

Supplemental Experimental Procedures

MRI Data Acquisition and Analysis

The task was performed during fMRI scanning using a Siemens 3T TIM Trio whole-body scanner with a 32-channel head coil. MRI scanning occurred at the Northwestern University Center for Translational Imaging, supported by the Northwestern University Department of Neurology. Head movement was minimized using padding. Visual stimuli were back-projected onto a screen and viewed through a mirror attached to the head coil. The resolution was 1280x1024 pixels. The projected display spanned 18.8° of visual angle vertically and 23.1° horizontally, at a refresh rate of 60 Hz. Responses were made with both hands using an optical, MRI-compatible response box (Current Designs, Philadelphia, PA). For fMRI, Learning trials lasted for 10-12 s ($M = 11$ s), with 4-12 s randomized inter-trial intervals (ISIs) ($M = 8$ sec), jittered in order to optimize estimated event-related signals for each condition. Test trials lasted 8.5 s, with 1-3 s ISIs ($M = 2$ sec). During ISIs, a gray fixation rectangle appeared every 4 s for a duration of 2 s. Subjects were instructed to blink twice at each appearance of the rectangle in order to provide an active baseline, and blinking was confirmed via eye-tracking.

Whole-brain BOLD EPI was used ($TR = 2000$ ms, $TE = 20$ ms, voxel size = $1.72 \times 1.72 \times 3$ mm³, $FOV = 22$ cm, flip angle = 80°). Learning phases included 308 volumes (10.3 min) and Test phases included 139 volumes (4.6 min), although only results from Learning are considered here. A structural image was acquired to provide anatomical localization (MPRAGE T_1 -weighted scans, $TR = 2400$ ms, $TE = 3.16$ ms, voxel size = 1 mm³, $FOV = 25.6$ cm, flip angle = 8°, 176 sagittal slices). All images were acquired in interleaved fashion in the axial plane starting at the top of the brain.

Preprocessing steps included motion correction, slice-timing correction to the first slice, functional/structural coregistration, stereotactic transformation using the Montreal Neurologic Institute (MNI) 305 template, and spatial smoothing with a 4-mm FWHM Gaussian kernel. We conducted group-level, random-effects analyses to identify regions significantly related to contextual information on a trial-by-trial level. Nuisance variables in the GLM fit to each subject included T_0 and T_1^* components of the MR signal and six-parameter movement estimates. Trials were modeled using a regressor of event onsets generated by convolving a boxcar function of 3-s duration with a canonical hemodynamic response function.

Eye-tracking Experimental Procedures

During fMRI acquisition, eye movements were recorded using an Eyelink 1000 remote tracking system (SR Research, Ontario, Canada) at a sampling rate of 500 Hz. The eye-tracking camera was focused on the right eye via the head coil-attached mirror used to view the projection screen. Continuous eye-movement records were transformed into a time series of fixations, saccades, and blinks using standard eye-tracking parameters. Motion (0.15 degrees), velocity (30 degrees per second), and acceleration (8000 degrees per second²) thresholds were used to identify saccades. Events in which the pupil size was very small were classified as blinks.

Saccade- and blink-free periods were categorized as fixations. The duration and time course of fixations in regions of interest (ROIs) corresponding to object and face locations were then imported and analyzed with custom scripts in Matlab (The MathWorks, Inc.). Fixations were assigned to the ROI with the nearest center within a maximum visual angle of 1.8° . Fixations not located within this distance from any ROI were discarded from further analysis.

We analyzed eye movements for Learning trials by considering the face-selection period. For the Active condition, the face-selection period was from stimulus onset until the face selection response was made (mean = 2.10 ± 0.33 s, sd.), whereas for the Passive condition, the entire 5-s face-selection period was considered. The face-selection period for each trial was subdivided into 20 equivalent intervals, and analyses were performed using the mean time spent fixating within ROIs for each interval (normalized for total viewing time within all ROIs combined).

Residual timeseries extraction

After estimating stimulus-evoked activity using an unconstrained impulse response model, the timeseries of residuals from the GLM for each voxel within each of the 6 ROIs were averaged spatially, separately for Active and Passive Learning blocks (for which BOLD images were collected separately). Residual timeseries were then band-pass filtered at 0.01 Hz – 0.1 Hz to remove linear trends and oscillations related to respiration and other nuisance factors. By using residual timeseries (i.e., with stimulus-evoked activity regressed out), we were able to assess connectivity due to sustained factors in the Active and Passive Learning blocks.