Perception & Psychophysics
2005, 67 (2), 185-194

Remembering "what" brings along "where" in visual working memory

INGRID R. OLSON
University of Pennsylvania, Philadelphia, Pennsylvania

and

CHRISTY MARSHUETZ
Yale University, New Haven, Connecticut

Does a behavioral and anatomical division exist between spatial and object working memory? In this article, we explore this question by testing human participants in simple visual working memory tasks. We compared a condition in which there was no location change with conditions in which absolute location change and absolute plus relative location change were manipulated. The results showed that object memory was influenced by memory for relative but not for absolute location information. Furthermore, we demonstrated that relative space can be specified by a salient surrounding box or by distractor objects with no touching surfaces. Verbal memory was not influenced by any type of spatial information. Taken together, these results indicate that memory for "where" influences memory for "what." We propose that there is an asymmetry in memory according to which object memory always contains location information.

One of the most influential findings in cognitive neuroscience is that there is a division of labor in visual processing into what are colloquially known as "what" and "where" processing. These types of information are carried by two different neural processing streams, one ventral, the other dorsal. This anatomical division is thought to govern the neural workings of visual perception: The ventral stream is used for perceiving the identity of items, and the dorsal stream for perceiving where items are located in space (Ungerleider & Mishkin, 1982; see also Milner & Goodale, 1995).

The ventral stream begins in primary visual cortex (V1) and travels through inferior temporal cortex. The dorsal stream also begins in V1, but it travels a more superior route through posterior parietal cortex (Livingstone & Hubel, 1988; Mishkin, Ungerleider, & Macko, 1983). It has been theorized that this anatomical and cognitive division continues to some degree into the frontal lobes and influences how the frontal lobes process working memory information. Single-unit recordings in the frontal lobes of nonhuman primates suggest that areas in dorsal prefrontal cortex respond selectively to the maintenance of location information in visual working memory (VWM). Neurons that are more lateral and ventral, on the inferior prefrontal convexity, respond selectively to the maintenance of object or featural information (Chafee & Goldman-Rakic, 1998; Funahashi, Bruce, & Goldman-Rakic, 1989, 1990; O'Scalaidhe, Wilson, & Goldman-Rakic, 1999; Wilson, O'Scalaidhe, & Goldman-Rakic, 1993).

These findings have been the impetus for numerous neuroimaging studies of working memory for location versus identity or featural qualities of items. Although this field has been popular for research in neuroimaging, the data are highly inconsistent. Many studies report a dorsal-ventral segregation of working memory processes (Baker, Frith, Frackowiak, & Dolan, 1996; Courtney, Ungerleider, Keil, & Haxby, 1996; Haxby, Petit, Ungerleider, & Courtney, 2000; Smith, Jonides, Koeppe, Awh, Schumacher, & Minoshima, 1995), but a significant number of experiments do not find such segregation (D'Esposito, Ballard, Zarahn, & Aguirre, 2000; Nystrom et al., 2000; Postle, Stern, Rosen, & Corkin, 2000).

Researchers have offered a number of explanations for the inconsistencies found in neuroimaging experiments. Some (e.g., Nystrom et al., 2000; Postle & D'Esposito, 2000) have suggested that some neural regions in prefrontal cortex are not organized by stimulus modality but rather by type of processing (i.e., by processing rather than by the nature of the information being processed; Stern et al., 2000). Others have argued that objects may be a "representational middle ground," (e.g., Marshuetz & Bates, 2004) containing both spatial and object properties, or that the object and spatial memory results may be clouded because sometimes objects or spatial locations are verbalizable (e.g., Fletcher & Henson, 2001). We argue that there may be yet another explanation for the ambiguity seen in the neuroimaging literature: Information about "what" inherently carries information about "where." In other words, spatial information may be incidentally encoded along with object information at the be-

Correspondence concerning this article should be addressed to I. R. Olson, University of Pennsylvania, Center for Cognitive Neuroscience, 3720 Walnut Street, Philadelphia, PA 19104-6241 (e-mail: iolson@psych.upenn.edu).

havioral level. In tasks lacking such behavioral divisions, then, one might also expect to see a lack of neural division.

If a consistent behavioral distinction between object and spatial working memory is observed, this would lend credibility to the hypothesis that the human brain has an anatomical distinction as well. What is the existing evidence for and against behavioral divisions between object and spatial working memory? Logie (1995) found that recall of a sequence of locations, but not of simultaneously presented colors, was disrupted when participants were asked to make unseen arm movements during the retention interval. Conversely, color memory, but not location memory, was poorer when participants were shown irrelevant pictures during the retention interval. From these findings, he concluded that spatial and object memory are separable.

Tresch, Sinnamon, and Seamon (1993) used a similar task in which participants were required to remember either one shape or one location. During the retention interval one of three things could happen: (1) nothing; (2) a color discrimination task in which participants were required to judge whether a color patch was more red or more blue; or (3) a motion discrimination task in which participants were required to find a stationary item among many moving items. After the retention interval, participants were required to identify either the location or the shape of the item being held in memory. The results of two experiments showed (1) that a motion discrimination task interfered with spatial but not object memory and (2) that color discrimination interfered with object but not spatial memory. Taken with the results of Logie (1995) and others (Hecker & Mapperson, 1997), these results provide convincing evidence that at least under some conditions, location memory and object memory are separate.

More recently, Jiang, Olson, and Chun (2000) showed that VWM for the location of items was not influenced by the color or shape of items. To show this, they tested VWM for location but changed one unattended feature of the memory image, such as color or shape. For instance, if the memory image had blue squares placed in various locations, on the probe image the squares would be changed to green. The results showed that color and shape changes were easily ignored and caused no memory interference.

However, like the neuroimaging literature, the behavioral literature is not entirely consistent. For instance, Downing (2000) required participants to hold a centrally located sample object in VWM. After 1.5 sec, two objects, one matching the sample object and the other novel, were presented simultaneously at right and left locations. These stimuli were task-irrelevant. After 40 msec, a target shape was presented at either the left or the right location, and the participant was required to make a speeded response. Responses were faster when the target shape appeared in the same location as the object matching the memorized object. This result showed that the location of the memorized object was incidentally encoded into memory and was used to guide attention to a location.

In addition, Jiang et al. (2000) found that incidental encoding of "where" influences VWM of "what." Participants were required to remember briefly presented colors or shapes over a short retention interval; memory for location was not required. When there was a mismatch between the locations of the memory and probe images, VWM for color or shape was impaired. Other experiments showed that this influence of "where" on "what" was based on the relative location of all items, not on absolute location. In the same set of experiments, Jiang et al. (2000) instructed participants to remember the color of randomly placed squares. Three conditions were tested by manipulating locations of the colored squares on the probe image. The results showed that performance remained unchanged from baseline when the configuration of colored squares expanded outward, so that absolute locations changed but relative locations were unchanged. Performance was harmed when the initial configuration of color squares changed to a new configuration (e.g., both absolute and relative locations changed). These results suggest that VWM for objects also has linked information about the relative location of objects.

How can the results of Hecker and Mapperson (1997), Logie (1995), and Tresch et al. (1993), suggestive of a division between object and spatial memory, be reconciled with those of Jiang et al. (2000)? One possibility lies in the displays used by Jiang et al.(2000). In multielement memory displays, any one of the items might be "changed" on the probe display, so participants need to remember all items as efficiently as possible. Participants may not have had time to fixate each object, but instead looked to the center of the display and mentally formed a configuration to sustain their memory for individual items. Thus, the spatial configurations employed by Jiang et al. (2000) may have been strategically encoded when participants were asked to remember color or shape, because there were many items to encode. In other words, perhaps spatial and object working memory are strictly divided, but that participants can choose to chunk a set of objects in a configuration in order to aid remembering.

To reconcile the divergent findings, in this article we further test the hypothesis that location information is linked to object information in VWM by testing memory for a single item, as did Tresch et al. (1993). We performed this experiment to rule out the hypothesis that participants in studies using multielement displays (e.g., Jiang et al., 2000) remember spatial location strategically when asked to remember color or shape because they have many items to encode. If memory for a single object is disrupted by location change, this result would help rule out the strategic hypothesis of spatial encoding and suggest, instead, that location memory is tightly linked to object memory.

In addition, we were interested in further exploring what type of location information is remembered when encoding object attributes. Jiang et al. (2000) provided some data relating to this issue. They found that memory for object attributes in a multielement display is differentially affected by absolute versus relative location information. These types of location changes are analogous to changes in different types of spatial reference frames. A spatial reference frame is a relational system consisting of located objects, reference objects, and the spatial relations that lie between the different object types (Shelton & McNamara, 2001). *Absolute* location change is defined as any change in position so that objects differ from the starting position. *Relative* location change is defined as a position change that affects the position of the target item relative to a background context or reference frame. Relative location changes are analogous to position changes relative to some environmental reference frame.

We tested the influence of location on object memory by instructing participants to remember a single face, the memory item. After a short delay, the probe image was presented. It contained either the same face or a different face, to which the participant made either a "same" or "different" response. In addition, location of the face was manipulated in two separate conditions: In the *local change condition*, the memory item changed position relative to a surrounding reference frame. This condition had both relative and absolute changes in location. In the *global change condition*, both the memory item and the reference frame changed positions, so that the memory item retained its position relative to the frame. This condition had only absolute location change. These conditions were compared to a *no-change* condition that had no location change of any sort (Figure 1). In all cases, participants were instructed to ignore location changes. If absolute location is incidentally encoded in conjunction with objects, then response times (RTs) should be slower in the local and global change conditions as compared to the no-change conditions. If relative location is incidentally encoded in conjunction with objects, then RTs should be slower in the local change condition only. If location is not encoded when participants are required to remember objects, RTs should be similar in all conditions.

EXPERIMENT 1

Experiment 1 tested VWM for a single item. Given that single-item memory tasks are easy, we were concerned about ceiling effects. Our pilot work showed memory for faces to be more difficult to retain than memory for colors or simple objects; thus, for Experiment 1 we used

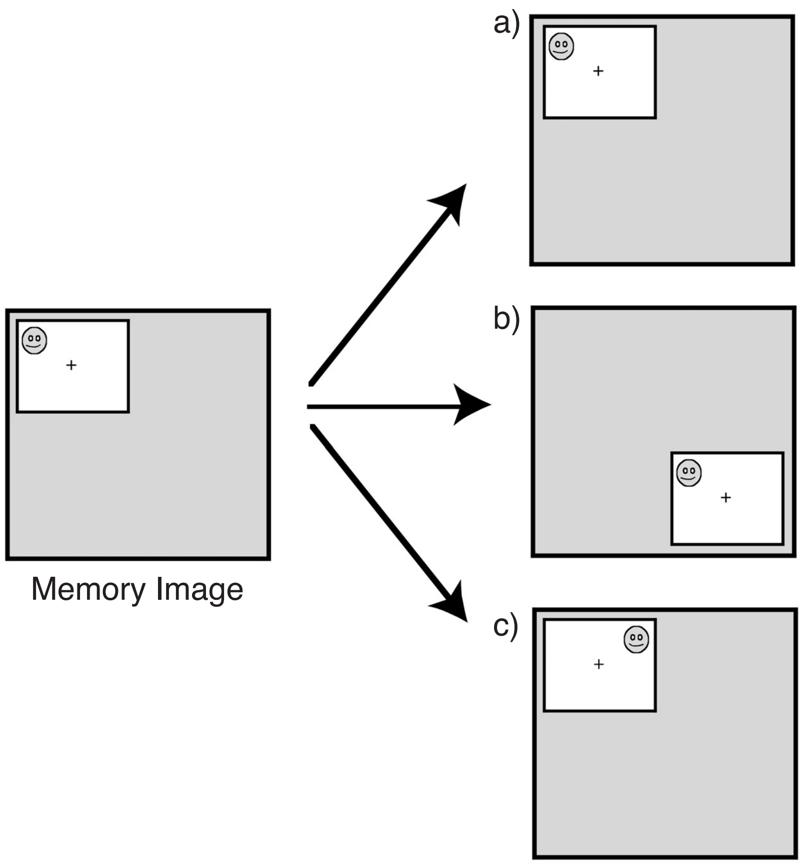

Figure 1. An illustration of the stimuli used in Experiment 1. The memory image was similar in all conditions and was followed by a 1,600-msec interstimulus interval and one of three types of probe image: (a) no-change condition; (b) global change condition; or (c) local change condition. Real rather than cartoon faces were used in the actual experiment.

faces as stimuli. Because participants only had to remember a single item and we expected accuracy to be high, RT was used as the dependent measure.

Method

Participants. Seventeen naive students and staff from Yale University participated in each experiment for payment. All had normal color vision and normal or corrected-to-normal visual acuity.

Stimuli. The face stimuli subtended 2º × 2.5º of visual angle. Faces had no hair or salient distinguishing features, such as large moles, and were grayscale. All faces had a neutral expression and were similarly posed. Faces appeared in one of four positions relative to a central fixation cross within a white box outlined in black. The box subtended 9º × 9º of visual angle. The box could appear in any of four locations on the uniformly gray background (subtending 18º × 18º of visual angle). Twelve faces were used.

Design. VWM of faces was tested in a probe change detection paradigm (Jiang, Chun, & Olson, 2004; Jiang et al., 2000; Olson & Jiang, 2002). Each trial was composed of a memory display and a probe display separated by a brief blank interval. The memory display contained one face presented at random locations within the white box and on the background. The probe display contained either the same face or a different face. The observers' task was to detect whether the face was the same or different, irrespective of position.

There were three types of probe displays: *no-change*, *local change*, and *global change*. In all conditions, the identity of the face changed on one half of all trials. In the no-change condition, the spatial location of the face and white box remained invariant. In the local change condition, the face moved to a new position within the white box on the probe image, although the position of the white box did not change. The horizontal and vertical movements were 5º, and the diagonal movements were about 7º. Thus, both the absolute and relative locations of the face changed in this condition. In the global change condition, the face and the white box moved to new positions on the screen, but the face remained in the same location relative to the white box. Thus, only the absolute location of the face changed.

Procedure. Participants initiated each trial by a barpress. Timing was set to the refresh rate of the computer. A red fixation point (0.2º × 0.2º) was presented for 507 msec, followed by a brief pause and the memory display of 267 msec. Participants were free to move their eyes. After a blank retention interval of 1,600 msec, the probe display was presented until a response key was pressed. Observers were instructed to memorize the face on the memory display and to detect whether the face on the probe display was the same as the previous face. They were told to ignore location, because the face would move about on the screen, and to answer as quickly as possible. Responses were entered via keypress: "Z" for same face and "X" for different face. A response prompter, located in the lower left corner of the screen, reminded participants of the proper keypress after each trial. Visual feedback concerning the accuracy of response was provided after each trial.

Each observer completed 10 practice and 120 experimental trials (i.e., 3 conditions × 2 probe types × 20 cases). Trials from all conditions were randomly intermixed and presented in a different random order for each participant.

Equipment. All observers were tested individually in a room with normal interior lighting. All experiments were carried out on a Macintosh computer with a 19-in. screen using MacProbe software (Hunt, 1994). The unrestricted viewing distance was approximately 57 cm, at which 1 cm on screen corresponds to 1º visual angle.

Results and Discussion

In all experiments, RTs exceeding 4,000 msec were discarded. Accuracy was high for all participants (>91%) in all conditions, and there was no difference in accuracy between conditions [$F(2,32) = 1.84, p = .18$].

Incorrect trials were discarded, and a repeated measures analysis of variance (ANOVA) tested the effects of condition on RT. Mean RT as a function of condition is plotted in Figure 2. The main effect of condition was significant [$F(2,32) = 20.01, p < .0001$]. Planned comparisons showed that there was no difference between the global change and no-change conditions, although there was a small trend in the direction of global change being slower [$t(16) = 1.70, p > .10$]. This result suggests that changes in only absolute location have minor effects on VWM. However, RTs in the local change condition were slower than those in the global change and no-change conditions [$t(16) = 4.43, p < .0001$; $t(16) = 6.13, p < .0001$]. Because absolute location changed in both the global and local change conditions, these findings suggest that VWM for faces is dependent on memory for relative location but is only negligibly dependent on memory for absolute location. In sum, these findings suggest that VWM encodes and stores the spatial location of items relative to a salient reference frame. Performance was harmed when an item changed position relative to the white box, even though it was not necessary to encode spatial position for the task.

EXPERIMENT 2

Faces are a peculiar stimulus class and have many unique properties, including the need to encode them configurally (Farah, Wilson, Drain, & Tanaka, 1998). Experiment 2 tested VWM for a single novel object in order to assess whether the effects observed in Experiment 1 generalize to other stimulus classes.

Method

Participants. Fifteen naive students or staff from Yale University participated for payment.

Stimuli and Procedure. Most aspects of the stimuli and procedure are similar to those reported for Experiment 1, with the fol-

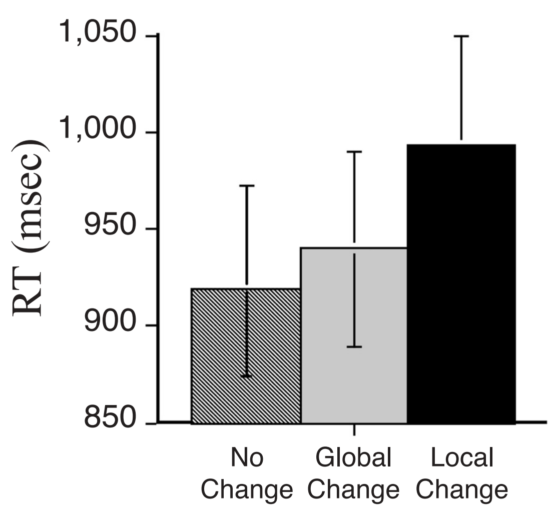

Figure 2. Results from Experiment 1. Response time (RT) as a function of condition.

lowing exceptions: One novel object appeared in one of four positions relative to a central fixation cross within a white box outlined in black on a gray background. Novel objects were similar to those used by Jiang et al. (2000). Novel objects were black and approximately the same size as the faces, subtending 2º × 2.5º of visual angle. Timing was the same as reported for Experiment 1. Each observer completed 10 practice and 90 experimental trials (i.e., 3 conditions × 2 probe types × 15 cases).

Results and Discussion

There was no difference in accuracy between conditions ($F < 1$, n.s.). Incorrect trials were discarded, and a repeated measures ANOVA tested the effects of condition on RT (Figure 3). The main effect of condition was significant [$F(2,28) = 3.54, p < .043$]. Planned comparisons showed that there was no difference between the global change and no-change conditions ($t < 1$), suggesting that changes in absolute location had little or no effect on VWM. However, RTs in the local change condition were significantly slower than those in the no-change condition [$t(14) = 2.63, p < .014$] and were also slower than those in the global change condition, although the difference fell short of significance [$t(14) = 1.66, p = .11$].

These results replicate the findings reported in Experiment 1, although the difference between the local and global change conditions is not as robust. Nevertheless, the results point to the same conclusion: VWM for objects includes incidentally acquired location information. Location relative to a salient reference frame is more critically tied to objects than is absolute location information.

EXPERIMENT 3

Experiments 1 and 2 showed that information about relative location was linked to object working memory. Recall that for our purposes, relative location was defined

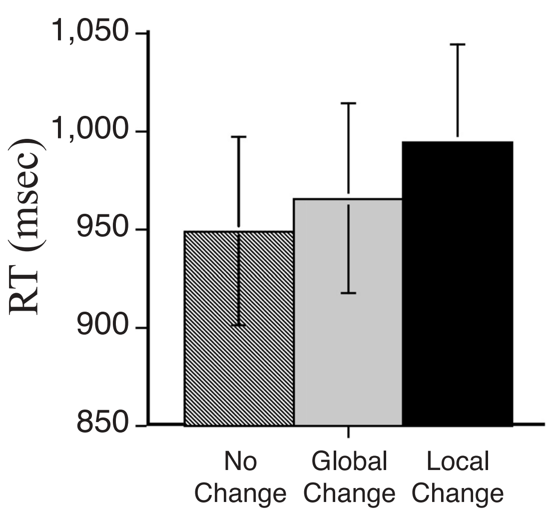

Figure 3. Results from Experiment 2, in which participants were required to remember one novel shape. Response time (RT) as a function of condition.

as location relative to a salient surrounding but task-irrelevant enclosing frame. However, one possibility is that an enclosing reference frame created, in effect, a new "object." Furthermore, it is important to characterize the boundary conditions for a frame of reference in order to begin to get a theoretical handle on what exactly a "reference frame" is. Experiments 3, 4, and 5 were designed to address the question of what constitutes a good reference frame for VWM and to rule out the possibility that it is necessary to form a new object via a salient enclosing boundary in order to observe our effects. To do this, we imposed different artificial reference frames on the display. Experiment 3 tested a minimal reference frame, a small fixation cross. The display did not have the salient white box provided in Experiments 1 and 2 and only provided a small fixation cross for spatial reference. Experiment 4 examined whether the size of the reference frame relative to the target object was critical. To test this, participants were shown a display that was similar to that used in Experiment 3, except that the central cross was very large relative to the stimuli. Experiment 5 examined whether other objects can serve as a reference frame.

Method

Participants. Thirteen naive students from Yale University participated for payment.

Stimuli and Procedure. Most aspects of the stimuli and procedure were similar to those reported in Experiment 1, with the following exceptions: Stimuli appeared on a plain white background (RGB 255). Faces appeared within one of four cells in an invisible box, with a small black cross marking the center of the invisible box. The small cross was the same size as the fixation cross used in Experiments 1 and 2. Timing was the same as reported for Experiment 1. Each observer completed 10 practice and 90 experimental trials (i.e., 3 conditions × 2 probe types × 15 cases).

Results and Discussion

A repeated measures ANOVA tested the effects of condition on accuracy. Accuracy was not affected by condition, although the effect approached significance [$F(2,24) = 2.73, p = .086$]. We investigated this result further; t tests showed that it was driven by an accuracy difference of the local condition as compared with the no-change condition, [$M = .93$ vs. $.88$; $t(12) = 2.23, p < .036$]. The comparison between local and global change conditions approached significance [$M = .93$ vs. $.89$; $t(12) = 1.73, p = .097$]. Thus, there is some indication that accuracy was higher in the local change condition, but because the accuracy analyses were post hoc, this finding must be interpreted with caution.

Incorrect trials were discarded, and a repeated measures ANOVA tested the effects of condition on RT performance. RT was not affected by condition ($F < 1$, n.s.). Data were further analyzed using planned comparisons that found no differences between conditions (all ts < 1).

These results suggest that the effect reported in Experiment 1 is dependent on a salient reference frame that is proximal to the stimuli (i.e., the computer monitor was not sufficient to serve as a frame of reference). They also demonstrate that a small central stimulus that does not aid the viewer in defining the boundaries of relative space cannot define relative location (Experiment 3). Thus, it appears that reference frame usage is highly flexible, yet it relies on at least two properties—relatively close spatial proximity and size relative to the stimuli.

EXPERIMENT 4

The results of Experiment 3 suggest that a small cross does not provide a good reference frame for encoding relative location, and we have suggested at least two potential properties that a "good" reference frame must have—proximity and a similar size scale to that of the other stimuli. In Experiment 4, we tested the latter hypothesis using a large cross at the center of the display rather than a small fixation cross.

Method

Participants. Nineteen naive students from Yale University participated for payment or credit.

Stimuli and Procedure. Most aspects of the stimuli and procedure were similar to those reported for Experiment 3, with the following exceptions: Stimuli appeared on a plain white background (RGB 255). Faces appeared within one of four cells in an invisible box, with a large black fixation cross in the center of the invisible box. The large cross subtended 9º × 9º of visual angle. Timing was the same as reported for Experiment 1. Each observer completed 10 practice and 90 experimental trials (e.g., 3 conditions × 2 probe types × 15 cases).

Results and Discussion

A repeated measures ANOVA tested the effects of condition on accuracy. Accuracy was unaffected by the condition manipulation [$F(2,36) = 1.25, p > .25$]. Incorrect trials were discarded, and a repeated measures ANOVA tested the effects of condition on RT. The main effect of condition was significant [$F(2,36) = 3.83, p < .031$]. Data were further analyzed using planned comparisons that showed slower RTs in the local versus the global change condition [$t(18) = 2.18, p < .036$] and in the local change versus the no-change condition [$t(18) = 2.57, p < .015$]. There was no difference between the global change and no-change conditions ($t < 1$, n.s.). Data are shown in Figure 4.

These results suggest that a large fixation cross makes a good reference frame and that the relative size of the frame matters.

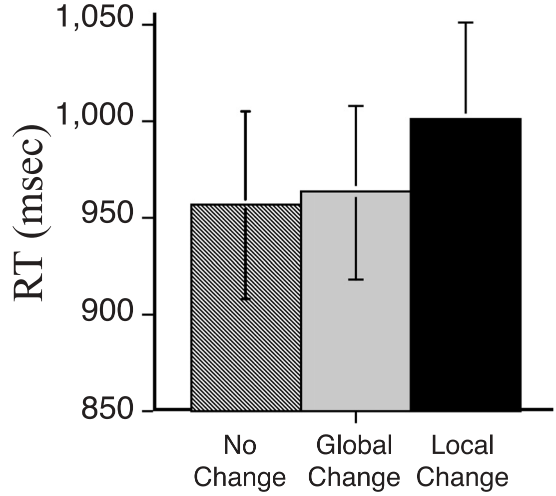

Figure 4. Results from Experiment 4. Participants were required to remember one face near a large fixation cross. Response time (RT) as a function of condition.

EXPERIMENT 5

Experiment 5 investigated whether distractor stimuli alone, without sharing any common spatial location or connected surface, could serve as a reference frame to encode the relative location of the face (see Figure 5). A cluster of small objects can form a small object under some circumstances, when items are arranged into a gestalt grouping. In order to avoid this sort of configural processing, we created displays with random placement of objects, devoid of any clear grouping cues. In Experiment 5, two novel shapes were presented in random locations near the face. Participants were told to ignore the shapes and to remember just the face.

Method

Participants. Eight naive students from Yale University participated for payment.

Stimuli. Many aspects of the stimuli and procedure were similar to those reported for Experiment 1. The background was white. Faces were drawn from the same set used in Experiments 1 and 2. Novel shapes were the same as those used in Experiment 2. Each item was located within one cell of an invisible 3 × 3 matrix, with the middle square excluded. This 3 × 3 matrix appeared in one quadrant of the computer screen.

Design. Face memory was tested in a change detection task. The memory image contained one face plus two novel shapes, which all appeared simultaneously. The probe image also contained one face and two novel shapes. As before, there were three types of probe displays: no-change, local change, and global change. In the local change condition, the location of the face changed relative to the novel shapes. In the global change condition, the entire configuration of face + novel shapes moved to a new location. In all conditions, the identity of the face changed on one half of all trials.

Procedure. Participants were instructed to remember the face and to ignore the shapes. They were also instructed to ignore locations of items because location was irrelevant to the task, and to respond as quickly as possible while still trying to be accurate. The memory image was shown for 534 msec, followed by an interstimulus interval of 1,600 msec and a probe image until response. Viewing time for the memory image was increased from that used in Experiments 1–4 because more stimuli were present on the memory image.

Each observer completed 10 practice and 90 experimental trials (i.e., 3 conditions × 2 probe types × 15 cases).

Results and Discussion

A repeated measures ANOVA tested the effects of condition on accuracy. Accuracy was unaffected by the condition manipulation ($F < 1$, n.s.). Incorrect trials and RTs greater than 4,000 msec were discarded. A repeated measures ANOVA tested the effects of condition on RT. The main effect of condition was significant [$F(2,14) =$

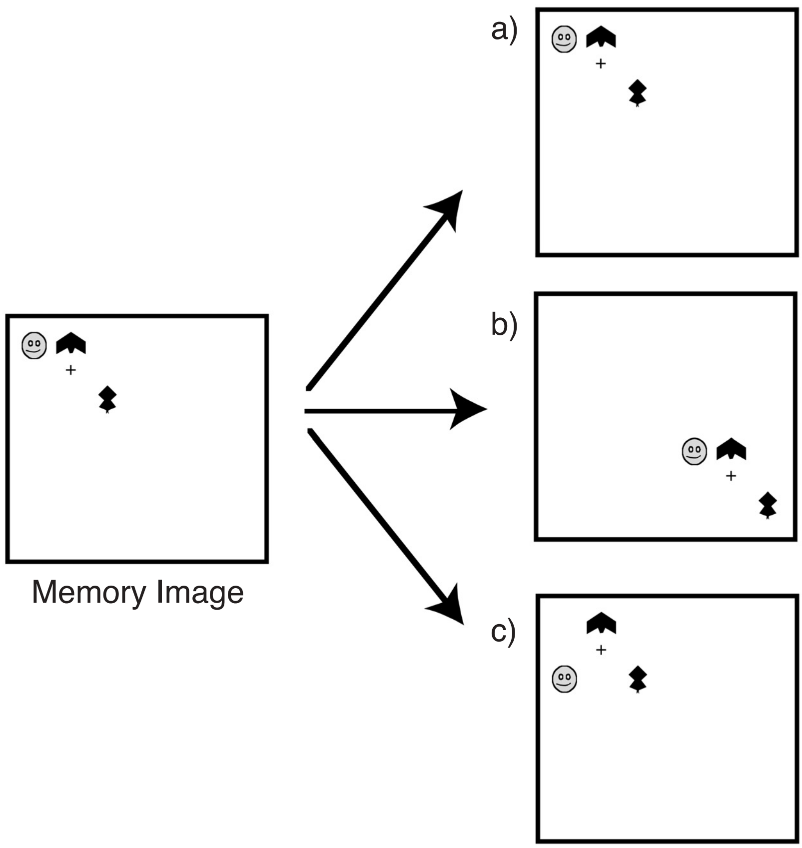

Figure 5. An illustration of the stimuli used in Experiment 5. The memory image was similar in all conditions and was followed by a 1,600-msec interstimulus interval and one of three types of probe image: (a) no-change condition; (b) global change condition; or (c) local change condition. Real rather than cartoon faces were used in the actual experiment.

5.42, $p < .018$]. Data were further analyzed using planned comparisons that showed slower RTs in the local change versus the global change condition [$t(7) = 2.94$, $p < .011$] and in the local change versus the no-change condition [$t(7) = 2.75$, $p < .016$]. There was no difference between the global change and no-change conditions ($t < 1$, n.s.). Data are shown in Figure 6.

These results suggest that objects that do not share spatial overlap and that are incidental to the memory task can serve as a reference frame for the to-be-remembered object. These results are somewhat contradictory of those reported in Experiment 4B by Jiang et al. (2000), who found that a combination of top-down and bottom-up cues allowed participants to ignore elements in the memory display and thereby to exclude them from the spatial reference frame that was encoded with the color information. This discrepancy is most likely due to the fact that the items that were ignored by participants in Jiang et al.'s study—12 white squares—could easily be grouped together into a separate configuration using the gestalt principle of similarity. Experiment 5 used two items with different shapes that were not easily grouped by most gestalt grouping principles (although grouping by proximity is arguable).

EXPERIMENT 6

The prior experiments showed that under a variety of conditions, VWM encodes and stores information about relative location. To understand the generality of this effect better, in Experiment 6 we tested whether verbal working memory encodes and stores information about relative location. Many researchers have suggested that VWM is separate from the verbal short-term store in the nature of the information it encodes (e.g., Baddeley, 1986; Logie, 1995). However, when verbal information is presented visually, it is possible that some visual information, such as spatial coding, is maintained in working memory. A strong version of the separation of the verbal and visual stores would predict that little or no visual information is maintained by verbal working memory. A weak version of the separate-store theory would predict that some visual information is preserved for visually presented verbal information.

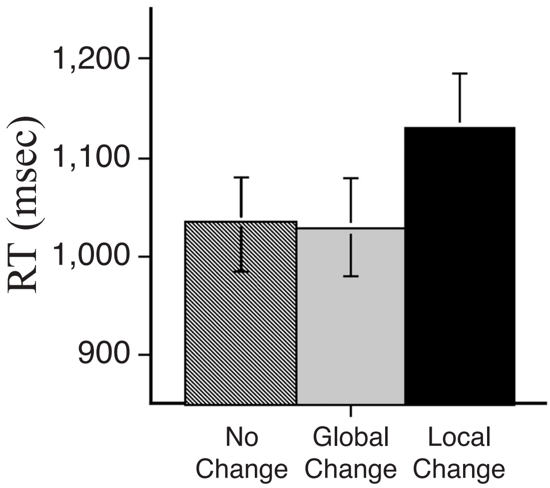

Figure 6. Results from Experiment 5. Response time (RT) as a function of condition.

Method

Participants. Twelve naive students from Yale University participated in this experiment for payment.

Materials. Most aspects of the stimuli and procedure were similar to those reported for Experiment 1, with the exception that three letters served as the memory stimuli instead of faces. It is well-known that the memory capacity for visual items is lower than that for verbal items (Luck & Vogel, 1997). Capacity for faces is around 2.5 items (Marshuetz & Olson, 2004), whereas that for letters is around 7 items. To equate difficulty levels, we used three letters rather than one. The memory image had three uppercase letters, randomly drawn from a set of nine consonant letters. The letters appeared in black Helvetica 22-point font. The probe image had three lowercase letters that were also in black Helvetica 22-point font. The three letters were grouped together and appeared in one of the four cells in a white box on a gray background. Participants were instructed to say the letters aloud as soon as they saw them, and to rehearse them aloud. These instructions ensured that a verbal, not a visual, strategy was used. As before, there were three types of probe displays: no-change, local change, and global change. Timing was the same as reported for Experiment 1. Each observer completed 10 practice and 90 experimental trials (i.e., 3 conditions $\times$ 2 probe types $\times$ 15 cases).

Results and Discussion

A repeated measures ANOVA found that accuracy was unaffected by the condition manipulation ($F < 1$, n.s.). Incorrect trials and RTs greater than 4,000 msec were discarded. A repeated measures ANOVA tested the effects of condition on RT. The main effect of condition was not significant [$F(2,22) = 1.37, p > .280$. Data were further analyzed using planned comparisons that showed no differences between conditions (all *p*s $> .14$). These results suggest that verbal working memory does not maintain representations of relative spatial location.

GENERAL DISCUSSION

Is object working memory separate from spatial working memory? This question has been debated by memory researchers (e.g., Goldman-Rakic, 2000; Logie, 1995; Miller, 2000). Although few behavioral studies have examined this question, those that exist have reported mixed findings: Some studies have reported dissociations between object and spatial VWM (Hecker & Mapperson, 1997; Tresch et al., 1993), whereas other studies have reported that object and spatial VWM are linked (Jiang et al., 2000). One common sense reason for a linkage between object and spatial VWM would be that "what" and "where" occur together in the real world. That is, objects are always at a location relative to the viewer and also relative to surrounding items in a scene. Because of this, the segregation of object and location in working memory tasks may reflect an artificial distinction (Rao, Rainer, & Miller, 1997). The spatial location of objects and object parts provides important information for our memory and perception: spatial location is needed to direct actions at objects and to assign meaning and significance. Our findings suggest that object memory and location memory are linked in that VWM for objects carries with it memory for the relative locations of the objects.

Experiment 1 showed that VWM for "what"—a face—incidentally carries with it information about a particular type of "where"—relative location. Relative location was manipulated by placing the face on a salient reference frame. Responses were slower when the face changed position relative to the reference frame. However, absolute location of the face was not maintained by VWM, as was demonstrated by the similar performance in the baseline condition and a condition in which the face and reference frame moved to an entirely new location but maintained the same internal spatial relationship. Experiment 2 generalized these results to nonface stimuli. The results of Experiment 2 were less robust than those found in Experiment 1, suggesting that faces may elicit this response more strongly. This may be the case because participants attend more to relative location information when the stimulus that must be remembered has salient configural information.

Experiments 3, 4, and 5 were designed to examine the question of what can serve as a reference frame for the to-be-remembered object. This question was addressed by testing face memory in the absence of a proximal, encompassing reference frame such as that used in Experiments 1 and 2. Experiment 3 provided only a small fixation cross for a reference frame, Experiment 4 provided a large cross, and Experiment 5 provided two distractor novel objects as references for relative spatial location. The results of Experiment 3 suggest that a small fixation cross is insufficient to serve as a space-defining reference. However, the results of Experiment 4 suggest that a larger cross can be a sufficient reference frame. In addition, the novel objects used in Experiment 5 were an excellent reference frame: RTs were slower when the face moved to a new position relative to the novel objects, but remained unchanged when the entire configuration of face + objects shifted to a new location. These experiments provide further evidence that under free viewing conditions, VWM for objects may not carry absolute but does carry relative location information. In the absence of a salient reference frame in Experiment 3, there was no difference between the small location changes of the local change condition and the larger location changes of the global change condition.

Experiment 6 showed that verbal working memory does not retain relative or absolute location information. When participants remembered three letters, there was no performance impairment when the letters moved either within the reference frame or in absolute terms. This finding suggests that location information may be selectively maintained by visual but not by verbal working memory.

Why didn't retinal or absolute location change affect memory performance? An explanation is offered by Dill and Fahle (1998). They used a visual memory task for simple stimuli and found that accuracy decreased with changes in retinal location. This effect was found only under certain conditions, and disappeared after rotation or contrast reversal of the to-be-remembered patterns, suggesting that positional specificity in memory depends on low levels of neural processing. Dill and Fahle suggested that tasks that rely on higher levels of processing will not show the effects of absolute location change because higher levels of the brain, such as inferior temporal (IT) cortex, have large receptive fields that are insensitive to retinal location. Thus, absolute or retinal location may not be encoded or remembered in memory tasks, such as ours, that require participants to remember complex stimuli that are presumably processed by regions in or around IT.

Our results conflict with some earlier studies that have reported a dissociation between object and spatial memory (Hecker & Mapperson, 1997; Tresch et al., 1993). We discuss some reasons for the contradictory findings in the following section. First, it is possible that Hecker and Mapperson and Tresch et al. did not see an effect of location interference on object VWM because this type of interference may only interfere with coding of absolute rather than relative location of items. Second, our task and dependent measure (e.g., RT) depart from those used previously to study the influence of location on object working memory. RT is considered to be a more sensitive measure of performance than accuracy. We suggest that the effects of relative location on object processing may be most apparent in VWM tasks that use an RT measure. Third, prior studies that reported a dissociation between object and spatial VWM required an explicit judgment or attention toward the spatial task (Hecker & Mapperson, 1997; Tresch et al., 1993), whereas studies that have reported a linkage between object and spatial VWM used incidental spatial manipulations (Jiang et al., 2000). Future studies should address how well the results outlined in our article generalize to other tasks and procedures.

VWM and the Brain

Location may be coded in different ways for different purposes in different neural subsystems. A viewer-centered or egocentric reference is particularly important when an immediate action, such as reaching or grasping, must be performed on a target (Milner & Goodale, 1995). A frame of reference based on other objects is useful when multiple objects are presented simultaneously. These objects form a spatial layout and can serve as anchors for one another (Jiang et al., 2000). When the spatial relationships among objects are part of the scene meaning, "where" may actually be part of "what." Such configural processing has been convincingly shown for facial features (Farah et al., 1998), and there is a great deal of evidence suggesting that faces are processed in the ventral visual stream (e.g., Kanwisher & Moscovitch, 2000). There is also evidence that other stimuli that are defined by the spatial relationships between features—scenes—are processed in the anterior extent of the ventral visual processing stream (Epstein & Kanwisher, 1998). We propose that when input to the ventral stream consists of a multielement visual pattern, a coherent representation of the entire pattern may be encoded for working memory. This encoding may occur because the spatial relationships are meaningful; alternatively, even if the relationships are not meaningful, attention may pick out one item for selective encoding, but excess attentional resources may nonetheless process some of the irrelevant distractors and their relative positions.

Our findings cannot answer questions concerning the functional divisions of the prefrontal cortex as it concerns working memory. However, these findings do suggest that a simple division of working memory areas along the lines of "what" and "where" is probably too simplistic, because the mnemonic encoding of location can be further divided into a variety of different coordinate-based systems.

Furthermore, these results suggest a reason for the discrepant findings seen in the neuroimaging literature: "What" and "where" may not appear to be cleanly divided at the neural level because they are not clearly divided at the behavioral level. It may be impossible to encode and remember *only* object attributes in object working memory tasks because location is incidentally encoded in most cases. The behavioral phenomena reported here are likely the results of rich interconnectivity between the "what" and "where" processing streams: Although numerous neurophysiological reports have convincingly demonstrated that object information and spatial information travel through a set of segregated pathways (e.g., Livingstone & Hubel, 1988; Mishkin et al., 1983), especially in posterior neural regions, the information is bound together via numerous ascending and descending connections between the two pathways (e.g., Barbas & Pandya, 1989; Bullier, Schall, & Morel, 1996; Pandya & Barnes, 1987). Our findings result from such rich interconnectivity.

REFERENCES

Baddeley, A. [D.] (1986). *Working memory*. Oxford: Oxford University Press.

Baker, S. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1996). Active representations of shape and spatial location in man. *Cerebral Cortex*, **6**, 612-619.

Barbas, H., & Pandya, D. N. (1989). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *Journal of Comparative Neurology*, **286**, 211-228.

Bullier, J., Schall, J. D., & Morel, A. (1996). Functional streams in occipito-frontal connections in the monkey. *Behavioural Brain Research*, **76**, 89-97.

Chafee, M. V., & Goldman-Rakic, P. S. (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *Journal of Neurophysiology*, **79**, 2919-2940.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex*, **6**, 39-49.

D'Esposito, M., Ballard, D., Zarahn, E., & Aguirre, G. K. (2000). The role of prefrontal cortex in sensory memory and motor preparation: An event-related fMRI study. *NeuroImage*, **11**, 400-408.

Dill, M., & Fahle, M. (1998). Limited translation invariance of human visual pattern recognition. *Perception & Psychophysics*, **60**, 65-81.

Downing, P. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, **11**, 467-473.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, **392**, 598-601.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, **105**, 482-498.

Fletcher, P. C., & Henson, R. N. (2001). Frontal lobes and human memory: Insights from functional neuroimaging. *Brain*, **124**, 849-881.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, **61**, 331-349.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1990). Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. *Journal of Neurophysiology*, **63**, 814-831.

Goldman-Rakic, P. (2000). Localization of function all over again. *NeuroImage*, **11**, 451-457.

Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage*, **11**, 380-391.

Hecker, R., & Mapperson, B. (1997). Dissociation of visual and spatial processing in working memory. *Neuropsychologia*, **35**, 599-603.

Hunt, S. M. J. (1994). MacProbe: A Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods, Instruments, & Computers*, **26**, 345-351.

Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, **66**, 446-453.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 683-702.

Kanwisher, N., & Moscovitch, M. (2000). The cognitive neuroscience of face processing: An introduction. *Cognitive Neuropsychology*, **17**, 1-11.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, **240**, 740-749.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hillsdale, NJ: Erlbaum.

Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.

Marshuetz, C., & Bates, J. E. (2004). Functional neuroimaging and the prefrontal cortex: Organization by stimulus domain? In S. Otani (Ed.), *Prefrontal cortex: From synaptic plasticity to cognition* (pp. 289-314). Boston: Kluwer.

Marshuetz, C., & Olson, I. R. (2004). *Face memory: Not so special after all*. Manuscript submitted for publication.

Miller, E. (2000). The prefrontal cortex: No simple matter. *NeuroImage*, **11**, 447-450.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. New York: Oxford University Press.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, **6**, 414-417.

Nystrom, L., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage*, **11**, 424-446.

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception & Psychophysics*, **64**, 1055-1067.

Ó Scalaidhe, S. P., Wilson, F. A. W., & Goldman-Rakic, P. S. (1999). Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: Evidence for intrinsic specialization of neuronal coding. *Cerebral Cortex*, **9**, 459-475.

Pandya, D. N., & Barnes, C. L. (1987). Architecture and connections of the frontal lobe. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 41-72). New York: IRBN.

Postle, B. R., & D'Esposito, M. (2000). Evaluating models of the topographical organization of working memory function in frontal cortex with event-related fMRI. *Psychobiology*, **28**, 132-145.

Postle, B. [R.], Stern, C. E., Rosen, B. R., & Corkin, S. (2000). An fMRI investigation of cortical contributions to spatial and nonspatial visual working memory. *NeuroImage*, **11**, 409-423.

Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, **276**, 821-824.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, **43**, 274-310.

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial versus object working memory: PET investigations. *Journal of Cognitive Neuroscience*, **7**, 337-356.

Stern, C. E., Owen, A. M., Tracey, I., Look, R. B., Rosen, B. R., & Petrides, M. (2000). Activity in ventrolateral and mid-dorsolateral prefrontal cortex during nonspatial visual working memory processing: Evidence from functional magnetic resonance imaging. *NeuroImage*, **11**, 392-399.

Tresch, M. C., Sinnamon, H. M., & Seamon, J. G. (1993). Double dissociation of spatial and object visual memory: Evidence from selective interference in intact human subjects. *Neuropsychologia*, **31**, 211-219.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.

Wilson, F. A. W., O'Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, **260**, 1955-1958.

(Manuscript received October 17, 2003;
revision accepted for publication April 16, 2004.)