Supporting Information

Zhang and Kourtzi 10.1073/pnas.1002506107

SI Materials and Methods

Participants. We tested 28 observers (mean age = 23 y). None of the observers had previous experience with the stimuli or experimental procedure. All observers had normal or corrected to normal vision, gave written informed consent, and were paid for their participation. The study was approved by the local ethics committee. Seven observers were tested in each of four groups: supervised training on orthogonal contours, exposure to orthogonal contours, supervised training on collinear contours, and exposure to collinear contours. Similar to our previous study (1), two observers trained on orthogonal contours and two observers exposed to collinear contours did not reach 80% performance after training on the maximum number of sessions. One of the observers showed enhanced behavioral performance after exposure to orthogonal contours (in contrast with the rest of the observers that did not show any learning effects). That is, learning occurred in 100% of the observers after supervised training on collinear contours, 71.43% of the observers after supervised training on orthogonal contours, 71.43% of the observers after exposure to collinear contours, and 14.29% of the observers after exposure to orthogonal contours. The small number of outlier participants in each group could not support conclusive analysis, and therefore, the data from these participants were excluded. Finally, the data from one of the participants trained with collinear contours were excluded because of excessive head motion during scanning.

Stimuli. For training or exposure sessions, contour stimuli contained four embedded contours to maximize the effect of training in contour detection. For test sessions and functional MRI (fMRI) scans, contour stimuli contained two embedded contours to ensure that the learning effect transferred to a more difficult condition and related to the perception for global contours rather than local cues. The minimal distance between any element and their nearest neighbor was 0.5° of visual angle. The wavelength of the Gabor elements was 0.2°, and the SD of their Gaussian envelope was 0.1°. The entire field subtended 9.7° of visual angle. All elements had the same contrast, which varied between 29.7% and 84.1% (up to 0.75 octaves from 50%). This variation was relevant only for the exposure sessions, but it was present across all sessions (training, exposure, and test) to ensure that stimulus contrast was similar across sessions. We also generated random stimuli by shuffling the local orientation of all of the elements in the display. Contour and random stimuli were matched for the position of the Gabor elements. This ensured that the distribution of orientations in the random stimuli was the same as for the stimuli-containing contours. This procedure guaranteed that if there were any local density cues in the stimulus, they would be similar for contour and random stimuli.

The contours (path lengths: test sessions = 7.4°; training/exposure sessions = 9.0°) were defined by elements that could be aligned along the invisible contour path (collinear contours) or perpendicular to the path (orthogonal contours). The distance between contours was 4.4–5.3° in test sessions and 2–2.4° in training sessions. The global orientation of the embedded contours could either be near the left (135°) or the right (45°) diagonal, with an orientation offset chosen at random from within ±15° of the diagonal. Furthermore, in psychophysical sessions, the alignment of contour elements relative to their mean orientation was jittered by a random amount. In pretest and posttest sessions, this jitter was ±0°, ±15°, ±30°, or ±45°. In training and exposure sessions, the jitter was randomly chosen between ±0° and ±15° in increments of 3°. In the quick-test sessions after training, we tested stimuli only at 0° and 45° jitter.

Psychophysical Test Sessions. The stimuli were presented on a computer screen (resolution = 1,280 × 1,024) at a distance of 65 cm in a darkened room. On each trial, observers were required to maintain fixation on a black cross (0.18°) in the center of the screen. On the first day, all observers were given three brief familiarization runs before commencing with the pretest session. In the first run, contour elements were rendered in white, whereas background elements were in red. This design was used to explain the task. Subsequently, they performed a run in which auditory feedback was given on incorrect responses and stimulus presentation lasted 1,000 ms. The third practice run was also with auditory feedback, but stimulus presentation lasted 200 ms (same as the actual experiment). Practice runs lasted for ∼20–40 trials each until observers were confident about the task.

In the test sessions, observers performed a two-interval forced choice (2IFC) task. In each trial, after a 400-ms fixation interval, observers were presented with a sequence of two Gabor field stimuli each for 200 ms. The interstimulus interval was 500 ms during which only the fixation cross was presented. One stimulus interval contained the target contours, whereas the other one contained random elements. After the second stimulus, observers were required to indicate which interval contained the contours by pressing one of two buttons. The time limit of response was 5,000 ms after the second stimulus offset. If there was no response within the time limit, the trial was recorded as an error. This task was self-paced [i.e., the next trial was initiated immediately (beginning with the fixation period) after the behavioral response or after 5,000 ms].

The pretest and posttest session comprised 320 trials (40 trials per jitter level per orientation), with resting breaks every 32 trials. For each group trained with collinear or orthogonal contours, observers were tested with the trained contour type at both sets of orientations. Each quick-test session comprised 100 trials (50 trials per jitter level), with resting breaks every 20 trials. Observers were only presented with the contour type and orientation used for training or exposure during the quick-test sessions.

Psychophysical Supervised Training and Exposure Sessions. For each observer, only one stimulus condition (collinear or orthogonal contours) with one global orientation (45 ± 15° or 135 ± 15°) was presented in all training or exposure sessions. Each supervised training session contained 480 trials, with rest breaks every 16 trials, and the task and procedure were similar to that of the test sessions. For the exposure sessions, observers were instructed to judge whether the contrast was higher or lower (by pressing one of two buttons) or the same (by withholding their response) as the reference contrast of 50%. Observers were made aware of the reference contrast by an example stimulus (no embedded contours and random field only) that was displayed during the resting breaks (every 100 trials). The contrast of all elements was varied using a two-down, one-up staircase procedure that converged on 70.7% correct. Auditory feedback (tone frequency of 600 Hz and duration of 0.15 s) was given on incorrect responses. The duration of each stimulus was 200 ms, and the interstimulus interval was 800 ms.

fMRI Scanning Sessions. Observers were scanned two times, one time before training (after the pretest psychophysical session) and one time after training (after the posttest psychophysical session). Each scanning session comprised eight experimental runs, each of which lasted 5 min 20 s. A run comprised fourteen 16-s long stimulus blocks, including the initial and final blocks, during which only the fixation cross was presented. The experimental blocks contained stimuli from six conditions: collinear contours near the left ($135 \pm 15°$) diagonal, collinear contours near the right ($45 \pm 15°$) diagonal, orthogonal contours near the left ($135 \pm 15°$) diagonal, orthogonal contours near the right ($45 \pm 15°$) diagonal, random-1, and random-2. For each observer, the contour type presented in the scanner was the same as in the psychophysical sessions. Random-1 and random-2 were two conditions of random stimuli (random fields without any embedded contours). For each stimulus in condition random-1, the 10 elements corresponding to the contour elements were presented at random positions and orientations. The set of stimuli presented in condition random-2 was generated by rotating these elements by 90°. Thus, stimuli in the two random conditions differed by 90° rotation of the local elements that matched the orientation difference between contours near the left and right diagonal.

Each of the six stimulus conditions was presented three times (three blocks per condition) in a counterbalanced order across runs. For each block, 20 stimuli were presented for 200 ms each and separated by a 600-ms interstimulus interval. Observers performed a target-detection task that required them to attend to the stimuli similarly across all conditions. That is, observers were instructed to detect collinear contours at cardinal orientations (0° or 90°). For these target stimuli, the contours were rendered more salient by making the contours longer (i.e., subtending the whole of the Gabor field; that is, up to 9.7° of visual angle) and the interelement spacing smaller (0.4°). Two target stimuli were randomly interspersed within each block of stimuli, with the constraint that two target stimuli could not appear in consecutive trials. Performance in this task was at 70.90% for 414 ms mean response time. No significant differences in performance were observed across training procedures [$F(1,23) = 0.22$, $P = 0.65$], sessions [$F(1,23) = 0.16$, $P = 0.69$], or conditions [$F(5,115) = 0.75$, $P = 0.59$], ensuring that observers engaged similarly with the task across all conditions and sessions.

fMRI Data Acquisition. The experiments were conducted at the Birmingham University Imaging Centre using a 3-T Philips Achieva MRI scanner. T2*-weighted functional and T1-weighted anatomical ($1 \times 1 \times 1$-mm resolution) data were collected with an eight-channel SENSE head coil. Echo planar imaging data (gradient echo-pulse sequences) were acquired from 32 slices (whole-brain coverage: repetition time = 2,000 ms; echo time = 35 ms, $2.5 \times 2.5 \times 3$-mm resolution).

fMRI Data Analysis. ***Data preprocessing.*** MRI data were processed using Brain Voyager QX (Brain Innovation BV). T1-weighted anatomical data were used for coregistration, 3D cortex reconstruction, inflation, and flattening. Preprocessing of the functional data involved slice scan-time correction, 3D head-movement correction, temporal high-pass filtering (three cycles), and removal of linear trends. No spatial smoothing was performed on the functional data used for the multivariate analysis. The functional images were aligned to anatomical data under careful visual inspection, and the complete data transformed into Talairach space. For each participant, the functional imaging data between the two sessions (before and after training) were coaligned, registering all volumes of each subject to the first functional volume. This procedure ensured a cautious registration across sessions.

Regions of interest. To define brain regions that are involved in the processing of the contour stimuli used in our study, we tested for cortical areas that responded significantly stronger to the contour than random stimuli after training (because contour-detection performance was poor before training) using random-effects general linear model (GLM) across all observers (random effect analysis, $P < 0.001$, cluster-size threshold corrected, 80 mm^2). We then localized these contour-responsive regions in individual observers using data from both scanning sessions (fixed-effects GLM, $P < 0.05$, cluster-size threshold correction). We also identified retinotopic visual areas by using standard mapping procedures (2–4).

fMRI multivoxel pattern analysis. For each subject, voxels in each region of interest (ROI; contour-responsive regions and retinotopic areas) were ranked according to their response (t statistic) to all stimulus conditions compared with fixation across both scan sessions. To enable comparisons across ROIs and observers, we selected the average number of voxels across ROIs and observers that had the strongest response to stimulus conditions rather than fixation ($P < 0.05$). This procedure resulted in the selection of 140 voxels per ROI, comparable with the dimensionality used in previous studies (5, 6). If any ROI in a subject had less than 140 active voxels (14.29% of cases across subjects and ROIs), we selected all voxels in that region in further analysis. The time course of each selected voxel was z score normalized for each experimental run and shifted by 4 s to account for the hemodynamic delay. We then obtained the average signal per block, resulting in three patterns per condition in each run (i.e., three blocks per condition). We applied a linear support vector machine (SVM) to classify patterns with different condition contrasts and calculated mean accuracies after a leave-one run-out cross-validation procedure.

Sharing the common approach with other linear discrimination techniques, SVMs assign a categorical class label $y_i \in \{\pm 1\}$ to a pattern $\mathbf{x}_i$ $(i = 1, \ldots, N)$(N is the number of the pattern) based on the output of the discriminant function

$$f(\mathbf{x}_i) = \mathbf{w}\mathbf{x}_i + b$$

with

$$y_i = \mathrm{sgn}(f(\mathbf{x}_i)),$$

where the weight vector w and bias b define the separating hyperplane between two classes. SVMs differ from standard linear-discrimination techniques in the derivation of the separating hyperplane from training patterns. SVMs maximize the margin (the distance of the nearest data point to the separating hyperplane) of separation $2/\|\mathbf{w}\|$ between two classes in feature space given that

$$y_i(\mathbf{w}\mathbf{x}_i + b) \geq 1 \text{ for all } i = 1, \ldots, N.$$

We used linear SVMs to avoid potential difficulties in the interpretation of the classification results associated with nonlinear mapping from the input pattern into the feature space (6, 7). SVMs implement soft margin classification for noisy signals by introducing a slack variable

$$\xi_i \geq 0 \text{ for all } i = 1, \ldots, N,$$

$$y_i(\mathbf{w}\mathbf{x}_i + b) \geq 1 - \xi_i \text{ for all } i = 1, \ldots, N.$$

The separating hyperplane is obtained by minimizing the following objective function

$$E = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_i \xi_i,$$

where $C > 0$ is a penalty factor that controls the tradeoff between margin maximization and training-error minimization. The support vectors (SVs) are defined as the data points critical for the classification (usually near the separating hyperplane) of the training data set. Labels are assigned to independent data by comparing these data with the SVs rather than the center of the two classes.

1. Schwarzkopf DS, Zhang J, Kourtzi Z (2009) Flexible learning of natural statistics in the human brain. *J Neurophysiol* 102:1854–1867.
2. Sereno MI, et al. (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
3. Engel SA, et al. (1994) fMRI of human visual cortex. *Nature* 369:525.
4. DeYoe EA, et al. (1996) Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci USA* 93:2382–2386.
5. Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8:686–691.
6. Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685.
7. Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) "brain reading:" Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19:261–270.

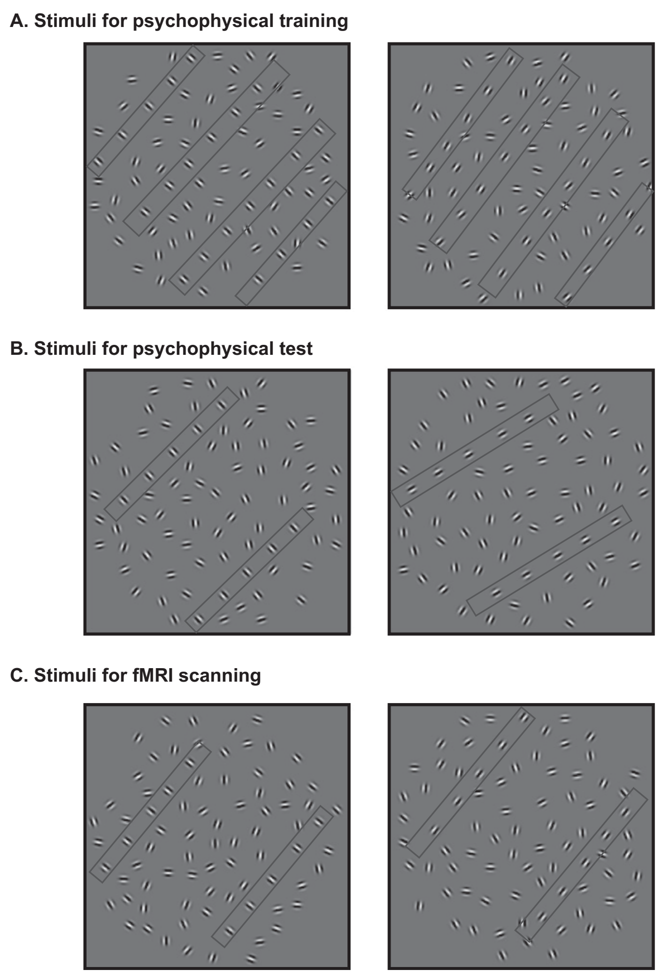

Fig. S1. Examples of stimuli for training and scan sessions. (*A*) Examples of orthogonal (*Left*) and collinear (*Right*) contours used for the supervised training and exposure sessions. Each stimulus consisted of four contour lines, illustrated by rectangles. (*B*) Examples of orthogonal (*Left*) and collinear (*Right*) contours used for the psychophysical test sessions. Each stimulus consisted of two contour lines. (*C*) Examples of orthogonal (*Left*) and collinear (*Right*) contours used for the scanning sessions. Each stimulus consisted of two contour lines.

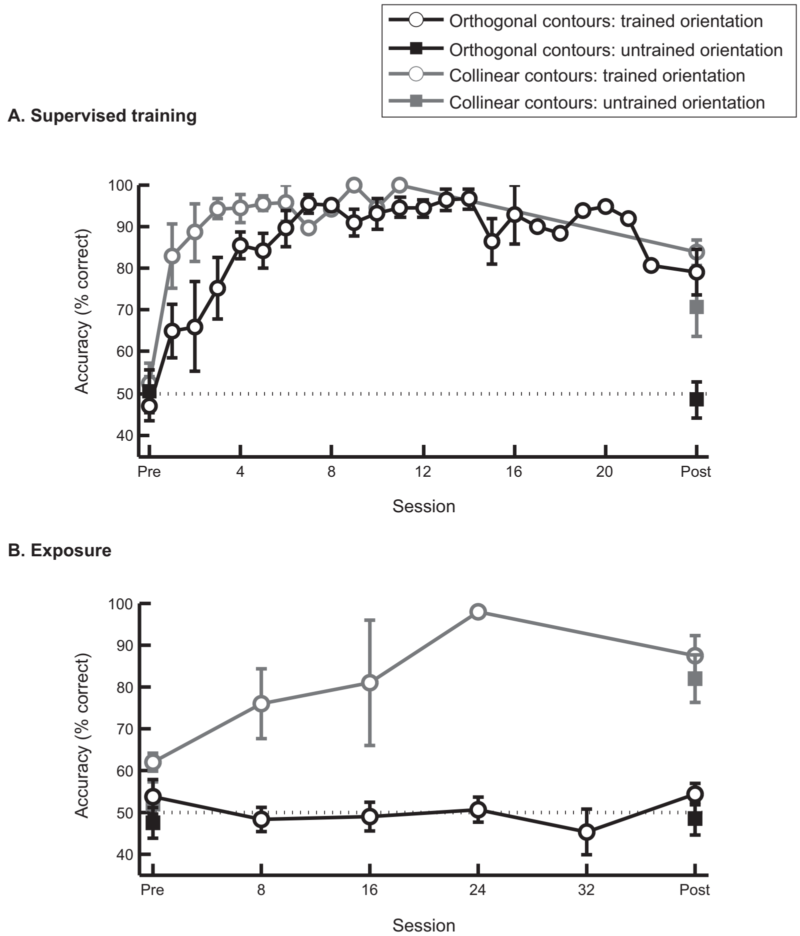

Fig. S2. Psychophysical data across sessions. Contour detection performance (percent correct) is plotted as a function of (*A*) supervised training sessions and (*B*) test sessions after each exposure session for collinear and orthogonal contours. Error bars denote SEM across observers. Comparing the number of training or exposure sessions required for improved detection showed that fewer training sessions were necessary for collinear than orthogonal contours [$F(1,21) = 18.11$, $P < 0.001$]. Furthermore, more sessions were required before reaching 80% performance in the exposure-based learning compared with supervised training [$F(1,21) = 28.52$, $P < 0.001$]. This result is consistent with the previously reported advantage for the detection of collinear contours (1–5).

1. Bex PJ, Simmers AJ, Dakin SC (2001) Snakes and ladders: The role of temporal modulation in visual contour integration. *Vision Res* 41:3775–3782.
2. Beck J, Rosenfeld A, Ivry R (1989) Line segregation. *Spat Vis* 4:75–101.
3. Field DJ, Hayes A, Hess RF (1993) Contour integration by the human visual system: Evidence for a local "association field," *Vision Res* 33:173–193.
4. Hess R, Field D (1999) Integration of contours: New insights. *Trends Cogn Sci* 3:480–486.
5. Geisler WS, Perry JS, Super BJ, Gallogly DP (2001) Edge co-occurrence in natural images predicts contour grouping performance. *Vision Res* 41:711–724.

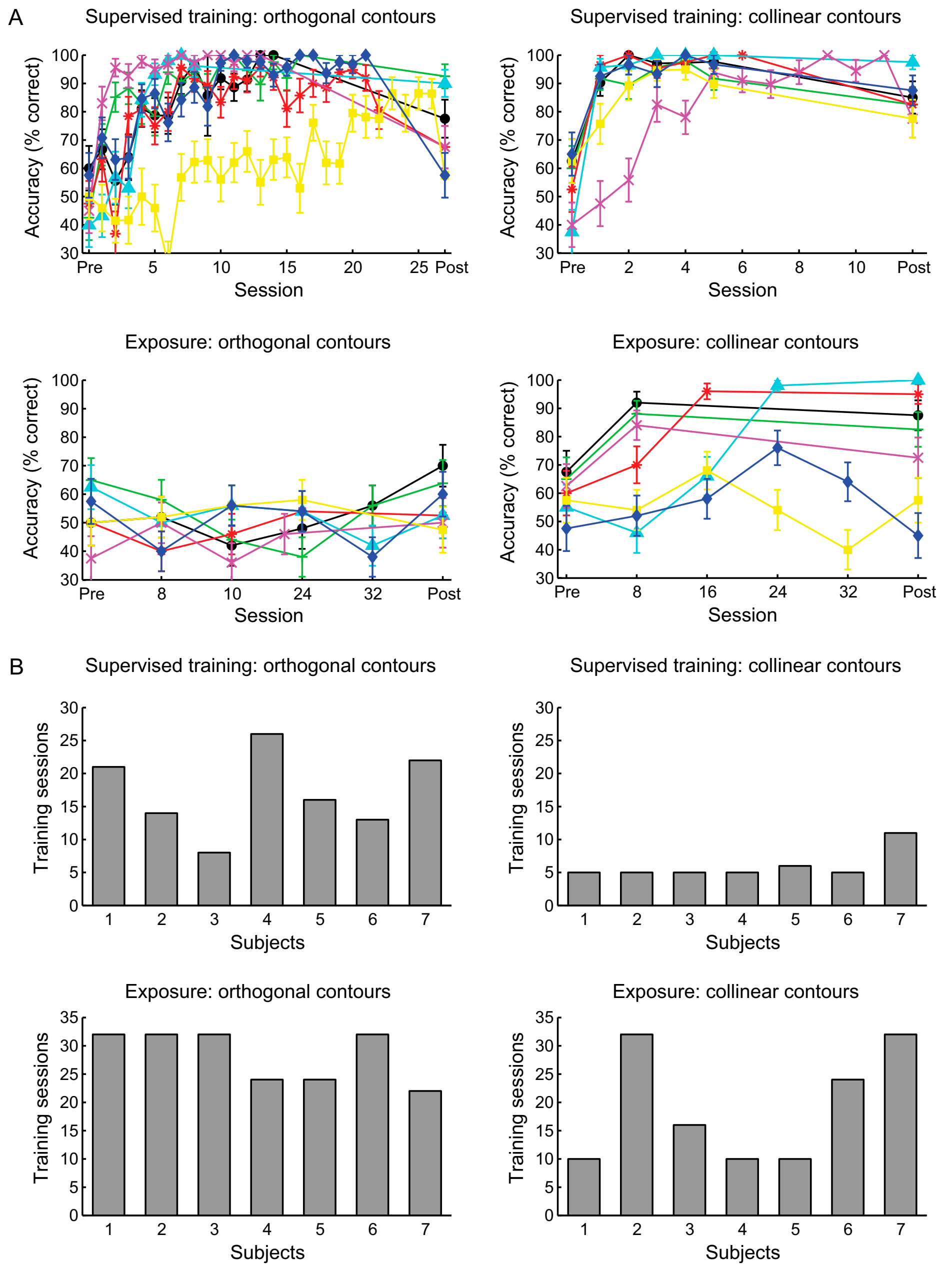

Fig. S3. Psychophysical data of individual subjects. (*A*) Contour-detection performance (percent correct) is plotted for individual subjects as a function of supervised training sessions and test sessions after each exposure session for collinear and orthogonal contours. Individual subject data are shown by different colors. The first and last data points on each curve denote the performance in the pretest and posttest, respectively. Error bars denote SEM across trials. (*B*) The number of training or exposure sessions completed by each individual subject in each group (supervised training on orthogonal contours, supervised training on collinear contours, exposure to orthogonal contours, and exposure to collinear contours).

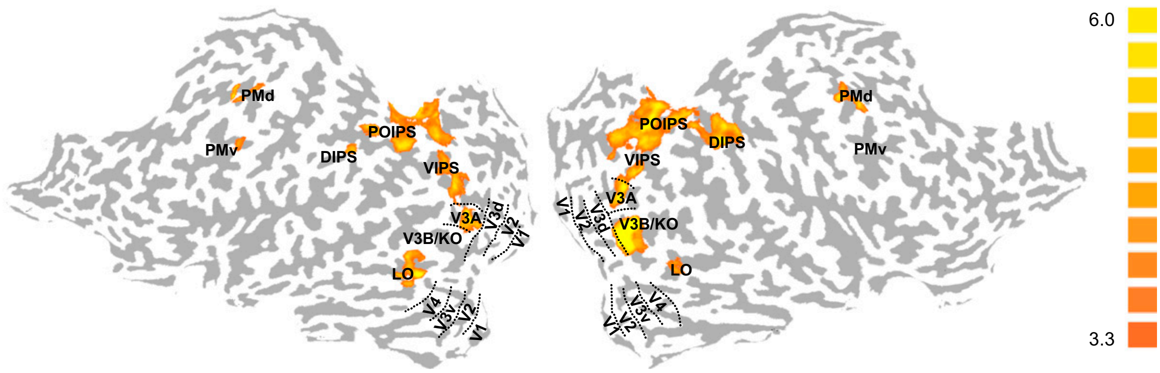

Fig. S4. Contour-responsive regions. Random-effects group GLM maps showing significantly higher fMRI responses to contour rather than random stimuli ($P < 0.001$, cluster-size threshold corrected, 80 mm^2). Data are presented on a flattened reconstruction of two cortical hemispheres. Talairach coordinates (mean, SD) for contour-responsive regions are given below.

ROI	Left hemisphere			Right hemisphere			Observers
	X	*Y*	*Z*	*X*	*Y*	*Z*	
V3A	−25.91 (4.68)	−90.32 (4.13)	4.14 (4.39)	21.14 (5.03)	−86.23 (5.68)	7.82 (5.41)	22
V3B/KO	−28.91 (6.32)	−87.73 (5.70)	−3.23 (5.45)	21.82 (3.72)	−88.00 (3.93)	−1.18 (4.91)	22
LO	−42.27 (5.31)	−72.32 (5.55)	−7.59 (5.42)	36.95 (4.28)	−67.68 (6.07)	−8.82 (5.25)	22
VIPS	−26.45 (6.25)	−74.50 (9.15)	17.91 (10.78)	22.09 (3.99)	−73.14 (6.90)	20.77 (9.52)	22
POIPS	−28.64 (6.05)	−62.23 (7.89)	36.59 (10.32)	22.09 (4.49)	−58.41 (5.34)	39.32 (6.54)	22
DIPS	−42.41 (8.80)	−41.50 (13.44)	35.50 (9.61)	36.68 (7.57)	−40.27 (7.11)	38.09 (4.95)	22
PMd	−35.32 (7.11)	−15.59 (14.99)	44.91 (11.95)	27.95 (8.04)	−11.91 (5.52)	44.50 (6.05)	22
PMv	−42.82 (4.83)	−9.32 (16.01)	30.05 (9.28)	40.09 (5.97)	−2.50 (7.79)	30.05 (8.14)	22

KO, kinetic occipital; LO, lateral occipital; VIPS, ventral intraparietal sulcus; POIPS, parieto-occipital intraparietal sulcus; DIPS, dorsal intraparietal sulcus; PMd, premotor dorsal; PMv, premotor ventral.

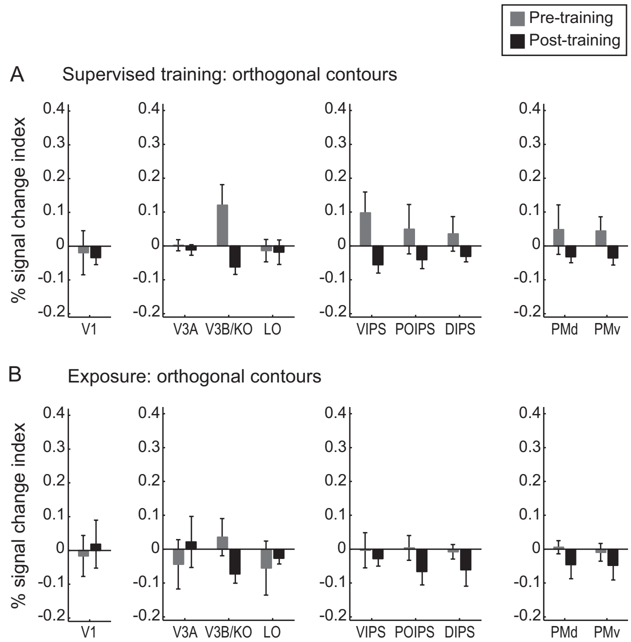

Fig. S5. fMRI responses for observers trained with orthogonal contours. Signal change index (percent signal change for orthogonal minus random contours) for each ROI. Data are shown for untrained contour orientations before (gray bars) and after (black bars) training for (*A*) supervised training and (*B*) exposure. Error bars denote SEM across observers.

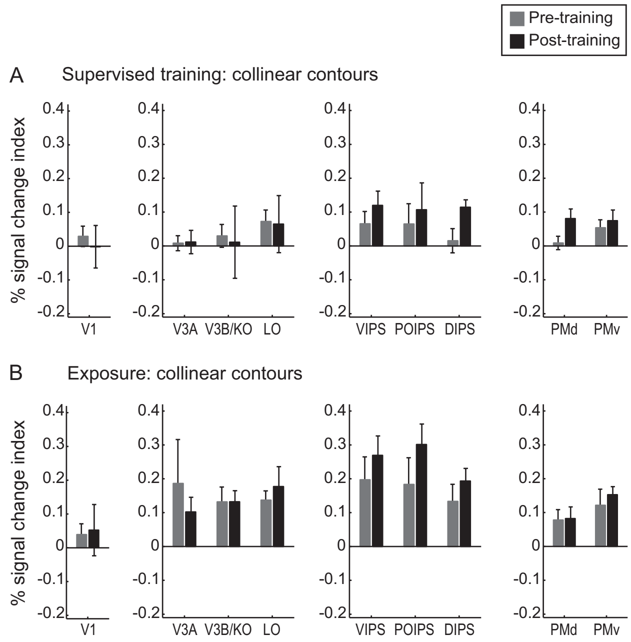

Fig. S6. fMRI responses for observers trained with collinear contours. Signal change index (percent signal change for collinear minus random contours) for each ROI. Data are shown for untrained contour orientations before (gray bars) and after (black bars) training for (*A*) supervised training and (*B*) exposure. Error bars denote SEM across observers. Statistical analysis (repeated-measures ANOVAs) compared the percent signal change for untrained orientations before and after training or exposure using the response to random stimuli as a baseline. For orthogonal contours (Fig. S5), no significant differences were observed for untrained orientations before vs. after supervised training in intraparietal regions [$F(1,4) = 2.42$, $P = 0.20$], occipitotemporal areas [$F(1,4) = 3.45$, $P = 0.14$], premotor areas [$F(1,4) = 1.70$, $P = 0.26$], or early visual areas [e.g., V1: $F(1,4) = 0.06$, $P = 0.82$]. Furthermore, no significant differences were observed for untrained orientations before vs. after exposure in intraparietal regions [$F(1,5) = 1.22$, $P = 0.32$], occipitotemporal areas [$F(1,5) = 2.16$, $P = 0.20$], premotor areas [$F(1,5) = 1.34$, $P = 0.30$], or early visual areas [e.g., V1: $F(1,5) = 0.16$, $P = 0.71$]. In contrast, for collinear contours (Fig. S6) presented at untrained orientations, we observed significantly higher fMRI responses after vs. before supervised training in intraparietal regions [$F(1,5) = 11.25$, $P < 0.05$] but not in occipitotemporal areas [$F(1,5) = 0.0002$, $P = 0.99$], premotor areas [$F(1,5) = 3.47$, $P = 0.12$], or early visual areas [e.g., V1: $F(1,5) = 0.40$, $P = 0.56$]. Furthermore, we observed significantly higher fMRI responses after vs. before exposure in intraparietal regions [$F(1,4) = 19.01$, $P < 0.05$] but not in occipitotemporal areas [$F(1,4) = 0.22$, $P = 0.66$], premotor areas [$F(1,4) = 1.27$, $P = 0.32$], or early visual areas [e.g., V1: $F(1,4) = 0.05$, $P = 0.84$].

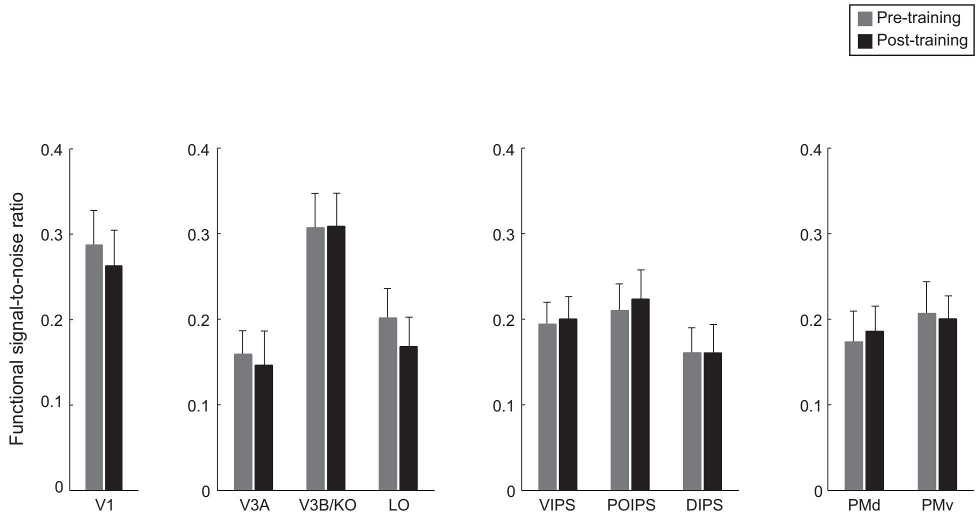

Fig. S7. Functional signal-to-noise ratio. Functional signal-to-noise ratio (fSNR) is shown for each ROI in each session. The fSNR is defined as the difference between the mean response to all stimuli and the response to fixation divided by the SD of the mean across all stimulus conditions and fixation. No significant differences were observed between sessions [$F(1,21) = 0.02$, $P = 0.89$].

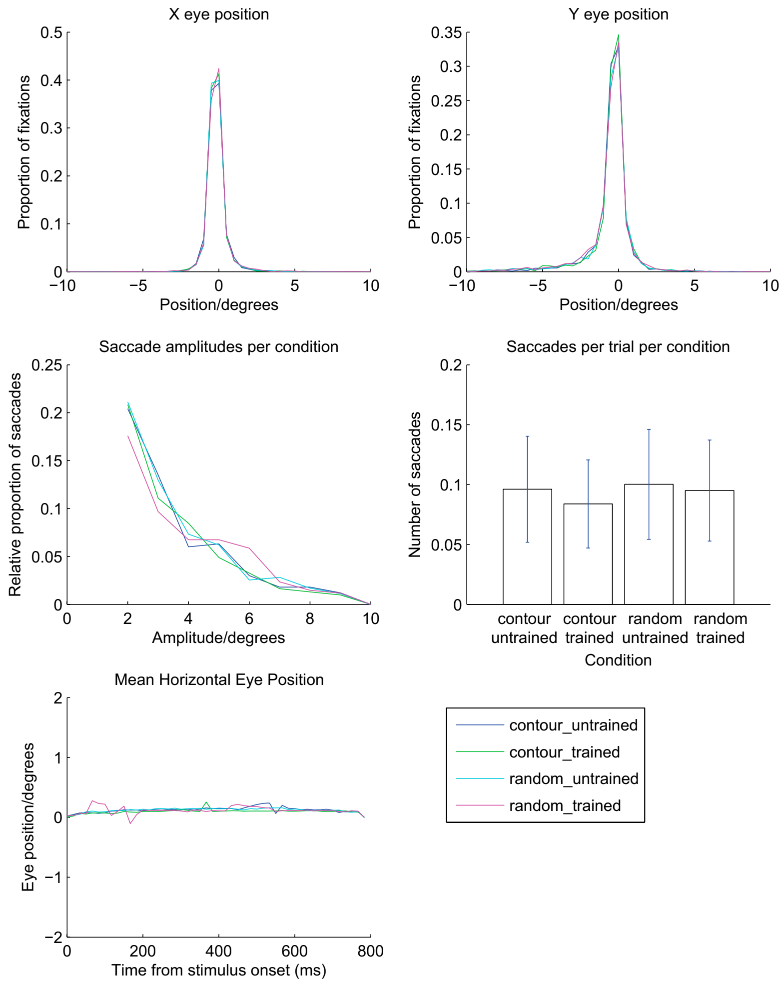

Fig. S8. Eye-movement analysis. We recorded eye movements before (this figure) and after (Fig. S9) training while observers performed the target-detection task in the scanner. Eye movements were recorded using the ASL 6000 Eye-tracker (Applied Science Laboratories). Eye-tracking data were preprocessed using the Eyenal software (Applied Science Laboratories) and analyzed using custom Matlab (Mathworks) software. For each scan session, we computed the horizontal (X) and vertical (Y) eye position, saccade amplitude, number of saccades per condition, and the event-related eye trace for each stimulus condition that shows the time course of the mean eye-position changes across trials. Data from each trial were brought to a common baseline (to remove drift) using the mean eye position over the 100-ms preceding stimulus onset. In both sessions, the horizontal and vertical eye positions for each stimulus type peaked and were centered on the fixation at 0°. A repeated measurement ANOVA indicated that there was no significant difference between stimulus conditions on mean horizontal eye position [before training: $F(1.16,9.26) = 0.54$, $P = 0.51$; after training: $F(1.46,7.32) = 0.34$, $P = 0.66$], mean vertical eye position [before training: $F(1.44,11.55) = 0.29$, $P = 0.69$; after training: $F(1.22,6.11) = 0.34$, $P = 0.62$], mean saccade amplitude [before training: $F(1.85,14.77) = 0.14$, $P = 0.86$; after training: $F(1.51,7.56) = 0.24$, $P = 0.74$], or the number of saccades per trial per condition [before training: $F(2.19,17.52) = 0.91$, $P = 0.43$; after training: $F(1.03,5.16) = 0.92$, $P = 0.39$]. In addition, no significant differences were observed between sessions (before vs. after training) for horizontal eye position [$F(1,5) = 1.34$, $P = 0.73$], vertical eye position [$F(1,5) = 0.44$, $P = 0.54$], mean saccade amplitude [$F(1,5) = 1.19$, $P = 0.33$], or number of saccades [$F(1,5) = 0.27$, $P = 0.53$]. These analyses suggest that it is unlikely that our results were significantly confounded by eye movements. Furthermore, plots of the event-related traces of mean horizontal eye position showed that mean deviations of eye position were very small during stimulus presentation, and there was no evidence of systematic differences between stimulus conditions in both sessions.

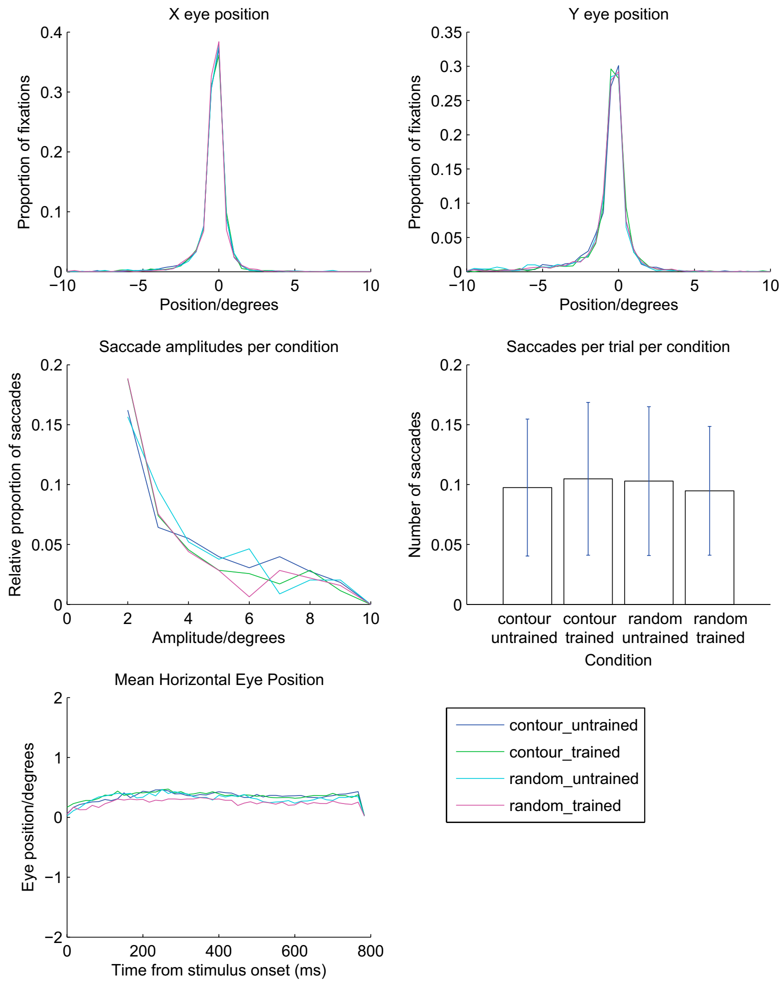

Fig. S9. Eye-movement analysis. We recorded eye movements before (Fig. S8) and after training (this figure) while observers performed the target-detection task in the scanner.