Neural systems underlying learning and representation of global motion

LUCIA M. VAINA*†‡, JOHN W. BELLIVEAU*†, ERIC B. DES ROZIERS*, AND THOMAS A. ZEFFIRO§

*Brain and Vision Research Laboratory, Department of Biomedical Engineering and Neurology, Boston University, 44 Cummington Street, Boston, MA 02215; †Massachusetts General Hospital NMR Center, Building 149, 13th Street, Charlestown, MA 02129; and §Sensor Systems Inc., Sterling, VA 20166

Communicated by Charles G. Gross, Princeton University, Princeton, NJ, August 14, 1998 (received for review June 24, 1998)

ABSTRACT We demonstrate performance-related changes in cortical and cerebellar activity. The largest learning-dependent changes were observed in the anterior lateral cerebellum, where the extent and intensity of activation correlated inversely with psychophysical performance. After learning had occurred (a few minutes), the cerebellar activation almost disappeared; however, it was restored when the subjects were presented with a novel, untrained direction of motion for which psychophysical performance also reverted to chance level. Similar reductions in the extent and intensity of brain activations in relation to learning occurred in the superior colliculus, anterior cingulate, and parts of the extrastriate cortex. The motion direction-sensitive middle temporal visual complex was a notable exception, where there was an expansion of the cortical territory activated by the trained stimulus. Together, these results indicate that the learning and representation of visual motion discrimination are mediated by different, but probably interacting, neuronal subsystems.

In previous psychophysical studies of perceptual learning of a direction-of-motion discrimination task we found that the subjects’ performance improves with practice over less than 300 trials in a single testing session, is retained over time, and is specific for the particular stimulus attributes (1). Given the time course of minutes over which this learning takes place, the neural substrates of the perceptual learning can be studied by using functional MRI (fMRI) as described here.

Until recently, the predominant view of plasticity of cortical maps and of functional properties of neurons in the early levels of cortical sensory processing was that they are specific to early development and are fixed in adulthood. Research over the past 15 years has amply demonstrated that this is not the case. It is now clearly established that cortical maps in the early stages of sensory processing in the adult animal are not fixed, but dynamic throughout life. Lesion studies in different adult mammalian species, including humans, have demonstrated significant neuronal plasticity. When a specific cortical area is deprived of its normalafferent inputs, it reorganizes so that it becomes responsive to inputs that were initially represented only by the surrounding cortex (for a review, see ref. 2). Another important form of neural plasticity, known as perceptual learning, relies on changes resulting from practicing stimulus discrimination. These two forms of cortical plasticity are complementary. In the former, there is a peripheral or central reduction of the input, whereas in the latter there is an enrichment of the input (3).

Currently, considerable research is aimed toward understanding the relationship between neuronal activity and performance on a task, particularly within the framework of perceptual learning. One proposal is that learning largely implies an increased representation of the trained stimulus, and thus the effects of learning may be manifested as recruitment of cortical regions involved in performing a task. Expansion of the cortical territory activated by a trained stimulus has been demonstrated in the somatosensory and auditory modalities in awake behaving monkeys (4–7). However, the reorganization of the cortical maps observed in these learning tasks occurred after a long period of training, up to several months.

Recent neurobiological (8, 9) and psychophysical studies (for reviews see refs. 10–12) have provided robust evidence for learning effects in the early stages of visual information processing. Interestingly, although learning in the visual system also occurs with training over days or weeks, several psychophysical studies have demonstrated that it can also occur rapidly, within a few trials in a single session.

The remarkable but still poorly understood phenomenon of rapid perceptual learning has been investigated psychophysically in several laboratories, including ours (1, 3, 13–16). In trying to characterize the rapid learning of direction discrimination in global-motion stimuli we investigated psychophysically the specificity of perceptual learning with respect to the stimulus attributes, visual field position, and the eye tested. The global-motion stimulus consisted of two-frame stochastic random-dot cinematograms in which a proportion of the dots (20%) moved coherently, in the same direction, and with the same speed, while the remainder moved in random directions and with random steps from one frame to the next, providing masking motion noise. On each trial subjects reported whether the motion field moved in one direction or in the opposite direction (e.g., right or left). Our results demonstrated rapid learning. Only a few minutes of a single session sufficed for human subjects to improve their performance from scoring close to chance to scoring almost perfect (1). Furthermore, when presented with a previously untrained direction discrimination (e.g., up or down), or when the same stimulus was displayed in a different location in the visual field, subjects’ scores immediately returned to chance levels. The stability and retention of the good performance, together with the time course of learning and the specificity to the trained direction of motion and spatial location of the stimulus, qualify it as perceptual learning. In the study described here, we used functional neuroimaging to determine the nature and location of the neuronal substrates for this short-term and stimulus-specific behavioral plasticity.

Electrophysiological studies have demonstrated that neurons in the middle temporal visual area (MT or V5) in macaque monkeys are particularly sensitive to direction in global-motion displays (17). Electrode recordings also show that the directional specificity of the MT and medial superior

Abbreviations: fMRI, functional magnetic resonance imaging; MT, middle temporal visual area; MST, medial superior temporal visual area; SC, superior colliculus.

†To whom reprint requests should be addressed. e-mail: vaina@enga.bu.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences

PNAS is available online at www.pnas.org.
temporal area (MST) neurons improved hand-in-hand with the animal’s progressively improving perceptual performance (18). Furthermore, electrical microstimulation of MT neurons while the monkey performed a psychophysical direction-discrimination task provided evidence for a causal relationship between neuronal activity and the perceptually judged direction of motion. Additional experiments in various laboratories, including ours, show that lesions of MT, both in macaque monkeys and in human patients, impair perception of the discrimination of direction in global-motion displays (19–21). Functional-imaging studies have reliably identified the location of the human homologue of MT and provide definitive evidence for its specific involvement in the analysis of global motion (22–24). Collectively, these findings support the hypothesis that the neuronal circuitry within area MT is a prime candidate for representing global motion and changes in its discrimination with practice.

There is also supporting data for the involvement of MT in learning. Zohary et al. (9) found a 13% increase in sensitivity of motion-sensitive cells in MT associated with a 19% improvement in the monkey’s ability to discriminate the direction of motion, and that the neuronal effects associated with learning transfer from a conditioned to an unconditioned part of the receptive field of an MT neuron. These findings indicate that the learning effect is likely to involve cortical regions with receptive fields at least as large as those of MT and to not involve earlier stages in cortical visual processing where neurons have smaller receptive fields. Although these results suggest that the same neuronal circuitry mediates both the representation and learning of global-motion discrimination, electrode and lesion studies have intrinsically limited spatial sampling and can provide only a partial answer. We therefore used the same global-motion task as previously described in conjunction with whole-head fMRI to explore the neural substrate of learning a visual direction discrimination. Preliminary reports of this work have appeared in refs. 25 and 26.

**MATERIALS AND METHODS**

**Experimental Paradigm.** The fMRI experimental paradigm consisted of two visual displays: a task condition, incorporating the global-motion stimulus (Fig. 1a), and a control condition, which contained a static random-dot display with the same statistical characteristics as the motion display except the dots were static. In the task condition, subjects performed a direction-of-motion discrimination task (i.e., left vs. right) and in the control condition, the central fixation square transiently (30 msec) turned into a T or an L, and subjects were asked to perform a letter-discrimination task.

For both conditions, subjects were instructed to maintain their gaze on the fixation mark at the center of the display and to enter their answers by key-press on predetermined keys on a keypad connected to a Macintosh computer, which generated the stimuli and collected the subject’s answers trial-by-trial. No feedback was provided.

Prior to the fMRI study all of the subjects were shown examples of the psychophysical stimulus and the tasks, but no practice was given. All subjects gave informed consent to participate in the study according to the Massachusetts General Hospital Human Subjects Committee requirements.

Each run lasted for 240 sec, composed of 60 sec of mean luminance (baseline fixation), 120 sec involving discrimination of either motion or letters, and another 60 sec of mean luminance (baseline fixation). Motion- and letter-discrimination runs were interleaved. For all functional runs, a time series of 100 images per slice was acquired from 20 slices that covered the whole brain.

Each subject underwent about eight consecutive runs, with learning-motion-discrimination runs interleaved with static control runs, yielding a total of roughly 400 motion-discrimination trials; this was amply sufficient for observing the fast learning demonstrated in our previous psychophysical studies.

**Data Acquisition and Experimental Setup.** Six subjects (mean age 26.2 years, range 21–37) were scanned on a 1.5-tesla General Electric Signa MRI system, retrofitted for echo planar imaging (Advanced NMR GE Systems, Wilmington, MA, Instascan), using a standard-receive-transmit headcoil (quadrature birdcage, receive only). An automated field-shimming procedure was performed to minimize magnetic-susceptibility distortions. A coplanar conventional volume was acquired by using 20 6-mm thick contiguous oblique slices (3.125 mm × 3.125 mm, in-plane) parallel to a line drawn between the anterior and posterior commissure sufficient to cover the whole brain. A flow series was obtained in the oblique planes selected for functional scanning to detect major blood vessels, followed by a spin-lattice relaxation time (T1)-weighted sagittal localizer series (TR = 25 sec, TI = 70 msec, NEX = 1, FOV = 24 cm, acquisition matrix 256 × 192). The scans were used to guide slice selection for the functional acquisitions. Functional images sensitive to changes in blood oxygenation state were obtained by using an ASE pulse sequence (TR = 2.5 sec, TE = 70 msec, τ offset = 25 msec, 100 images per slice). For each subject a high-resolution conventional structural scan (124-slice sagittal volume) was also acquired (SPGR: FOV = 24 cm, acquisition matrix 256 × 192, voxel size: 1 × 1 × 1.5 mm; TE = 4 msec, TR = 25 msec).

During the experiment, the room was darkened to reduce unwanted activation of visual cortex. Before entering the MRI chamber, subjects were fitted with earplugs. Subjects were supine, foam pads were put around the ears tightly to hold the head still, and an adjustable bite-bar minimized head motion. During fMRI scanning, visual stimuli were rear-projected onto an acrylic screen providing an activated visual field of 40° × 25°. Stimuli subtending 79 square degrees were projected onto the screen by a Sharp 2000 color LCD projector, through a collimating lens.

**Data Analysis.** To determine the areas of brain activation specific to the training on the motion-discrimination condition, image analysis and visualization was performed with MEdX 2.1 image analysis software (Sensor Systems, Sterling, VA). The image data were motion-corrected and ratio-normalized. For
neuroanatomical localization, each functional volume was registered onto the high-resolution three-dimensional images translated into Talairach space. Planned contrasts between the motion-learning and fixation conditions were computed by using the $t$ statistic and converted into $Z$ scores. The time series for each significant peak was examined to verify the presence of task-related signal-intensity modulation. To isolate the activation specific to the motion task, the activation elicited by the static stimulus compared with the baseline of mean luminance was subtracted from the activation in the motion task compared with the mean luminance.

A standard correlation analysis between psychophysical performance and activation was performed for each subject.

**RESULTS**

As in our previous psychophysical study, in the block of trials subjects scored close to chance and then showed rapid improvement in the subsequent runs. Fig. 1b shows averaged data of the motion runs from the four subjects. Each data point corresponds to approximately 100 trials (for each subject) obtained during a motion run (120 sec).

Fig. 1c shows that when subjects were asked to discriminate an untrained direction of motion (up-down), scores fell to close-to-chance again (as expected from our previous studies), indicating the directional specificity of the learning. On the letter-discrimination task all subjects scored almost 100% correct, demonstrating that they were able to maintain attention to tasks throughout the imaging session.

Because the psychophysical stimulus used for training incorporated global motion, we expected that the pattern of learning-related activity in the cerebral cortex would be centered on the MT complex. The average activity in four subjects at the location of peak intensity in MT in the first run, (Early) before subjects practiced with the task, is illustrated in Fig. 2B, and in the fourth run, after subjects’ psychophysical performance indicate that they (Late) have learned the task and performance was stable, in Fig. 2C. The results of correlation analysis indicated that the extent of the most activated cortical region, corresponding to MT, was directly correlated with learning ($r = 0.8$), increasing 5-fold between the two comparisons. Consistent with Fig. 2B, in the early runs there was additional and more distributed activation, including higher level extrastriate areas in the dorsal-motion processing pathway, including foci of activity in the posterior and medial parietal lobes (Table 1).

Concurrent with the expansion of the area of cortical activation in the MT complex in the late runs was a significant reduction of activity in the other extrastriate areas. This suggests that one effect of learning is the development of a more “focused” representation of the motion stimulus and/or the elimination of activity that is irrelevant to performance of the task.

The activation in the superior colliculus (SC) observed in the first run (Fig. 2B, slice 31) was totally absent from the late runs (Fig. 2C), in which subjects’ psychophysical performance (close-to-perfect score) indicated that they learned the global-motion-discrimination task. Although the SC has been implicated in oculomotor control (for a review see ref. 27), it is quite certain that changes in eye movements do not provide a plausible explanation for our findings. First, although the activation was specific only to the direction-discrimination task, the motion stimulus was on for only 90 msec, which is not sufficient for the initiation of smooth pursuits. Moreover, the activation in the SC was present only during learning, and not after the subjects learned the task or during the control task (letter discrimination). The SC is strongly connected with the MT, and its contribution via the subcortical (retinotectal) motion pathway to the motion-specific neural sensitivity of MT is well established (28, 29). It is thus plausible that SC may be

Table 1. Regions specifically activated by the motion learning task

<table>
<thead>
<tr>
<th>Area</th>
<th>Talairach coordinates, mm</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Inferior occipital gyrus (BA19)</td>
<td>−44 −70 −8</td>
<td>6.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Lingual gyrus (BA18)</td>
<td>14 −82 −8</td>
<td>4.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Inferior occipital gyrus (BA18)</td>
<td>−10 −82 −6</td>
<td>4.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Inferior occipital</td>
<td>10 −82 −6</td>
<td>5.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L MT</td>
<td>−42 −70 −6</td>
<td>6.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L GTM (BA37/39)</td>
<td>−34 −48 10</td>
<td>6.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Temporal gyrus (BA37)</td>
<td>−52 −56 −6</td>
<td>5.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Inferior temporal gyrus (BA20)</td>
<td>−42 −36 −16</td>
<td>5.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior colliculus</td>
<td>0 30 −6</td>
<td>6.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior colliculus</td>
<td>0 −28 −2</td>
<td>6.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Anterior cingulate</td>
<td>4 30 34</td>
<td>4.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Anterior cingulate</td>
<td>−4 32 20</td>
<td>5.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum vermis</td>
<td>0 −52 −20</td>
<td>5.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Cerebellum (QuP and ScS)</td>
<td>−37 −63 −22</td>
<td>5.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Cerebellum (QuP and ScS)</td>
<td>34 −50 −22</td>
<td>6.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Cerebellum (QuP and ScS)</td>
<td>30 −50 −26</td>
<td>7.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Lingual gyrus (18)</td>
<td>14 −82 −8</td>
<td>4.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Inferior occipital (BA37)</td>
<td>−10 −82 −6</td>
<td>4.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L MT</td>
<td>−44 −66 −6</td>
<td>6.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R GTM (BA37)</td>
<td>−42 −58 6</td>
<td>5.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Cerebellum (QuP and ScS)</td>
<td>40 −48 −24</td>
<td>4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Vermis</td>
<td>4 −58 −4</td>
<td>4.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Areas of evoked activity during the task condition (learning direction discrimination) compared with the control condition (letter discrimination). Only areas that reach $P < 0.0005$ and $Z > 4$ after correction for multiple comparisons are reported. X, Y, Z refer to distances right (−) or left (+) from the midline, anterior (+) or posterior (−) from the anterior comissure, and below (−) or above (+) the anterior comissure-posterior comissure line. BA refers to the Broadmann areas. QuP, posterior quadrangular globule; ScS, superior semilunar lobule. L and R refer to the side of the brain. Early refers to the first run, and Late refers to the fourth run.
involved in modulating the activity of MT neurons during learning.

Another region selectively activated during the learning task, and only when subjects’ performance was still improving with training, is the anterior cingulate. The role of this region in specific aspects of attention is well established.

By far, in all subjects the most dramatic learning-related activation was observed in the cerebellum. The involvement of the cerebellum in learning the motion-discrimination task is illustrated through 12 coronal slices in Fig. 3. Correlation analysis between the regional extent of significant fMRI signal and psychophysical performance showed a strong inverse correlation ($r = 0.9$); the extensive activation of the cerebellum in the first run (Fig. 3B) decreased 93% in activation area after learning occurred (the fourth run; Fig. 3C). Most intriguing, when subjects are presented with a novel (nontrained) direction discrimination (up-down), scores return to chance levels, and the strong fMRI signal in the cerebellum is reinstated (Fig. 4D). The role of the cerebellar activity in learning the task is also indicated by the absence of activation during the control task of letter discrimination, which clearly did not involve learning (Fig. 4E).

DISCUSSION

Two major learning-related changes in activation were detected in all subjects participating in this study: the changes occurred over a few minutes (300 trials or less) within a single session and were stimulus attribute-specific.

First, the 5-fold increase in the activated area centered on the MT complex, a cortical region particularly well suited for

---

**Fig. 3.** Changes in cerebellar activity during learning. (A) Position of the twelve coronal slices through the cerebellum, shown on a midsagittal anatomical image. Slice 1 is the most anterior slice. (B) Functional map illustrating the location of the average activity in four subjects during the first run. The most extensively and intensely ($Z > 6$) activated region during the learning phase is in the left superior posterior cerebellum, with the approximate peak activation corresponding to Talairach coordinates ($-37, -63, -22$), an area previously shown to be involved in visual attention (33). (C) Functional map illustrating the location of the average activity in four subjects during the last run. There is a 93% decrease in the total area activated, and in intensity.

**Fig. 4.** (A) Position of the twelve coronal slices through the cerebellum in one subject, shown on a midsagittal anatomical image. Slice 1 is the most anterior slice. (B) Functional map illustrating the location of the average activity for training on left–right direction-of-motion discrimination in one subject during the first run. (C) Activity in the same slices in the final run, after learning occurred. (D) Functional map illustrating the return of cerebellar activity for an unpracticed direction of motion (up–down). The return of activity corresponds to the subject’s close-to-chance scores. (E) Cerebellar activity is absent when the subject performs the static-letter-discrimination task, which does not require learning.
representing the global-motion stimulus we used, suggests a learning-induced cortical recruitment. In humans, training-dependent cortical plasticity, specifically cortical recruitment, has been reported previously in functional-imaging studies of motor learning tasks (30), which also revealed an increase in the region of activation. Similarly, magnetoencephalography studies of Braille readers (31) and of string-instrument players (32) have shown an increase in the somatosensory representation of the particular finger of the hand used in these tasks as compared with the representation for the corresponding fingers of the other hand or of different fingers of the same hand in control subjects. Taken together with previously reported training-dependent changes in the somatosensory and auditory cortical map organization in behaving animals, these results support the view that training-induced cortical plasticity in adult animals is critical for some forms of perceptual learning.

We observed constriction in the extent of spatial distribution of cortical activation in several higher extrastriate areas as a result of learning the motion-discrimination tasks. This suggests that a plausible important effect of perceptual learning may be a more spatially compact and efficient representation of the practiced stimulus.

A second important result of this study is a very significant decrease (93%) in the cerebellar activation with learning. This pattern of cerebellar activity is consistent with recent functional neuroimaging studies of motor and nonmotor learning tasks (33–35). For example, in a language learning task Raichle et al. (33) found that in novel trials, specific activation was seen in several cortical areas and the lateral cerebellum, but with less than 15 min of practice these areas became inactive, and other areas previously inactive became activated.

Immediately relevant to our findings is a recent study by Allen et al. (36). By using fMRI to examine the differential involvement of the cerebellum in a motor task and in a visual-attention task that had no motor component, they found a double dissociation of function between different areas of the cerebellum. During the visual-attention task the most common site of activation was in the left superior portion of the cerebellum (the posterior quadrangular lobule and the superior semilunar lobule). However, in the motor task the most common site of activation was in a different, nonoverlapping region in the right anterior cerebellum (the anterior vermis, the central lobule, and the anterior portion of the quadrangular lobule). The region of the cerebellum that we observe to be most strongly modulated during perceptual learning corresponds to the area that Allen et al. reported to be involved in modulation of visual attention. Unlike Allen et al., we observed bilateral cerebellar modulation.

The role of attention during learning is also supported by the training-related activation in the superior colliculus and the anterior cingulate cortex. Both were significantly active in the early but not the late stage of performing direction discrimination in the global-motion task. It is therefore plausible that activation of the superior colliculus and anterior cingulate during learning has a modulatory role. A recent combined fMRI and psychophysical study elegantly demonstrated the involvement of the SC in modulation of motion-related activity by attentional load (37). As we noted earlier, it is likely that the SC is also involved in modulation of visual attention. The role of the anterior cingulate in attention, specifically in the selection among competing, complex contingencies has been confirmed in both functional neuroimaging studies (33, 38) and single-unit recording work in monkeys (39). We suggest that the learning-dependent activation in the cerebellum, the superior colliculus, and the cingulate provides an attentional neuronal network that is active during learning, when allocation of attention to the stimulus is necessary. When the task becomes “automatic” (or learned) there is no longer a role for this modulatory attentional network, and as a consequence we saw that the neuronal activity in these areas almost disappeared.

This hypothesis that attention modulates the performance of the neuronal network involved in stimulus processing is consistent with recent results from psychophysical studies of the nature of visual perceptual learning. These studies (40–44) converge on the view that this learning involves plastic changes to early neural processing levels that are stimulus-related and are dependent for their induction and consolidation on “the general behavioral state of the subject,” such as attention (10).

The results from the fMRI study reported here suggest, in addition, that the central role of the attentional mechanisms diminishes significantly once the task is learned. After learning occurs, the “scaffolding” is not required (S. Peterson, personal communication), and what remains is the representation specific to the task. Whether the restricted pattern of activations remains tightly correlated with efficient performance on a particular perceptual task after several other versions of the task have been mastered, or even after the passage of much longer time periods, is currently unknown.

This work was supported in part by National Institutes of Health Grant 2EY07861–8 and and National Science Foundation Professional Opportunities for Women in Research and Education Grant SBR-9753009 (to L.M.V.). J.W.B. was supported in part by the Whitaker Foundation.