

Stimulus Specificity and Temporal Dynamics of Working Memory for Visual Motion

Tatiana Pasternak and Daniel Zaksas

Department of Neurobiology and Anatomy and Center for Visual Science, University of Rochester, Rochester, New York 14642

Submitted 1 May 2003; accepted in final form 9 June 2003

Pasternak, Tatiana and Daniel Zaksas. Stimulus specificity and temporal dynamics of working memory for visual motion. *J Neurophysiol* 90: 2757–2762, 2003. First published June 11, 2003; 10.1152/jn.00422.2003. When asked to compare two moving stimuli separated by a delay, observers must not only identify stimulus direction but also store it in memory. We examined the properties of this storage mechanism in two macaque monkeys by sequentially presenting two random-dot stimuli, sample and test, in opposite hemifields and introducing a random-motion mask during the delay. The mask interfered with performance only at the precise location of the test, 100–200 ms after the start of the delay, and when its size and speed matched those of the remembered sample. This selective interference suggests that the representation of the motion stimulus in memory preserves its direction, speed, and size and is most fragile shortly after the completion of the encoding phase of the task. This precise preservation of sensory attributes of the motion stimulus suggests that the neural mechanisms involved in the processing of visual motion may also be involved in its storage.

INTRODUCTION

As we navigate the visual world and images appear and disappear on our retina, we are often faced with the task of comparing images present at different locations in space at different points in time. Thus the ability to briefly store visual information is fundamental to performing appropriate visually guided behaviors and to maintaining continuity of visual experience. This capacity to briefly retain task-related sensory information is often referred to as working memory (Baddeley 1986). Visual motion is one of the more important features of objects in the environment, and the mechanisms underlying the processing of visual motion have been extensively studied, both psychophysically and physiologically (for recent review see Pasternak et al. 2003). In contrast, less is known about the way visual motion is stored and subsequently used to guide motor action. A few studies that did examine this question showed that the information about visual motion retained in memory is spatially localized (Zaksas et al. 2001) and can be retained in memory for many seconds with good fidelity (Blake et al. 1997; Magnussen and Greenlee 1992).

In this study we focused on the properties of the remembered motion information. We used a novel approach in which we spatially separated the encoding and retrieval/comparison phases of the working memory task. This separation allowed us to use a random-motion mask to examine the attributes of the

remembered stimulus as well as to measure the spatial and temporal characteristics of the underlying mechanism. Our results demonstrate that the memory representation is highly specific, retaining the information not only about stimulus direction and its location but also about its size and speed. Furthermore, our results show that this specific representation of visual motion is most fragile immediately after encoding. This faithful preservation of all attributes of the remembered motion stimulus suggests active involvement of motion processing cortical areas in its storage.

METHODS

SUBJECTS. Two adult male macaque monkeys (*Macaca nemestrina*), weighing 8–9 kg, were used. On weekdays, water was restricted and the daily water ration was provided during the behavioral testing. On weekends, the monkeys were not tested behaviorally and received 100 ml/kg water per day. Food was continually available in the home cage and monkeys received supplements of fresh fruit and vitamins daily. Body weights were measured at least three times per week to ensure good health and normal growth. The monkeys were implanted with scleral search coils and head restraint devices to monitor their eye position (see Rudolph and Pasternak 1999 for description). Experiments were carried out in accordance with the guidelines published in the National Institutes of Health Guide for the Care and Use of Laboratory Animals (National Institutes of Health publication No. 86–23, revised 1987).

Stimuli and behavioral procedures

STIMULI. The stimuli consisted of moving dots within a stationary circular aperture, 3° or 4° in diameter, repeatedly displaced in a direction of motion chosen randomly from a uniform distribution of directions (Fig. 1A). Such stimuli were first introduced by Williams and Sekuler (1984) and subsequently used extensively by this and other laboratories (e.g., Blake et al. 1997; Pasternak et al. 1990; Watamaniuk et al. 1989). In most experiments, the lifetime of an individual dot was equal to the duration of the stimulus presentation (300 ms). A few measurements (in Monkey 2) were performed with dot lifetime set to 50 ms. The dots were displaced by a constant step size (Δx) and temporal interval ($\Delta t = 13$ ms) and the dot density was kept at 4.7 dots/deg². In all experiments, all three

Address for reprints and other correspondence: T. Pasternak, Department of Neurobiology and Anatomy, Box 603, University of Rochester, Rochester, NY 14642 (E-mail: tania@cvs.rochester.edu).

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

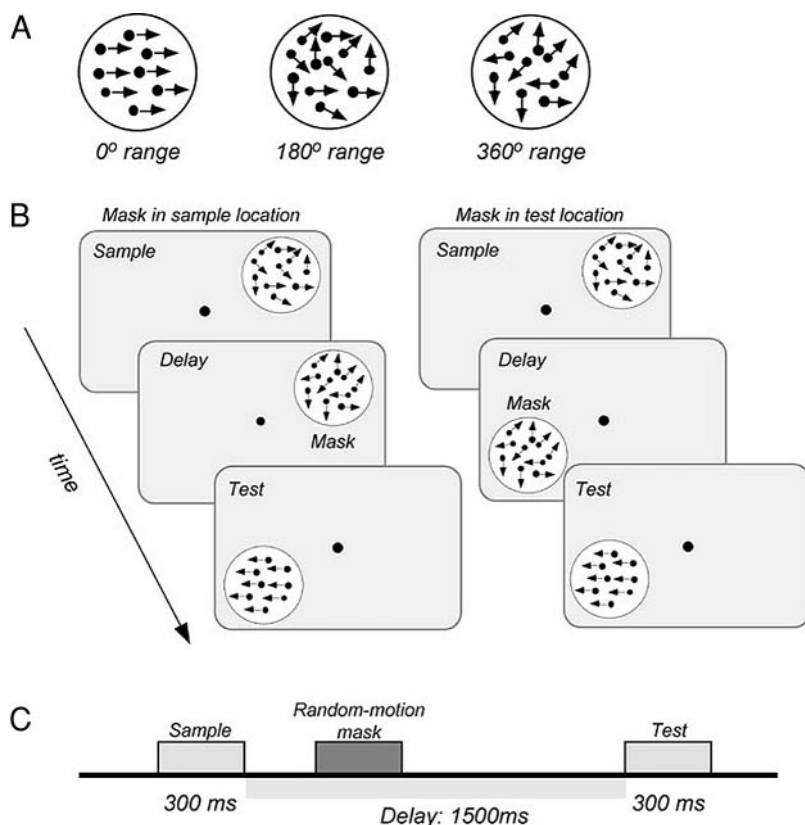


FIG. 1. *A*: direction range stimuli. The sample stimulus consisted of randomly distributed dots moving in a predetermined range of directions, resulting in a net direction of motion for the stimulus. On each trial the net motion of the sample was selected at random and set to in 1 of 8 cardinal directions. *B*: behavioral task. Monkeys initiated each trial by fixating a small dot for 1000 ms, after which the two comparison stimuli, sample and test, separated by a 1500-ms delay were presented. The duration of sample and test was set to 300 ms. In most experiments a third intervening stimulus, the mask, was added during the delay. The mask consisted of dots with direction range set to 360° (no net motion). The sample and test stimuli were placed in noncorresponding locations of opposite hemifields (sample: 5° to the right and 5° up from the fovea; test: 5° to the left and 5° down from the fovea). The mask appeared either at the location of the preceding sample or at the location of the upcoming test. The trials were run in blocks making the appearance of test and the mask highly predictable. At the end of the trial the monkeys indicated whether the 2 stimuli moved in the same or different directions by pressing 1 of 2 response buttons. The monkeys were required to maintain fixation throughout the duration of the trial. *C*: temporal parameters of the trial. Sample and test, each lasting 300 ms were separated by a 1500-ms delay. The random-motion mask was presented at various times during the delay and its duration was predetermined. Here, the mask is presented early in the delay and its duration is set to 300 ms.

stimuli (sample, mask, and test) were always presented at an eccentricity of 7° in the upper or lower portions in the noncorresponding quadrants of the visual field (see Fig. 1*B*).

The range of the distribution of directions set to 0° renders a stimulus with all the dots displaced in the same direction, while the distribution of directions of 360° contains only local random motion of individual dots and no net motion. When the distribution was narrower than about 320–340°, the dots appeared to flow in the direction of the mean of the distribution. The sample stimulus consisted of dots repeatedly displaced in a direction chosen randomly from a uniform distribution of directions (Fig. 1*A*). The test stimulus always consisted of dots moving coherently (direction range = 0°) in the same or in the opposite direction to that in the sample. Sample and test were either presented at the same spatial location or in opposite, noncorresponding portions of the visual field, but always at equal distance from the fovea (Fig. 1*B*).

BEHAVIORAL TASK. During each testing session, the monkeys compared the directions of two sequentially presented random-dot stimuli, sample and test, separated by a 1500-ms delay. A tone was presented for the duration of the trial and the monkeys were required to maintain fixation throughout that period. On each trial, the direction of net flow in the sample was chosen at random from eight equally spaced directions and the test moved in a direction that was the same as, or opposite from, that of the sample. The monkeys judged the directions as the same or different by pressing one of two adjacent buttons. Incorrect responses resulted in a 3- to 5-s tone and no reward. The temporal sequence during each trial is shown in Fig. 1*C*. At various times during the 1500-ms delay, a random-motion stimulus (direction range = 360°), the mask, was introduced either at the location of the sample or at the location of the upcoming test. During each testing session, the

location of all stimuli was the same. Thus on each trial the monkeys knew where the sample, the test, and the mask would appear. Each session consisted of 400–800 trials, separated by a 3-s intertrial interval.

THRESHOLD MEASUREMENTS. Performance was evaluated by measuring direction range thresholds in a staircase procedure. In this procedure, three consecutive correct responses resulted in an increase in direction range, while a single incorrect response resulted in a decrease in direction range. The data were fitted with a maximum likelihood Weibull function (Weibull 1951) and the threshold was defined as the stimulus value at which the animal performed at 75% correct. Three to five or more threshold determinations were performed for each stimulus condition. The significance of the effects of each manipulation was determined using a one-way ANOVA.

RESULTS

We examined the effect of a random-motion masking stimulus, introduced during the delay separating the sample and the test by placing it either at the location of the sample or at the location of the upcoming test in the opposite hemifield. We reasoned that a disruptive effect of the mask on performance would suggest interference with the representation of the preceding sample in memory and the selectivity of such interference would reveal the properties of this representation.

Spatial localization of the remembered sample

We found that the mask interfered with performance under a limited set of spatial and temporal conditions. The data in Fig. 2 show that the mask had a marked detrimental effect on

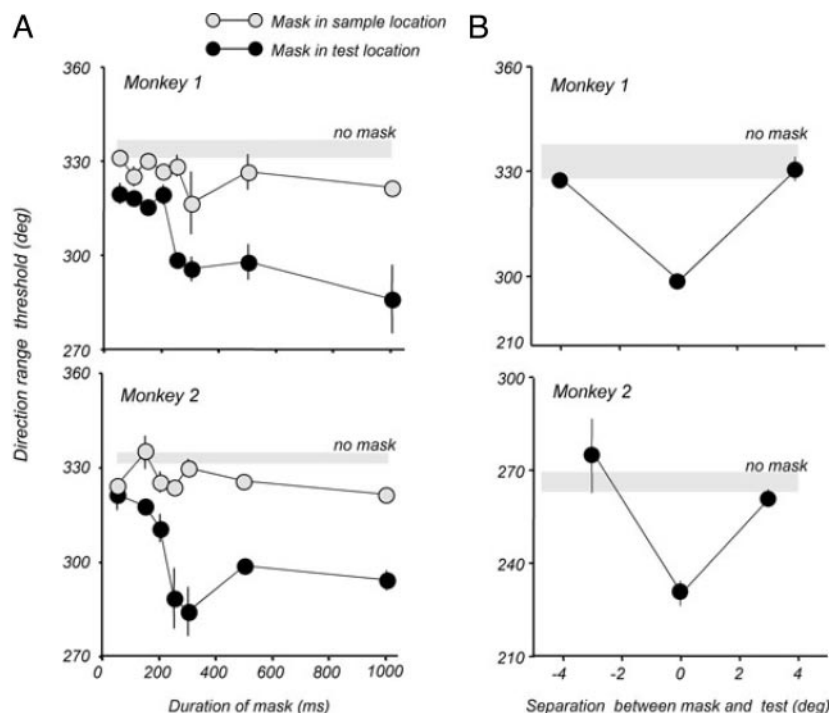


FIG. 2. *A*: effect of mask duration. Direction range thresholds are plotted as a function of mask duration as the mask is placed in the location of the preceding sample (gray points) or that of the upcoming test (black points). Range thresholds measured without the mask are indicated by the vertical position of the gray bar, where the width represents mean performance \pm SE. The mask was presented 100 ms after the sample offset. At all mask durations the threshold was significantly different from the “no mask” condition ($P < 0.05$, ANOVA), when the mask appeared in the test location. Performance was generally worse when the mask appeared in the test rather than the sample location, and this difference was significant for mask durations > 200 ms ($P < 0.05$, ANOVA). Dots lifetime was set to 300 ms for both monkeys. *B*: spatial specificity of mask effect. In this experiment, the mask was placed either at the location of the test (indicated by 0° on the X-axis) or at 2 locations on either side of the upcoming test. Thresholds in the “no mask” condition are shown by the gray bar (see Fig. 2). All targets were always presented isoeccentrically, 7° from fovea. Spatial separation indicates the center-to-center distance between the mask and test stimuli. The thresholds for monkey 2 were measured with dot lifetime set to 50 ms and for that reason were lower than thresholds for monkey 1 whose thresholds were measured with 300 ms lifetime. The mask had a disruptive effect on performance only when the mask appeared in the precise location of the upcoming test ($P < 0.0001$, ANOVA, for both monkeys). Error bars represent \pm SE.

performance only when it was placed at the location of the upcoming test. Maximal disruption was achieved when the mask was present at that location for ≥ 250 –300 ms. At shorter durations, the effect of the mask in the test location was much less pronounced. On the other hand, the effect of the mask at the sample location was minimal, irrespective of its duration.

We also examined the extent to which this representation is confined to the test location by placing the mask at several distances from the upcoming test. The data in Fig. 2*B* show that the mask lost its detrimental effect if it were displaced only a few degrees from the location of the upcoming test. This spatial specificity of the mask effect provides new evidence supporting the notion of spatial localization of the remembered stimulus proposed by Zaksas et al. (2001). That study demonstrated spatial specificity by manipulating the spatial separation between the sample and the test. In the present study, by manipulating the location of the mask we were able to more precisely pinpoint the site of the remembered stimulus to the location where the perceptual decision was to take place.

What is remembered?

To learn more about the nature of the sample representation in memory we determined whether, in addition to motion direction, the information about its size and speed was also retained. We examined the properties of the remembered stimulus by manipulating the size and the local speeds of the mask and using it as a probe.

We found that the mask was most disruptive when its size approximated the size of the remembered sample (Fig. 3), suggesting that the size of the sample was likely to have been preserved in memory.

To determine whether the speed of the sample was also preserved in memory, we manipulated the spatial displacement of individual dots (Δx) in the mask. This manipulation resulted in a mismatch between the local speeds in the mask and the

local speeds of the sample. A detrimental effect of the mask despite the mismatch in temporal properties would be indicative of the sample speed not being retained. The results in Fig. 4 show that when sample and test moved at $10^\circ/\text{s}$ ($\Delta x = 0.13$, $\Delta t = 13$ ms), the effect of the mask was most pronounced if the dot displacement in the mask and the sample matched precisely.

When is the mask most effective?

To determine the point in time when the representation of the remembered sample is most susceptible to disruption, the

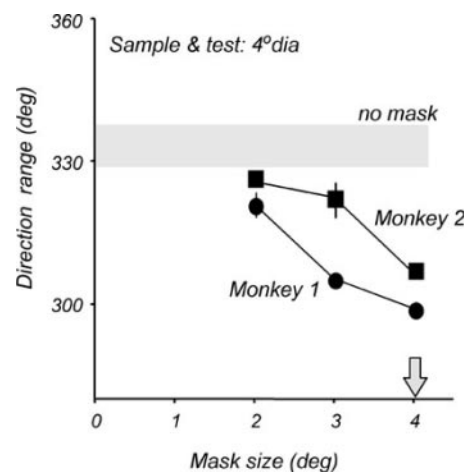


FIG. 3. Effect of mask size. Direction range thresholds plotted as a function of the size of the mask. The mask always appeared at the location of the test, 100 ms after the sample offset. The “no mask” condition for both monkeys is plotted as a gray bar. The arrow indicates the size of the sample stimulus. There was a significant loss in thresholds for all 3 mask sizes relative to the “no mask” condition, although the mask of the same size as the sample and test had the largest effect ($P < 0.05$, $P < 0.001$, $P < 0.0001$, ANOVA, for increasing mask size, respectively, for both monkeys). Error bars represent \pm SE. The arrow points to the size of sample and test stimuli.

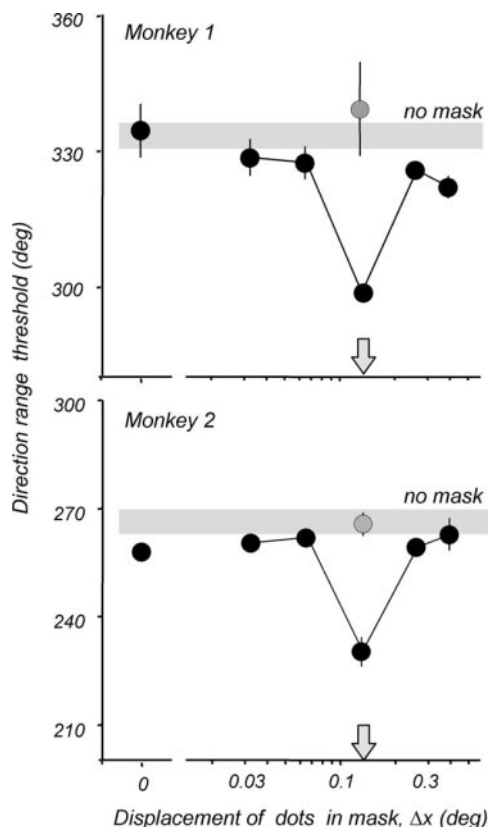


FIG. 4. Effect of mask speed. Thresholds measured with sample and test moving at 10 °/s ($\Delta x = 0.13^\circ$; $\Delta t = 13$ ms) as function of Δx of the dots in the mask (see METHODS for detail). In both monkeys, the mask caused a large and significant decrement in performance only when it appeared in the test location (black points) and its Δx (local speed) matched that of the sample and the test ($P < 0.0001$, ANOVA). A mask with stationary dots ($\Delta x = 0$, black symbol on far left of each plot) did not reveal any performance decrement. Performance with the mask appearing in the sample location (in gray) and for the “no mask” condition (gray bar) is shown for comparison.

random-motion mask was introduced at various points of the delay. The results of these measurements are shown in Fig. 5. In both monkeys, the masking was most pronounced when the mask was introduced into the test location about 100–150 ms after the start of the delay.

The effectiveness of the mask decreased later in the delay and by 200–300 ms into the delay thresholds with the mask were nearly identical in sample and in test locations, and similar to those measured without the mask. This extreme temporal specificity of the mask effect suggests that very early in the delay, about 100 ms after the offset of the sample, the neural representation of the stimulus has already been moved to the test location and is most susceptible to disruption.

DISCUSSION

We found that the introduction of the random-motion mask during the delay interfered with performance in a spatially and temporally selective way. The interference was most pronounced when the mask stimulus matched the size and temporal properties of the remembered motion stimulus, if it was introduced early in the delay, and if its location matched precisely the location of the upcoming comparison test.

Interfering effect of the mask

The specific nature of the mask effect is indicative of interference with the remembered stimulus rather than the nonspecific disruption in behavior due to distraction. If this effect was due to simple confusion we would expect to see a decrease in performance that was not related to the nature of the mask or to the remembered sample. Furthermore, the effect of the mask was absent when the remembered stimulus moved coherently (0° range)(Zaksas et al. 2001).

Temporal dynamics of storage

In this experiment the location and the time of appearance of the sample, the mask, and the test were highly predictable. Thus the monkeys knew where and when the mask and the test were going to appear. The spatially selective effect of the mask suggests that, in anticipation of the test appearing in a remote location, the monkeys transferred information about the remembered stimulus to that location. Particular susceptibility of performance to the mask during a relatively narrow temporal window early in the delay points to the possible time of this

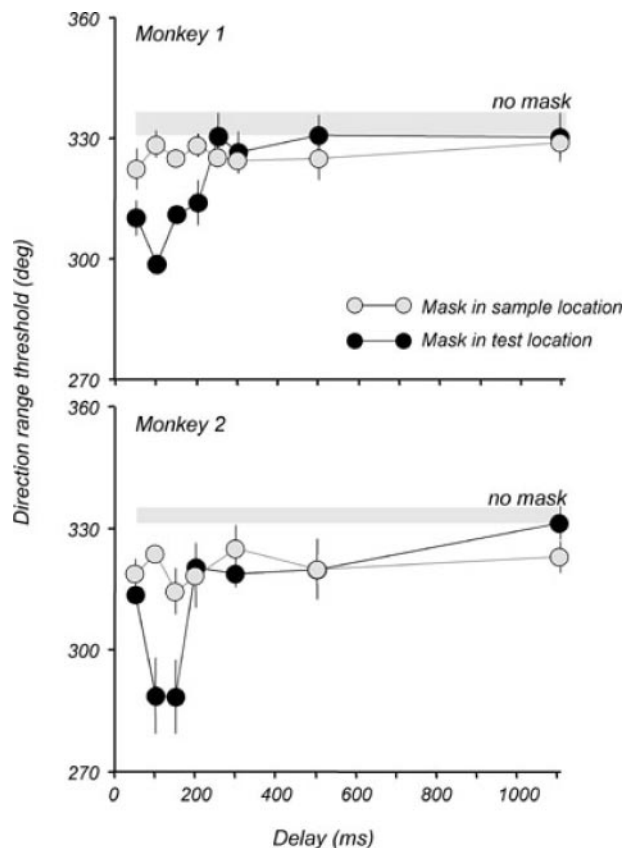


FIG. 5. Effect of mask onset time. Thresholds measured with the 250-ms random-motion mask presented at different points in the delay. The thresholds with mask at the test and sample locations appearing at different times were measured in separate sessions. When the mask appeared in the test location (black points), there was a specific and significant decrement in performance only when the mask appeared at the test location very early in the delay, around 100 ms after sample offset ($P < 0.001$, ANOVA, for both monkeys). No significant spatial specificity was observed if the mask appeared 200 ms or more after sample offset. Thresholds for “no mask” condition are shown for comparison by the gray bar. Dots lifetime was set to 300 ms for both monkeys. Error bars represent \pm SE.

transfer and to the vulnerability of memory representation early in the delay. We speculate that, around the time of transfer, the representation of the sample in memory may be less stable and thus more susceptible to noise introduced by the mask. It is interesting that effectiveness of the mask appears to be most pronounced at the time of its introduction and is less dependent on its duration (see Fig. 2). It remains to be seen whether the vulnerability of performance 100–150 ms after the start of the delay is the result of neural events taking place at time of mask onset or is related to the time it takes for the mask to reach its maximal effect or both.

The idea of signal instability early in the delay is supported by a recent network model of working memory (Compte et al. 2000; Wang 2001). In this model, sensory information is fed into a network of neurons designated for parametric storage by way of stable recurrent activity. Wang and his colleagues have shown that a distracter signal had its maximal effect when it varied little from the original stimulus along the parameter being retained by such recurrent networks (Compte et al. 2000). Furthermore, it takes a relatively long time for these recurrent circuits to achieve stability because they depend on relatively slow receptor dynamics and on inhibitory circuits with time constants on the order of several hundred milliseconds (Compte et al. 2000; Wang 2001). The properties of the proposed recurrent networks are consistent with our data, which show the maximal disruption by a noisy mask occurring very shortly after the sensory signal in the task-relevant spatial location. A working memory signal therefore appears to be disrupted by parameter-specific noise injected into the system prior to stabilization.

Our data show that the mask had to endure for ≥ 200 ms to maximally disrupt performance. This may be indicative of the relatively long temporal integration characteristic of the mechanisms responsible for processing complex motion (Watamaniuk and Sekuler 1992). This observation further strengthens the notion that mechanisms involved in processing of motion stimuli also participate in retaining these stimuli in memory (Bisley and Pasternak 2000; Fuster 1997).

Nature of the remembered stimulus

We used the random-motion mask to reveal specific characteristics of the remembered stimulus. Masking stimuli have previously been used in a number of experiments examining perceptual memory. In these “memory masking” experiments the mask was applied while subjects performed delayed discriminations of speeds (Magnussen and Greenlee 1992) and of spatial frequencies (Bennett and Cortese 1996; Lalonde and Chaudhuri 2002; Magnussen et al. 1991). The masks in these experiments consisted of stimuli that bore different degrees of similarity to the stimuli being discriminated. The general finding of these experiments was that the less the mask resembled the discriminated visual stimuli, the greater the detrimental effect of the mask. Since these masking effects were selective for specific stimulus dimensions, it was concluded that perceptual memory for different stimulus attributes is served by parallel, higher-order perceptual mechanisms (e.g., motion, spatial frequency) (Magnussen and Greenlee 1999).

In our experiments, the mask contained only random motion and thus provided no motion information relevant to the task.

It was used to inject noise into the process underlying the retention of the sample (Ball and Sekuler 1979). We reasoned that a detrimental effect of such a “memory mask” would be a reflection of its interference with the representation of the remembered stimulus. Indeed in the previous study we used such masks to demonstrate that this representation is spatially localized to the site of the upcoming test (Zaksas et al. 2001). In that paper we also provided evidence that the sample containing a broad range of directions was more likely to be retained as a veridical representation of the stimulus than a simple vector representing the mean direction. The present findings provide strong support for this notion. The dependence of the mask’s effects on its size and local speed suggests that the representation of the sample in memory retains not only the information about its direction as required by the task, but also about its size and speed.

One of the striking observations of our study is the speed-dependent specificity of the mask effect. When the sample moved at a lower speed, the mask was effective only when its local speeds (i.e., Δx) matched precisely the local speeds of the remembered sample. Together with other reports of visual masks selectively interfering with sensory information (Ball and Sekuler 1979; Bennett and Cortese 1996; Magnussen and Greenlee 1992), these results suggest that the storage of visual signals is likely to be accomplished, at least in part, by the very mechanisms that are involved in their encoding.

Possible neural mechanisms

A number of recent studies from this laboratory have implicated cortical middle temporal (MT) area in the performance of the memory for motion task. Bisley and Pasternak (2000) have shown that deficits produced by lesions of MT and middle superior temporal area (MST) MT/MST were greatest at longer delays separating sample and test stimuli. In a subsequent study, Bisley et al. (2001) reported a severe disruption in performance of the task by microstimulation of MT during the delay. Furthermore, in a recent psychophysical study, we have demonstrated that the spatial scale of the mechanisms underlying the performance of our task closely matches the spatial scale of MT neurons (Zaksas et al. 2001).

The involvement of MT neurons in remembering visual motion is also supported by our recent recordings from MT neurons carried out in the same animals performing the task identical to that used in the present study (Bisley, Zaksas, Droll, and Pasternak, unpublished data). These recordings revealed a pattern of activity early in the delay that may be related to the phenomenon observed here. Specifically, many MT neurons showed transient activation during the first few hundred milliseconds of the delay and the properties of this activation reflected the direction of the preceding sample. While it remains to be seen whether this activation is linked directly to the ability to remember visual motion, its existence in the absence of a visual stimulus implicates area MT in the circuitry subserving memory for visual motion. The striking parallel between the time of maximal effectiveness of the mask and the appearance of early activation in MT suggests a possibility that the mask may be interfering with this early delay activity in MT. While we do not have direct evidence to support this possibility, the present study strongly implicates

cortical area(s) with machinery to process complex visual motion in the ability to remember it.

We thank M. Mancarella and D. Moore for excellent technical assistance, B. Singer for software development, and M. Gira for help with electronics. We also thank B. Merigan for comments on the manuscript.

DISCLOSURES

This work was supported by National Institutes of Health Grants, RO1 EY-11749 to T. Pasternak and T32 NS-07489 to D. Zaksas and in part by P30 EY-01319.

REFERENCES

- Baddeley A.** *Working Memory*. Oxford, UK: Clarendon Press/Oxford, 1986, p. 289.
- Ball K and Sekuler R.** Masking of motion by broadband and filtered directional noise. *Percept Psychophys* 26: 206–214, 1979.
- Bennett PJ and Cortese F.** Masking of spatial frequency in visual memory depends on distal, not retinal, frequency. *Vis Res* 36: 233–238, 1996.
- Bisley JW and Pasternak T.** The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion. *Cereb Cortex* 10: 1053–1065, 2000.
- Bisley JW, Zaksas D, and Pasternak T.** Microstimulation of cortical area MT affects performance on a visual working memory task. *J Neurophysiol* 85: 187–196, 2001.
- Blake R, Cepeda NJ, and Hiris E.** Memory for visual motion. *J Exp Psychol Hum Percept Perform* 23: 353–369, 1997.
- Compte A, Brunel N, Goldman-Rakic PS, and Wang XJ.** Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10: 910–923, 2000.
- Fuster JM.** Network memory. *Trends Neurosci* 20: 451–459, 1997.
- Lalonde J and Chaudhuri A.** Task-dependent transfer of perceptual to memory representations during delayed spatial frequency discrimination. *Vis Res* 42: 1759–1769, 2002.
- Magnussen S and Greenlee MW.** Retention and disruption of motion information in visual short-term memory. *J Exp Psychol Learn Mem Cogn* 18: 151–156, 1992.
- Magnussen S and Greenlee MW.** The psychophysics of perceptual memory. *Psychol Res* 62: 81–92, 1999.
- Magnussen S, Greenlee MW, Asplund R, and Dyrnes S.** Stimulus specific mechanisms of visual short-term memory. *Vis Res* 31: 1213–1219, 1991.
- Pasternak T, Albano JE, and Harvitt DM.** The role of directionally selective neurons in the perception of global motion. *J Neurosci* 10: 3079–3086, 1990.
- Pasternak T, Bisley JW, and Calkins D.** Visual information processing in the primate brain. In: *Biological Psychology*, edited by Gallagher M and Nelson RJ. New York: Wiley, 2003, chapt. 6, p. 139–185.
- Rudolph K and Pasternak T.** Transient and permanent deficits in motion perception after lesions of cortical areas MT and MST in the macaque monkey. *Cereb Cortex* 9: 90–100, 1999.
- Wang XJ.** Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci* 24: 455–463, 2001.
- Watamaniuk SN and Sekuler R.** Temporal and spatial integration in dynamic random-dot stimuli. *Vis Res* 32: 2341–2347, 1992.
- Watamaniuk SN, Sekuler R, and Williams DW.** Direction perception in complex dynamic displays: the integration of direction information. *Vis Res* 29: 47–59, 1989.
- Weibull W.** A statistical distribution function of wide applicability. *J Appl Mech.* 18: 292–297, 1951.
- Williams DW and Sekuler R.** Coherent global motion percepts from stochastic local motions. *Vis Res* 24: 55–62, 1984.
- Zaksas D, Bisley JW, and Pasternak T.** Motion information is spatially localized in a visual working-memory task. *J Neurophysiol* 86: 912–921, 2001.