Learning to see: experience and attention in primary visual cortex

Roy E. Crist, Wu Li and Charles D. Gilbert

The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA
Correspondence should be addressed to C.D.G. (gilbert@rockefeller.edu)

The response properties of neurons in primary sensory cortices remain malleable throughout life. The existence of such plasticity, and the characteristics of a form of implicit learning known as perceptual learning, suggest that changes in primary sensory cortex may mediate learning. We explored whether modification of the functional properties of primary visual cortex (V1) accompanies perceptual learning. Basic receptive field properties, such as location, size and orientation selectivity, were unaffected by perceptual training, and visual topography (as measured by magnification factor) was indistinguishable between trained and untrained animals. On the other hand, the influence of contextual stimuli placed outside the receptive field showed a change consistent with the trained discrimination. Furthermore, this property showed task dependence, only being manifest when the animal was performing the trained discrimination.

Cortical plasticity is likely involved in the normal cognitive processing of adult animals, and is important even in the functioning of primary sensory areas. Whereas plasticity may underlie functional recovery following CNS lesions, it may also mediate functional changes accompanying experience and learning. Several characteristics of perceptual learning suggest the involvement of early stages in sensory processing, perhaps even as early as a stage of primary sensory cortex. Evidence in support of this idea has been found in the somatosensory and auditory systems. Primary sensory cortex is a useful model for learning because its underlying mechanisms—its circuitry, functional architecture and receptive field (RF) structure—can be accessibly studied.

Various cortical changes both are associated with improvement in perceptual performance and occur in primary sensory cortices, which suggests that the mechanisms of learning may be general to the neocortex as a whole. Independent studies of plasticity of RF properties and functional architecture of primary visual cortex (V1) heighten the possibility that plasticity associated with perceptual learning might occur there. Whereas certain RF properties such as ocular dominance are mutable only during a limited critical period early in postnatal development, a number of other properties, most notably, visual topography and RF size, can be influenced by visual experience throughout life.

In the visual system, psychophysical evidence shows that training can improve discrimination stimulus attributes, including position, depth, orientation, motion, texture, spatial phase and hyperacuity. In a previous study, we showed that practice with a particular visual discrimination task, three-line bisection, produced a substantial improvement that is specific to the trained stimulus. The specificity for position and orientation suggests that the early stages of visual processing are involved in the learning of this task. To determine whether V1 is involved in the learning of such visual discriminations, we trained two macaque monkeys to perform the same bisection task used with human subjects, and we recorded from cells in area V1. We examined the RF properties and the map of visual space in trained animals for changes that might relate to perceptual training. In addition to its classical response properties, a cell’s response to a stimulus within the RF is modulated by the presence of additional stimuli around the RF, and this modulation depends strongly on the geometric relationship of the stimulus elements. We explored the tuning of cells to shifts in the lateral placement of two parallel lines with positions analogous to those of lines in the bisection task. Furthermore, we examined the influence of the behavioral state of the animal on these interactions.

**Results**

Monkeys improve bisection discrimination with training

Both monkeys showed substantial improvement in bisection performance with training. Over the course of 30 weeks of training, monkey 1 showed a threshold reduction greater than a factor of three (Fig. 1c), and monkey 2 showed a 57% threshold reduction (Fig. 1d).

Monkeys were trained with a series of tasks that approximated the final bisection task. The initial threshold (Fig. 1b) was measured during the first week of training on the bisection task itself, and some learning may have occurred before this time. During the recording phase, the monkey had fewer opportunities to practice the bisection discrimination, but the threshold measured several months after electrophysiological recordings began shows that the improved performance was maintained over this period (Fig. 1b).

Cortical magnification does not change with training

One of the most striking effects of perceptual training in the somatosensory and auditory modalities is the remapping of the cortical representation of the sensory surface such that a much larger region of cortex is responsive to stimulation of the trained region. To examine whether a remapping of visual space accompanied the learning of the bisection discrimination task described above, we constructed a topographic map of the cortex of the trained monkeys from recordings of superficial layer cortical cells. The location of the RF center of an isolated unit was determined by recording the response to the flashing of a small bar in a pattern


correspondence should be addressed to C.D.G. (gilbert@rockefeller.edu)
ing bisection trials, many cells were strongly facilitated by a parallel bar. In a study of contextual interactions in monkeys viewing passively, cells in untrained animals did not show facilitation for this pattern \(^{21}\). The absence of such facilitation in untrained animals strengthens the conclusion that these interactions were the result of bisection training. A substrate for these interactions exists in V1 in the axon collaterals of superficial layer pyramidal cells, which extend over several millimeters. These cells connect units of similar orientation preference \(^{25-33}\) and are therefore well positioned to mediate interactions between the elements of the bisection stimulus used in our study. The ability of long-range horizontal connections to provide both excitatory and inhibitory input to their targets \(^{34-36}\) is also suggestive of a role in the interactions described above. The implementation of perceptual learning by a selective modulation of subsets of horizontal connections allows for the specificity of the learning for the details of stimulus configuration. A mechanism of cortical recruitment would not have this property, and further, might be expected to lead to a decline in performance in the untrained portions of visual space represented by the adjacent cortical regions. This 'robbing' is not seen in the perceptual learning experiments.

The contextual influence that best represents the attribute involved in the bisection task is the modulation in the cells' responses to a second, parallel line placed at varying separation from a line placed centrally within the RF. The observed effects were specific for this property, because they were not seen for a different contextual influence, that of a colinear line at varying offsets. To make a stronger connection between the modulation index and the trained task, one can ask how the scale of the sensitivity to line separation measured for an individual neuron compares to the threshold in the bisection task. At the point at which neurons show their greatest sensitivity (the steepest part of the slope of the contextual tuning curves), the amount of change in the distance between the parallel lines that gives one standard deviation change in the firing rate is 9.75 minutes of arc, on average. The threshold in the task, after training, was four minutes of arc. The sensitivities are therefore roughly of the same scale, but the differences argue that some amount of pooling in the activity of the neurons is required to achieve the level of behavioral performance.

Perhaps the most striking aspect of the findings reported here is the apparent ability of the cortex to dynamically modify the processing of visual information according to immediate behavioral requirements. The monkeys in this study were trained to perform two different tasks: a simple dimming task and the three-line bisection discrimination. Lateral interactions in trained monkeys depended on the task the monkey was performing at the time. The task dependence of the contextual interaction permits the same neurons to mediate entirely different perceptual functions that may require opposing neuronal mechanisms. The inhibition of responses by parallel lines under the fixation task has been suggested to be involved in surface segmentation \(^{21}\). As shown here, this inhibition can switch to facilitation during the performance of the three-line bisection task. One would not want to design a system in which three-line bisection training would disrupt the subject's ability to segment the visual scene. To allow the performance of both tasks, one would either have to segregate the neurons mediating the task into separate functional compartments, or allow the same neurons to multiplex their function in a task-dependent fashion. The results of the current study suggest the latter solution. The further implication of this idea is that at the same time the cells change their RF structure, they change their line label, such that modulation in their firing is interpreted differently by the rest of the nervous system.

As a mechanism underlying the improvement in the bisection task, we propose a change in the strength and excitatory/inhibitory balance of a subset of horizontal inputs to V1 neurons. This modulation of the horizontal input would then vary according to the separation between the source and target neurons, therefore providing a greater modulation of the tuning to the separation of parallel lines. The contextual modulation would in turn be modulated by top-down influences, presumably mediated by feedback connections from higher-order cortical areas, to generate its task dependence. Thus, contextual influences within a particular cortical area may come not just from lateral connections within that area, but as an interaction between local circuits and feedback connections from higher-order cortical areas, thus providing a mechanism for both the stimulus selectivity and task dependence of the cortical responses to trained stimulus patterns.

**Methods**

Two macaque monkeys (Macaca mulatta) were trained and used to collect the physiological data reported here. During recording and training, monkeys were seated in a primate chair facing a computer monitor. The monkey's head was restrained using a surgically implanted stainless steel post. Eye movements were monitored using a scleral search coil system \(^{13}\) (CNC Engineering, Seattle, Washington). During stimulus presentation, monkeys were required to maintain fixation within a 0.75–1.0° rectangular window; trials were aborted and reward was withheld if an eye movement greater than 0.5° was made. The actual variability in eye position from trial to trial was much less than that allowed by the fixation window. The mean position and standard deviation for typical experiments (based on a sample of data obtained during the recording of 21 units), relative to the fixation point, was 0.07 ± 0.03° in azimuth and -0.02 ± 0.05° in elevation. Quality of fixation was the same during both fixation and bisection trials: the mean difference in eye position between fixation and bisection trials for typical sessions was 0.03 ± 0.04° in azimuth and 0.01 ± 0.02° in elevation. 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 review boards.

**Training.** Monkeys were initially trained to perform a simple fixation task (Fig. 1a). The animal was taught to initiate trials by pulling a lever attached to the primate chair. When the lever was pulled, a small bright spot was displayed on the screen, and the animal was required to maintain fixation on the spot until it was dimmed. If the monkey released the lever within a brief period of time following the dimming of the fixation spot, a small drop of juice was given as a reward. To teach the monkeys to perform the bisection discrimination, we first trained them to perform a series of tasks designed to lead to bisection performance. The initial thresholds for bisection performance reflected the first thresholds measured for the final bisection task. In the bisection task, the monkey was presented with a set of three parallel horizontal lines (Fig. 1a). The monkey's task was to determine whether the central line was nearer to one or the other flanks. During the presentation of the stimulus, the monkey was required to maintain fixation on a small spot. Following the presentation of the bisection stimulus, the fixation spot was extinguished and the monkey indicated its response by making an eye movement to one of two small spots presented at the top and bottom of the screen. Correct responses were rewarded with a drop of juice. The responses of the monkey to the bisection task were recorded and used to calculate the threshold of performance using the method of probits.

**Electrophysiological recording.** After training was complete, monkeys were surgically implanted with a steel chamber (inner diameter, 22 mm) enclosing a craniootomy over a portion of V1. Surgical procedures were done under aseptic conditions with pentobarbital sodium anesthesia. Recordings were made with glass-coated platinum iridium microelectrodes \(^{16}\) with impedances between 1.0 MΩ and 3.0 MΩ. Using a stepping motor microdrive (Narisihige, Tokyo, Japan), penetrations were made at typical intervals of 0.5 mm through the duras.
Recordings from single units were conducted daily in 2 to 4 hour sessions. After penetrating the dura, a rough RF map was obtained while the animal performed a fixation task. All recordings were made from the oculopausal surface of V1 at RF eccentricities ranging from 1.5° to 5°. Granule layers were identified by the characteristics such as high levels of spontaneous activity and brisk on/off response. When such activity was encountered, the electrode was retracted to restrict recording to the superficial 600 μm of the cortex.

Neuronal activity was recorded over 600-ms epochs spanning the presentation of the stimulus. The level of background activity was measured for 200 ms, and a stimulus was then presented for 100 ms. During each trial, three to five 600-ms recording periods were conducted. For each cell, a time window was set within a range of 50–250 ms after stimulus onset depending on the latency and length of the response. The mean firing rate within this response window and the spontaneous firing rate was calculated from the number of spikes obtained in the first 200 ms of the recording epoch, was used to determine the magnitude of the evoked response. The t-test was used to evaluate the significance of the evoked response.

Each recording session began by characterizing the RF extent and orientation preference while the monkey performed fixation trials. RF extent was determined by the minimum response field technique in which a small bar (typically 0.2–0.25° in length and 3 inches in width) was presented in steps (0.1–0.25° apart) either along the principal orientation axis of the cell to determine its length, or orthogonal to the orientation axis to determine its width. The distance between the outermost points that elicited a significant response was defined to be the size of the RF. For the comparison between RF profiles during fixation and bisection trials in trained and untrained hemisphere (Fig. 7), RF extent was determined by fitting the responses obtained at each position with a Gaussian and taking the standard deviation to be the size of the RF. Subsequently, contextual interactions were examined while the animal performed either fixation or bisection trials. During bisection trials, experimental stimuli were presented simultaneously with the bisection stimulus. The fixation and bisection trials were collected in separate blocks, and we collected 10 trials per stimulus condition. The tuning curves were calculated, and are shown along with standard error bars.

**Acknowledgements**

This work was supported by NIH grants EY07968 (C.D.G.) and GM 07524 (R.E.C.).

**Received 12 October 2000; accepted 28 February 2001**