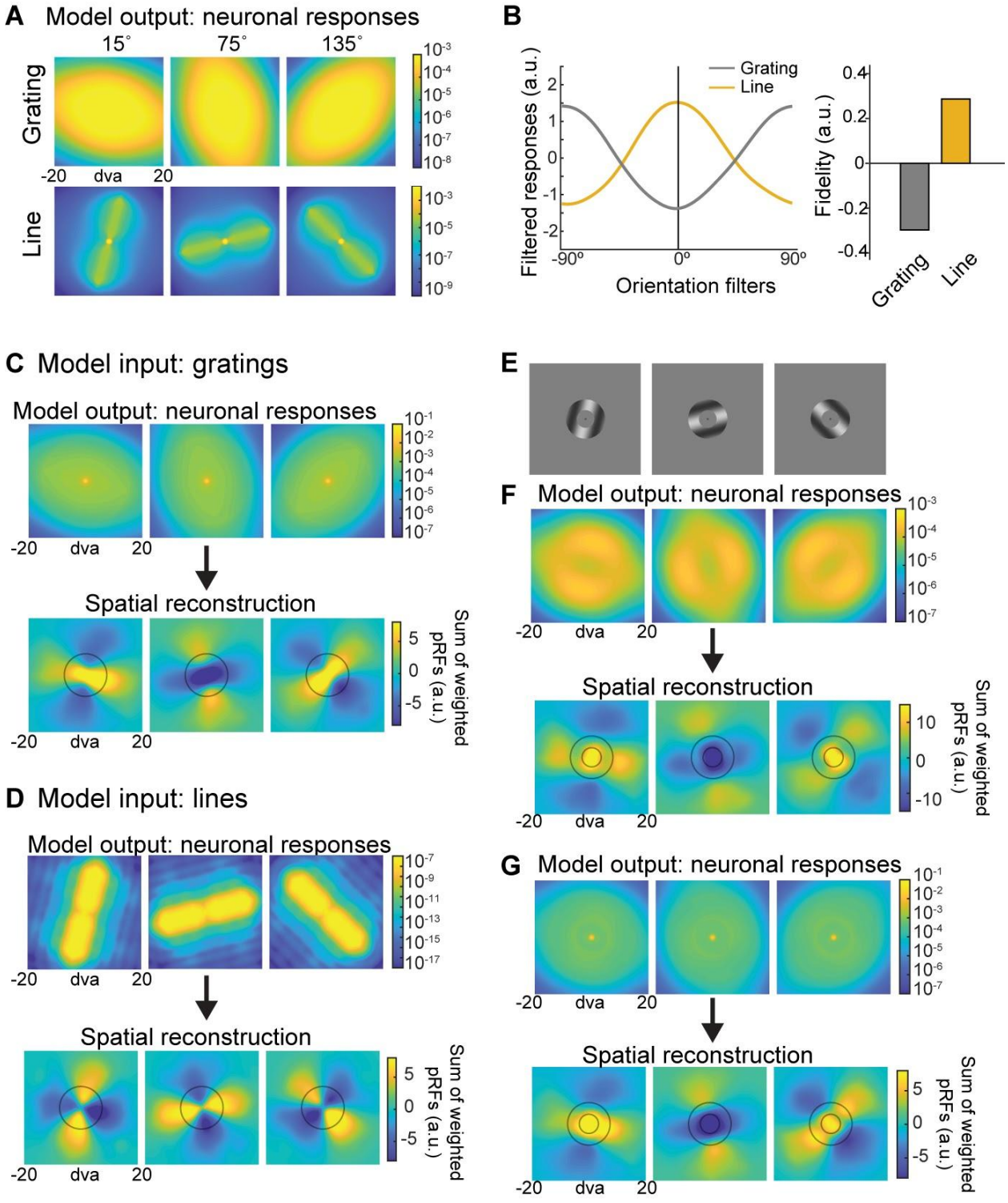

Figure S3. Spatial reconstruction analysis with simulated neuronal responses from the image-computable model of V1, related to Figure 4.

(A) The output neuronal responses from the image-computable model for grating and line images are plotted in log scale. Inputs to the model were grating and line images

oriented 15º, 75º, 135º clockwise from vertical (Figure 4A). For grating images, we present the sum of two subbands of which the center spatial frequencies are the closest to the spatial frequency of the grating stimulus. One of the two chosen subbands corresponds to the subband with the maximal response. For line images, we present the sum of all subbands to be more conservative. For completeness, we also present the sum of all subbands for grating images in (C) and the maximal subband for line images in (D) (See STAR Methods for details on the model). These model output neuronal responses were later used to simulate V1 voxel amplitudes by summing up the neuronal responses within each voxel's pRF. These simulated voxel responses were fed into the spatial reconstruction analysis in Figure 3A which resulted in the reconstruction maps in Figure 4B.

(B) Filtered responses and associated fidelity were computed from the spatial reconstruction maps in Figure 4B for quantification. The filtered responses when input to the model were line images (bottom row of (A) and Figures 4A-B) peak at the true orientation, suggesting that simulated V1 responses are highest along the actual remembered orientation (positive fidelity value). Filtered responses computed from the grating image reconstruction maps (top row of (A) and Figures 4A-B) show an opposite pattern (negative fidelity value). Since the output model responses were identical for all participants and only the pRF parameters differed, statistical analyses were not performed on the model simulations.

(C-D) Neuronal output responses and the reconstruction maps for the sum of all subbands when inputs to the model were grating images in (C) and for the maximal subband when inputs were simple line images in (D). These results resemble the model outputs in (A) and the reconstruction maps in Figure 4B. Therefore, our results are not an artifact of the chosen subband.

(E) For sanity check, we modified our grating stimulus images by increasing the size of the inner aperture (6.75º diameter; corresponding to the inner black circle in the reconstruction maps) to test whether stimulus aperture influences the spatial reconstruction maps. The model outputs and spatial reconstructions maps are shown in (F-G).

(F) The sum of two subbands with the spatial frequency centered on the spatial frequency of the grating stimulus. The spatial reconstruction maps exhibited a slightly different pattern from those in Figure 4B as they were a mixture of patterns parallel and orthogonal to the remembered orientation.

(G) The sum of all subband levels. A larger inner aperture did not change the pattern of reconstruction maps, as it was still orthogonal to the remembered orientation, similar to those from our original stimuli (Figure 4B).

All neuronal output response plots are presented in log scale. For all reconstruction maps, the size of each map is from -20° to 20° of visual angle (dva), and the stimulus size (inner and outer apertures) is shown in black circles for reference.

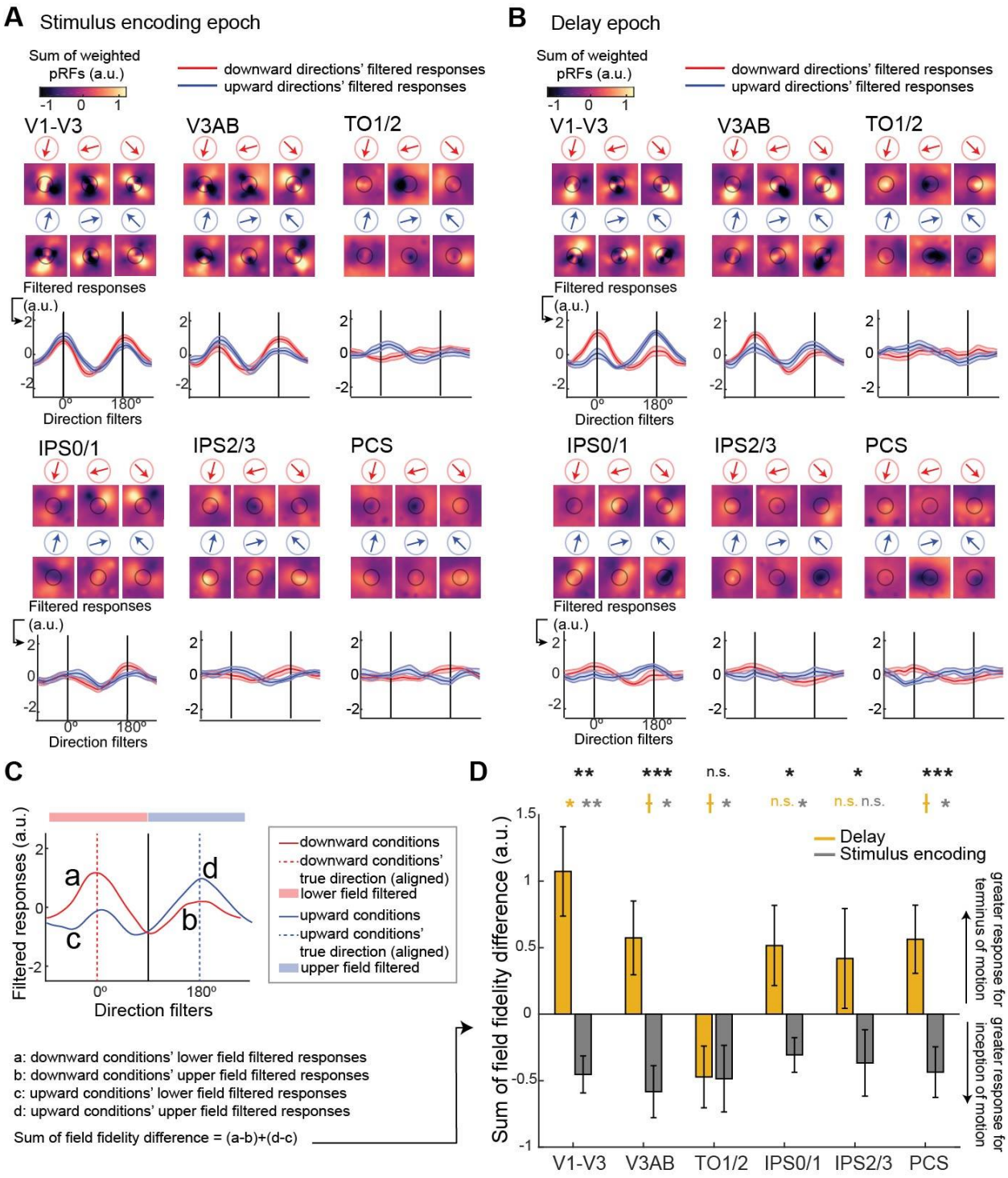

Figure S4. Distinct formats for perceptual and working memory representations of motion direction, related to Figures 3 and 4. Spatial reconstruction was performed on 6 direction conditions separately, as the previously reported aperture bias in perception makes different predictions for opposite motion directions with the same orientation axis (Wang et al., 2014).

(A-B) Reconstruction maps and associated filtered responses for the stimulus encoding in (A) and delay epochs in (B). The direction conditions from which the reconstruction maps are generated are shown in circles (red for downward directions and blue for upward directions). The size of all reconstructed maps is from -20° to 20° of visual angle. Each ROI's filtered response curves were generated by aligning and averaging the individual filtered responses for each of the 6 direction conditions. The first and second bumps at 0º and 180º in the filtered response curves correspond to the information captured by direction filters spanning the lower and upper visual field (360º of polar angle, to quantify whether the representation of remembered direction differs in strength between locations in the topographic map corresponding to the terminus and inception of motion direction). See (C) for a detailed description.
(C) Interpretation of filtered response curves in (A-B). Schematics is shown for the pattern of filtered response curves for most visual maps in the delay epoch in (B). For example, the red curve shows that memory representations for downward direction conditions form a line in topographic space corresponding to the orientation axis of the remembered direction (two bumps), but the portion of the line representation in the lower visual field (inception of moving dots for downward direction conditions; a) is stronger than that of the upper visual field (terminus of the moving dots for downward direction conditions; b). The blue curve shows that memory representations for upward direction conditions show higher response in the upper visual field (inception of moving dots for upward direction conditions; d), compared to the lower visual field (terminus of moving dots for upward direction conditions; c). We additionally computed the sum of field fidelity differences which quantifies the difference in activation between the portion of the topographic maps corresponding to the terminus and the inception of dot motion. Taking into account how the sum of field fidelity differences is computed ((a-b)+(d-c)), larger values indicate greater activation near the inception of moving dots compared to the terminus of moving dots.
(D) As previously reported during dot motion perception (Wang et al., 2014), the line-like patterns we observed early and time-locked to the visible motion stimulus (stimulus encoding epoch) were biased with more activation in the portion of the topographic map near the inception of the moving dots (gray bars and asterisks). During memory (delay epoch) we found the opposite; in most visual maps, there was greater activation near the terminus (yellow bars and asterisks). The difference between these biases during the stimulus encoding and delay epochs was statistically significant in many of the ROIs (black asterisks), demonstrating differences in the representational formats of perceiving and remembering dot motion. The only exception was TO1/2 where the reconstruction patterns observed during stimulus encoding and delay were similar. *p < 0.05, **p < 0.01, ***p < 0.001, n.s. not significant, corrected for multiple comparisons; †p < 0.05 uncorrected (p values in Table S1).

Supplementary Table

Table S1.

Decoding accuracy for the delay epoch[1].						
3-way ANOVA. Decoding type x ROI x Stimulus type						
Decoding type df: (1,10)	***F* = 7.061** ***p* = 0.017** **η_p^2 = 0.414**					
ROI df: (5,50)	***F* = 9.965** ***p* = 0** **η_p^2 = 0.499**					
Stimulus Type df: (1,10)	*F* = 4.001 *p* = 0.063 η_p^2 = 0.286					
Decoding type x ROI df: (5,50)	***F* = 3.039** ***p* = 0.020** **η_p^2 = 0.233**					
Decoding type x Stimulus Type df: (1,10)	*F* = 0.691 *p* = 0.439 η_p^2 = 0.065					
ROI x Stimulus Type df: (5,50)	***F* = 3.576** ***p* = 0.006** **η_p^2 = 0.263**					
Decoding type x ROI x Stimulus Type df: (5,50)	***F* = 2.987** ***p* = 0.023** **η_p^2 = 0.230**					
Post-hoc 2-way ANOVA. Decoding type x Stimulus type for each ROI (FDR thres = 0.005 for decoding type, 0 for stimulus type and decoding type x stimulus type)[2].						
	V1-V3	V3AB	TO1/2	IPS0/1	IPS2/3	PCS
Decoding type df: (1,10)	*F* = 5.505 *p* = 0.040 η_p^2 = 0.351	***F* = 13.948** ***p* = 0.005** **η_p^2 = 0.582**	*F* = 1.508 *p* = 0.232 η_p^2 = 0.131	*F* = 2.015 *p* = 0.194 η_p^2 = 0.168	*F* = 4.254 *p* = 0.065 η_p^2 = 0.298	*F* = 6.470 *p* = 0.022 η_p^2 = 0.393
Stimulus type df: (1,10)	*F* = 6.059 *p* = 0.049 η_p^2 = 0.377	*F* = 6.969 *p* = 0.028 η_p^2 = 0.411	*F* = 0.708 *p* = 0.405 η_p^2 = 0.066	*F* = 1.685 *p* = 0.220 η_p^2 = 0.144	*F* = 0.618 *p* = 0.428 η_p^2 = 0.058	*F* = 0.046 *p* = 0.832 η_p^2 = 0.005
Decoding type x Stimulus type df: (5,50)	*F* = 1.477 *p* = 0.263 η_p^2 = 0.129	*F* = 3.614 *p* = 0.088 η_p^2 = 0.265	*F* = 0.006 *p* = 0.941 η_p^2 = 0	*F* = 0.033 *p* = 0.865 η_p^2 = 0.003	*F* = 1.807 *p* = 0.195 η_p^2 = 0.153	*F* = 0.104 *p* = 0.762 η_p^2 = 0.010
One-sample *t* tests against chance (FDR thres = 0.006 for within-stimulus, 0.029 for cross-stimulus)[3].						

	V1-V3	V3AB	TO1/2	IPS0/1	IPS0/2	PCS
Within-stimulus df: 10	***t* = 4.138** ***p* = 0** ***d* = 1.354**	***t* = 4.880** ***p* = 0.001** ***d* = 1.196**	***t* = 3.537** ***p* = 0.005** ***d* = 1.324**	***t* = 3.591** ***p* = 0.004** ***d* = 1.471**	***t* = 3.117** ***p* = 0.006** ***d* = 1.066**	***t* = 3.432** ***p* = 0.003** ***d* = 1.083**
Cross-stimulus df: 10	***t* = 3.584** ***p* = 0.002** ***d* = 1.081**	***t* = 3.203** ***p* = 0.002** ***d* = 0.966**	***t* = 2.097** ***p* = 0.029** ***d* = 0.632**	***t* = 2.767** ***p* = 0.008** ***d* = 0.834**	***t* = 2.261** ***p* = 0.015** ***d* = 0.682**	***t* = 2.309** ***p* = 0.020** ***d* = 0.696**
3-way ANOVA. Decoding type x ROI x Stimulus type, after matching the training and testing procedures across decoding types.						
Decoding type df: (1,10)	***F* = 7.061** ***p* = 0.027** **η_p^2 = 0.414**					
ROI df: (5,50)	***F* = 9.965** ***p* = 0** **η_p^2 = 0.499**					
Stimulus Type df: (1,10)	*F* = 4.000 *p* = 0.077 η_p^2 = 0.286					
Decoding type x ROI df: (5,50)	***F* = 3.039** ***p* = 0.019** **η_p^2 = 0.233**					
Decoding type x Stimulus Type df: (1,10)	*F* = 0.691 *p* = 0.419 η_p^2 = 0.065					
ROI x Stimulus Type df: (5,50)	***F* = 3.576** ***p* = 0.007** **η_p^2 = 0.263**					
Decoding type x ROI x Stimulus Type df: (5,50)	***F* = 2.987** ***p* = 0.018** **η_p^2 = 0.230**					
Decoding accuracy for the stimulus encoding epoch[4].						
3-way ANOVA. Decoding type x ROI x Stimulus type.						
Decoding type df: (1,10)	***F* = 19.957** ***p* = 0.002** **η_p^2 = 0.666**					
ROI df: (5,50)	***F* = 6.825** ***p* = 0** **η_p^2 = 0.406**					

Stimulus type df: (1,10)	F = 1.126 p = 0.314 η_p^2 = 0.101	
Decoding type x ROI df: (5,50)	***F* = 10.876** ***p* = 0** **η_p^2 = 0.521**	
Decoding type x Stimulus type df: (1,10)	F = 1.503 p = 0.273 η_p^2 = 0.131	

Post-hoc 2-way ANOVA. ROI x Stimulus type for each decoding type (FDR thres = 0 for ROI, Stimulus type, ROI x Stimulus type)[5].

	Within-stimulus	Cross-stimulus				
ROI df: (5,50)	***F* = 15.011** ***p* = 0** **η_p^2 = 0.600**	F = 0.753 p = 0.605 η_p^2 = 0.070				
Stimulus type df: (1,10)	F = 1.319 p = 0.279 η_p^2 = 0.117	F = 0.015 p = 0.908 η_p^2 = 0.002				
ROI x Stimulus type df: (5,50)	***F* = 9.007** ***p* = 0** **η_p^2 = 0.474**	F = 2.192 p = 0.067 η_p^2 = 0.180				

One-sample t tests against chance for cross-stimulus decoding accuracy (FDR thres = 0)[6].

	V1-V3	V3AB	TO1/2	IPS0/1	IPS2/3	PCS
Cross-stimulus df: 10	t = 0.304 p = 0.378 d = 0.092	t = −0.057 p = 0.510 d = −0.017	t = −1.945 p = 0.956 d = −0.586	t = 0.264 p = 0.413 d = 0.079	t = 0.122 p = 0.473 d = 0.037	t = 0.772 p = 0.244 d = 0.233

Reconstruction fidelity for the delay epoch[7].

Orientation and direction trials combined (FDR thres = 0.006).

	V1-V3	V3AB	TO1/2	IPS0/1	IPS2/3	PCS
Fidelity df: 10	***t* = 16.597** ***p* = 0** ***d* = 5.004**	***t* = 9.742** ***p* = 0** ***d* = 2.937**	***t* = 3.133** ***p* = 0.006** ***d* = 0.945**	***t* = 4.379** ***p* = 0.001** ***d* = 1.320**	t = 1.686 p = 0.067 d = 0.508	t = −0.605 p = 0.278 d = −0.182

Reconstruction fidelity on the sum of field reconstruction fidelity differences, for the direction trials[8].

2-way ANOVA. Epoch x ROI.

	Sum of field fidelity differences	
Epoch df: (1,10)	***F* = 33.953** ***p* = 0** **η_p^2 = 0.773**	
ROI df: (5,50)	*F* = 1.990 *p* = 0.088 η_p^2 = 0.166	
Epoch x ROI df: (5,50)	***F* = 3.380** ***p* = 0.007** **η_p^2 = 0.253**	

One-sample *t* tests on sum of field reconstruction fidelity differences, separately for delay epoch and stimulus encoding epoch (FDR thres = 0.002 for delay epoch, 0.036 for stimulus encoding epoch).

	V1-V3	V3AB	TO1/2	IPS0/1	IPS2/3	PCS
Delay epoch df: 10	***t* = 3.190** ***p* = 0.002** ***d* = 0.962**	*t* = 2.068 *p* = 0.048 *d* = 0.623	*t* = −0.232 *p* = 0.030 *d* = −0.613	*t* = 1.710 *p* = 0.055 *d* = 0.516	*t* = 1.117 *p* = 0.147 *d* = 0.337	*t* = 2.194 *p* = 0.028 *d* = 0.662
Stimulus encoding epoch df: 10	***t* = −3.251** ***p* = 0.001** ***d* = −0.980**	***t* = −2.985** ***p* = 0.009** ***d* = −0.900**	***t* = −1.935** ***p* = 0.036** ***d* = −0.583**	***t* = −2.354** ***p* = 0.020** ***d* = −0.710**	*t* = −1.470 *p* = 0.079 *d* = −0.443	***t* = −2.288** ***p* = 0.017** ***d* = −0.690**

Paired-sample *t* tests on sum of field reconstruction fidelity differences: delay epoch vs stimulus encoding epoch (FDR thres = 0.015).

	V1-V3	V3AB	TO1/2	IPS0/1	IPS2/3	PCS
Difference between epochs df: 10	***t* = 3.558** ***p* = 0.002** ***d* = 1.073**	***t* = 5.293** ***p* = 0** ***d* = 1.596**	*t* = 0.069 *p* = 0.474 *d* = 0.021	***t* = 2.779** ***p* = 0.007** ***d* = 0.838**	***t* = 2.395** ***p* = 0.015** ***d* = 0.722**	***t* = 4.421** ***p* = 0** ***d* = 1.333**

Table S1. Non-parametric *p* values.

[1] Significant tests ($p < 0.05$, FDR corrected if applicable) are marked in bold. Factors are decoding type (within-/cross-stimulus), ROI (V1-V3, V3AB, TO1/2, IPS0/1, IPS2/3, PCS), and stimulus type (train on orientation/train on direction). To generate the null distribution, classifiers were trained on data shuffled across both the trial label and voxel dimensions. If applicable, *p* values were FDR corrected across the horizontal dimension of each table.

[2] A post-hoc 2-way ANOVA with decoding type and stimulus type as factors was conducted to examine whether the two stimulus types (train on orientation/train on direction) could be combined for further analyses.

[3] To quantify whether orientation/motion direction information could be decoded from ROIs (within-stimulus) and whether the two features share the same neural representation during WM (cross-stimulus), decoding accuracy, averaged across the orientation/direction trials, was compared against the null distribution.

[4] Significant tests ($p < 0.05$, FDR corrected if applicable) are marked in bold. Factors are decoding type (within-/cross-stimulus), ROI (V1-V3, V3AB, TO1/2, IPS0/1, IPS2/3, PCS), and stimulus type (train on orientation/train on direction). To generate the null distribution, classifiers were trained on data shuffled

across both the trial label and voxel dimensions. If applicable, *p* values were FDR corrected across the horizontal dimension of each table.

5 The main purpose of this analysis was to determine whether cross-stimulus decoding accuracies could be combined across the two stimulus types (train on orientation/train on direction). As there was neither the main effect of stimulus type nor an interaction for cross-stimulus decoding, data were combined for further analyses.

6 To examine whether the representational formats for orientation and direction were shared during stimulus encoding, cross-decoding accuracy averaged across the train-on-orientation and train-on-direction conditions was compared with the null distribution.

7 Significant tests ($p < 0.05$, FDR corrected if applicable) are marked in bold. To generate the null distribution, fidelity values were calculated from reconstruction maps computed from shuffled beta coefficients (shuffled across trial label and voxel dimensions). If applicable, *p* values were FDR corrected across the horizontal dimension of each table.

8 Significant tests ($p < 0.05$, FDR corrected if applicable) are marked in bold. For the null distribution, fidelity values were calculated from reconstruction maps generated from shuffled beta coefficients (shuffled across trial label and voxel dimensions). If applicable, *p* values were FDR corrected across the horizontal dimension of each table.