

1

2 **Supplementary Information for**  
3 **Bottom-up saliency and top-down learning in the primary visual cortex of monkeys**

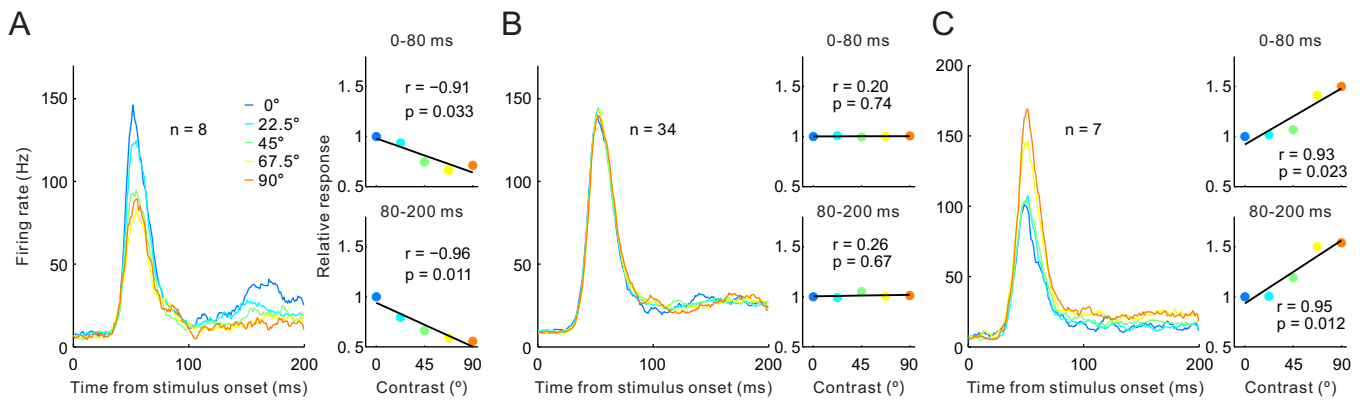
4 Yin Yan, Li Zhaoping and Wu Li

5 Wu Li  
6 E-mail: [liwu@bnu.edu.cn](mailto:liwu@bnu.edu.cn)

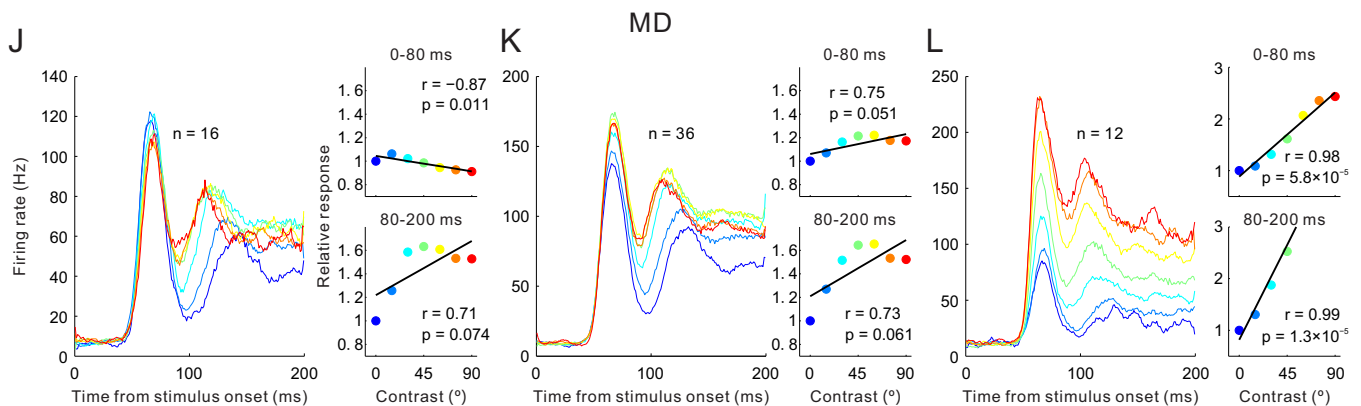
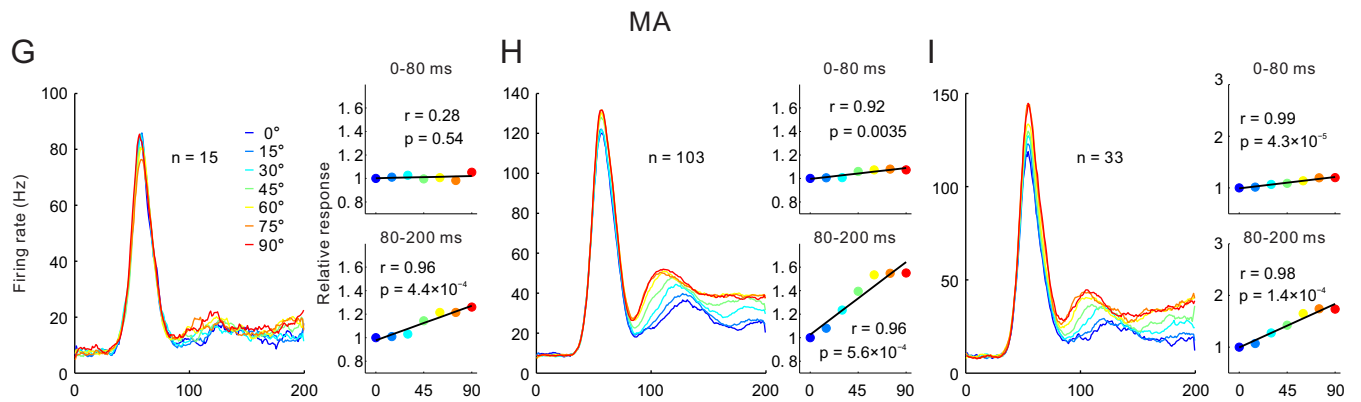
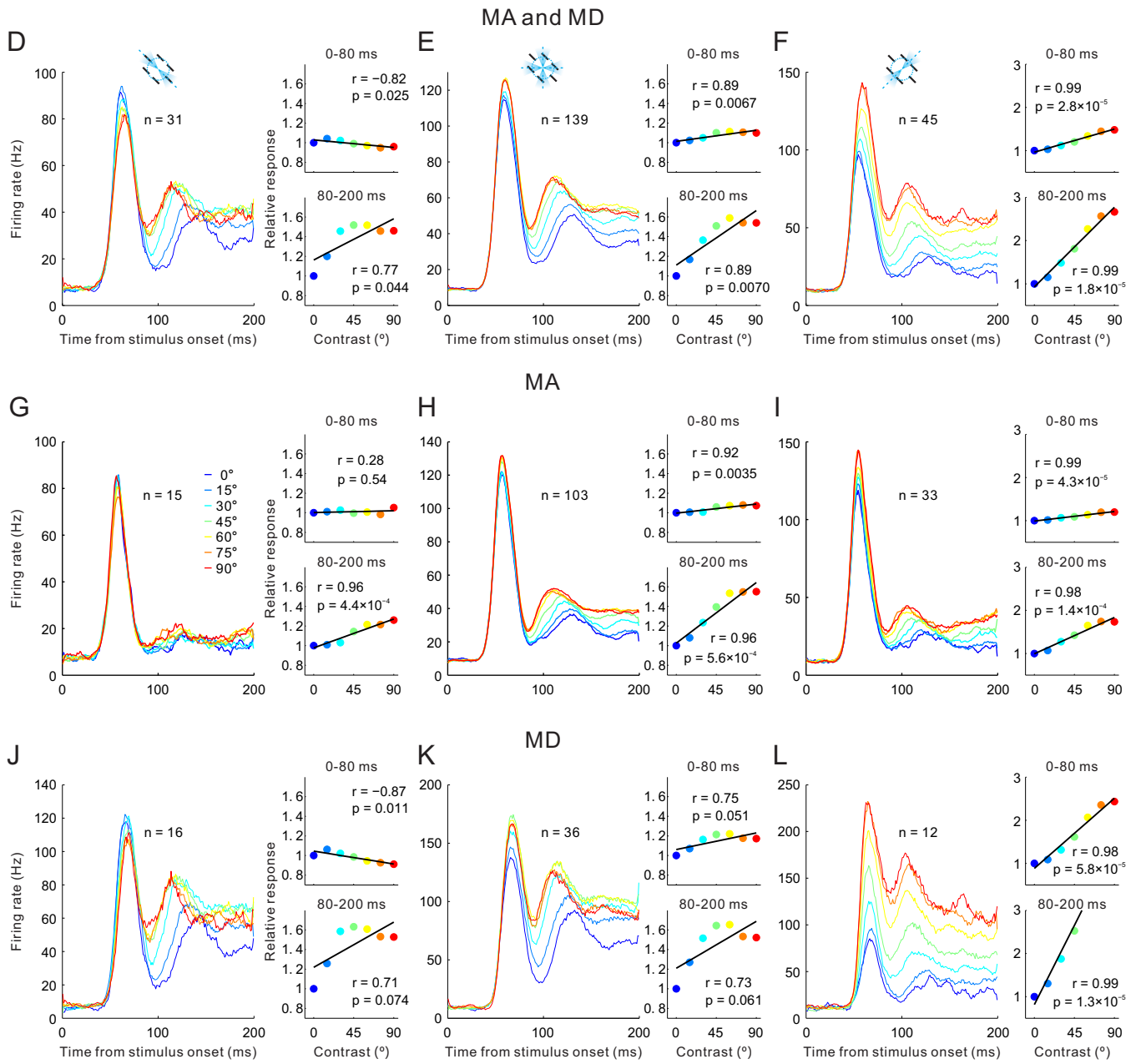
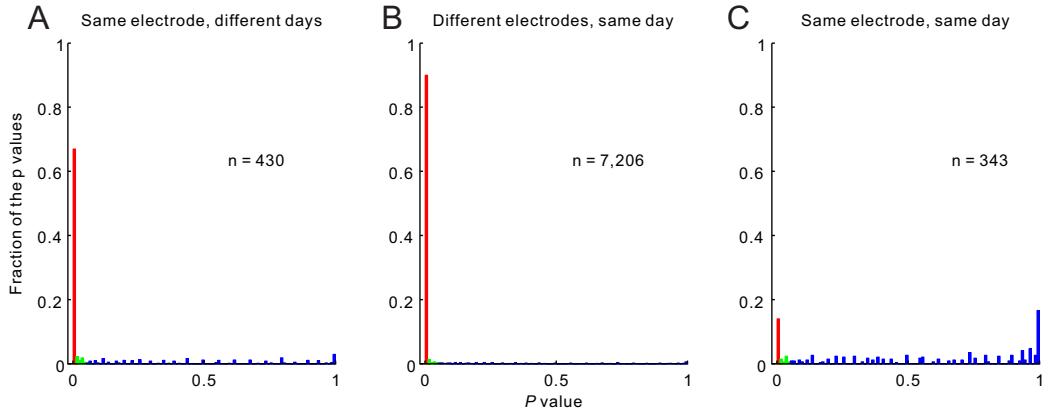
7  
8 Li Zhaoping  
9 E-mail: [z.li@ucl.ac.uk](mailto:z.li@ucl.ac.uk)

10 **This PDF file includes:**

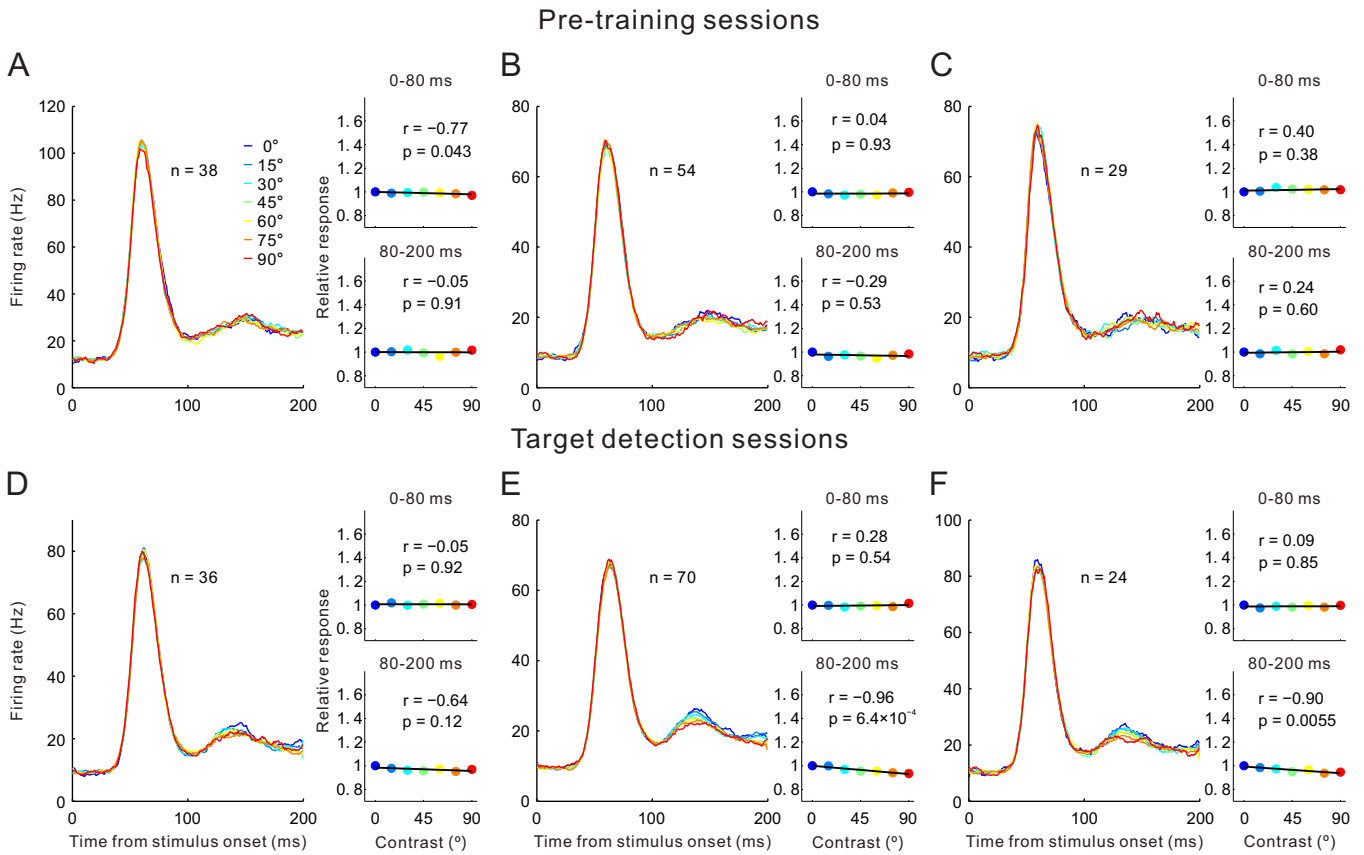
11 Figs. S1 to S8



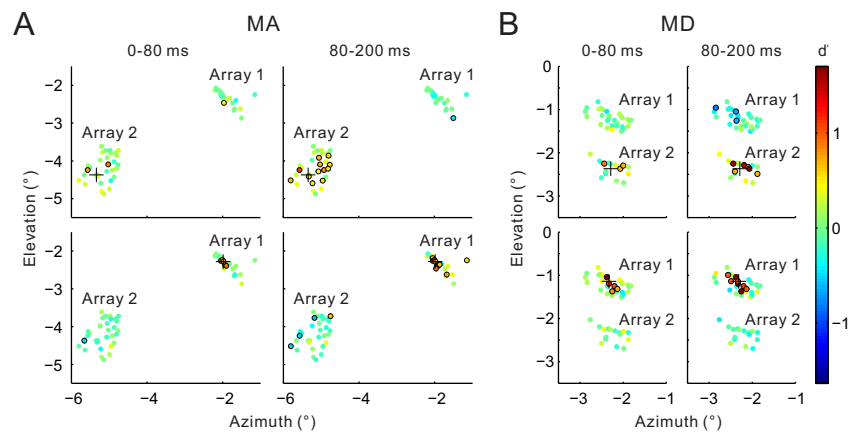
**Fig. S1.** Population averaged V1 responses to gratings having various orientation deviations from an invisible reference orientation. The reference orientation is that of the background bars in the singleton detection task. The V1 sites are separated into 3 groups according to the deviations of their preferred orientations from the reference orientation: within  $22.5^\circ$  (A), between  $22.5^\circ$  and  $67.5^\circ$  (B), or larger than  $67.5^\circ$  (C). The PSTHs in A-C are plotted in the same format as those shown in Fig. 2 D-F, respectively, and using exactly the same V1 sites, except that the neural responses were collected during the daily RF mapping sessions when the monkeys were doing the fixation task (Materials and Methods). Consistent with Fig. 2, the angles  $0^\circ$ ,  $22.5^\circ$ ,  $45^\circ$ ,  $67.5^\circ$ ,  $90^\circ$  (blue to red) are the differences (clockwise or anticlockwise) between the grating's orientation and the reference orientation defined above (with clockwise and anticlockwise deviations pooled). Note that the averaged population responses in B are largely invariant to these different angles. Similar results were obtained using data from the pretraining days. See Fig. 2 legend for more details.



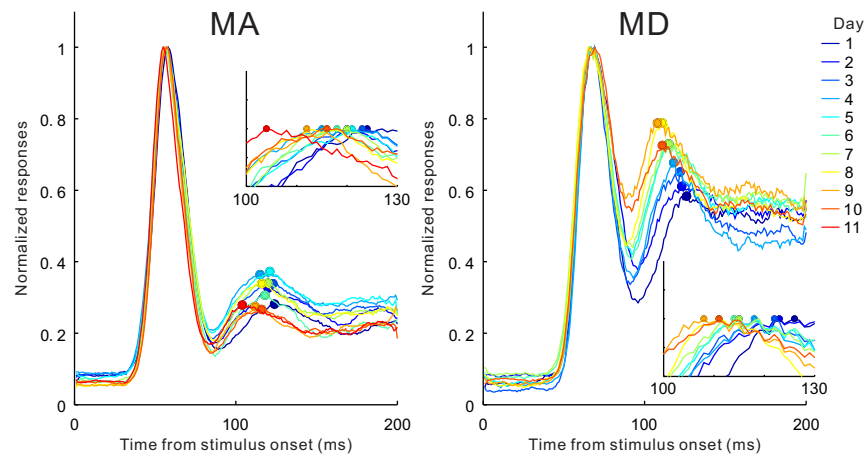
**Fig. S2.** The same data shown in Fig. 2 *D–F* were reanalyzed by treating recordings from the same electrode in different days as independent samples. (*A–C*) Statistical test showing significant differences in the distributions of V1 activities recorded in different days. The horizontal axes mark the  $p$  values, each is the probability that two distributions of spike counts are statistically the same using the Kolmogorov-Smirnov (K-S) test. We calculated the fractions of these  $p$  values in each bin (bin size 0.01). *A*: The two distributions used for each K-S test are from the same electrode, one distribution from one recording day and the other from another day. The  $p$  values are pooled from all electrodes in both monkeys, and from all pairwise combinations of recording days. This plot shows that the majority of electrodes picked up statistically different neuronal responses in different days. *B*: Similar to *A*, but the two distributions for each K-S test involve spike counts from two different electrodes respectively in the same monkey on the same day. The  $p$  values are pooled from all pairwise combinations of electrodes, and pooled from both monkeys and all recording days. This plot shows that different electrodes picked up statistically different neuronal responses on the same day. *C*: Similar to *A* and *B*, but the two distributions for each K-S test involve spike counts from the first and second half of blocks of trials for an individual electrode on the same day. The  $p$  values are pooled from all electrodes in both monkeys, and from all recording days. This plot shows that, for the majority of electrodes, each one picked up statistically similar neural signals within the same recording session/day. Red, green, and blue mark  $p < 0.01$ ,  $0.01 \leq p < 0.05$ , and  $p \geq 0.05$ , respectively. In each distribution, each sample of spike count was calculated during 0-200 ms after the onset of a singleton stimulus pattern in a single trial of the pretraining fixation task. All the stimulus conditions were pooled (3 target locations  $\times$  12 target orientations, randomly interleaved; see the main text for more details). Note that in *A* and *B*, we only used the first half of the trials in each condition, so that each spike-count distribution contained the same total number of trials as those in *C*. (*D–L*) Data shown in Fig. 2 *D–F* are replotted here in the same format, except that V1 responses recorded by the same electrode in different days were treated as coming from different sites (justification in *A*), which contributed to the total sample size  $n$  in each plot. *D–F*: V1 sites pooled from both monkeys; *G–I*: V1 sites from MA only; *J–L*: V1 sites from MD. As in Fig. 2, the V1 sites were separated into 3 groups according to the deviations of their preferred orientations from the orientation of the background bars: within  $22.5^\circ$  (*D*, *G*, *J*), between  $22.5^\circ$  and  $67.5^\circ$  (*E*, *H*, *K*), or larger than  $67.5^\circ$  (*F*, *I*, *L*). Qualitatively similar results were obtained when the data from the two animals were pooled or separated, although the sampled V1 sites in monkey MD were overall more sensitive to the orientation contrast, especial in the late response components. See Fig. 2 legend for more details.



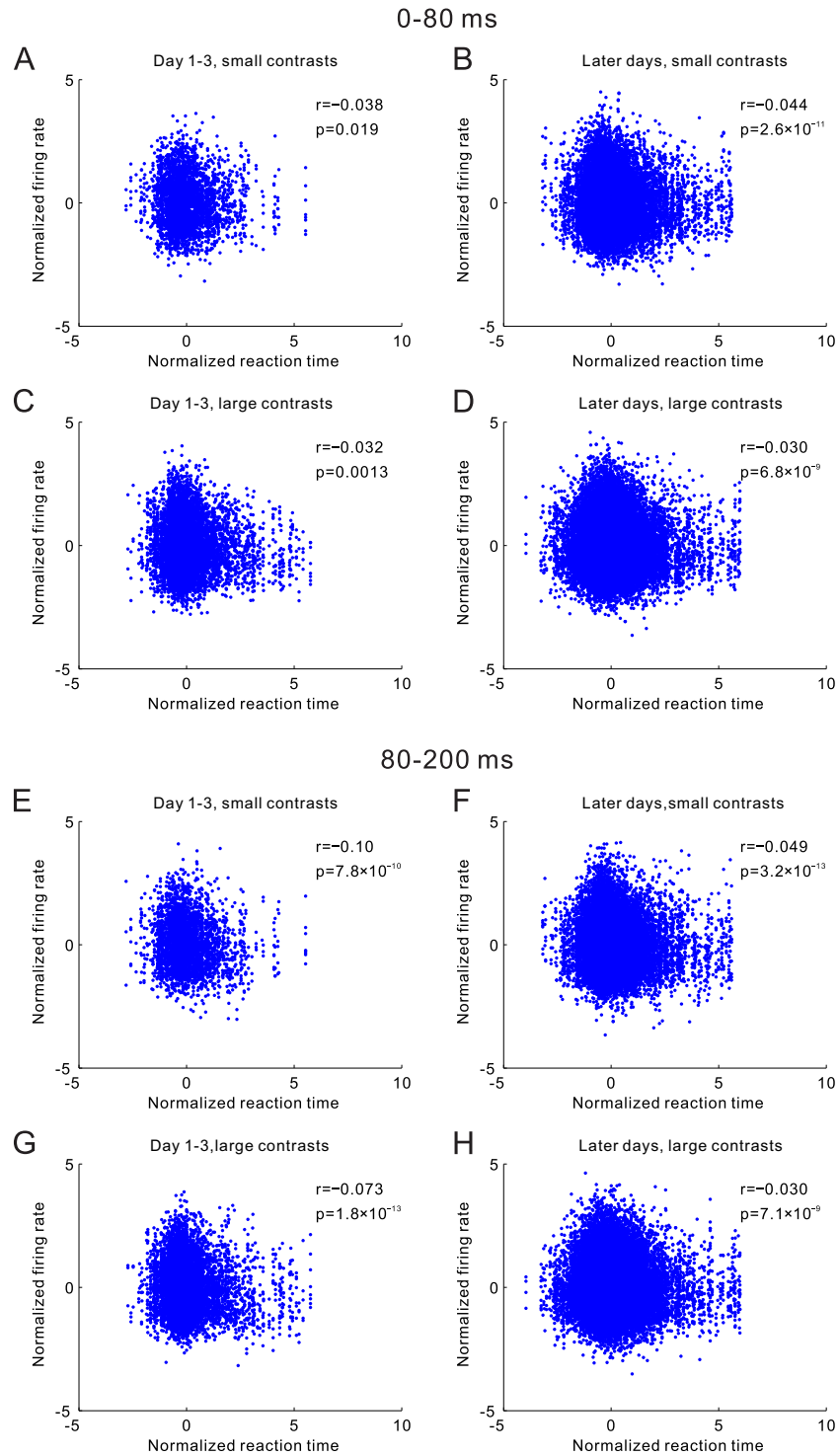
**Fig. S3.** Population averaged responses of V1 sites on the iso-orientated background bars. Same format as Fig. 2, except that the responses were from V1 sites whose RF centers were  $> 1^\circ$  away from the center of the singleton bar (the three inset cartoons shown in Fig. 2 A–C are removed here). (A–C) Averaged PSTHs in the fixation task before detection training. The V1 sites from both monkeys were pooled and separated into 3 groups according to the deviations of their preferred orientations from the orientation of the background bars: within  $22.5^\circ$  (A), between  $22.5^\circ$  and  $67.5^\circ$  (B), or larger than  $67.5^\circ$  (C). Each PSTH corresponding to an orientation contrast ( $0^\circ$  to  $90^\circ$ ) is the average of  $n$  PSTHs, each in turn is the average of the corresponding PSTHs from a single electrode across the pre-training days. The preferred orientation associated with an electrode was the average of the preferred orientations measured across the days. For these V1 sites on the background, their pre-training responses—both early and late components—were independent of the orientation contrasts of the singleton outside the RFs. See Fig. 2 A–C legend for more details. (D–F) Same as A–C, but during the singleton detection task. Practicing singleton detection induced a negative correlation between the orientation contrast and the late, but not the early, responses. In the extreme case when the target was in the quadrant opposite to the RFs, the late responses also became invariant to the target's orientation contrast.



**Fig. S4.** Visual field maps of the recorded V1 sites in response to singletons with  $90^\circ$  orientation contrast for monkey MA (A) and MD (B) on the last training day. For each V1 site, we used the  $d'$  to quantify its target detectability in a given condition:  $d' = (\mu_1 - \mu_2) / (\frac{1}{2}(\sigma_1^2 + \sigma_2^2))^{1/2}$ , where  $\mu_1$  and  $\mu_2$  are the mean spike counts within the specified time window across the trials in which the singleton was, respectively, within the pattern covering the RFs or within the pattern in the opposite visual field (i.e., two mirror conditions);  $\sigma_1$  and  $\sigma_2$  are the corresponding standard deviations of the spike counts. Different panels show  $d'$  results from different time windows (0-80 ms or 80-200 ms after stimulus onset) and different target locations marked by '+' (within array 2 or array 1). Each dot marks the center of a recorded RF in the visual field, showing its ability (color-coded  $d'$  value) to differentiate the two mirror conditions. Black circles highlight the sites showing  $d'$  significantly different from 0 (two-tailed t-test,  $p < 0.05$ ).

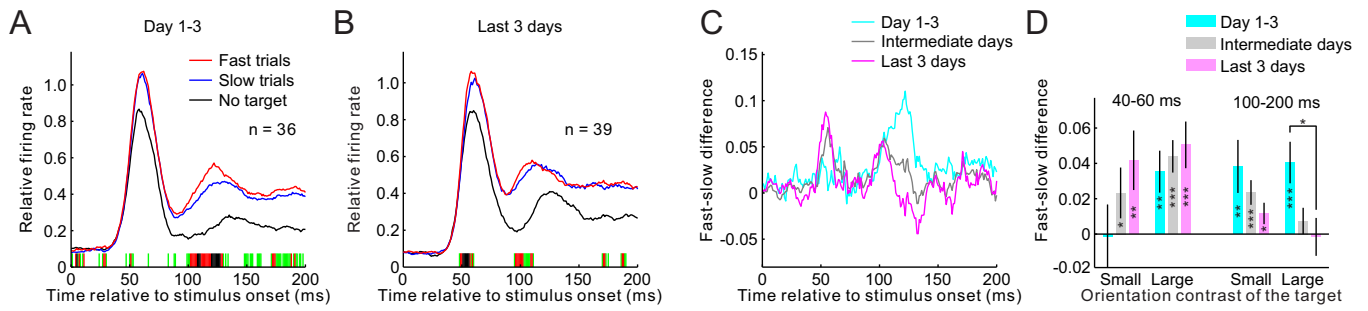


**Fig. S5.** Learning-induced changes in V1 late response components. The time-to-peak of the late response components moved progressively earlier with training days in both monkeys (MA and MD). For a given day and a monkey, each PSTH was the average of the  $PSTH(\theta, s)$  across different V1 sites  $s$  covering the target location and across different target orientations  $\theta$  relative to the background orientation. The averaging was carried out in two steps: first across all the non-zero  $\theta$  of the target covered by a V1 site, and then across all the V1 sites  $s$  that covered any target location. The peaks of the early responses ( $< 80$  ms) were scaled to 1. The insets show zoomed views around the peaks of the late responses, which were scaled to unity to better illustrate the progressive changes of the time-to-peak (see also Fig. 3C).

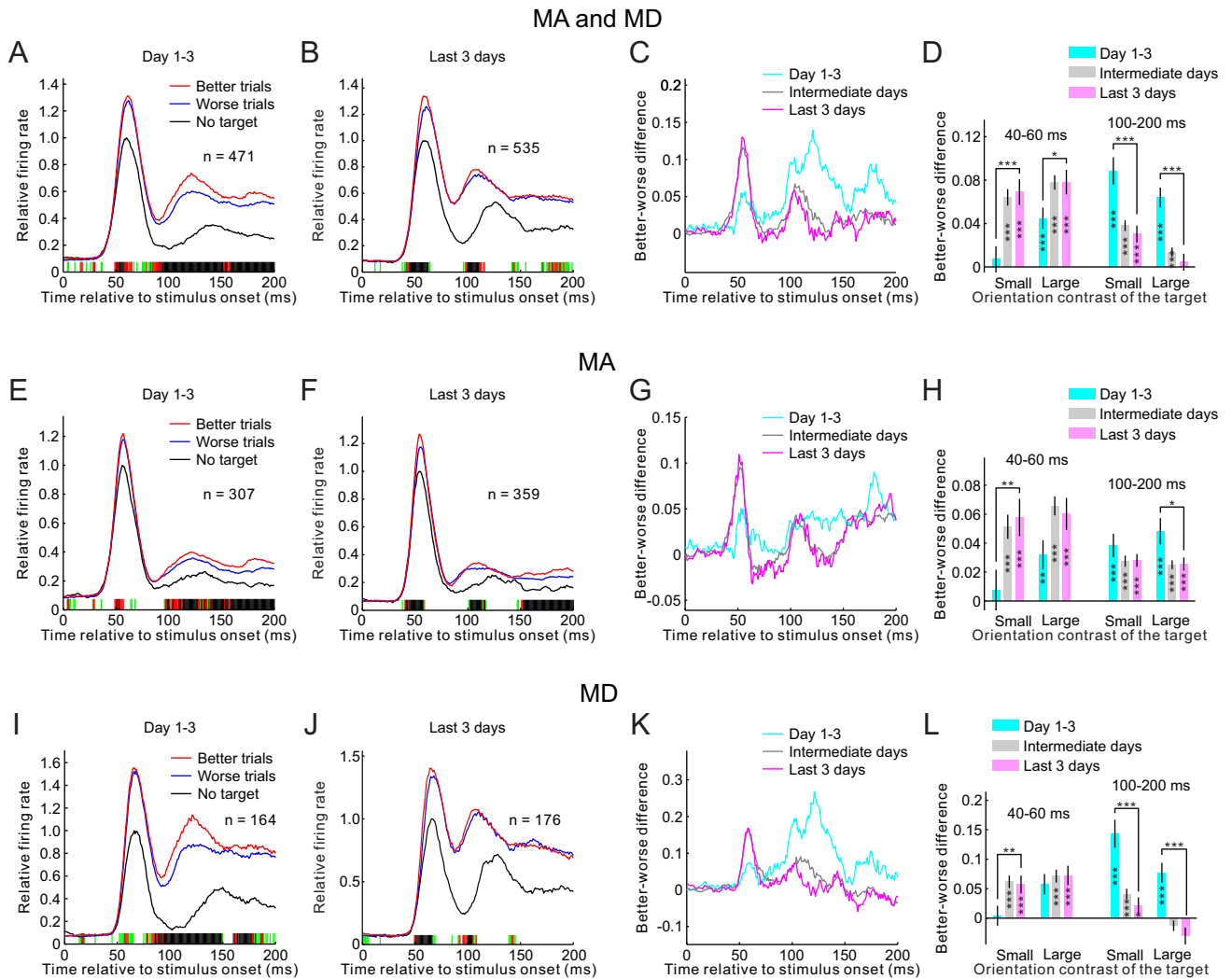


**Fig. S6.** Trial-by-trial correlations between behavioral reaction times and early (A–D) or late (E–H) V1 responses. Given a target condition, a recording site, and a training day, the raw firing rate  $r_i$  in the  $i^{\text{th}}$  trial is normalized as a z-score  $z_i = (r_i - \bar{r})/\sigma$ , where  $\bar{r}$  and  $\sigma$  are the mean and standard deviation of the firing rates across trials in this condition and on this day. Similarly, the corresponding raw reaction times were also normalized as their z-scores. After these normalizations, only the intra-condition fluctuations were left and the inter-condition differences were removed. These z-scores were pooled separately for the first 3 training days (A, C, E, G) and later days (B, D, F, H), and for small orientation contrasts ( $15^\circ$ ,  $30^\circ$ , and  $45^\circ$ ; A, B, E, F) and large contrasts ( $60^\circ$ ,  $75^\circ$ , and  $90^\circ$ ; C, D, G, H). The correlation coefficients were Spearman's rank correlations.





**Fig. S7.** Another version of Fig. 4 by replacing the better-worse division with the fast-slow division of trials. For each target condition, the miss trials were excluded in the analysis and all the hit trials were divided evenly into the fast and slow trials (with shorter and longer reaction times). The analyses here involved exactly the same electrodes and the same conditions as those in Fig. 4. See Fig. 4 legend for details.



**Fig. S8.** The same data shown in Fig. 4 were reanalyzed by treating recordings from the same electrode in different days as independent samples (cf. Fig. 2 legend). (A–D) Data were pooled from both monkeys, and the results are presented in the same format as those shown in Fig. 4 A–D respectively. (E–H) Data from MA only; (I–L) Data from MD. The sample size  $n$  in this figure involves the same electrodes as those used in Fig. 4; but the V1 responses recorded by a single electrode in different days and in different target conditions were treated as independent samples, which contributed to the  $n$ . For other details see Fig. 4 legend.