Inferior Temporal Mechanisms for Invariant Object Recognition

Andreas Lueschow, Earl K. Miller, and Robert Desimone

Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, Maryland 20892

The specific size and retinal location of an object are readily perceived, yet recognition of an object's identity is hardly affected by transformations of its size or location. To explore how such stimulus transformations are treated by known mechanisms for visual short-term memory in inferior temporal (IT) cortex, IT cells were recorded in monkeys performing a delayed matching-to-sample task. The stimuli were pictures of complex objects, and the monkeys ignored differences in size and retinal location when matching the test items to the sample held in memory. The sensory information communicated by cells was assessed in their responses to the sample stimuli, and mnemonic information was assessed in their responses to the test stimuli. In the sensory domain, the ordering of relative stimulus preferences for nearly all cells was invariant over changes in size or location; however, some cells nonetheless preferred stimuli of a given size or location. In the mnemonic domain, the responses of many cells were modulated according to whether the test stimulus matched the sample held in memory, and these memory effects were invariant over the relative sizes and locations of the stimuli. Thus, IT neuronal populations may mediate not only the recognition and memory of object identity, which are invariant over size and location, but also the perception of the transformations themselves.

One of the most powerful features of human memory is the ability to recognize objects that have been transformed in some way from the time of their first appearance. Having seen an object in one particular view, for example, we later recognize it as the same object even though its image may differ in size or location on our retinas. What are the memory mechanisms that underlie this size and location invariance?

A likely site for such memory mechanisms is inferior temporal (IT) cortex, as lesions of this region severely impair visual memory (Gross, 1973; Mishkin, 1982; Horel et al., 1987; Zola-Morgan et al., 1989; Gaffan and Murray, 1992). Several studies have shown that items in memory influence the responses of IT cells to current stimuli (Gross et al., 1979; Baylis and Rolls, 1987; Miller et al., 1991, 1993; Riches et al., 1991; Eskandar et al., 1992; Fahy et al., 1993; Li et al., 1993; Miller and Desimone, 1994; Vogels and Orban, 1994). We have found that the responses of about half the cells in IT cortex communicate sensory information alone, whereas the responses of the other half are jointly determined by the current stimulus (e.g., whether the stimulus has the appropriate shape for a shape-selective cell) and stored memory traces (Miller et al., 1991, 1993). When studied in conventional delayed matching-to-sample (DMS) tasks, most of these mnemonic cells are suppressed according to the similarity of the current stimulus to traces of stimuli in memory. Thus, memory functions as a type of filter, through which incoming stimuli pass, and it preferentially passes visual information that is different from that present in the recent past. We have termed this property "adaptive mnemonic filtering" (Miller et al., 1991).

The nature of the stimulus representations upon which adaptive mnemonic filtering operates is not yet clear. Mnemonic effects might be due, on the one hand, to a literal, or "pixel-by-pixel," comparison of the current stimulus to previously seen ones. On the other hand, the comparison might be made on more abstract stimulus representations, that would be invariant over stimulus transformations that preserve object identity. To begin to explore the role of stimulus transformations in the processing of both sensory and mnemonic information in IT cortex, we recorded from IT neurons while monkeys performed a DMS task. The sample and test stimuli in the DMS task were transformed in either size or retinal location.

Cerebral Cortex Sep/Oct 1994;5:523-531; 1047-3211/94/$4.00

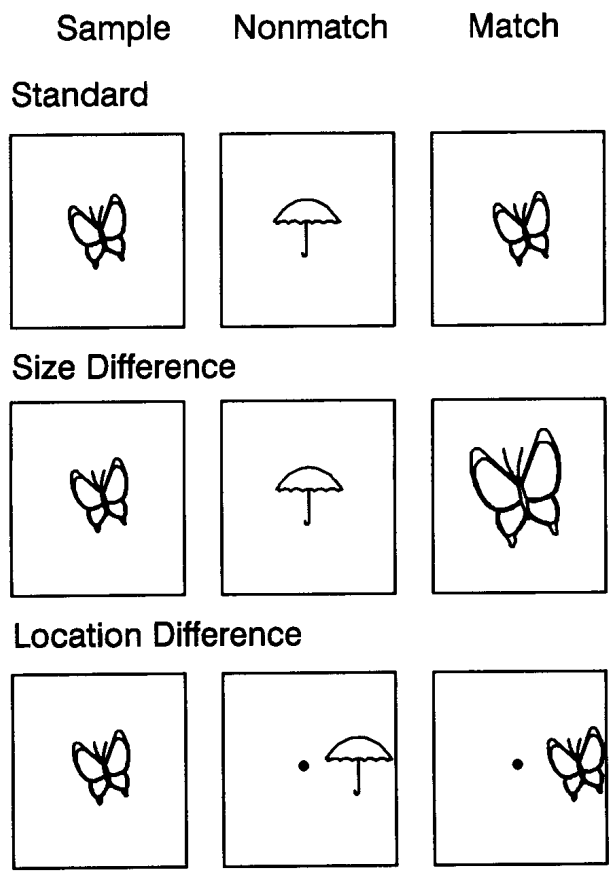

Figure 1. Examples of trials from the DMS task. Shown are examples of trials in which the match stimuli were (1) the same size and presented at the same location as the sample (*top*), (2) a different size than the sample (*middle*), or (3) presented at a different location than the sample (*bottom*). The number of intervening stimuli in a trial actually varied in number from zero to two in the location task and from zero to one in the size task.

Materials and Methods

Subjects and Localization of Recording Sites

Three rhesus monkeys (*Macaca mulatta*) were used. All surgical and experimental procedures have been described in detail previously (Miller et al., 1993) and will be only briefly summarized here. Prior to the implantation of the recording chamber, the monkeys were placed in a plastic and aluminum stereotaxic machine and scanned with magnetic resonance imaging (MRI). The MRI images were used to position the chamber over the target area in IT cortex and to localize the subsequent electrode penetrations. The target area was the same region from which we recorded in our previous studies of short-term memory in IT cortex (Miller et al., 1991, 1993), namely, the anterior-ventral portion between the anterior middle temporal and rhinal sulci.

Stimuli and Behavioral Task

The task (Fig. 1) was a modified version of DMS, similar to that used in previous studies (Miller et al., 1991, 1993). At the start of each trial, the monkey grasped a metal bar after which a small (0.1°) fixation target appeared. The monkey was required to maintain fixation of the target, which remained on throughout the trial. The stimuli used were multicolored pictures, digitized from real objects or magazine photos, presented on a computer display. Some were identifiable objects, such as faces or fruit, and others were multicolored patterns.

The first stimulus presented on each trial was the sample, followed by one to three test stimuli. When one of the test stimuli matched the sample, the animal released the bar to receive a drop of orange or apple juice. A stimulus used as the sample-match stimulus on one trial was used as a nonmatch stimulus on a different trial. The match stimulus was the last stimulus presented on each trial, and it was extinguished as soon as the animal made its response, about 350 msec after stimulus onset.

Location Experiment

All stimuli were shown for 400 msec, with zero to two intervening nonmatch stimuli between the sample and final matching stimulus. Most cells were recorded from with 1500 msec interstimulus delays and the remainder with 2000 msec delays.

A fixed set of 24 stimuli, 1–3° in size, was used. Each neuron was initially tested with the entire set of stimuli presented at the center of gaze, incorporated into the DMS task. Based on these initial responses, two stimuli that elicited clear responses and one that was ineffective were chosen for the test of location invariance.

Three different stimulus configurations were used in the location test: (1) sample and test stimuli all positioned at the center of gaze, (2) all stimuli centered at a location 5° into the visual field contralateral to the recording site (on the horizontal meridian), or (3) sample stimuli presented at the center of gaze and test stimuli at the 5° location in the contralateral field. All cells were tested in the first configuration, and the large majority were tested in all three. Each of the three location configurations was run in short blocks of about 36 trials each, with all combinations of sample and test stimuli randomly interleaved within the block. The three types of blocks with the different location configurations were presented in an interleaved fashion, with three or four repetitions per block per cell. The responses of each cell were recorded until the animal had accumulated six to eight correct trials per individual stimulus combination.

Size Experiment

All stimuli were shown for 500 msec, with 700 msec interstimulus delays. To simplify the task, we presented a maximum of one intervening nonmatch stimulus between the sample and final match.

Each cell was tested with a fixed set of stimuli consisting of either two or six different digitized pictures. Each of these pictures was digitized at both a small size (typically about 2° wide) and a large size that was twice the small one in linear dimension. The total set size was therefore either four or 12 stimuli per cell. Human observers readily recognized the same object or pattern in the pictures at the two different sizes, although the size difference was easily perceived.

The small and large patterns appeared randomly as

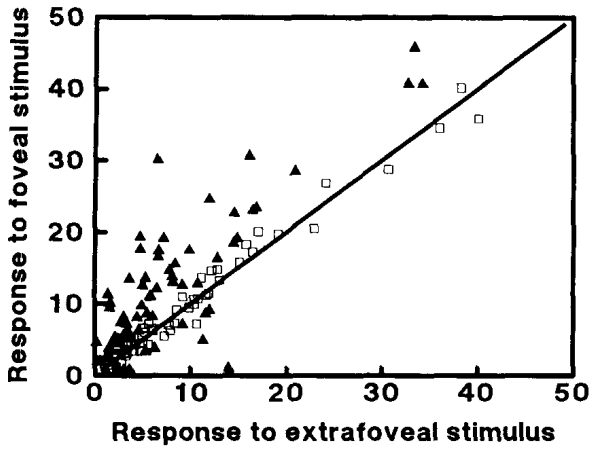

Figure 2. Comparison of responses to the same stimuli presented foveally and extrafoveally. *Diagonal* indicates equal responses. *Triangles* indicate responses that were significantly different depending on the location of the stimulus. *Squares* indicate responses that were not significantly different.

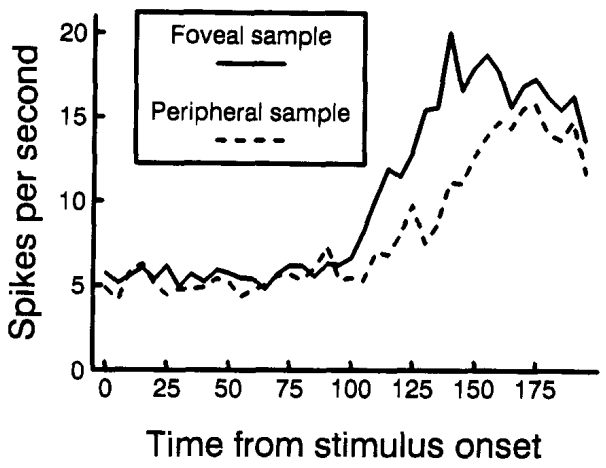

Figure 3. Responses to foveally and extrafoveally presented samples, averaged across all cells that elicited a response at both locations.

sample and test stimuli, in a balanced design for each cell. Thus, on some trials the animal was required to match the small size of a given picture (sample) to a large size of the same picture (match), on other trials a large size to a small size, on other trials a small to a small, and so on. Likewise, the small and large pictures appeared randomly as the nonmatching item as well.

Data Analysis

For most analyses, responses to the sample and test stimuli were calculated over a 200–225 msec time interval beginning 75 msec following stimulus onset. The beginning of the time interval was chosen to coincide with the earliest response latencies of IT neurons, and the end was chosen to occur before the animal's behavioral response to the match stimulus. Because responses to peripheral stimuli were sometimes weak, responses to the sample stimuli in the location study were averaged over a somewhat longer (325 msec) interval in order to classify cells as responsive or not. Responsiveness was assessed by comparing the response to the sample and the preceding baseline activity, using a paired t test ($P < 0.05$). All other results were assessed statistically with either analysis of variance (ANOVA) or t tests, evaluated at the $P < 0.05$ level of significance.

Results

Location Experiment

A total of 71 cells were tested with stimuli in one or more locations in the visual field in two of the three monkeys. Of these, 62 (87%) were visually responsive to one or more stimuli.

To quantify the relationship between stimulus and location selectivity, we computed a two-way ANOVA on the responses to the sample stimuli for each cell individually. The three stimuli were one factor and the two locations (foveal and extrafoveal) were the other. Based on this analysis, 90% (56 of 62) of the cells were stimulus selective. Although the high incidence of stimulus selectivity is consistent with results of previous studies (e.g., Miller et al., 1993), it was virtually guaranteed in this case, as both preferred and nonprefered stimuli were selected for use in the DMS task on the basis of pretesting each cell with a larger set of stimuli.

In addition to their stimulus selectivity, many cells were also selective for location. Of the 58 responsive cells tested with stimuli at both the foveal and extrafoveal locations, about a third (18 of 58, or 30%) responded about the same to foveal and extrafoveal stimuli, and the remainder (40 of 58, or 69%) significantly preferred one location over the other. Of the latter cells, a large majority (32 of 40, or 80%) responded better to stimuli located at the fovea than extrafoveally; a few (8 of 40, or 20%) responded better at the extrafoveal location. Thus, whereas the responses of some cells are invariant across retinal location, at least for the locations tested here, other cells seem to communicate significant information about the location of preferred stimuli.

This variation in response with location appeared to be due to a change in the gain of response to all stimuli, rather than a reordering of stimulus preferences. Only seven (of 58, or 12%) cells showed a significant interaction between stimulus and location selectivity and changed their rank order of responses to the different stimuli at the two locations.

These conclusions based on individual cells were extended by an analysis of responses to individual stimuli. We compared the relative responses to foveal and extrafoveal stimuli using a t test (evaluated at $P < 0.05$). Of the 128 items that elicited a response at either location, over half (70 of 128, or 55%) evoked significantly greater responses at one location or the other. Of these, 59 items elicited better responses at the fovea and 11 showed the opposite behavior. A scatterplot of responses to foveal and extrafoveal stimuli is shown in Figure 2. Responses tend to lie near the equal-response diagonal, with a shift toward greater responses to stimuli located at the fovea.

Foveal stimuli elicited responses with a slightly shorter latency than extrafoveal stimuli. Figure 3 shows histograms of responses averaged across cells,

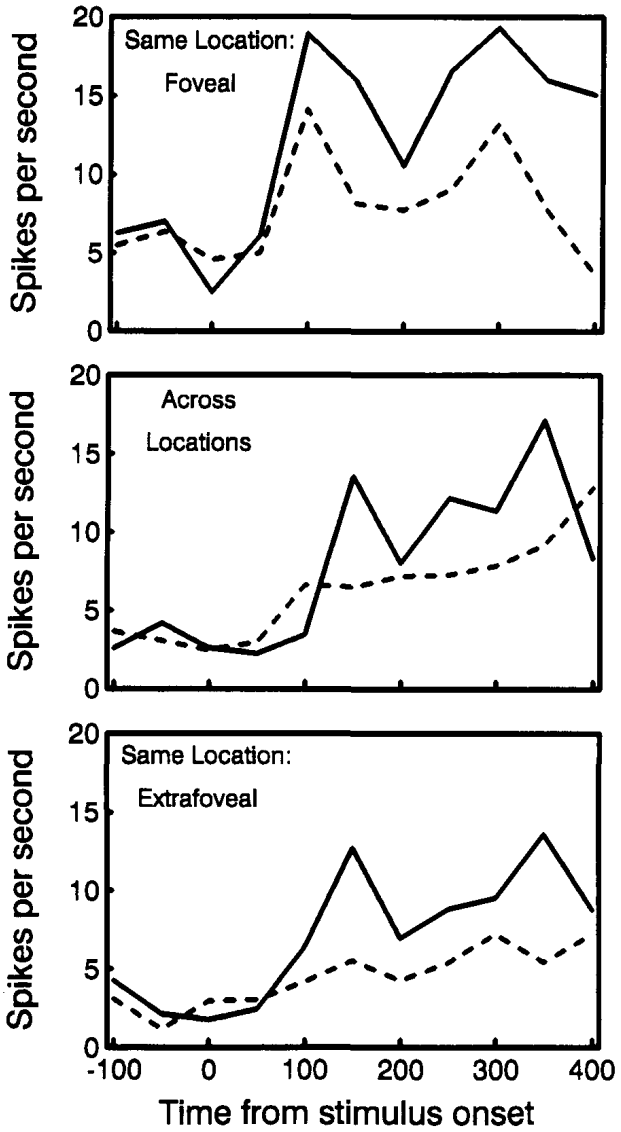

Figure 4. Responses of an IT neuron to a single stimulus appearing as a nonmatch and as a match when both the sample and the test stimuli appeared at the fovea (*top*), when the sample was presented at the fovea and the test stimuli extrafoveally (*middle*), and when all stimuli were presented extrafoveally (*bottom*). *Solid line* indicates nonmatch; *dashed line* indicates match.

for the 80 stimuli that elicited a response at both foveal and extrafoveal locations. We determined the response latency of each population histogram by comparing the average firing rate in each bin to the baseline (prestimulus) firing rate, evaluated at $P < 0.05$. The population response to foveal stimuli began slightly earlier (105 msec) than the response to extrafoveal stimuli (110 msec). The two locations appeared to have an even greater difference in their latency to peak response.

Taking all these results together, it is clear that some IT cells show considerable generalization of response across retinal translation and that nearly all cells preserve their relative ordering of response to different stimuli across retinal translation. However, some spatial (retinal) information is retained in IT cortex, in that many cells respond best and more quickly to foveal stimuli.

Effects of Memory on Test-Stimulus Responses

Our analyses of memory effects are limited to the 45 cells that were tested in all three spatial configurations and that responded to stimuli at both locations. To determine the effects of memory for the sample on the test stimulus responses, we computed a three-way ANOVA on the responses of each cell. The three test stimuli were one factor, the matching-nonmatching status of the stimuli was another, and the three spatial configurations of stimuli (foveal, extrafoveal, and across locations) were the third.

For about a quarter of the cells (12 of 45, or 27%), the responses to test stimuli varied significantly depending on whether or not they matched the previous sample, even when the test stimuli were presented at a different location from the sample (i.e., there was a significant main effect of match-nonmatch status, with no significant interaction with spatial configuration). Most (10 of 12, or 83%) of these cells were also stimulus selective, and most (10 of 12) gave weaker responses to test stimuli that matched the sample than to nonmatching ones. This is the same "adaptive mnemonic filtering" found in previous studies (Miller et al., 1991, 1993). An example of one such cell is shown in Figure 4. Thus, the responses of a subpopulation of IT cells are suppressed when the current stimulus matches a previously seen sample, even when the sample was seen at a different location. Interestingly, about half of the cells with invariant memory effects (7 of 12, or 58%) nonetheless showed a sensory preference for one location or the other, as measured by responses to the samples (e.g., see Fig. 4).

For only three cells (of 45, or 7%) were the match-nonmatch effects not invariant over location (significant interaction between match-nonmatch and location factors). These cells tended to respond better to stimuli at one location or the other and showed a disproportionately large match-nonmatch effect for test stimuli at this location.

We extended these analyses based on individual cells with an analysis of responses to individual stimuli. A two-way ANOVA was calculated on the responses to each stimulus for each cell. The match-nonmatch status of the stimulus was one factor, and the spatial configuration of the stimuli (foveal, extrafoveal, cross-location) was the other.

Of the 82 stimuli that elicited significant responses at both locations, the responses to 15 (18%) had significant match-nonmatch effects that were invariant across spatial configuration. Responses to all but one of these were weaker to matching than to nonmatching stimuli. Only three stimuli (of 82, or 4%) showed a significant interaction between match-nonmatch status and location. The average responses to the 14 items showing suppression effects are shown in Figure 5. As can be seen in this figure, responses to matching stimuli are suppressed in all three spatial configurations.

Size Experiment

A total of 59 cells was studied in two of the three monkeys. Of these, 47 cells were studied with a fixed

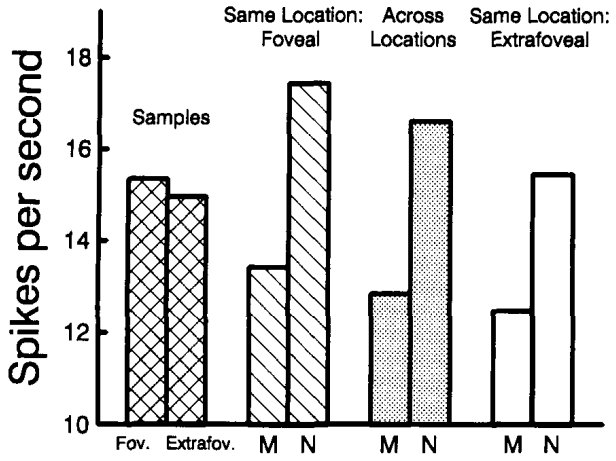

Figure 5. Average responses across cells to the same set of stimuli appearing as samples and as matches and nonmatches. *Fov.*, foveal stimulus presentation; *Extrafov.*, extrafoveal stimulus presentation; *M*, match; *N*, nonmatch.

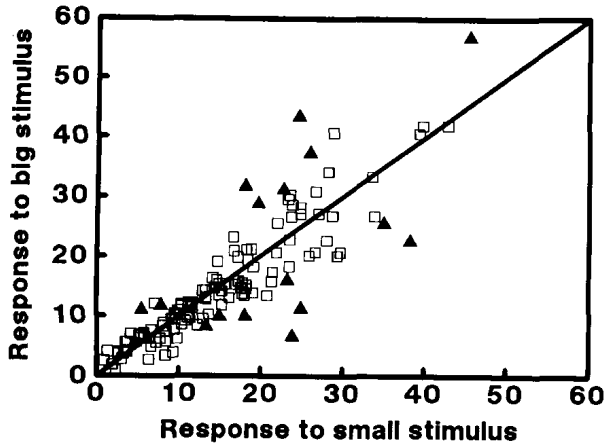

Figure 6. Comparison of responses to the same stimuli presented at two different sizes. *Diagonal* indicates equal responses. *Triangles* indicate responses that were significantly different depending on the size of the stimulus. *Squares* indicate responses that were not significantly different.

set of two stimuli, each at two different sizes, and 12 were studied with a fixed set of six stimuli, each at two different sizes. There were 37 visually responsive cells, of which 31 were in the former group and six in the latter.

An interesting behavioral observation was made at the beginning of the experiment. Before starting the size experiment, both monkeys had only matched stimuli in the DMS task that were identical in size. However, when they were first presented with test stimuli that differed in size from the sample, they matched the test and sample items without error. That is, they recognized that the stimuli were the same patterns, despite their difference in size.

To evaluate stimulus and size selectivity quantitatively, we computed a two-way ANOVA on the responses to the sample stimuli for each cell individually. The different stimuli were one factor and the two sizes of each stimulus (small and large) were the other. Based on this analysis, 51% (19 of 37) of the cells had significant stimulus selectivity. Unlike in the location study, the stimuli used were not chosen individually for each cell and therefore provide an unbiased measure of selectivity. The incidence of significant stimulus-selective cells was somewhat smaller in this study than in previous ones (Miller et al., 1991, 1993), most likely because most cells were tested with fewer stimuli than in previous studies, affording fewer opportunities for selective responses.

There was nearly an even split between cells whose responses to stimuli were invariant across size (21 of 37, or 56%) and cells whose responses varied significantly with size (16 of 37, or 43%). Of the latter cells, some (7 of 16, or 44%) preferred all stimuli at the larger or smaller size (significant main effect of size, with no interaction with stimulus), whereas others (9 of 16, or 56%) responded best to a particular stimulus at a particular size (significant interaction between stimulus and size). Only two (of 37, or 5%) cells showed a significant interaction between size and stimulus selectivity and changed their rank order of stimulus preferences at the different sizes. Larger stimuli were preferred by 10 cells, and smaller stimuli by six cells.

Size preference was also assessed on an item-by-item basis. Of the 133 (of 166) stimuli that elicited a response, 17 (13%) elicited a significant size preference, of which nine elicited better responses at the larger size and eight at the smaller size. The smaller percentage of stimuli than of cells showing significant effects of size is probably due to the smaller amount of data available for an individual stimulus in the statistical test. The distribution of responses to small and large stimuli is shown in Figure 6. Thus, the responses of IT cells to patterns are largely invariant with stimulus size, at least for the limited range of sizes tested here. However, at least some information about size is available in IT cortex, as some cells prefer one size over another.

Effects of Memory on Test-Stimulus Responses

To determine the effects of memory for the sample on the test stimulus responses, we computed a three-way ANOVA on the test-stimulus responses of each cell individually. The different stimuli used were one factor, the matching-nonmatching status of the stimuli was another, and the size relationship between the sample and test stimuli (same or different sizes) was the third.

The responses to test stimuli often varied significantly depending on whether or not they matched the previous sample, even when the sample and test stimuli were different sizes. Based on the ANOVA, 38% (14 of 37) of the cells had significant match-nonmatching effects (significant main effect of match-nonmatch status), without any significant dependence on the relative sizes of the sample and test stimuli (no significant interaction between match-nonmatch and size factors). An example of one such cell is shown in Figure 7. This size invariance of the memory effect is confirmed in the item-by-item analysis presented below.

Of the cells showing size-invariant effects, 78% (11 of 14) gave smaller responses to match compared to

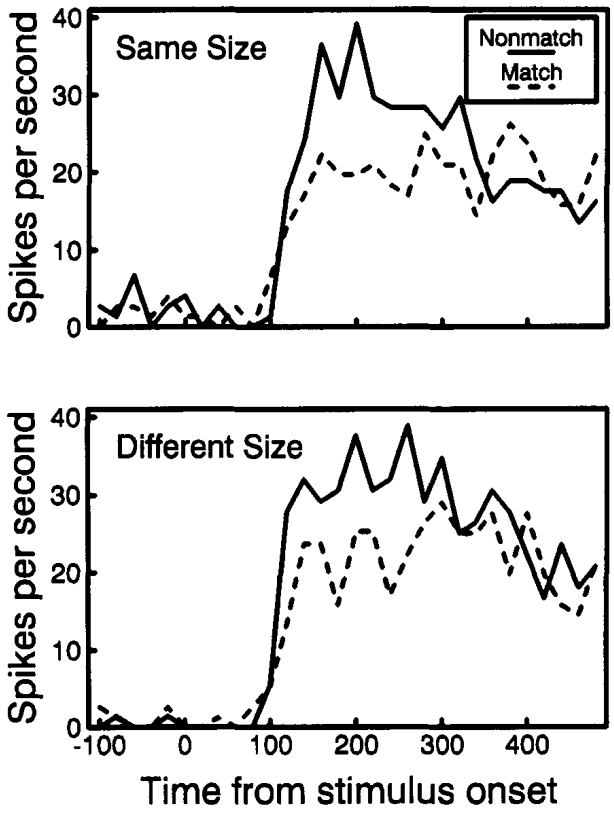

Figure 7. Responses of an IT neuron to a single stimulus appearing as a nonmatch and as a match when both the sample and the test stimuli were the same size (*top*), and when the sample and test stimuli were different sizes (*bottom*).

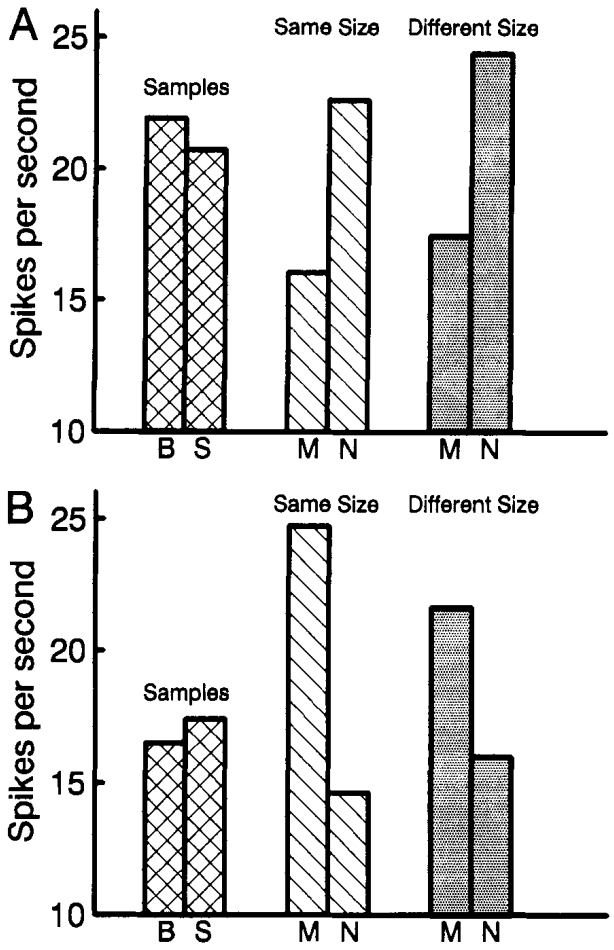

Figure 8. Average responses across cells to the same set of stimuli appearing as samples and as matches and nonmatches. *A*, Average responses for stimuli for which match responses were weaker than nonmatch responses (termed match suppression). *B*, Average responses for which match responses were stronger than nonmatch responses (termed match enhancement). B, large size sample; *S*, small size sample; *M*, match; *N*, nonmatch.

nonmatch stimuli, consistent with previous studies showing a preponderance of match suppression in IT cortex. Likewise, the majority (9 of 14 or 64%) of the cells showing size-invariant memory effects were also stimulus selective (significant main effect of stimulus), consistent with previous studies showing that IT cells are both stimulus selective and influenced by stored memories.

In addition to the cells showing match–nonmatch effects that were invariant across size, three cells (8%) showed memory effects whose magnitude varied significantly depending on the relative sizes of the two stimuli. These cells all gave the largest match–nonmatch effect for sample and test stimuli that were the same size.

As in the location experiment, we extended the cell-by-cell analysis to individual stimuli. A two-way ANOVA was computed on the responses to each individual stimulus for each cell. The match–nonmatch status of the stimulus was one factor, and the relative sizes of the sample and test stimuli (same or different size) were the other. Of the 133 stimuli that elicited significant responses to stimuli of both sizes, the responses to 20% (26 of 133) of them were significantly different depending on whether or not they matched the sample. For 73% (19 of 26) of these items, responses were suppressed when the test stimulus matched the sample, compared to only 27% (7 of 26) of the items with the opposite effect.

The average responses to all items eliciting either match suppression or match enhancement are shown in Figure 8. For both groups, the magnitude of the match–nonmatch effect was nearly equivalent when the sample and test stimuli were of different sizes as when they were the same size. Thus, consistent with the analysis of the cell-by-cell data, the results show that IT cells detect stimuli that match items in memory, even when the stimuli have been transformed in size.

Discussion

IT cortex has long been thought to play a role in invariant object recognition. IT neurons are sensitive to global features of objects, such as their shape, and have extremely large receptive fields (Gross et al., 1972; Desimone and Gross, 1979; Schwartz et al., 1983; Desimone et al., 1984; Tanaka et al., 1991; Fujita et al., 1992). Gross and Mishkin (1977) proposed many years ago that the large receptive fields of IT neurons mediate the perceptual equivalence of stimuli over retinal translation and, more generally, contribute to the constancy of object perception over transformations that preserved object identity. Consistent with this, it was later found that at least some shape- and orientation-selective cells in IT cortex maintain their shape

and orientation tuning over transformations of size or location (Schwartz et al., 1983; Sary et al., 1993), and that face-selective cells maintained their selectivity for faces over changes in rotation, translation, and contrast (Desimone et al., 1984; Rolls and Baylis, 1986; Hasselmo et al., 1989). Interestingly, a recent simulation model of neuronal properties in IT cortex also exhibits size and location invariance (Gochin, 1994, pp 532–543 in this issue). Some IT neurons maintain their selectivity for shape even when the local cues that define the shape's borders are switched from luminance to texture gradients (Sary et al., 1993).

Behavioral studies have also suggested a role for IT cortex in object invariance. Monkeys with IT lesions are impaired in their ability to make visual discriminations across stimulus transformations that preserve object identity, including transformations of size (Weiskrantz and Saunders, 1984) and location (Seacord et al., 1979). In learning object discriminations, monkeys with IT lesions typically show the greatest impairment on object pairs that normal monkeys find difficult to discriminate. An exception is learning to discriminate a pattern from its mirror-image (or any rotation greater than 60°), which normal monkeys find difficult but which pose no special difficulty for monkeys with IT lesions (Gross, 1978; Holmes and Gross, 1984). Gross (1978) suggested that normal monkeys find the discrimination difficult because they view the rotated pattern as equivalent while monkeys with IT lesions view them as different patterns because the monkeys are missing "rotational invariance." Due to this loss of invariance, the rotated patterns are easier to discriminate for the monkeys with lesions, which compensates in part for their general pattern discrimination deficit. The loss of invariances is not absolute, however, as monkeys with IT lesions can learn to transfer a learned object discrimination to the same objects shown at a different size (Holmes and Gross, 1984). Taken together, physiological and behavioral studies indicate that IT cortex plays an important role in maintaining object constancy in the face of transformations that alter local visual features but preserve object identity.

The present results help clarify the role of IT neurons in perceptual invariance and suggest that they play a similar role in recognition within short-term memory. In the perceptual domain, we found that the responses of some cells are invariant in their absolute magnitude over changes in stimulus location and size. For other cells the magnitude of response changed with size or location. However, because responses to all stimuli changed proportionally (or at least maintained their preferred order) for these cells, even they may contribute to perceptual invariance. One could, in principle, extract information about invariant features such as shape from these cells by comparing the responses (e.g., taking their ratio) of cells tuned to different shape features.

In the mnemonic domain, we found in this and previous studies that the responses of IT neurons to a current stimulus are jointly determined by the sensory features of that stimulus as well as by its similarity to items stored in memory (Miller et al., 1991, 1993). The present results show that for most cells, this interaction of the current stimulus with the memory trace is invariant over size and retinal translation. The interaction must take place at a high level of object representation rather than at the level of local object features. It is presumably due, at least in part, to cells such as these that one recognizes a recently seen object that is changed in size or location from its original appearance.

Sensory information about the size and locations of objects is not entirely absent in IT cortex, however. It was found in this and other studies that most IT receptive fields have a "hot spot" at the fovea, most IT cells respond better to stimuli in the contralateral than in the ipsilateral visual field (Gross et al., 1972; Gross, 1973; Desimone et al., 1984), and many cells also respond differentially to stimuli located in the upper versus lower contralateral fields (Chelazzi et al., 1993 and unpublished observations). In other words, the hot spot of IT receptive fields is often a "hot streak," beginning at the fovea and extending into the upper or lower contralateral field. Likewise, we found in the present study that a significant minority of IT neurons respond selectively according to object size. For some cells, responses to all objects increased or decreased with size, while other cells responded only to a specific object at a specific size. These results suggest that retinal location and size are treated like other object features by some IT neurons. That is, some cells may be coarsely tuned to location and size the same way that they are tuned to color, shape, and so on. In principle, then, both types of information are available from different cell populations in IT cortex, namely, information about specific object features like retinal location and size, as well as information about object identity, which is invariant across changes in location and size. Which type of information, either size- and location-specific on the one hand, or invariant on the other, is used by the organism at any given time may depend on task demands.

Enhancement and Suppression Mechanisms

In previous studies, we found two different short-term memory mechanisms in IT cortex. One, a suppressive mechanism, is engaged by simple stimulus repetition (Miller et al., 1991, 1993). The other, an enhancement mechanism, is engaged when the animal must actively hold a stimulus "in mind" and compare incoming stimuli to it (Miller and Desimone, 1994). So far, we have distinguished these two mechanisms only in a special version of the DMS task termed "ABBA." In the ABBA version, some of the intervening "nonmatch" items match each other. The monkey must ignore these irrelevant stimulus repetitions and signal only when a stimulus matches the sample item presented at the beginning of the trial. This task forces the monkey to actively maintain the memory of the sample and compare test items to it. By contrast, in conventional DMS, monkeys can solve the task by responding automatically to any stimulus repetition within the trial. Indeed, we found that when monkeys trained in conventional

DMS were suddenly confronted with ABBA trials, they responded inappropriately to the repeated nonmatches.

The ABBA task revealed that cells showing match suppression are equally suppressed by the intervening items that match each other (i.e., an irrelevant stimulus repetition), whereas cells showing match enhancement are only enhanced by the one item matching the sample held in memory. The former type of cell may mediate the automatic detection of stimulus repetition, while the latter type may mediate, in part, a more cognitive "working memory" (Baddeley, 1986). Both the suppressed cells and the small number of enhanced cells we found in this study seemed to exhibit size and location invariance. Thus, at least some degree of object invariance is likely to be characteristic of both neural mechanisms.

Because match suppression occurs automatically with stimulus repetition, regardless of whether the stimuli are behaviorally relevant, it is tempting to speculate it may have some relationship to repetition priming (e.g., Schacter, 1992). In this regard, it is interesting that priming effects, as assessed by reaction times to repeated stimuli, occur despite changes in size, location, or orientation between two successive presentations of the same object (Ellis et al., 1989; Biederman and Cooper, 1991, 1992).

Finally, the fact that these neural memory mechanisms exhibit size and location invariance may shed some light on their origin in the cortex. We previously found that suppression and enhancement both begin at the very onset of the visual response, around 80 msec. That is, the effects occur practically as soon as visual information reaches IT cortex. This appeared to rule out "on-line" feedback from other structures as a cause of the effects in IT cortex. Logically, however, the suppression and enhancement might have been generated prior to IT cortex. Indeed, we have found that some V4 cells exhibit match suppression in the DMS task (Miller, Li, and Desimone, unpublished observations). However, V4 receptive fields are not large enough to extend from the center of gaze to a location 5° into the periphery, which is the range over which we found suppression with repetition in IT cortex. Furthermore, unlike IT cells, there is no evidence that V4 cells exhibit size invariance. It seems likely, therefore, that size- and location-invariant memory effects are generated intrinsically within IT cortex, and are not simply passed on from earlier visual areas.

Notes

We thank Rebecca Hoag for help with the monkeys, Thomas White for writing the computer programs, and Mortimer Mishkin for his support throughout all phases of the study. Peter DeWeerd and Leonardo Chelazzi provided helpful comments on the manuscript. The study was supported in part by the Human Frontier Science Program Organization.

Correspondence should be addressed to Robert Desimone, Laboratory of Neurophysiology, National Institute of Mental Health, Building 49, Room 1B80, Bethesda, MD 20892.

References

Baddeley A (1986) Working memory. Oxford: Clarendon.

Baylis GC, Rolls ET (1987) Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. Exp Brain Res 65:614–622.

Biederman I, Cooper EE (1991) Evidence for complete translational and reflectional invariance in visual object priming. Perception 20:585–593.

Biederman I, Cooper EE (1992) Size invariance in visual object priming. J Exp Psychol [Hum Percept] 18:121–133.

Chelazzi L, Miller EK, Duncan J, Desimone R (1993) A neural basis for visual search in inferior temporal cortex. Nature 363:345–347.

Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. Brain Res 178:363–380.

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051–2062.

Ellis R, Allport DA, Humphreys GW, Collis J (1989) Varieties of object constancy. Q J Exp Psychol 41A:775–796.

Eskandar EN, Richmond BJ, Optican LM (1992) Role of inferior temporal neurons in visual memory: I. Temporal encoding of information about visual images, recalled images, and behavioral context. J Neurophysiol 68:1277–1295.

Fahy FL, Riches IP, Brown MW (1993) Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior and rhinal cortex. Exp Brain Res 96:457–472.

Fujita I, Tanaka K, Ito M, Cheng K (1992) Columns for visual features of objects in monkey inferotemporal cortex. Nature 360:343–346.

Gaffan D, Murray EA (1992) Monkeys (*Macaca fascicularis*) with rhinal cortex ablations succeed in object discrimination learning despite 24-hr intertrial intervals and fail at matching to sample despite double sample presentations. Behav Neurosci 106:30–38.

Gross CG (1973) Visual functions of inferotemporal cortex. In: Handbook of sensory physiology (Jung R, ed), pp 451–482. Berlin: Springer.

Gross CG (1978) Inferior temporal lesions do not impair discrimination of rotated patterns in monkeys. J Comp Physiol Psychol 92:1095–1109.

Gross CG, Mishkin M (1977) The neural basis of stimulus equivalence across retinal translation. In: Lateralization in the nervous system (Harned S, Doty R, Jaynes J, Goldberg L, Krauthamer G, eds), pp 109–122. New York: Academic.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J Neurophysiol 35:96–111.

Gross CG, Bender DB, Gerstein GL (1979) Activity of inferior temporal neurons in behaving monkeys. Neuropsychologia 17:215–229.

Hasselmo ME, Rolls ET, Baylis GC, Nalwa V (1989) Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. Exp Brain Res 75:417–429.

Holmes EJ, Gross CG (1984) Stimulus equivalence after inferior temporal lesions in monkeys. Behav Neurosci 98: 898–901.

Horel JA, Pytko-Joiner DE, Voytko ML, Salsbury K (1987) The performance of visual tasks while segments of the inferotemporal cortex are suppressed by cold. Behav Brain Res 23:29–42.

Li L, Miller EK, Desimone R (1993) The representation of stimulus familiarity in anterior temporal cortex. J Neurophysiol 69:1918–1929.

Miller EK, Desimone R (1994) Parallel neuronal mechanisms for short-term memory. Science 263:520–522.

Miller EK, Li L, Desimone R (1991) A neural mechanism for working and recognition memory in inferior temporal cortex. Science 254:1377–1379.

Miller EK, Li L, Desimone R (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. J Neurosci 13:1460–1478.

Mishkin M (1982) A memory system in the monkey. Philos Trans R Soc Lond [Biol] 298:83–95.

Riches IP, Wilson FA, Brown MW (1991) The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. J Neurosci 11:1763–1779.

Rolls ET, Baylis GC (1986) Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. Exp Brain Res 65:38–48.

Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. Science 260:995–997.

Schacter DL (1992) Priming and multiple memory systems: perceptual mechanisms of implicit memory. J Cognit Neurosci 4:244–256.

Schwartz EL, Desimone R, Albright TD, Gross CG (1983) Shape recognition and inferior temporal neurons. Proc Natl Acad Sci USA 80:5776–5778.

Seacord L, Gross CG, Mishkin M (1979) Role of inferior temporal cortex in interhemispheric transfer. Brain Res 167: 259–272.

Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol 66:170–189.

Vogels R, Orban GA (1994) Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. J Neurophysiol 71:1428–1451.

Weiskrantz L, Saunders RC (1984) Impairments of visual object transforms in monkeys. Brain 107:1033–1072.

Zola-Morgan S, Squire LR, Amaral DG, Suzuki WA (1989) Lesions of perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. J Neurosci 9:4355–4370.