nature human behaviour

Supplementary information https://doi.org/10.1038/s41562-025-02320-4

Mesoscale cortical mechanisms of perceptual conflict resolution in binocular rivalry

In the format provided by the authors and unedited

Supplementary Methods

Method S1. Detailed parameters for MRI data acquisition

MRI data were acquired on a 7T scanner (Siemens MAGNETOM) using a 32-channel receive single-channel transmit head coil (NOVA medical) in the Beijing MRI Center for Brain Research. A bite bar was used to reduce head motion. In Experiment 1, T2*w BOLD signals from the occipital and parietal cortices were acquired with a 2D GE-EPI sequence (0.8 mm isotropic voxels, 31 oblique-coronal slices, FOV = 128×128 mm2, TE = 23 ms, TR = 2000 ms, nominal flip angle = 80°, bandwidth = 1157 Hz/pixel, partial Fourier = 6/8, GRAPPA = 3; except for S01, 0.845×0.845×0.85 mm3, 24 oblique-coronal slices, FOV = 120×120 mm2, TE = 24.6 ms, bandwidth = 927 Hz/pix). The author C.Q. was also scanned with a 2D passband bSSFP sequence to acquire T2w BOLD signals (voxel size 0.5 × 0.5 × 1.5 mm3, 2 oblique-coronal slices, FOV = 96 × 96 mm2, volume acquisition time = 2400 ms for localizer and 1600 ms for rivalry/replay, TR = 5.64 ms, TE = 2.82 ms, nominal flip angle = 29° or 30°, bandwidth = 521 Hz/pix, no in-plane acceleration for localizer and GRAPPA = 2 for rivalry/replay). 3D passband bSSFP sequence (0.8-mm isotropic voxels, 10 oblique-coronal slices, FOV = 102 × 102 mm2, volume acquisition time = 6 s, TR = 5.54 ms, TE = 2.77 ms, nominal flip angle = 15°, bandwidth = 471 Hz/pix, partial Fourier = 7/8 in both phase and slice direction, GRAPPA = 2) was also used in a separate localizer session to evaluate the robustness of V1 ODC pattern. The same 2D GE-EPI sequence was used, albeit with different parameters in Experiment 2 (1.2 mm isotropic voxels, FOV = 180×180 mm2, TE = 22 ms, flip angle = 78°, bandwidth = 1587 Hz/pix, GRAPPA = 2. 62 oblique-transversal slices were acquired for 8 of 15 participants with a multi-band factor of 2. 31 slices were acquired for the rest of participants without multi-band acquisition). EPI volumes with reversed phase encoding and readout directions were also acquired for susceptibility distortion correction. For all Experiments, T1w anatomical volumes were acquired using an MP2RAGE sequence (0.7-mm isotropic voxels, FOV = 224×224 mm2, 256 sagittal slices, TE = 3.05 ms, TR = 4000 ms, TI1 = 750 ms, flip angle = 4°, TI2 = 2500 ms, flip angle = 5°, bandwidth = 240 Hz/pix, partial Fourier = 7/8, GRAPPA = 3).

Method S2. Simulation experiment for pattern correlation analysis (correspond to Extended Data Fig. 6a-f)

The absolute value of the instantaneous correlation coefficient is very small (on the order of 0.1) because a single submillimeter fMRI measurement is very noisy, and each r value is derived from only a single volume (although the ODC map is an average of 4-6 runs or about 1000 volumes). Only a small amount of variance in the activation pattern at a single time point may be accounted for by the ODC structure, the rest is largely thermal noise. However, the shape (or the width) of the r distribution is an aggregated quantity from many measurements and can become more reliable as we collect more data, even though the absolute r values in the distribution will remain low. The key issue here is, in such a high-noise regime, if the synchronization level in the instantaneous activation pattern changed across conditions or layers, does the pattern correlation method have enough (statistical) power to detect that change. The answer is yes, as shown in the following simulation.

We simulated an fMRI experiment with 12 participants, each had 512 volumes per condition with 2-s TR, which is identical to Experiment 1. The stimulated V1 comprised 500 left-eye voxels and 500 right-eye voxels. For each voxel, the BOLD signal was simulated by a variable length boxcar train whose intervals were sampled from a gamma distribution with a mean between 5-10 s, convolved with a canonical hemodynamic function and then added with Gaussian white noise (to simulate thermal noise which is dominant in 7T). Two types of asynchrony were simulated. Most of the time, the voxel was in the activated state when its preferred eye took dominance and rested at zero if not. But in any given interval, a small fraction of voxels was randomly chosen to be in the opposite state (e.g., rest at zero even in its preferred interval). For the "Sync" condition, 20% of voxels were set to be out-of-phase at any interval (in other words, each voxel would be out-of-phase 20% of the time). For the "Async" condition, this fraction was 30%. We also simulated another type of asynchronization, in which the transition of the voxel activation state is not exactly locked to the transition of the perceptual state but has a small temporal jitter. For each voxel and each transition, the jitter was independently sampled between 0-2 s uniformly, either negative or positive, so the mean absolute jitter was 1 s. We then computed the per-TR correlation coefficient across space between the instantaneous BOLD activation pattern and the ODC pattern (i.e., the eye preference of each voxel) and constructed the r distribution, the same way as in Experiment 1. The signal to noise ratio was chosen to be 0.1, so that the distribution looks similar to (actually even narrower than) our experiment result (Figure 3b), in a regime where the typical correlation coefficient was smaller than 0.1 (Extended Data Fig. 6a). Nevertheless, the difference between Sync and Async conditions in r-distribution width was highly significant (Extended Data Fig. 6b).

If we repeat the above simulated experiment many times (we choose 200 here to reduce computation burden) where we know there is an effect, and count the number of times that a significant result is detected, we can assess the power of this method. We asked how much asynchrony there must be for the pattern correlation method to be able to detect it at least 80% of the time (i.e., with >80% power), by systematically varying the asynchronization level in the Async condition. The results showed that the difference in asynchrony was reliably detected (>80% power) when the mean absolute temporal jitter was 1.5 s (Extended Data Fig. 6c), or the percentage of out-of-phase voxels increased 5 points (Extended Data Fig. 6d), compared with the baseline condition (Sync) with no jitter and 20% out-of-phase voxels. The asynchrony detection power also increased with the number of TRs (Extended Data Fig. 6e) and the number of voxels in the activation pattern (Extended Data Fig. 6f). For Async condition with 1 s jitter and 30% out-of-phase voxels, 80% power was achievable using just 40 TRs with 1000 voxels or 180 voxels with 512 TRs, whereas in Experiment 1 we used 512 TRs and 527 V1 voxels on average.

In summary, simulation suggests that by comparing difference in r-distribution width, our pattern correlation analysis has enough power to detect moderate change in BOLD response synchrony, even if the signal to noise ratio is low and the instantaneous correlation coefficient is on the order of 0.1.