Learning to Link Visual Contours

Wu Li[1,2], **Valentin Piëch**[1], and **Charles D. Gilbert**[1]

[1]*The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA*

[2]*State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China*

SUMMARY

In complex visual scenes, linking related contour elements is important for object recognition. This process, thought to be stimulus driven and hard wired, has substrates in primary visual cortex (V1). Here, however, we find contour integration in V1 to depend strongly on perceptual learning and top-down influences that are specific to contour detection. In naive monkeys the information about contours embedded in complex backgrounds is absent in V1 neuronal responses, and is independent of the locus of spatial attention. Training animals to find embedded contours induces strong contour-related responses specific to the trained retinotopic region. These responses are most robust when animals perform the contour detection task, but disappear under anesthesia. Our findings suggest that top-down influences dynamically adapt neural circuits according to specific perceptual tasks. This may serve as a general neuronal mechanism of perceptual learning, and reflect top-down mediated changes in cortical states.

INTRODUCTION

Parsing visual scenes into different objects involves grouping processes, including contour integration, whereby contour elements belonging to the same object are perceptually linked and segregated from other scene components. Contour integration is generally characterized as a bottom-up process conforming to the rules of natural scene geometries. From the point of view of Gestalt psychology, the visual system has built-in functionality to connect line elements that form continuous and smooth contours (Wertheimer, 1923; Field et al., 1993; Li and Gilbert, 2002). This rule of continuity in perceptual organization is ecologically correlated with two statistical regularities commonly seen in natural scenes—collinearity and co-circularity (Geisler et al., 2001; Sigman et al., 2001). With respect to the underlying cortical circuitry, the long-range horizontal connections formed by the axons of pyramidal cells in the primary visual cortex (area V1) tend to link cells with non-overlapping receptive fields (RFs) but with similar orientation preference (Gilbert and Wiesel, 1979; Rockland et al., 1982; Gilbert and Wiesel, 1983, 1989; Schmidt et al., 1997; Stettler et al., 2002). This hard-wired intra-cortical connectivity is ideally suited for mediating contour integration, both in terms of its orientation specificity and its spatial extent (Li and Gilbert, 2002; Stettler et al., 2002). The involvement of V1 in contour integration is supported by physiological evidence that collinearly arranged line segments can facilitate V1 neuronal responses (Kapadia et al., 1995; Polat et al., 1998; Bauer and Heinze, 2002; Li et al., 2006). Computational models that simulate interconnected V1 neurons also demonstrate the capability of extracting global visual contours out of complex

backgrounds without the intervention of top-down influences (Ullman, 1992; Li, 1998; Yen and Finkel, 1998; VanRullen et al., 2001; Ernst et al., 2004). As hardware encoding of contour integration is given particular emphasis in the literature, the roles of top-down influences and perceptual learning are largely overlooked.

Within a background of randomly oriented lines (Figure 1A), those discrete line segments following the Gestalt law of "good continuation" are easily grouped together, forming a global visual contour. The perceptual saliency of contours in a complex environment depends on the spatial arrangement of contour and background elements (Field et al., 1993; Kovács et al., 1999; Li and Gilbert, 2002). A contour made up of more collinear lines is easier to detect, or more salient, than a shorter contour within the same background (compare Figure 1A with Figure 1B); and the same array of collinear lines forming the contour appear less salient when they are spaced further apart (compare Figure 1B with Figure 1C). On the other hand, contour detectability depends not only on the geometry of visual stimuli but also on perceptual learning. In particular, the contours that are originally hidden or less salient (Figure 1C) become increasingly easier to detect with training due to an enhancement in interactions between contour elements (Kovács et al., 1999; Li and Gilbert, 2002). It has also been shown that perceptual learning leads increased strength and effective distance of facilitation between collinearly arranged targets (Polat and Sagi, 1994).

In a recent study we have shown that in monkeys performing a contour detection task, responses of V1 neurons are predictive of the animals' behavioral performance on contour detection, and therefore are closely correlated with perceptual saliency of contours (Li et al., 2006). The correlation is seen in the facilitation of neuronal responses by collinear line segments lying outside the classical RF within a background of randomly oriented lines. This facilitation is stronger with more salient contours, and is present in virtually all orientation selective neurons in superficial layers of V1, indicating that V1 is intimately involved in linking visual contours and in mediating contour saliency. However, the facilitation of neuronal responses by the same contours is significantly weakened if the animals perform a different task irrelevant to contour detection. This implies a degree of top-down influence on V1 responses, but the nature of this influence is not obvious. The top-down effects could be due to differences in either the spatial locus of attention or the task of contour detection itself.

In the current study, in order to examine whether the neural substrate of the improvement in contour detection with practice is found in area V1, and also to explore the nature of top-down influences on contour integration, we trained monkeys on a succession of tasks, beginning with fixation, then attending to the location of a contour, and then detecting the contour embedded in a complex background. Neuronal responses were recorded from the same V1 regions during different experimental stages. This design enabled us to follow the emergence of contour-related responses in V1 as the animals learned to detect contours. These experiments revealed striking experience- and task-dependent changes in V1 responses associated with training on contour detection.

RESULTS

Two naïve adult monkeys (*Macaca mulatta*, male, 6–7 kg) were used in all experiments. As the results from the two animals were qualitatively similar (see Figure S1 in the Supplemental Data), data from both subjects were pooled in most analyses to guarantee fair sample size in statistical tests. A total of 275 cells were recorded from V1 in the four hemispheres of the two animals.

Perceptual learning in contour detection

For simple discrimination and detection tasks, repeated execution of the same task usually leads to a remarkable improvement in a subject's performance on the task. This perceptual phenomenon, known as perceptual learning, has been reported in contour detection tasks in human subjects (Kovács et al., 1999; Li and Gilbert, 2002). The two animals used in the current study showed similar learning effects.

The visual stimuli used in the contour detection task consisted of a series of collinear line segments embedded in the center of an array of randomly oriented and positioned lines (Figures 1A and 1B). In addition to this contour pattern, a second pattern without any embedded contour (referred to as the noise pattern) was simultaneously presented as a distracter in the visual-field quadrant opposite to the contour pattern (Figure 2A). The perceptual task was to indicate which of the two stimulus patterns contained a straight contour. From trial to trial the contour pattern and the noise pattern were randomly interchanged so that the overall detection ratio simply by guessing was 50%. Within a block of trials five stimulus conditions with visual contours made up of 1, 3, 5, 7, 9 collinear lines were interleaved, with line spacing kept constant. To eliminate any possible cues, other than the embedded contour, for the two stimulus patterns shown in each trial, all the non-contour elements and the very central line segment in the contour pattern were made identical to those in the noise pattern (compare the corresponding line segments between the two stimulus patterns shown in Figure 2A). With this design, when the embedded contour consisted of only a single line segment, the contour pattern and the noise pattern were identical. This particular condition was included as a baseline control. After 476 ms exposure of the stimuli, the contour pattern and the noise pattern were replaced with two dots. The animals indicated the location of the pattern containing the embedded contour by making a saccadic eye movement towards either of the two dots.

In order to let the animals understand and execute this two-alternative-forced-choice (2AFC) task, a preliminary procedural or operational training was conducted before data collection. The main training procedure is as follows. At the training outset, the complex background surrounding the contour elements, as well as the whole noise pattern, was hidden by setting the luminance of line segments to that of the display background, leaving the collinear lines shown in isolation at either of the two locations. It took a couple of days, about 1500 trials a day, for the animals to learn, by trial and error for a reward, to associate the location of the stand-alone contours with the direction of saccades. Subsequently, the hidden line elements were rendered visible by gradually increasing their luminance above the display background, but these non-contour elements were kept noticeably dimmer than the contour. It took another several days for the animals to reliably choose, between the two stimulus patterns, the one with a highlighted contour of different lengths at any given orientation. Then the luminance difference between contour and background lines was removed and data collection was started. For each session of the contour detection task, the orientation of the embedded contour was determined by the preferred orientation of the recorded V1 cell; therefore, the contour orientation was the same within a session but could be different across sessions.

Daily training and recordings began with the lower-left and upper-right quadrants in the visual field (Figure 2A; defined as Location 1 throughout the text). Both animals' performance on contour detection was improved with practice and reached different plateaus for different contour lengths, a typical effect of perception learning (Figures 2B and 2C).

A prominent characteristic of perceptual learning is its specificity to visual-field location. The learning-induced improvement is usually restricted to the area where the task stimuli are displayed (for example see Crist et al., 1997; Sigman and Gilbert, 2000). The improved contour detection ability in these two animals was also specific to the trained visual-field location. After the animals' performance had reached plateaus at Location 1, we moved the stimulus patterns

to untrained areas, the lower-right and upper-left quadrants of the visual field (Figure 2D; defined as Location 2 throughout the text). Before data collection we had ensured that the animals still understood the detection task at the new stimulus location. This procedure was done with highlighted contours as in the initial operational training described previously, but it was much quicker and took only a couple of hundred trials.

Compared with the final performance level at the trained Location 1, the animals' performance on contour detection dropped substantially in the first couple of days at the new location for the same contour length (compare Figure 2E with Figure 2B, and Figure 2F with Figure 2C, for the two animals, respectively). Subsequent training greatly enhanced the animals' contour detection ability (Figures 2E and 2F), indicating a repetition of the perceptual learning process at the new visual field location. Note that the initial performance at Location 2 could be even worse than the initial performance at Location 1, suggesting that a certain degree of perceptual learning had already taken place during the preliminary procedural training at the first location (Location 1). But in any event, the changes in the animals' behavioral responses along with the location specificity of these changes clearly demonstrated perceptual learning in contour detection. As the animals learned to detect the embedded contours, recordings from area V1 showed parallel changes in neuronal responses that were related to contour integration.

Absence of contour information in untrained V1

As a first stage of the experimental series and before training the monkeys on the contour detection task, we recorded responses of V1 neurons to the embedded contours when the naïve animals were only trained on a simple central fixation task. In this task the animals were required to detect the dimming of a small fixation target. They indicated the dimming by releasing a lever that they held from the outset of the trial. The visual stimuli were identical to those described previously in the contour detection task, except that the contour patterns were always displayed over the RF location

While the monkeys performed this central-dimming task and were passively exposed to the stimuli in the periphery, responses of single orientation-selective neurons to the contour patterns were recorded from superficial layers of V1. The orientation of the contour was adjusted to the preferred orientation of the recorded cell, and the central line element of the contour was centered in the RF. Although the visual contours consisting of different numbers of collinear lines were displayed at the RF location, the mean neuronal responses to these stimuli were indistinguishable and were identical to the responses to the noise pattern (Figure 3A; Kruskal-Wallis test, $p = 0.987$). This experiment showed that the information about the embedded contours was absent in the firing rates of V1 neurons in monkeys naïve to the contour detection task.

In the central-dimming task, the animals were not required to pay attention to the spatial location where the contours were embedded. To examine whether spatial attention by itself would differentiate neuronal responses to contours of different lengths, in the second stage of experiment the monkeys were trained to maintain their fixation at the fixation point while attending to the middle of the contour pattern, where the central line element of the contours was slightly dimmed at times randomly distributed between 476 and 1476 ms after stimulus onset. The animals had to respond to the dimming by releasing the lever within 600 ms. We refer to this task as the peripheral-dimming task, as distinguished from the central-dimming task in the previous stage. Even though the animals attended to the center of contour patterns and to the RF location, no significant difference was observed in the mean V1 responses to contours of different lengths (Figure 3B; Kruskal-Wallis test, $p = 0.997$). Therefore, for monkeys that had never been trained to detect the embedded visual contours, V1 responses were independent of the presence and length of contours, regardless of whether or not the animals' attention was directed to the target location.

Learning-induced V1 changes related to contour integration

In an earlier study we observed a close correlation between V1 responses and contour lengths in monkeys well trained on contour detection (Li et al., 2006). The absence of contour-related responses here in animals naïve to the contour detection task suggests that either the process of training on contour detection, or the specific action of conducting the contour detection task, results in neuronal responses to the visual contours. To differentiate between these possibilities, in the next stage of the experiments the animals were trained to perform the contour detection task (Figures 2A–2C). After the animals' performance had reached plateaus, and when the animals were actively detecting the embedded contours, a late response component associated with contour length emerged—the longer the contours, the stronger the neuronal responses (Figure 3C). These contour-related responses started about 100–120 ms after stimulus outset. Statistical tests showed that the contour length became a significant factor influencing neuronal responses (Kruskal-Wallis test, $p = 0.000019$), and that the responses in the 9-line condition were significantly stronger than the 1-line condition (Wilcoxon rank-sum test, $p = 0.000018$). Note that the initial response peak in the peristimulus-time histogram (PSTH) did not systematically change with the contour length.

With the monkeys already trained on the contour detection task, we recorded again neuronal responses to the contour patterns while the animals performed the central (Figure 3D) and peripheral (Figure 3E) dimming tasks in which the visual contours were task-irrelevant. Longer contours still evoked stronger neuronal responses: Contour length remained a significant influencing factor (Kruskal-Wallis test, $p = 0.0038$ in Figure 3D; $p = 0.045$ in Figure 3E); responses in the 9-line condition were significantly stronger than the 1-line condition (Wilcoxon rank-sum test, $p = 0.00054$ in Figure 3D; $p = 0.012$ in Figure 3E). This is in sharp contrast with the experimental stages in which the animals did the same dimming tasks but had never been trained to perform the contour detection task (compare Figure 3D with Figure 3A, and Figure 3E with Figure 3B). On the other hand, the amount of facilitation with contour length was significantly weaker in both dimming tasks than the contour detection task (compare Figures 3D and 3E with Figure 3C; Wilcoxon rank-sum test, $p < 0.05$ for all 5-, 7-, and 9-line conditions).

To examine the changes in neuronal responses in relation to contour integration quantitatively, receiver-operating-characteristic (ROC) analysis (Tolhurst et al., 1983; Parker and Newsome, 1998) was used to calculate the probability that an ideal observer can distinguish the contour pattern from the noise pattern simply based on a single neuronal spike count. By calculating this probability as a function of the number of collinear lines forming the contour, we obtained a neurometric curve, which indicates how well the contour and noise patterns could be differentiated at different contour lengths based on neuronal responses. In the central- and peripheral-dimming tasks before training on contour detection, the neurometric curves averaged across all recorded cells were at the chance (50%) level, independent of the contour length (Figure 3F); therefore, the embedded contours could not be detected simply based on V1 firing rates. After training, however, the discrimination ability of V1 neurons increased with contour length, and this emergent response property was most robust when animals did the contour detection task.

Specificity of learning-induced changes to visual field location

The specificity of perceptual learning to visual-field location suggests learning-induced changes in early visual cortical areas where the visual field is represented at the finest resolution by a topographic map. To examine whether the emergence of V1 responses to the camouflaged contours was specific to the trained retinotopic region, in the next experimental stage, we flipped the contour and noise patterns around the vertical meridian, changing their display locations to the untrained lower-right and upper-left quadrants (Location 2; Figure 2D). We

first repeated the central- and peripheral-dimming experiments by recording from the V1 region representing the untrained lower-right visual field. The differences in the mean neuronal responses to contours of different lengths were significantly smaller than those at the trained Location 1 for the same central- or peripheral-dimming tasks (compare Figure 4A with Figure 3D, and Figure 4B with Figure 3E). Statistical tests showed that, at the new visual-field location, contour length was no longer a significant factor affecting neuronal responses in both the central-dimming (Kruskal-Wallis test, $p = 0.42$) and peripheral-dimming (Kruskal-Wallis test, $p = 0.74$) tasks. Subsequent training on contour detection in this untrained area induced the profound changes in neuronal responses that were comparable to those observed in the previously trained area (Figures 4C–4F). In detail, the contour length became again a significant modulatory factor in the contour detection (Kruskal-Wallis test, $p = 0.0000017$), central-dimming ($p = 0.00018$), and peripheral-dimming ($p = 0.0019$) experiments. Moreover, the contour-related responses in the 9-line condition were significantly stronger than the 1-line condition in the contour detection (Wilcoxon rank-sum test, $p = 0.000012$), central-dimming ($p = 0.000089$), and peripheral-dimming ($p = 0.00028$) tasks. However, the amount of facilitation of neuronal responses with contour length was significantly weaker in both central- and peripheral-dimming tasks than the contour detection task (Wilcoxon rank-sum test, $p < 0.05$ for all 3-, 5-, 7-, and 9-line conditions).

Parallel effects of training on perception and neuronal responses

In perceptual learning, the improvement in perceptual skills evolves with practice and time (Figure 2; see also Crist et al., 1997;Crist et al., 2001;Li et al., 2004). To examine the parallel effects of training on the animal's ability in contour detection and on V1 responses, we arbitrarily divided the data collected from the contour detection task into two chronological groups: those collected before the animals' performance reached plateaus (referred to as the early phase) versus those collected afterwards (referred to as the late phase; the dividing time point is indicated by the vertical dashed lines in Figure 2). A quantitative comparison of the results from the late and early training phases showed that the animals' performance on contour detection was significantly improved in the late phase, resulting in a steepening of the psychometric curve that indicates the animals' detection performance as a function of the number of collinear lines (Figure 5A; unpaired Student's t-test, $p = 0.43, 0.12, 0.0037, 0.0042,$ and 0.00040, respectively, for the 1-, 3-, 5-, 7-, and 9-line conditions). The neurometric curves averaged across the cells recorded within the same periods of time showed parallel changes, increasing steepness with training.

To examine whether the parallel improvement in both psychometric and neurometric functions was restricted to the trained area, we compared the averaged psychometric and neurometric functions at Location 1 in the late training phase with those at Location 2 in the early phase of training (Figure 5B). When the stimuli were moved from the trained (Location 1) to untrained (Location2) visual field location, the animal's performance on contour detection, the psychometric curve, was significantly lower (unpaired Student's t-test, $p = 0.99, 0.11, 0.027, 0.0087,$ and 0.0018, respectively, for the 1-, 3-, 5-, 7-, and 9-line conditions). The contour detectability based on ROC analysis of V1 responses, the neurometric function, was also lowered. These results indicate a lack of generalization of the learning effects to the new visual-field location and to the corresponding retinotopic region in V1.

Disappearance of learning-induced changes in V1 under anesthesia

The experiments described previously indicate that contour-related responses in V1 are driven by a number of factors, including visual stimuli, training experience, and top-down control. To obtain a measure of responses that are wholly independent of behaviorally mediated top-down influences, in the last experimental stage we recorded neuronal responses to the embedded contours in the trained V1 regions when the same animals anesthetized. The late

response components associated with the embedded contours completely disappeared (Figure 6; Kruskal-Wallis test, p = 0.997), suggesting that even in the trained V1 area the stimulus-driven process by itself does not suffice to extract the embedded visual contours. In addition to the absence of contour-related responses, the overall neuronal responses to the complex stimuli were much weaker in anesthetized state than in awake state, as has been reported earlier (Lamme et al., 1998).

DISCUSSION

It has been shown that training monkeys on some visual discrimination tasks can increase the selectivity of V1 neurons for the trained visual stimuli (Crist et al., 2001; Schoups et al., 2001; Li et al., 2004). Moreover, training can enhance V1 responses to familiar targets within distractors (Lee et al., 2002; Kourtzi et al., 2005). Our current study is consistent with these observations, supporting the notion of adaptive cortical processing of visual information even at the earliest stage. The most important finding of the current study is that even the process of contour integration in V1, which is generally believed to be a hard-wired process, can strongly depend on perceptual learning and top-down influences.

Task-Specific Top-Down Influences and Experience-Dependent Changes

Psychophysical studies have shown that for very similar (Fahle and Morgan, 1996; Fahle, 1997; Sigman and Gilbert, 2000) or even completely identical (Shiu and Pashler, 1992; Ahissar and Hochstein, 1993; Saffell and Matthews, 2003) visual stimuli, a subject's discrimination ability for a given stimulus attribute can be improved only if the attribute is attended and related to the trained task. Repeated passive exposure to the stimuli rarely induces any learning effect. Our current study demonstrated that top-down influences specific to the task of contour detection were responsible for the emergence of contour-related responses in V1, as passive exposure to the embedded contours did not alter V1 responses, nor did attending to the contour location.

As different forms of top-down influences could be mingled in any perceptual task, it is arguable that the late response components that we observed in V1 of the trained animals could be attributed to attentional load or task difficulty rather than the contour integration task itself. One would then, however, expect a negative correlation between V1 responses and the contour length, because longer or more salient contours are much easier to detect, and therefore less attentionally demanding, than short ones. To the contrary, we show that longer contours are associated with better detection performance and stronger neuronal responses. Moreover, in the periphery dimming task, which is more difficult than simple fixation (the central dimming task), neuronal responses were actually somewhat weaker (compare Figure 3D with Figure 3E, and Figure 4D with Figure 4E). Our findings therefore suggest that top-down influences are not limited to spatial attention, but can convey much more information, including information about specific perceptual tasks. By using perceptual learning to engage different components of top-down control, first spatial attention and then task, we were able to separately see the effects of these components. Consistent with our previous studies showing task-specific modifications of V1 response properties (Crist et al., 2001; Li et al., 2004), our findings have important implications for the neuronal mechanisms underlying the task specificity of perceptual learning, in that learning may involve the way top-down influences engage certain cortical areas. A recent fMRI study has shown that task-specific top-down influences can trigger a large-scale reorganization of cortical activity that can account for the improvement with training in detecting a particular shape within similar distracters (Sigman et al., 2005). That study is consistent with the current study in that both the overall behavioral performance and V1 activity increase significantly with perceptual learning.

The parallel changes of psychometric and neurometric functions indicate that the improvement in the animal's performance on contour detection is associated with the enhancement of V1 signals for the embedded contours. On the other hand, the neurometric curve was significantly below the psychometric curve over the same training period (Figure 5), indicating that the prediction of contour-detection performance based on single spike count is generally worse than the animals' actual performance. A possible explanation is that several V1 neurons with RFs lying along the contours contribute to detection of an extended contour. Interestingly, our earlier study has shown that, for monkeys extensively trained on contour detection, the difference between psychometric and neurometric functions can be much smaller (Li et al., 2006). Taken together, these observations suggest that, although both the overall behavioral and neuronal responses are gradually enhanced in parallel with training, the behavioral performance reaches a plateau earlier than the average changes in V1. It may be that different subsets of neurons exhibit different learning speeds, some more rapid, and thus more contemporaneous with the behavioral change, than the others. Alternatively, perceptual learning may first involve engagement of a cortical area by feedback to that area, and then a consolidation and increased efficiency in the representation of the trained stimulus.

Once the animals had been trained, a measure of contour-related responses was observed even in the central- and peripheral-dimming tasks in which the contours were task-irrelevant. This indicates that the emergence of contour-related responses in V1 is associated with the animals' particular experience with the contour detection task. It is likely that contour integration is turned into a rather automatic process due to repetitive practice of the detection task. This might be interpreted in one of two ways. Either a component of the contour related responses becomes mediated entirely by local circuits without top-down control, or that the contours become more salient, attracting attention to involuntarily engage the animal in the contour detection task. Nevertheless, the facilitation of neuronal responses by the embedded contours was strongest only if the animals actively performed the contour detection task, suggesting an important role of task-specific top-down influences in contour integration at all stages of training. This is strongly supported by our another observation that contour-related responses were abolished, not merely attenuated, by anesthesia, a finding that is consistent with the observation that V1 responses associated with figure-ground segregation based on texture differences are absent under anesthesia (Lamme et al., 1998). The removal of contour-related responses under anesthesia suggests that the integration process cannot operate without a measure of top-down control. This is not inconsistent, however, with the idea that local circuits, such as the long range horizontal connections, are involved in contour integration. Rather, we propose that it is the interaction between local circuits and cortical feedback that results in V1 responses to salient contours, and that the operation of local circuits is gated by feedback connections. Supporting evidence for such an interaction is suggested by the findings that higher-order cortical areas, such as the lateral intraparietal area (LIP), are involved in generating saliency maps about visual stimuli according to both stimulus features and behavioral goals (Ipata et al., 2006). The recurrent processing within and across cortical areas may account for the delay of contour-related responses in V1, and is in agreement with the notion of visual routines (Ullman, 1984) and incremental grouping (Roelfsema, 2006) by which the bottom-up representations of basic stimulus attributes are shaped by processes involving higher-level cognitive influences, resulting in extraction and enhancement of behaviorally relevant information based on the knowledge of specific tasks.

The complete absence of contour-related responses in monkey V1 before training on contour detection seems to be at variance from psychophysical observations that collinear facilitation in perception is present in naïve human subjects. It is worth noting that when a naïve subject is performing a task as instructed, the appropriate task-specific top-down influences have already taken effect. In fact, immediately after the initial operational training on contour detection and in the first couple of sessions in the early perceptual learning phase, we were

already able to find V1 cells showing significant responses to the embedded contours. This is consistent with the psychophysical finding that a proper deployment of attentional resources to collinear stimuli triggers collinear facilitation effects on perception (Freeman et al., 2001; Shani and Sagi, 2005).

Neural Mechanisms of Perceptual Learning

Our findings, which support the attention-gated reinforcement learning model (Roelfsema and van Ooyen, 2005), have broader implications for the general neural mechanisms of perceptual learning. Different visual cortical areas are involved in processing different attributes of visual stimuli. For those cortical areas that are intimately involved in processing the information related to the current task, top-down influences specific to the task can dynamically alter the local circuitry to meet the requirements of the perceptual task. During repeated execution of the same task, state changes in cortical circuits occur under task-specific top-down control. Perceptual learning may be a process by which the feedback connections associate with the appropriate intrinsic connections during the encoding of the learned information, and the same set of relationships are utilized during the recall of this information.

EXPERIMENTAL PROCEDURES

Visual Stimuli

Visual stimuli, generated by a VSG2/5 stimulator (Cambridge Research Systems) on a CRT monitor, were composed of white (19.5 cd/m^2) line segments on a gray (6.5 cd/m^2) background. Each stimulus display contained two circular patterns (Figure 2A), a contour pattern with 1, 3, 5, 7, or 9 collinear lines embedded in the center of an array of randomly oriented lines, and a noise pattern without any embedded contour. The noise pattern was identical to the contour pattern if the number of collinear lines in the contour was 1. To construct the stimulus patterns, a circular area of $5.2°$ in diameter was divided into 13 by 13 compartments, each containing a randomly oriented line segment of $0.2°$ by $0.05°$ whose position was slightly jittered within its compartment. This arrangement defined an average line spacing of $0.4°$ in the stimulus pattern. The orientation and position jitter of each line was re-randomized in each trial. By aligning some adjacent line segments in a collinear, evenly spaced arrangement, a straight contour was generated in the center of the contour pattern. The center-to-center distance between two adjacent collinear lines was fixed at $0.44°$ (for details on the geometry of stimulus generation see Li et al., 2006).

Behavioral Tasks

A trial began when a lever was pulled by the animal. A $0.08°$ fixation point (FP) was displayed in the CRT center. Eye positions were sampled at 30 Hz by an infrared tracking system (Matsuda, K. *et al., Soc. Neurosci. Abstr* **26**, 744.2, 2000), which offered a spatial resolution of up to $0.05°$ (for details see Li et al., 2006). Within 600 ms after FP presentation the animal was required to fixate within an invisible circular window of $0.5°$ in radius around the FP. After the animal maintained its fixation for 191 ms, the contour pattern and noise pattern were presented. Subsequent timing and behavioral requirements in a trial were different for different tasks, which are described in detail when each task is introduced in the Results section. The time window allowing the animal to make a response in each trial of a task was 600 ms. Different stimulus conditions in a task were randomized, and each stimulus condition was repeated 10–40 times.

Electrophysiological Recordings

Recordings in all experiments were confined to the same V1 region for each hemisphere, about 5 by 5 mm in area, which corresponded to eccentricities $3.0°$ to $4.8°$. All recordings were

restricted to the first 500 µm from entering the cortex. To further ensure the electrode was in the superficial layers, background spontaneous activity was monitored. Superficial-layer neurons were typically marked by a lack of spontaneous activity (seethe population PSTHs before stimulus onset in Figure 3 and Figure 4; also refer to von der Heydt and Peterhans, 1989;Snodderly and Gur, 1995). Entering from superficial layers into layer 4 is indicated by a remarkable increase in background activity. The RFs were quantitatively mapped using the same methods described elsewhere (Li et al., 2006).

For the experiments that required the animals to perform certain tasks, daily recordings were carried out four days a week, and each experimental stage lasted about one month for each animal. In the last experiment in which the animals were anesthetized, venous cannulation and tracheal intubation were conducted following an initial induction with ketamine hydrochloride (10 mg/kg). Anesthesia and paralysis were maintained using intravenous infusion of Nembutal (pentobarbital sodium, 1.5 mg/kg/hr) and Norcuron (vecuronium bromide, 0.1 mg/kg/hr). The animals were artificially respired. EKG, expired CO_2, and rectal temperature were monitored continuously to ensure that the anesthesia was not very deep. All procedures were conducted in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and under approval of Institutional Animal Care and Use Committee at Rockefeller University.

Data Analysis

The spike counts within a 400 ms window, beginning from 110 after stimulus onset, was used to calculate the neurometric curve with ROC analysis (Tolhurst et al., 1983; Parker and Newsome, 1998). In brief, the ROC curve for a given number of collinear lines in the contour detection task was calculated by plotting the hit rate (the percentage of spike counts larger than a given decision threshold when the contour pattern was presented to the RF) against the false alarm rate (the percentage of spikes larger than the same threshold when the noise pattern was shown) while varying the decision threshold from the smallest to the largest spike count. The area under the ROC curve corresponds to the probability that an ideal observer can correctly identify, based on a single spike count, whether the contour or noise pattern was presented for the given contour length, and thus gives a data point on the neurometric curve.

To estimate meaningful error bars and statistical significance of the averaged data, the cell population was resampled by bootstrapping method (Effron and Tibshirani, 1993), and then, the spike counts in each trial from a cell were resampled as well (10^3 resampling iterations were used in both cases). This resampling method accounts for both the variability due to the limited number of trials per stimulus display, and due to the difference between individual cells.

Supplemental data

The Supplemental Data include one supplemental figure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

Ahissar M, Hochstein S. Attentional control of early perceptual learning. Proc. Natl. Acad. Sci. U. S. A 1993;90:5718–5722. [PubMed: 8516322]

Bauer R, Heinze S. Contour integration in striate cortex. Classic cell responses or cooperative selection? Exp. Brain Res 2002;147:145–152. [PubMed: 12410329]

Crist RE, Kapadia MK, Westheimer G, Gilbert CD. Perceptual learning of spatial localization: specificity for orientation, position, and context. J. Neurophysiol 1997;78:2889–2894. [PubMed: 9405509]

Crist RE, Li W, Gilbert CD. Learning to see: experience and attention in primary visual cortex. Nat. Neurosci 2001;4:519–525. [PubMed: 11319561]

Effron, B.; Tibshirani, RJ. An Introduction to the Bootstrap. New York: Chapmann and Hall; 1993.

Ernst UA, Mandon S, Pawelzik KR, Kreiter AK. How ideal do macaque monkeys integrate contours? Neurocomputing 2004;50–60:971–977.

Fahle M. Specificity of learning curvature, orientation, and vernier discriminations. Vision Res 1997;37:1885–1895. [PubMed: 9274774]

Fahle M, Morgan M. No transfer of perceptual learning between similar stimuli in the same retinal position. Curr. Biol 1996;6:292–297. [PubMed: 8805246]

Field DJ, Hayes A, Hess RF. Contour integration by the human visual system: evidence for a local "association field". Vision Res 1993;33:173–193. [PubMed: 8447091]

Freeman E, Sagi D, Driver J. Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. Nat. Neurosci 2001;4:1032–1036. [PubMed: 11559851]

Geisler WS, Perry JS, Super BJ, Gallogly DP. Edge co-occurrence in natural images predicts contour grouping performance. Vision Res 2001;41:711–724. [PubMed: 11248261]

Gilbert CD, Wiesel TN. Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. Nature 1979;280:120–125. [PubMed: 552600]

Gilbert CD, Wiesel TN. Clustered intrinsic connections in cat visual cortex. J. Neurosci 1983;3:1116–1133. [PubMed: 6188819]

Gilbert CD, Wiesel TN. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. J. Neurosci 1989;9:2432–2442. [PubMed: 2746337]

Ipata AE, Gee AL, Gottlieb J, Bisley JW, Goldberg ME. LIP responses to a popout stimulus are reduced if it is overtly ignored. Nat. Neurosci 2006;9:1071–1076. [PubMed: 16819520]

Kapadia MK, Ito M, Gilbert CD, Westheimer G. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. Neuron 1995;15:843–856. [PubMed: 7576633]

Kourtzi Z, Betts LR, Sarkheil P, Welchman AE. Distributed neural plasticity for shape learning in the human visual cortex. Plos Biology 2005;3:1317–1327.

Kovács I, Kozma P, Feher A, Benedek G. Late maturation of visual spatial integration in humans. Proc. Natl. Acad. Sci. U. S. A 1999;96:12204–12209. [PubMed: 10518600]

Lamme VA, Zipser K, Spekreijse H. Figure-ground activity in primary visual cortex is suppressed by anesthesia. Proc. Natl. Acad. Sci. U. S. A 1998;95:3263–3268. [PubMed: 9501251]

Lee TS, Yang CF, Romero RD, Mumford D. Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. Nat. Neurosci 2002;5:589–597. [PubMed: 12021764]

Li W, Gilbert CD. Global contour saliency and local colinear interactions. J. Neurophysiol 2002;88:2846–2856. [PubMed: 12424317]

Li W, Piech V, Gilbert CD. Perceptual learning and top-down influences in primary visual cortex. Nat. Neurosci 2004;7:651–657. [PubMed: 15156149]

Li W, Piech V, Gilbert CD. Contour saliency in primary visual cortex. Neuron 2006;50:951–962. [PubMed: 16772175]

Li Z. A neural model of contour integration in the primary visual cortex. Neural Comput 1998;10:903–940. [PubMed: 9573412]

Parker AJ, Newsome WT. Sense and the single neuron: Probing the physiology of perception. Annu. Rev. Neurosci 1998;21:227–277. [PubMed: 9530497]

Polat U, Mizobe K, Pettet MW, Kasamatsu T, Norcia AM. Collinear stimuli regulate visual responses depending on cell's contrast threshold. Nature 1998;391:580–584. [PubMed: 9468134]

Polat U, Sagi D. Spatial interactions in human vision - from near to far via experience-dependent cascades of connections. Proc. Natl. Acad. Sci. U. S. A 1994;91:1206–1209. [PubMed: 8108388]

Rockland KS, Lund JS, Humphrey AL. Anatomical binding of intrinsic connections in striate cortex of tree shrews (Tupaia glis). J. Comp. Neurol 1982;209:41–58. [PubMed: 7119173]

Roelfsema PR. Cortical algorithms for perceptual grouping. Annu. Rev. Neurosci 2006;29:203–227. [PubMed: 16776584]

Roelfsema PR, van Ooyen A. Attention-gated reinforcement learning of internal representations for classification. Neural Comput 2005;17:2176–2214. [PubMed: 16105222]

Saffell T, Matthews N. Task-specific perceptual learning on speed and direction discrimination. Vision Res 2003;43:1365–1374. [PubMed: 12742106]

Schmidt KE, Goebel R, Lowel S, Singer W. The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. Eur. J. Neurosci 1997;9:1083–1089. [PubMed: 9182961]

Schoups A, Vogels R, Qian N, Orban G. Practising orientation identification improves orientation coding in V1 neurons. Nature 2001;412:549–553. [PubMed: 11484056]

Shani R, Sagi D. Eccentricity effects on lateral interactions. Vision Res 2005;45:2009–2024. [PubMed: 15820518]

Shiu LP, Pashler H. Improvement in line orientation discrimination is retinally local but dependent on cognitive set. Percept. Psychophys 1992;52:582–588. [PubMed: 1437491]

Sigman M, Cecchi GA, Gilbert CD, Magnasco MO. On a common circle: natural scenes and Gestalt rules. Proc. Natl. Acad. Sci. U. S. A 2001;98:1935–1940. [PubMed: 11172054]

Sigman M, Gilbert CD. Learning to find a shape. Nat. Neurosci 2000;3:264–269. [PubMed: 10700259]

Sigman M, Pan H, Yang YH, Stern E, Silbersweig D, Gilbert CD. Top-down reorganization of activity in the visual pathway after learning a shape identification task. Neuron 2005;46:823–835. [PubMed: 15924867]

Snodderly DM, Gur M. Organization of striate cortex of alert, trained monkeys (Macaca fascicularis): ongoing activity, stimulus selectivity, and widths of receptive field activating regions. J. Neurophysiol 1995;74:2100–2125. [PubMed: 8592200]

Stettler DD, Das A, Bennett J, Gilbert CD. Lateral connectivity and contextual interactions in macaque primary visual cortex. Neuron 2002;36:739–750. [PubMed: 12441061]

Tolhurst DJ, Movshon JA, Dean AF. The Statistical Reliability of Signals in Single Neurons in Cat and Monkey Visual-Cortex. Vision Res 1983;23:775–785. [PubMed: 6623937]

Ullman S. Visual routines. Cognition 1984;18:97–159. [PubMed: 6543165]

Ullman S. Low-level aspects of segmentation and recognition. Philos. Trans. R. Soc. Lond. B. Biol. Sci 1992;337:371–378. [PubMed: 1359588]

VanRullen R, Delorme A, Thorpe SJ. Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. Neurocomputing 2001;38:1003–1009.

von der Heydt R, Peterhans E. Mechanisms of Contour Perception in Monkey Visual-Cortex .1. Lines of Pattern Discontinuity. J. Neurosci 1989;9:1731–1748. [PubMed: 2723747]

Wertheimer M. Untersuchungen zur Lehre von der Gestalt. Psychol. Forsch 1923;4:301–350.

Yen SC, Finkel LH. Extraction of perceptually salient contours by striate cortical networks. Vision Res 1998;38:719–741. [PubMed: 9604102]
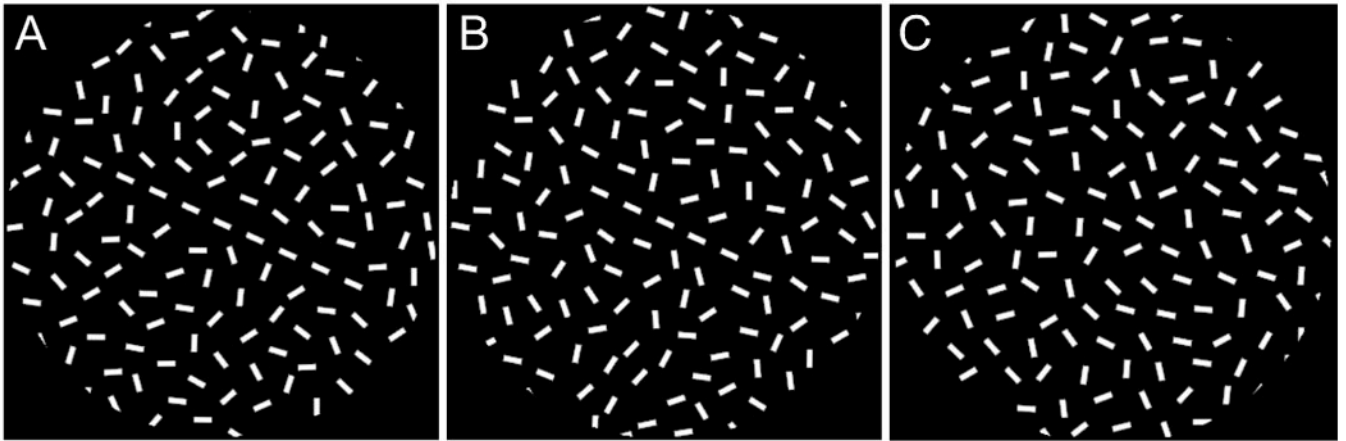
Figure 1. Visual contours formed by collinear line segments embedded in a background of randomly oriented lines
Within the same background, perceptual saliency of contours varies with the number of collinear lines (**A** and **B**) and with the spacing between them (**B** and **C**).
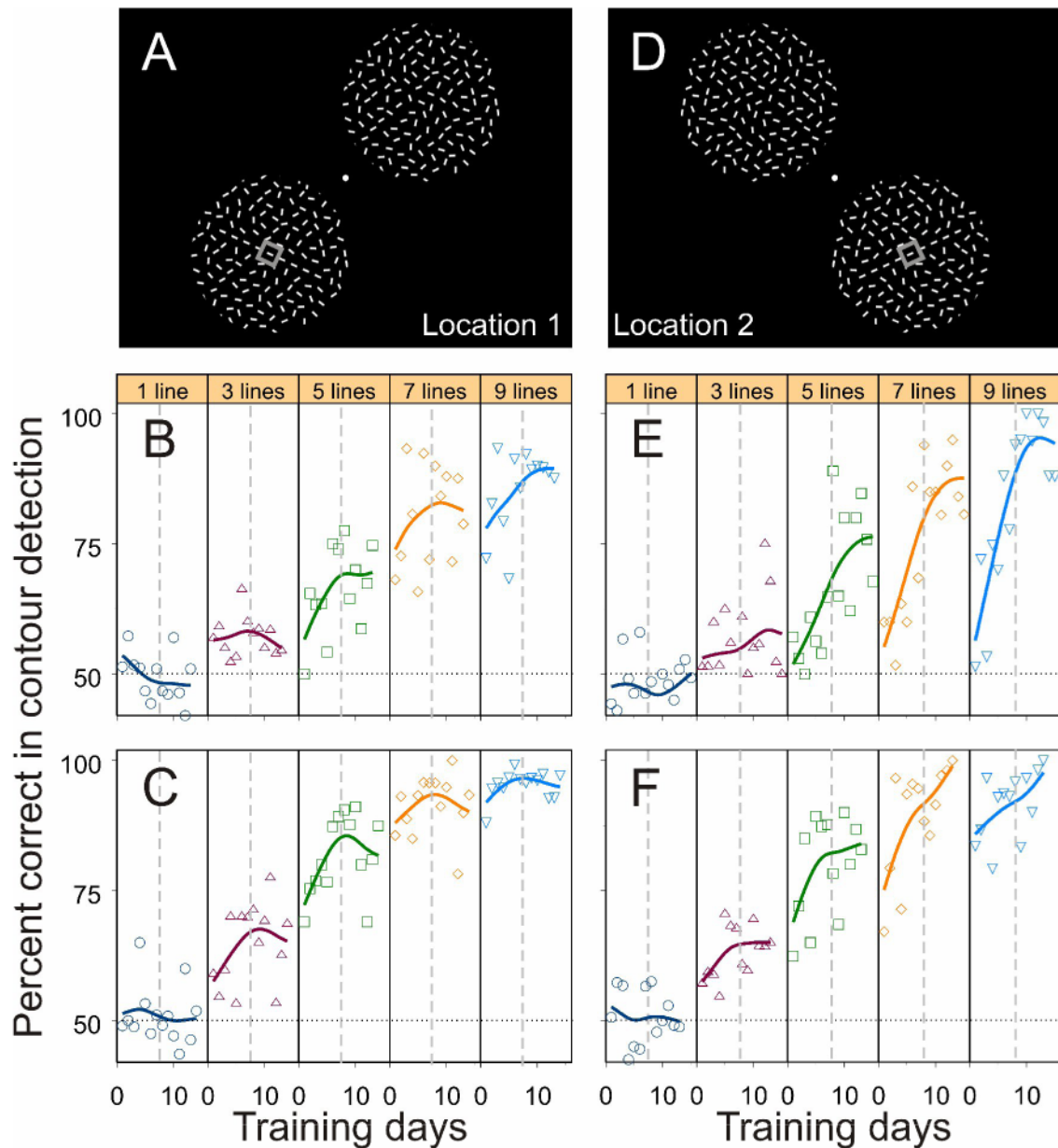
Figure 2. Learning curves in the contour detection task
(A) The visual field locations trained first (named Location 1). The animal maintained its fixation at the central spot throughout the trial. The positions of the contour and noise patterns were randomly interchanged from trial to trial. The animal indicated the location of the contour pattern by making a saccadic eye movement to either of two target dots displayed at the end of the trial. The small grey square denotes the receptive field during V1 recordings.
(B and **C)** The learning curves of the two animals, respectively, at visual-field Location 1. The five panels correspond to the five stimulus conditions in which the number of collinear lines forming the embedded contour was 1, 3, 5, 7, 9, respectively. Each dot represents the mean performance within a day based on 20–160 trials (with a median of 90 trials), and each curve was generated by a cubic spline with 3 degrees of freedom. The horizontal dotted line denotes the chance level (50%); the vertical dashed lines indicate the arbitrary division of the whole training period into the early phase and the late phase.

(D) The new visual field location (named Location 2) tested after all the experiments at Location 1 were completed.

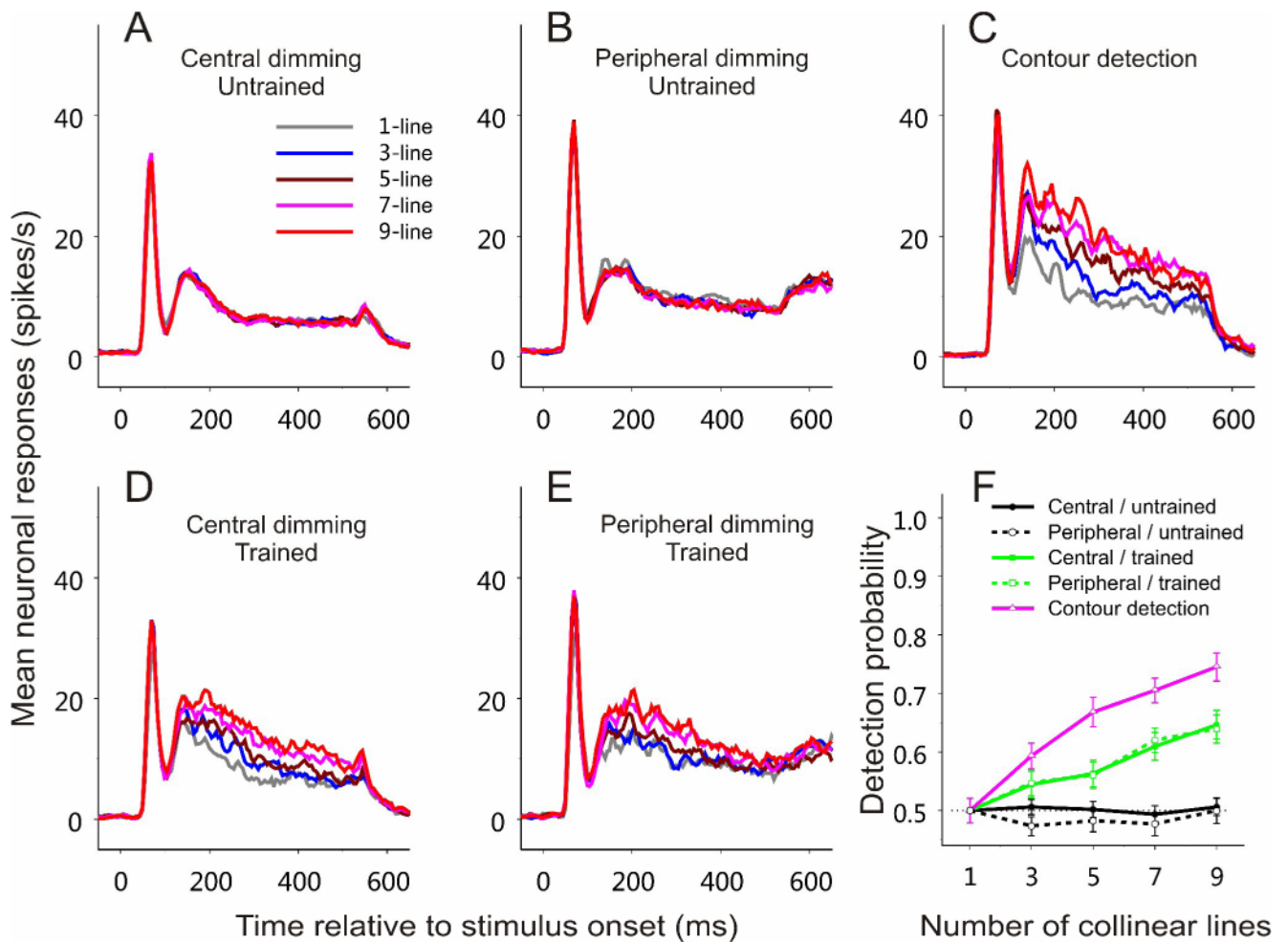**(E** and **F)** The learning curves of the two animals, respectively, at Location 2.

Figure 3. Emergence of contour-related responses in V1 with training on contour detection
(A–E) Peristimulus time histograms (PSTHs) were constructed by binning neuronal spikes at
5 ms resolution over time and by averaging across all recorded cells for each experimental
condition. The PSTHs were smoothed using a rectangular window of 20 ms (boxcar filter).
Time 0 indicates stimulus onset. Five curves in each panel represent neuronal responses to
contours made up of 1, 3, 5, 7, 9 collinear lines respectively. Different panels show different
experimental stages. All data were collected from a 5×5 mm region in the right hemisphere,
and the visual stimuli were presented at visual field Location 1 (Figure 2A).
(A) Central-dimming task before training on contour detection. The task was to respond to a
slight dimming of the fixation point. n = 73 (52 and 21 cells, respectively, from the two
animals).
(B) Peripheral-dimming task before training on contour detection. The task was to maintain
fixation at the fixation point and respond to a dimming of the line segment in the RF, which
was also the central element of the collinear contour. n = 45 (25 and 20, respectively, from the
two animals).
(C) Contour detection task. The animal had to indicate the location of the embedded contour
in a two-alternative-forced-choice task. Neuronal responses recorded in both correct and error
trials during the late training phase are included (refer to Figure 2 for the division of early and
late phases). n = 32 (13 and 19, respectively, from the two animals).
(D) Central-dimming task after training on contour detection. n = 47 (24 and 23, respectively,
from the two animals).

(E) Peripheral-dimming task after training on contour detection. n = 48 (26 and 22, respectively, from the two animals).

(F) Neurometric curves based on ROC analysis of the data shown in **(A)–(E)**, respectively. Each curve was averaged across all the recorded cells for each condition. Error bars represent ± SEM calculated by the method of bootstrapping (see Experimental Procedures).
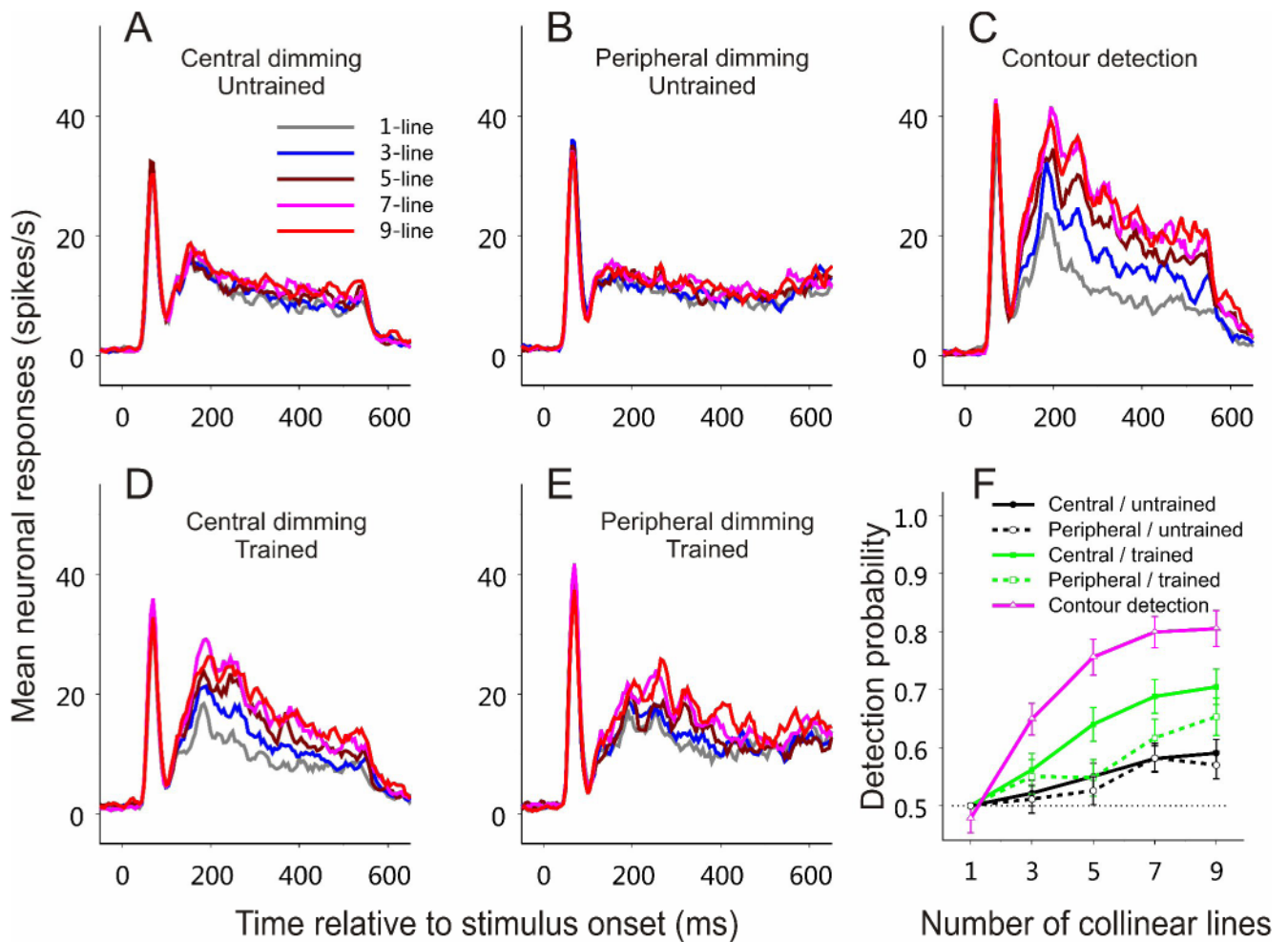
Figure 4. Results from a new visual field location
After all the experiments shown in Figure 3 had been completed at visual-field Location 1
(Figure 2A), the same series of experiments were repeated at a new stimulus location (Location
2; Figure 2D). All data were collected from a 5×5 mm V1 region in the left hemisphere.
(A–E) PSTHs based on averaged neuronal responses to contours consisting of 1, 3, 5, 7, 9
collinear lines in different experimental stages.
(A) Central-dimming task before training on contour detection. n = 50 (26 and 24 cells,
respectively, from the two animals).
(B) Peripheral-dimming task before training on contour detection. n = 47 (24 and 23,
respectively, from the two animals).
(C) Contour detection task during the late training phase. n = 28 (16 and 12, respectively, from
the two animals).
(D) Central-dimming task after training on contour detection. n = 37 (22 and 15, respectively,
from the two animals).
(E) Peripheral-dimming task after training on contour detection. n = 36 (22 and 14, respectively,
from the two animals).
(F) Averaged neurometric curves based on ROC analysis of the data shown in **(A)–(E)**,
respectively.
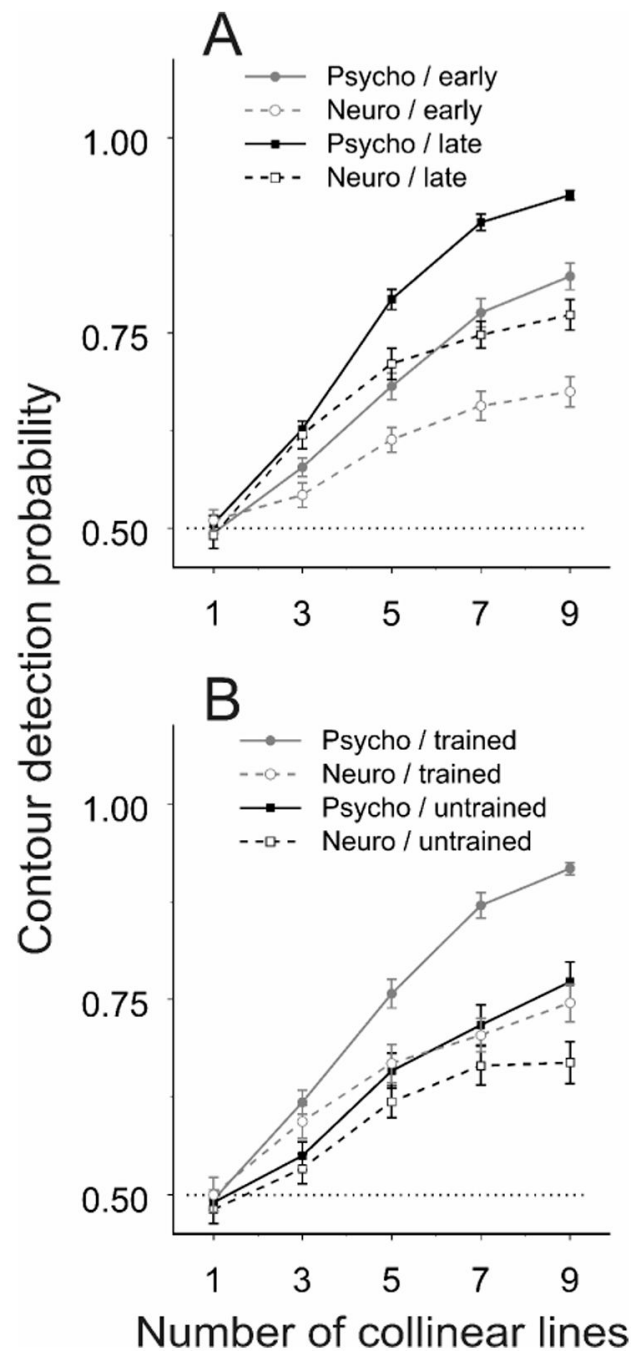Refer to Figure 3 for more details about each panel.

Figure 5. Parallel effects of learning on behavioral and neuronal responses
(**A**) The neurometric curves (averaged across all cells) and the psychometric curves (averaged over the corresponding recording sessions) are shown for two different training phases: the early phase before the animal's performance reached plateaus (n = 66; 31 and 35 cells from the two animals, respectively) and the following late phase (n = 60; 29 and 31 from the two animals, respectively). The dividing time point between the early and late phases is indicated by the vertical dashed lines in Figure 2. Data were pooled from the four hemispheres of the two animals within each training phase. A significant steepening in both psychometric and neurometric curves was observed with training. Error bars represent ± SEM calculated by the method of bootstrapping (see Experimental Procedures).

(B) Data collected in the early training phase at visual-field Location 2 (untrained; n = 32, 13 and 19 cells from the two animals, respectively) are compared with those in the late training phase at the previously trained Location 1 (trained; n = 34, 15 and 19 from the two animals, respectively). A significant drop in steepness of psychometric and neurometric curves was seen in the untrained area relative to the trained.
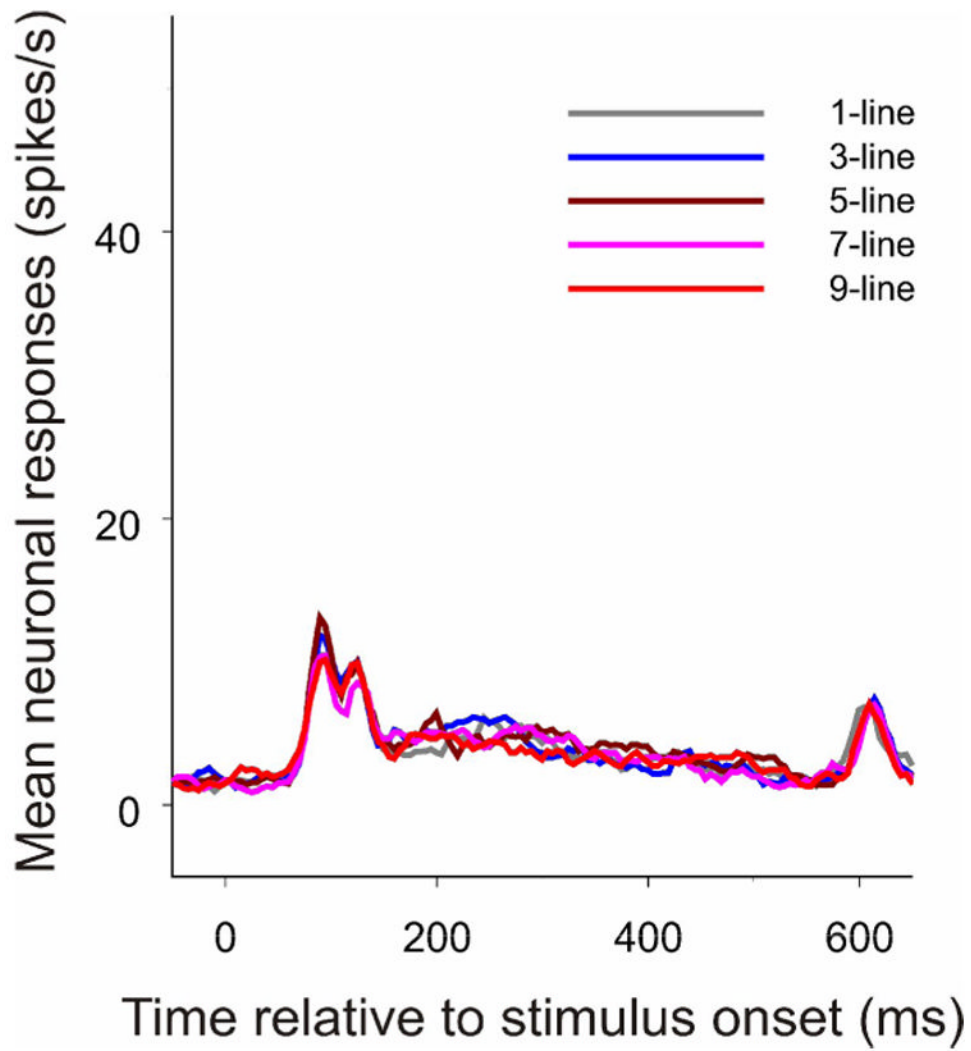
Figure 6. The effect of anesthesia
When all behaviorally related top-down influences were eliminated by anesthesia, contour-related responses disappeared in the trained V1 regions (n = 22). The PSTHs were constructed as described in Figure 3.