Science Advances
AAAS

Supplementary Materials for

Alpha-frequency feedback to early visual cortex orchestrates coherent naturalistic vision

Lixiang Chen *et al*.

Corresponding author: Lixiang Chen, lixiang.chen@fu-berlin.de; Daniel Kaiser, danielkaiser.net@gmail.com

Sci. Adv. **9**, eadi2321 (2023)
DOI: 10.1126/sciadv.adi2321

This PDF file includes:

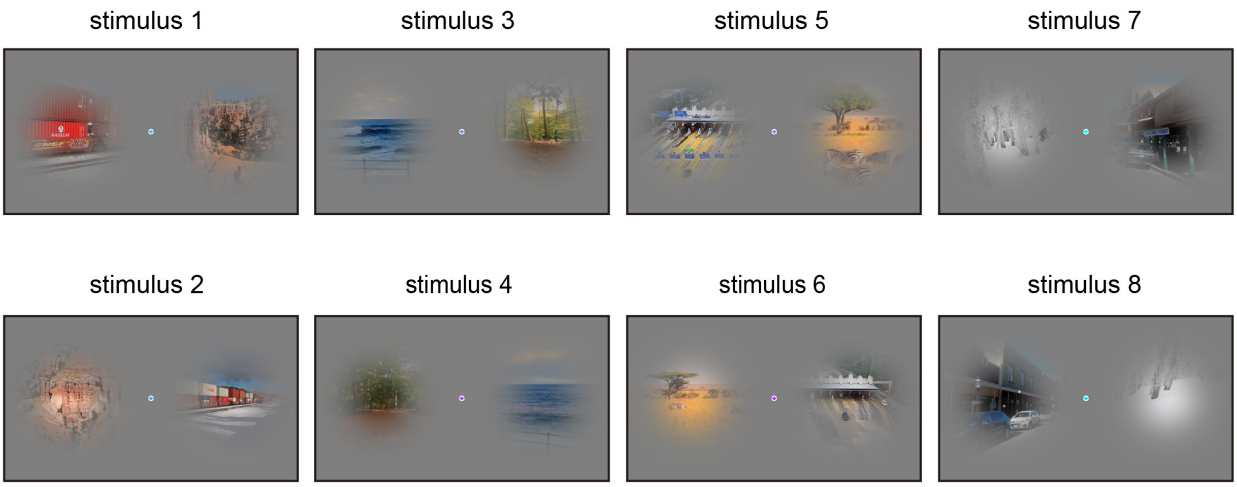

Fig. S1.
Still images from the incoherent video stimuli.

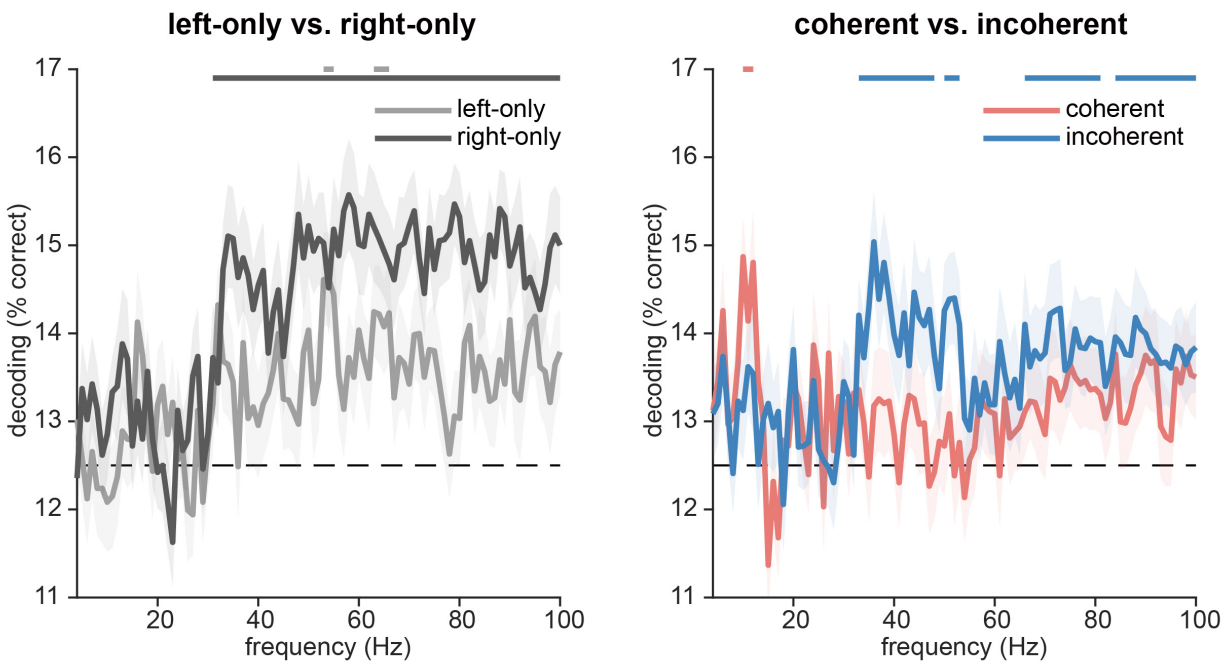

Fig. S2.

EEG frequency-resolved decoding analysis on spectral power patterns using sliding windows. We decoded between the eight video stimuli at each frequency from 4 to 100 Hz using a sliding window approach with a 5-frequency resolution, separately for each condition. The incoherent and single video stimuli were decodable from the γ frequency band, whereas coherent stimuli were decodable from the α frequency band. Line markers denote significant above-chance decoding ($p < 0.05$; FDR-corrected).

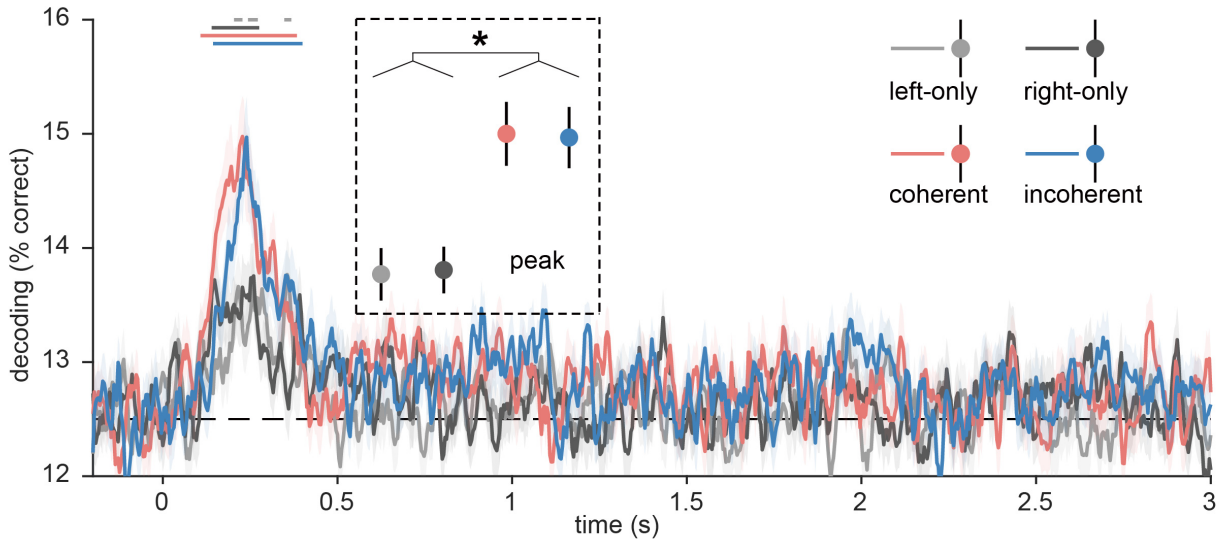

Fig. S3.

EEG time-resolved decoding on evoked response patterns. We performed decoding analysis on time-resolved broadband responses across channels to discriminate the eight video stimuli at each time from -200 ms to 3,000 ms relative to the onset of the stimulus, separately for each condition. The obtained decoding timeseries for each condition were smoothed by the moving average algorithm (6 time points). We extracted the peak decoding accuracy for each condition and then compared the decoding difference between conditions using paired t-tests. The results revealed a sustained representation of the video stimuli across the first 500 ms of processing, with stronger peak responses to two video conditions (coherent/incoherent) than to single video conditions (right-/left-only), but no differences between the coherent and incoherent conditions. Decoding onsets did not differ between the coherent and incoherent video stimuli (permutation test, $p = 0.176$). Error bars represent standard errors. Line markers denote significant above-chance decoding ($p < 0.05$; FDR-corrected). *: $p < 0.05$.

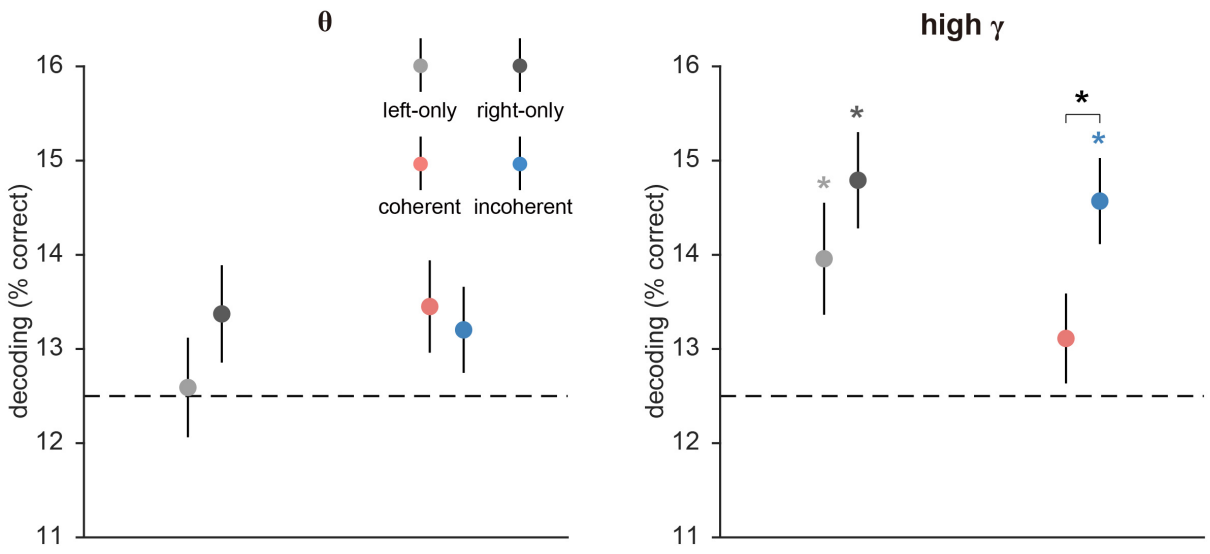

Fig. S4.

EEG decoding analysis on spectral power patterns separately for the theta (4–7 Hz) and high-gamma (71–100 Hz) frequency bands. For each frequency band (theta and high-gamma), we extracted the power of the frequencies included in that band across all channels from the power spectra, and then used the resulting patterns across channels and frequencies to classify the eight video stimuli in each condition. In the theta band, we did not find any significant above-chance decoding. In the high-gamma band, we found significant above-chance decoding for both single video stimuli and incoherent stimuli, but not for coherent stimuli. As in the 31-70 Hz gamma range, the incoherent stimuli were also decoded better than coherent stimuli. *: $p < 0.05$.

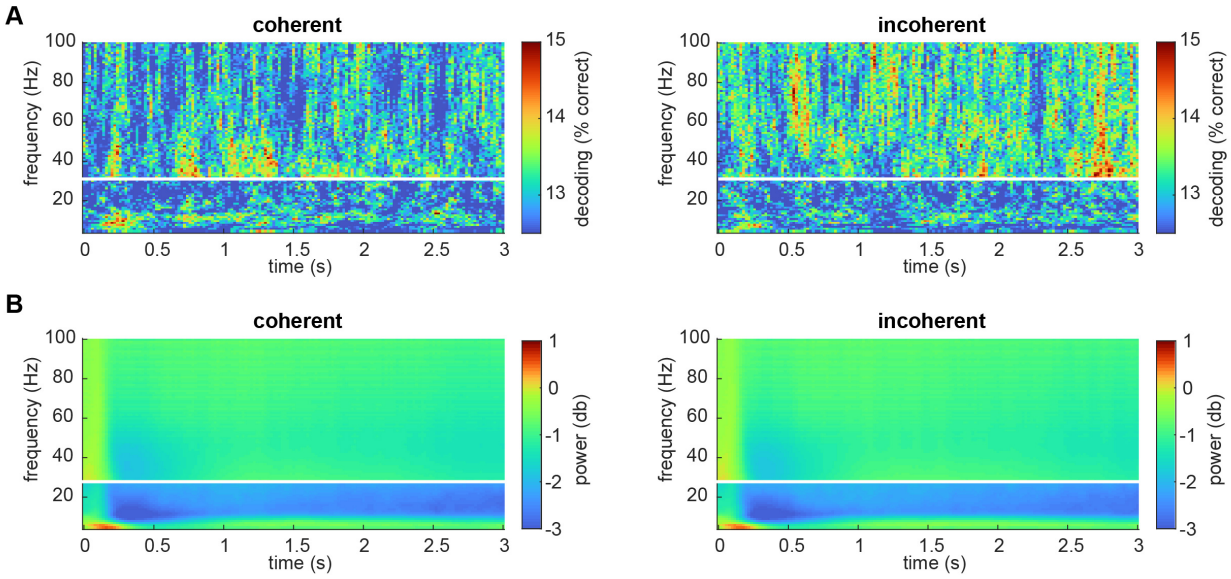

Fig. S5.

EEG time-frequency analysis. We first performed a time-frequency analysis to estimate the power at each frequency (4–100 Hz) and each time point (0–3 s) separately for each channel. As in the powerspectrum analysis (see Materials and Methods for details), we used single tapers for the low frequencies (4–30 Hz), and multitapers for high frequencies (31–100 Hz). **A)** We performed the decoding analysis at each time-frequency combination. We did not find significant differences in decoding performance between the coherent and incoherent conditions ($p < 0.05$; FDR-corrected; top two panels). The results suggested that we had insufficient statistical power for concurrently resolving the data across time and frequency, given the signal-to-noise ratio of our data. **B)** We transformed the power values to dB relative to the baseline to obtain the event-related spectral perturbation (ERSP). No significant differences were found in ERSP between the coherent and incoherent conditions ($p < 0.05$; FDR-corrected; bottom two panels). This suggests that the shift in representation from gamma to alpha dynamics is not accompanied by large-scale changes in the univariate spectral power over visual cortex.

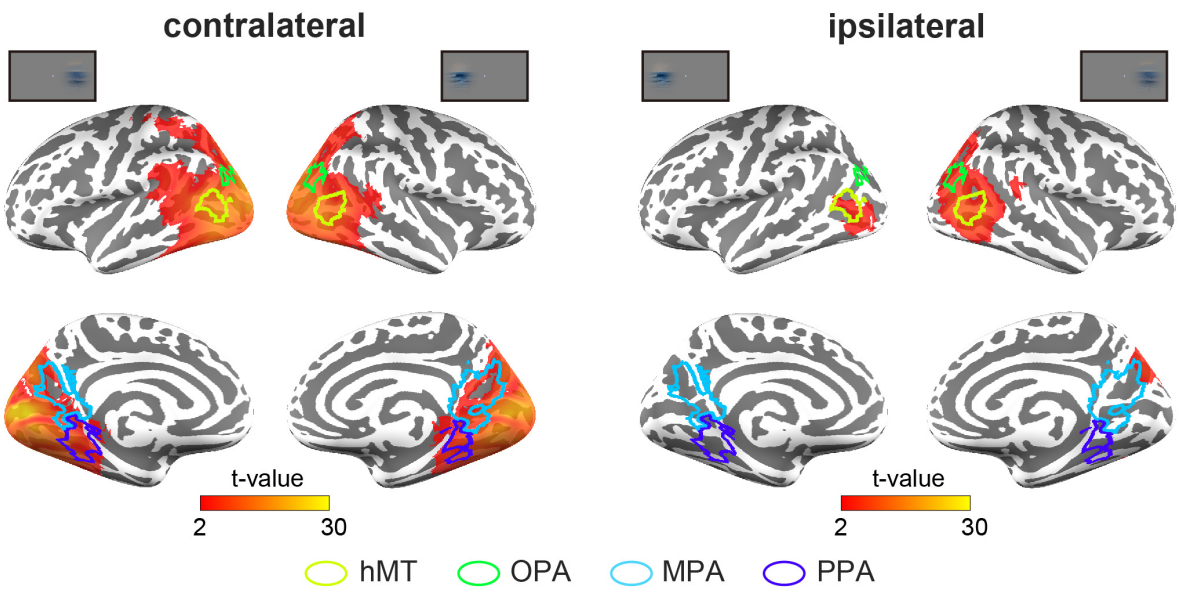

Fig. S6.

fMRI searchlight decoding analysis for right- and left-only conditions. Similar to the ROI decoding results, single video stimuli were decodable across the visual cortex in the contralateral hemisphere. In the ipsilateral hemisphere, the stimuli were primarily decodable in the parietal lobe including hMT. Multiple comparison correction was performed using GRF (voxel-level $p < 0.005$, cluster-extent $p < 0.05$).

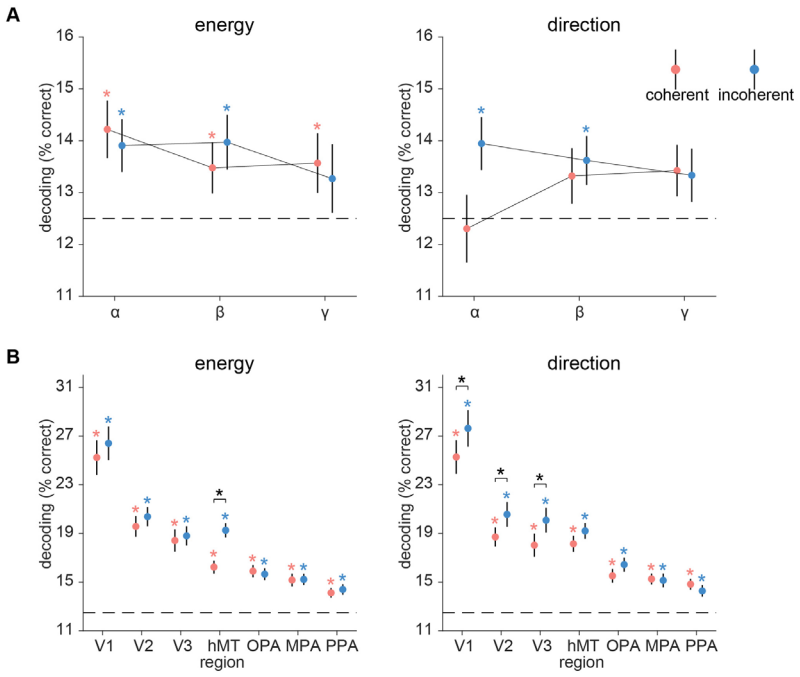

Fig. S7.

EEG and fMRI decoding analyses (grouping stimuli based on motion coherence). For each of the coherent and incoherent video stimuli, we first quantified inter-frame motion energy using the Motion Energy Analysis (MEA) software (https://psync.ch/mea/) separately for the left and right apertures, and then estimated its motion energy coherence by calculating the Pearson correlation between the left and right motion energy time-series. We split the stimulus set into two halves based on the values of motion energy coherence. Eight stimuli that are more coherent in motion energy were grouped into the motion coherent group and the other 8 stimuli were grouped into the motion incoherent group. In addition, we also grouped the stimuli based on motion direction coherence. For each stimulus, we estimated inter-frame optical flow using the Computer Vision Toolbox implemented in MATLAB and estimated its motion direction coherence by calculating the Pearson correlation between the mean motion direction time-series of the left and right apertures. Equivalent as described above, based on the values of motion direction coherence we split half the stimuli into a motion direction coherent and a motion direction incoherent group. **A)** We performed EEG frequency-resolved decoding analysis between the 8 stimuli in each frequency band (alpha, beta, gamma), separately for the newly formed coherent and incoherent groups. For the motion energy grouping, we found significant decoding for coherent stimuli in the alpha, beta, and gamma frequency bands, and significant decoding for incoherent stimuli in the alpha and beta frequency bands. For the motion direction grouping, the incoherent stimuli were decodable in the alpha and beta bands. **B)** We performed fMRI ROI decoding analysis to classify the stimuli in each ROI (V1, V2, V3, hMT, OPA, MPA, PPA), separately for the newly formed coherent and incoherent groups. For the motion energy grouping, both coherent and incoherent stimuli were decodable in all seven ROIs, and the incoherent stimuli were decoded better than coherent stimuli in the hMT. For the motion direction grouping, both coherent and incoherent stimuli were decodable in all the ROIs, and the incoherent stimuli were more decodable than coherent stimuli in V1, V2, and V3. *: $p < 0.05$ (FDR-corrected).

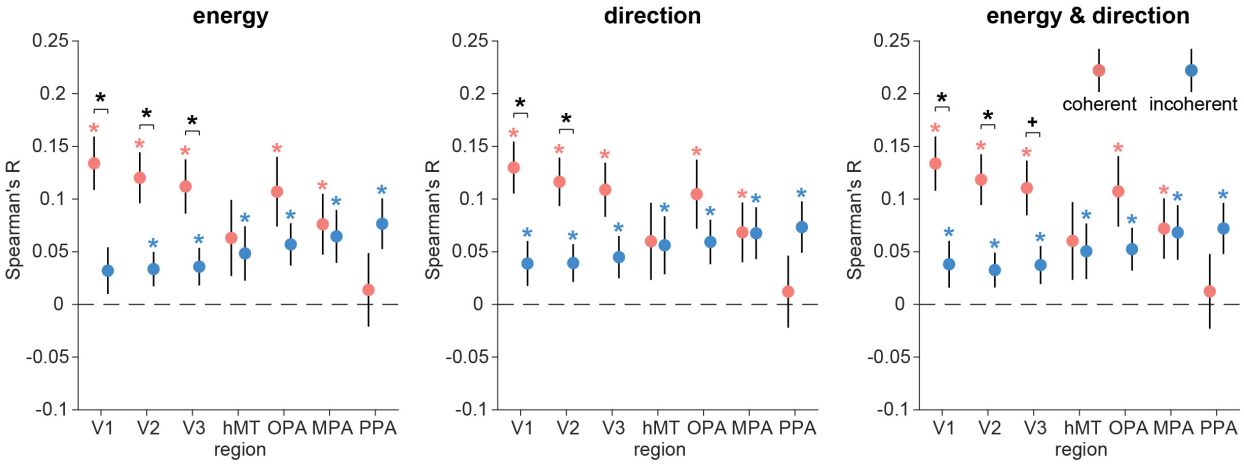

Fig. S8.

EEG-fMRI fusion analysis while controlling for motion coherence. For each of the coherent and incoherent video stimuli, we first quantified the inter-frame motion energy using the Motion Energy Analysis (MEA) software (https://psync.ch/mea/) separately for the left and right apertures, and then estimated its motion energy coherence by calculating the Pearson correlation between the left and right motion energy time-series. To construct the motion representational dissimilarity matrices (RDMs), we used the absolute difference in motion energy coherence as the distance between stimuli separately for coherent and incoherent conditions. In addition, we also constructed a motion RDM based on motion direction coherence. Specifically, for each stimulus, we estimated the inter-frame optical flow using the Computer Vision Toolbox (in MATLAB) and estimated its motion direction coherence by calculating the Pearson correlation between the mean motion direction time-series of the left and right apertures. Next, we constructed a motion direction RDM based on the absolute difference in motion direction coherence between stimuli separately for the coherent and incoherent conditions. In the EEG-fMRI fusion analysis, we calculated partial correlations between the participant-specific EEG RDMs in each frequency band (α, β, γ) and the group-averaged fMRI RDMs in each region (V1, V2, V3, hMT, OPA, MPA, PPA) that control for the motion energy RDM, motion direction RDM, or both motion RDMs, separately for the coherent and incoherent conditions. Similar to the main fusion results, we found that representations in the alpha band corresponded more strongly with representations in the early visual cortex when the videos were presented coherently, rather than incoherently, in all three analyses. Error bars represent standard errors. *: $p < 0.05$ (FDR-corrected), $^{+}$: $p < 0.05$ (uncorrected).

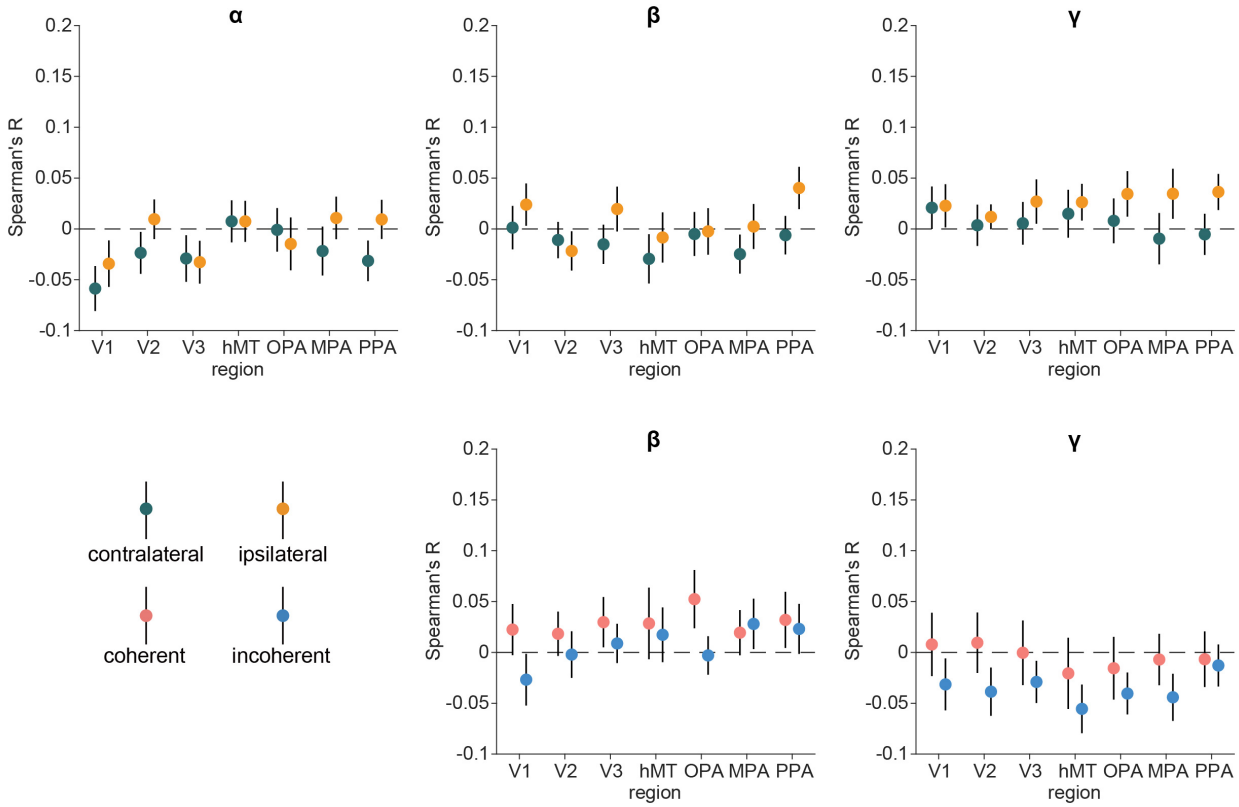

Fig. S9.

EEG-fMRI fusion analysis separately for each frequency band. For each condition, EEG representational dissimilarity matrices (RDMs) for each frequency band (α, β, γ) and fMRI RDMs for each region of interest (V1, V2, V3, hMT, OPA, MPA, PPA) were first obtained using pair-wise decoding analyses. To assess correspondences between spectral and regional representations, we calculated *Spearman*-correlations between the participant-specific EEG RDMs in each frequency band and the group-averaged fMRI RDMs in each region, separately for each condition. For the right- and left-only conditions, there was no significant correspondence between EEG responses in each frequency band and fMRI activations in each region, either for contralateral or ipsilateral presentations. For the coherent and incoherent conditions, there were no significant correspondences between β/γ responses and fMRI activations. Error bars represent standard errors.

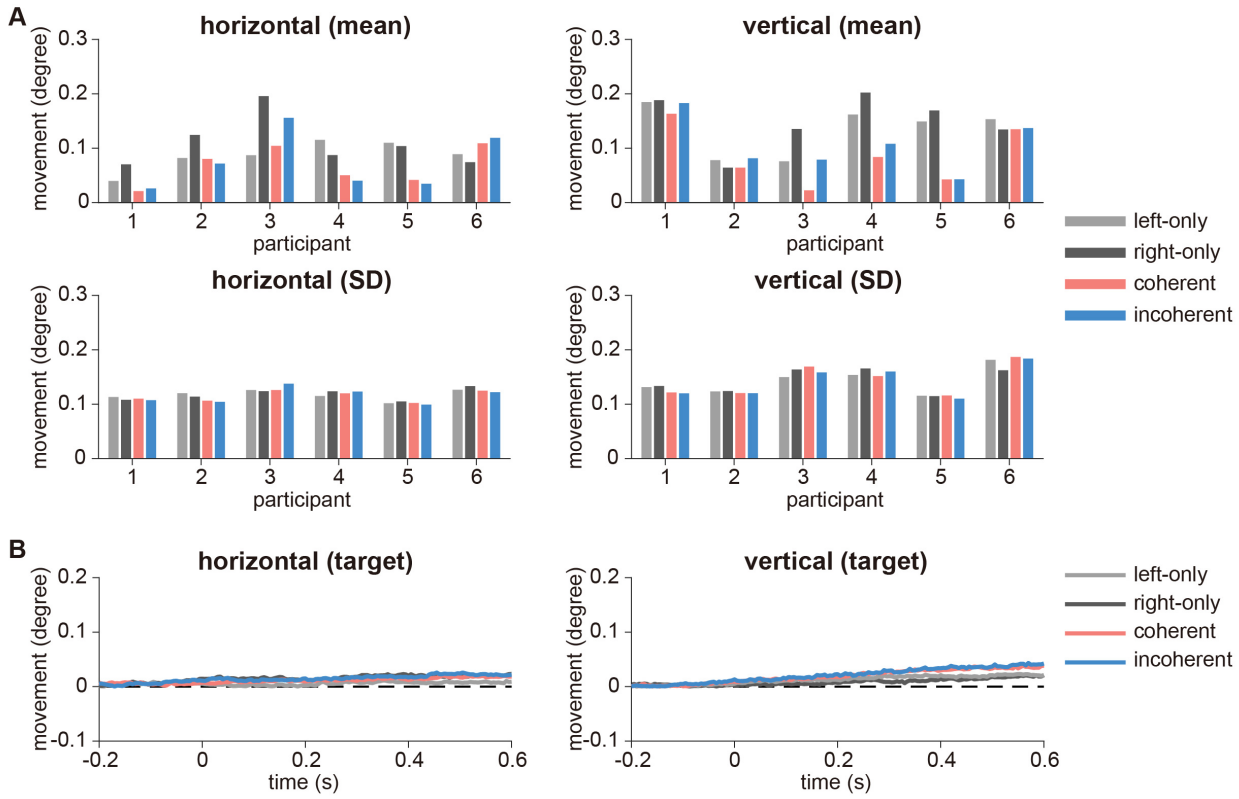

Fig. S10.

Eye-tracking data. Six participants (gender: 2 M/4 F, age: 26.5 ± 1.2 years) took part in the eye-tracking experiment using the same paradigm and the same stimuli as in the EEG and fMRI experiments. Eye movements were recorded monocularly (left eye) with an Eyelink 1000 Tower Mount (SR Research Ltd., Mississauga, Ontario, Canada) using the Psychophysics and Eyelink Toolbox extensions at 1000 Hz. Eye tracking data were segmented into epochs from -0.5 to 3.5 s relative to the onset of the stimulus, downsampled to 200 Hz, and baseline corrected. The data were then transformed from their original screen coordinate units (pixels) to visual angle units (degrees). **A)** To check fixation patterns during video presentation, we calculated the mean and standard deviation (SD) of the horizontal and vertical eye movement across time (0–3 s) in each trial and then averaged the mean and SD values across trials separately for each condition. For all participants, we found means of eye movement lower than 0.3 degrees (top two panels), and SDs of eye movement lower than 0.2 degrees (middle two panels), indicating stable central fixation. **B)** To determine whether eye movements occurred once the fixation color was detected, we extracted eye tracking data from -200 to 600 ms relative to the onset of the target color. We found no significant eye movement deviations after the target was presented, indicating that participants did not disengage from fixation after the target was presented.

Table S1.

Behavioral accuracy and response time for the color discrimination task in both EEG and fMRI experiments. Notes: Means ± standard errors (SE); N.A., not available.

	Left-only	Right-only	Coherent	Incoherent
Accuracy (EEG)	93.27 ± 1.70%	93.27 ± 1.58%	93.13 ± 1.66%	93.45 ± 1.71%
Accuracy (fMRI)	91.58 ± 1.18%	91.09 ± 1.19%	91.58 ± 1.24%	91.51 ± 1.21%
Response Time (EEG)	N.A.	N.A.	N.A.	N.A.
Response Time (fMRI)	521.9 ± 19.3 ms	521.7 ± 18.6 ms	525.2 ± 19.4 ms	526.8 ± 18.5 ms