

# Location and binding in visual working memory

ANNE TREISMAN

*Princeton University, Princeton, New Jersey*

and

WEIWEI ZHANG

*University of Iowa, Iowa City, Iowa*

Visual working memory (VWM) was explored separately for features and for their binding. Features were better recognized when the probes retained the same binding as in the original display, but changing the locations had little effect overall. However, there were strong interactions of location with binding and with matching or new features, suggesting that, when objects are attended, features and locations are spontaneously integrated in VWM. Despite this, when the locations are changed, features can also be accessed with little decrement, perhaps from separate feature maps. Bindings, on the other hand, are more vulnerable to location changes, suggesting that locations play a central role in the early maintenance and retrieval of bound objects as well as in their initial encoding, at least when verbal coding is prevented. The results qualify past claims about the separation of locations and objects in VWM.

We live in a world of objects and events, and it is this world, not a world of separate features and locations, with which we must interact. However, there is evidence that, on the way to forming the representations that enter our consciousness and control our behavior, we do separately analyze features and locations in specialized feature maps (Treisman & Gelade, 1980). Thus we are faced with the problem of binding these features and locations together to form integrated objects and then maintaining the bound representations in visual working memory (VWM) over short intervals of time while we make decisions about what action to take. Feature integration theory (FIT) proposes that focused attention is directed to one filled location at a time to bind the features the location contains. The temporary episodic representations formed in this way, called *object files* (Kahneman, Treisman, & Gibbs, 1992), are addressed by their spatiotemporal coordinates, and they maintain the identity and perceptual continuity of objects in dynamic displays as they move around and change. It seems plausible that these representations are also the perceptual elements that are stored as object tokens in VWM (Irwin & Andrews, 1996).

Independent evidence that bound objects are stored in VWM comes from experiments by Luck and Vogel (1997). They showed patterns of 1–12 colored squares, followed after 900 msec by either the same display with one feature changed or the identical display. They found accurate recognition memory for up to four colors. Using bars that varied in size, orientation, color, and the

presence or absence of gaps, they found the same four-object limit, despite a fourfold increase in the number of features. They concluded that VWM holds up to four bound objects with at least four features each, a claim for which Irwin and Andrews (1996) had also previously found evidence in a series of experiments on transsaccadic memory.

Earlier findings, however, raise some questions about the idea that bound objects are the only units encoded in VWM. Treisman (1977) showed that, at least under time pressure, memory for features is much more accurate than memory for conjunctions. Participants had an error rate of more than 30% in successive same–different judgment tasks when both features matched but the binding was changed. This far exceeded the error rate obtained when new features were introduced. Stefurak and Boynton (1986) tested memory for color–shape conjunctions across longer intervals of 3 and 15 sec and found that recognition was at chance for the bindings when verbal labeling was prevented, although the features were quite well recognized. These results suggest that features may be retained in an unbound form as well as in bound object representations.

Wheeler and Treisman (2002) used the Luck and Vogel (1997) paradigm to compare memory for separate features and memory for their bindings, using a new color, location, or shape as foils in the feature tests and a new pairing of previously presented features as foils in the binding tests. Change detection was significantly worse for changes of binding than for new features when a whole display was presented at test, whether all of the items were potential targets or whether a single target item had been postcued. However, when single item probes were presented at test, memory was as good for the bindings as for the features. This implies that the bindings were encoded and stored just as well as the features, but that they were

---

This research was supported by Grant P50 MH62196 from the NIH Conte Center, Grant 1000274 from the Israeli Binational Science Foundation, and Grant 2R01 MH 058383-04A1 visual coding and the deployment of attention from the NIH. Correspondence should be addressed to A. Treisman, Department of Psychology, Princeton University, Princeton, NJ 08544 (e-mail: treisman@princeton.edu).

more vulnerable to interference from the whole display probes. Wheeler and Treisman suggested that VWM can draw information both from separate feature stores, each with its own independent capacity, and from integrated object files but that forming and maintaining the latter require additional attentional resources. When a whole display is used to probe recognition for the bindings, each probe object must be selectively attended, drawing attention away from items stored in VWM. When a single probe is presented, its features can be bound by default and attention remains with the memorized items. Up to four integrated object files can be retained as long as attention to them is maintained. We attributed the better performance on feature changes than on binding changes to additional memory traces of unbound features.

Another type of information that can also be retained in VWM is the set of spatial locations occupied by the presented objects. How are these stored? Are they integrated with the objects, or held in a separate specialized store? These questions have been tackled using a variety of experiments. Initially the main tests devised by Baddeley and his colleagues were based on selective interference in dual task paradigms (1992). If memory could be selectively disrupted for one kind of material by one kind of activity and for another by a different kind of activity, the two kinds of material were likely to be separate (see, e.g., Baddeley, 1992; Hecker & Mapperson, 1997; Logie & Marchetti, 1991; Magnussen, 2000; Tresch, Sinnamon, & Seaman, 1993; Zimmer, Speiser, & Seidler, 2003). More recently, neuroscientists using PET, fMRI, or ERPs have looked for selective activation for different kinds of material in different brain areas (e.g., Courtney, Ungerleider, Keil, & Haxby, 1996; Haxby, Petit, Ungerleider, & Courtney, 2000; Munk et al., 2002; Postle & D'Esposito, 1999; Smith et al., 1995). Evidence can also be sought in the selective deficits of patients with localized lesions (see, e.g., Carlesimo, Perri, Turriziani, Tomaiuolo, & Caltagirone, 2001; Hanley, Young, & Pearson, 1991; Nunn, Polkey, & Morris, 1998) or in the localized effects of TMS (see, e.g., Mottaghy, Gangitano, Sparing, Krause, & Pascual-Leone, 2002).

Although evidence that storage for locations is separate from storage for objects certainly exists, this conclusion is not universally accepted. For example, Postle and D'Esposito found memory for locations and memory for objects intermixed in frontal areas. Rizzuto, Mamelak, Sutherling, Fineman, and Andersen (2005) found that the ventral prefrontal area, which was assumed to be specialized in object working memory (Wilson, O'Scalaidhe, & Goldman-Rakic, 1993), is also active in spatial working memory. Prabhakaran, Narayanan, Zhao, and Gabrieli (2000) found frontal areas that specialized in coding integrated representations, and Haxby et al. (2000) emphasized that the segregation is "a matter of the degree of participation of different regions, not the discrete parcellation of different functions to different modules" (p. 145). Single unit activity in monkeys also suggests an initial separation of object and spatial storage, but a reintegration of the two in frontal cortex where "over half of the PF

[prefrontal] neurons with delay activity showed both what and where tuning" (Rao, Rainer, & Miller, 1997, p. 821).

A final source of evidence about the organization of VWM comes from the observation that recognition of one aspect of a stimulus can be disrupted by irrelevant variation in another aspect. The theory is that if recognition is disrupted by task-irrelevant changes, we can infer that the memory trace has integrated them with the relevant features. Changes in the physical stimuli can affect performance only if the changes are also present in the memory representation that mediates the response. The first display is no longer present when the comparison with the probe is made, thus the nature of the encoding is critical. The present experiments explore this method. It was previously used by Jiang, Olson, and Chun (2000) and by Woodman, Vogel, Luck, and Hollingworth (2007) to compare memory for the colors of objects when the locations or the orientations of the objects changed versus when they remained the same. Jiang et al. found some decrement when the locations changed. Woodman et al. (2007), however, found none and explained the discrepancy by suggesting that Jiang et al.'s participants were cued to the relevant item with a spatial location and therefore were encouraged to encode the conjunctions.

Woodman et al. (2007) collected further evidence for the separation of objects and locations in VWM; they found that presenting the colors sequentially in the same single location did not impair memory. They found no effects of either spatial shifts or scrambling, although memory for color was impaired when the orientations changed, showing that the method was sensitive enough to reveal interference when integration had occurred. Finally, Woodman et al. (2007) found better memory for three colors and three locations than for six colors or six locations, whereas there was no such benefit to dividing the items between colors and orientations. They suggested that features are bound in VWM, but that their locations are separately stored: "Representations in visual working memory are not tightly bound to either absolute or relative spatial locations."

The general, if not unanimous, consensus, then, seems to be that spatial and object information are held in distinct stores. None of the experiments mentioned, however, distinguished the effect of location changes on memory for features from memory for the features' bindings, although some tested both at once (see, e.g., Munk et al., 2002). It seemed worth investigating further how tightly locations are bound to features on the one hand and to objects on the other in VWM. Because locations seem to be crucial to establishing the bindings in the first place and because we have evidence that object files are addressed by their locations, our prediction was that location would be more important to VWM for bound objects than for features. However, it is of course possible that locations are used in binding features only during perception, and that, once established, the object representations become independent of location.

The present studies also explore some further questions. Under what conditions do we store bound objects and when do we also, or only, store their features? How

flexible are the strategies used in encoding visual information? Can features and locations be separately stored when required, or is binding automatic when attention is directed to an object? FIT (Treisman & Gelade, 1980) claimed that attention is necessary for accurate binding; however, it does not follow that attention is sufficient—i.e., that attention makes binding unavoidable. Evidence of interference when the bindings are changed would suggest that binding is automatic for attended objects, even when only the separate features of the objects are relevant to the task.

In all of our experiments, we presented three colored shapes, followed either by three colored shapes (whole display) or by just one (single probe). In Experiments 1 and 2, participants were asked to decide whether the features in the probe exactly matched those in the preceding display, disregarding any changes in binding or in location. The features in the probe display either exactly matched those in the previous display or differed from them in one feature, either color or shape; for example, a red T would become a blue T or a red O (when blue or O had not been present in the previous display). In Experiments 3, 4, and 5, the task was to decide whether the bindings were the same as in the previous display, again regardless of any location changes. In these binding tests, the probe display always contained the same features as the original display, and either the bindings also matched or one feature (shape or color) was exchanged between two of the objects, so that the bindings were changed; for example, a red T and a green X would become a green T and a red X.

## EXPERIMENT 1

### Visual Working Memory for Features

In the first two experiments, we explicitly tested memory only for features—that is, shapes and colors—but we looked for indirect evidence that bindings and locations were also stored. Luck and Vogel (1997) claimed that VWM contains only bound objects. Wheeler and Treisman (2002) suggested instead that VWM stores both separate features and integrated objects, with attentional resources required to produce and maintain the bindings. The new question is whether there is any voluntary control over what is stored. When only features are relevant to a task, can they be stored without forming integrated object files that are bound to their locations, or is binding automatic when attention is directed to the items in the display? The distinction between attention and intention is often blurred, since both are contrasted with automaticity. In the present experiments, we ensured that participants would attend to the stimuli by explicitly testing memory for the stimuli's features; the stimuli's locations and bindings, however, were supposedly irrelevant, making any memory for these properties unintentional. If features are bound in working memory only when the task requires it, there should be no effect of changing locations and/or bindings between presentation and test when only the features are relevant. On the other hand, if irrelevant features are integrated automatically with the relevant

ones, the changes might facilitate “different” trials and interfere with “same” trials relative to the baseline case with no irrelevant changes (as in Luck & Vogel, 1997, and most other VWM experiments). But because the irrelevant features are orthogonal to the task, they also introduce uncertainty about the objects, which may cancel any facilitation on “different” trials.

If binding is automatic when the objects are attended, then a further question is: What exactly is bound? Are feature–feature bindings retained only indirectly, through links to shared locations? If so, they should be lost when the locations are changed, making any switch of binding irrelevant. But if the features, once selected through their shared location, are bound directly to each other, the resulting object files might become independent of location. If so, a change of bindings might interfere with feature memory even when the locations are new. Finally, the features might be bound both to their locations and to each other, resulting in interference in feature recognition both from changed locations and from changed bindings.

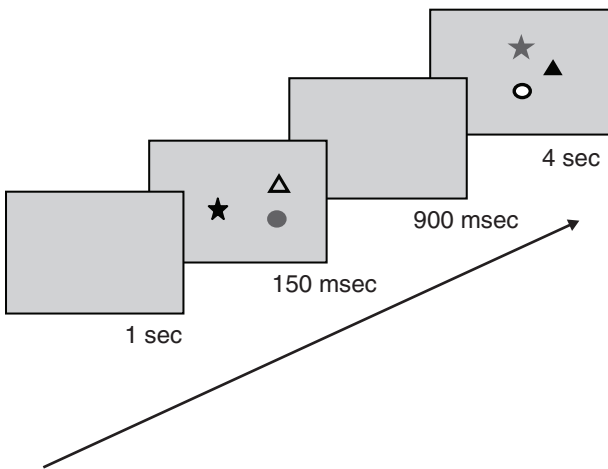
In displays of three bound items, participants were asked to remember only the features, and they were tested with displays that either preserved the original locations or used new locations that either preserved the feature–feature bindings or switched them. We tested colored shapes on one group of 8 participants and colored letters on another group of 8 participants. Because color often contributes to the identification of objects but is normally irrelevant for symbols, we thought that binding might be more automatic for shapes than for letters.

## Method

**Participants.** Sixteen Princeton undergraduates (13 women and 3 men) participated in the 1-h study for class credit. Eight were tested with letters and 8 with shapes. All were naive to the experimental paradigm and reported normal or corrected-to-normal visual acuity and color vision.

**Stimuli.** Stimuli were presented on a light gray background on the 15-in. screen of a Power Macintosh G3 computer, using MATLAB with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The viewing distance was 60 cm. There were six possible values on each dimension. The colors, chosen to maximize discriminability, were red, blue, yellow, green, brown, and violet. The shapes were a circle, with diameter subtending  $1^{\circ}40'$ ; a square,  $1^{\circ}30'$ ; an equilateral triangle,  $1^{\circ}60' \times 1^{\circ}40'$ ; a heart,  $1^{\circ}55'$ ; a star,  $1^{\circ}45' \times 1^{\circ}50'$ ; and a cross,  $1^{\circ}40'$ . The letters were S, X, T, C, Q, and Z, in 60-point, bold Arial font, subtending  $1^{\circ}20'$ . On each trial, three shapes (or three letters) were presented in randomly chosen locations in a  $3 \times 3$  invisible grid subtending  $16^{\circ}8' \times 12^{\circ}7'$ . Neither colors nor shapes (nor letters) were repeated within a display, either in the initial presentation or at test. Figure 1 shows an example of a trial sequence.

**Design and Procedure.** Two experiments, identical except that one used colored shapes and one used colored letters, were run with different participants. The following three variables were manipulated within each experiment. (1) Feature: On match trials, all of the probe features were also present in the original display. For the group that was tested with colored shapes, one randomly chosen color–shape combination had either its color or its shape replaced with a new color or a new shape; color and shape were tested equally often. The same procedure was used for the group that was tested with color–letter combinations. (2) Location: The probe display occupied either the same three locations as the original display (old locations) or three previously empty locations in the same grid (new



**Figure 1.** Schematic description of a match trial in the whole display condition of Experiment 1 with colored shape stimuli tested in new locations with changed bindings.

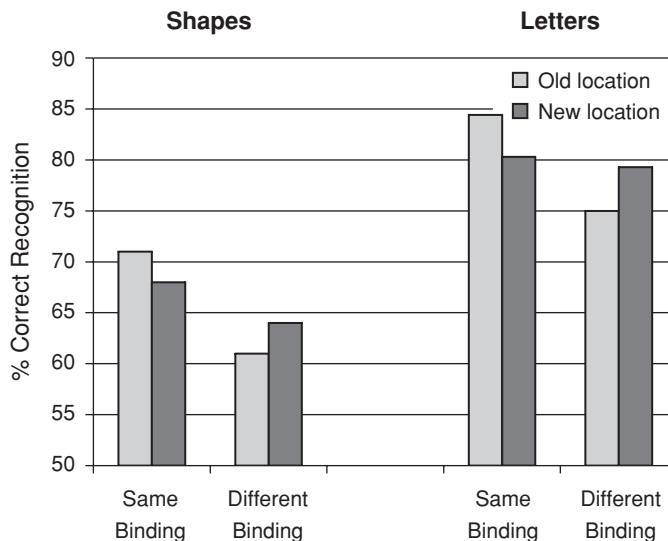
locations). The location factor was blocked, and the order was counterbalanced among participants. (3) Binding: In 50% of trials within each location condition, the test display kept the same binding of colors with shapes (or colors with letters), whereas in the remaining 50%, the binding was changed for all three objects in the display. Trials of each type were randomly mixed within blocks. When the locations were the same, half of the changed binding trials changed the locations of the colors, and half changed the locations of the shapes or letters.

Each participant completed 32 trials in each combination of binding and location match or mismatch. On each trial the sequence of events was as follows: A small black cross was presented at center screen for 1 sec, then an initial display of three items flashed for 150 msec, followed by a 900-msec blank interval, and finally a probe display that remained for 4 sec or until a response was made, whichever came first (Figure 1). Participants indicated by a keypress

whether there was a new feature in the probe display or whether the display contained all the original features. In the new locations condition, they were told that the test items would appear in new locations. They were also told that in all the trials the features might be recombined but that neither the bindings nor the locations were relevant to their task. If a new color, shape, or letter appeared, the correct response was "different." Otherwise they should answer "same." The ratio of match to new feature trials was 1:1, and this was balanced across all levels of the other two factors. Participants were instructed to aim for accuracy; if they made an error, they received a tone as feedback. Articulatory suppression was used to prevent verbal labeling of the stimuli: Participants repeated "Coca-Cola" throughout each trial. At the beginning of the experiment, they practiced 16 trials of each block type and were given a chance to ask questions. At the conclusion, they were given an opportunity to comment and ask questions and then were debriefed.

## Results and Discussion

Figure 2 shows the mean recognition accuracy in each condition. Letters yielded much better performance than shapes: 80% correct versus 66% [ $F(1,14) = 15.73, p < .005$ ]. Letters may have been more discriminable or more familiar. However, the effects of the other variables were very similar for both shapes and letters. There was a significant main effect of changing the binding [ $M = 6\%$ ;  $F(1,14) = 20.27, p < .001$ ], suggesting that features are bound even when the task does not require it. Surprisingly, there was no overall effect of changing locations (0%). One might infer that location is not stored or not retrieved in VWM for features, but this would clearly be wrong, since there was a strong interaction between location and binding [ $F(1,14) = 17.87, p < .001$ ]. Repeating the locations improved participant performance on same binding trials but worsened it on different binding trials. Mean accuracy did not differ between match and new feature trials, suggesting no overall response bias. However, there were highly significant interactions of



**Figure 2.** Accuracy in feature recognition for colored shapes and colored letters in Experiment 1 as a function of preserved or changed bindings and locations.

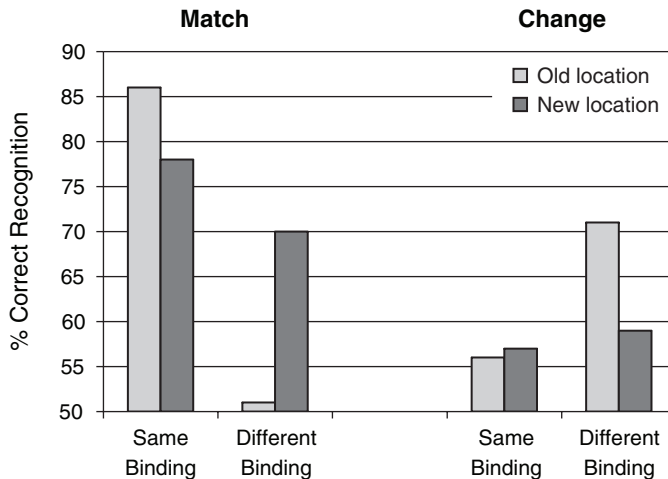


Figure 3. Accuracy on colored shapes on match and new feature trials, from Experiment 1 (with whole display probes).

match/new feature with binding [ $F(1,14) = 24.67, p < .001$ ] and with location [ $F(1,14) = 10.78, p < .01$ ] and a significant three-way interaction with binding and location [ $F(1,14) = 31.75, p < .001$ ]; see Figure 3. This interaction resulted from a bias to respond “same” except on trials with different bindings in old locations. The default seemed to be to assume no change unless there were feature–location mismatches between the original display and the probe. Finally, match trials showed an interesting asymmetry of interference between binding changes and location changes: Changed bindings in the old locations were much more likely to evoke false positive change detections than changed locations with the same binding (35% difference in accuracy compared with 8%).

Discussion follows the description of Experiment 2.

### EXPERIMENT 2 Feature Memory With Single Probe Versus Whole Display Probes

Wheeler and Treisman (2002) found that when the whole display was presented at test, memory was worse for the bindings than for the features themselves, but when a single probe was used, memory for the two was equally accurate. What should we predict when the bindings are irrelevant? Since bindings were better remembered with a single probe in the earlier study, we might expect more interference with feature memory from a changed binding when the single probe is presented. On the other hand, more items changed their bindings in the whole display test, which might have increased the total conflict.

#### Method

The displays and procedure were the same as in the shape condition of Experiment 1, except that at test a single probe item was presented, either in the same location or in a new, previously empty location. Again, the location factor was blocked and match versus

new feature trials were randomly mixed, as were same and different bindings in the match trials. On new feature trials in the single probe condition, the binding could neither match nor differ from the original because only one of the original features was present. Nine new Princeton undergraduate volunteers (7 women and 2 men) participated in the 1-h study for class credit.

#### Results

We analyzed the data from Experiment 2, which tested 9 participants on single probe recognition, together with the whole display data from the 8 participants in Experiment 1 who were tested on shapes rather than letters. The results are shown in Table 1. In the first ANOVA and in Figure 4, we collapsed across same and different bindings, since these cannot be distinguished in the new feature trials of the single probe condition.

Performance was better when participants were tested with the single probe (75% accuracy) than with the whole display (66%) [ $F(1,15) = 5.65, p < .05$ ]. The single probe condition has two possible advantages, both of which could play a part: One is the fact that only one item must be checked, reducing the decision load relative to the whole display. Wheeler and Treisman (2002) ruled out this account in their experiments by presenting the whole display but cuing just one item at test. The cue did not improve performance relative to the whole display condition (78% vs. 82%), although it narrowed down the relevant items to a single one. The other potential advantage of the single probe condition is that interference from multiple items in the probe is eliminated.

Again there was no main effect of changing the locations and no interaction of location with probe type. Location did interact with match versus new feature [ $F(1,15) = 10.82, p < .005$ ], with the new locations actually giving 4% better performance than the old locations when the features matched the original display. This probably resulted from the decrement that occurred when the features and locations were the same but the binding was changed.

**Table 1**  
**Mean Accuracy (in Percentages) in Shape Memory**  
**With Whole Display and Single Probe Tests**

Location	Match			Mean
	Same Binding	Different Binding	New	
Whole Display (Experiment 1)				
Same	86	51	63	66
Different	78	70	58	66
Mean	82	60	61	
Single Probe (Experiment 2)				
Same	78	72	73	74
Different	82	74	71	75
Mean	80	73	72	

In order to see the effects of changed versus preserved binding, we also did an ANOVA on both experiments on the match trials only, since changes of binding were impossible in the new feature trials. Here the effects of probe type were not significant, suggesting that the main decrement with the whole display took the form of change blindness rather than false change detections. The match trials showed a highly significant effect of binding: 81% for same binding versus 67% for different binding [ $F(1,15) = 28.92, p = .0001$ ]; a significant interaction of binding with probe type, with a decrement of 21% for whole display probes and 7% with the single probe [ $F(1,15) = 7.00, p < .05$ ] and of binding with location [ $F(1,15) = 8.69, p = .01$ ]; and a three-way interaction of binding with location and with probe type [ $F(1,15) = 12.37, p < .005$ ].

## Discussion

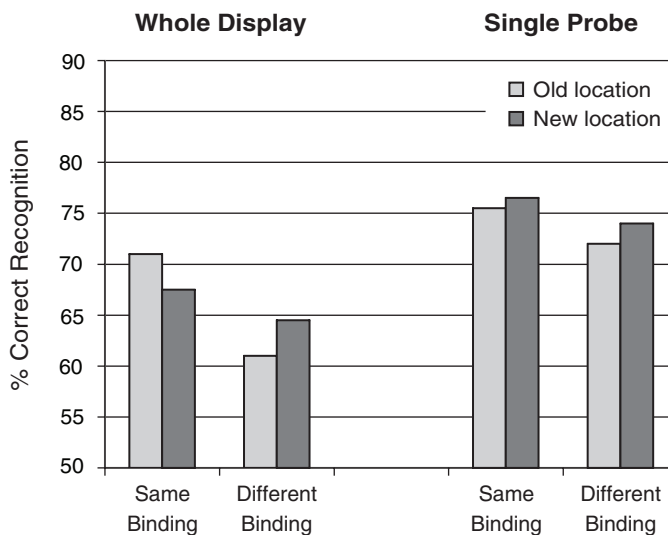
The results of Experiments 1 and 2 throw some light on the nature of representations in VWM. We argued that if the features from a single object are stored in an in-

tegrated format, participants should be more accurate in responding to test items that preserve the original binding than to test items that re-pair features from separate objects. Moreover, if memory for binding is mediated through links to locations, there should be differences between probes in old locations and those in new ones. On the other hand, if features can be stored separately, there should be no differences.

The main findings can be summarized as follows.

1. Memory for features showed clear effects of the original binding, even though this was irrelevant to the task. When the locations were new, or when the probe was a single item, changing the binding had a smaller but still significant effect. This effect was not present on letter trials, suggesting weaker feature–feature bindings with letters than with shapes. We can draw these conclusions: First, it seems that feature–feature bindings are mediated at least in part through location–feature bindings. However, some additional automatic binding of feature to feature seems to survive the loss or reduction of location information. The findings suggest that features are bound automatically when the object is attended. No intentional effort is required, and there may be little voluntary control.

2. Retrieval of features was as efficient overall when the locations were changed between encoding and test as when they were preserved. However, we should not conclude that the memory traces are independent of location. There were strong interactions of location with binding and with match versus new displays. The old locations had a clear advantage both on match trials when the binding also matched and on change trials when the binding also differed. When the binding was changed but the features matched, the new locations actually yielded better recognition, presumably because in the old location the change of binding was mistaken for a change of feature. These



**Figure 4.** Accuracy in feature recognition in Experiment 2 for shapes and colors with single probes or whole display probes.

results suggest that when a previously filled location is reactivated, retrieval of its previous contents (if they survive) is automatic. Within each location, participants compared the present features to the past ones. If either color or shape was new, they tended to respond “different,” even if the feature in question had been present in another location in the initial display. Thus feature–location bindings seem to be automatically encoded.

3. The location interactions mostly disappeared with single probes, suggesting that the locations were represented as spatial patterns—i.e., as relative rather than absolute locations. The effect of binding changes was also reduced on single probe match trials, most likely as a result of the loss of relative location information; however, as with the whole display probes, there was still some decrement when the features were rebound, even without the support of the feature–location bindings.

### EXPERIMENT 3 Memory for Binding: Effects of Location and Delay

How do the effects of preserved versus new locations in the probe differ when we test memory for the bindings rather than for the features? We used the same set of stimuli as in Experiment 2, and again we tested with both whole display and single probes to compare the effects of relative versus absolute locations. Each display again contained three colored shapes drawn from sets of six colors and six shapes. In the present experiment, the features always matched between the initial display and the probe. The task was to decide whether the bindings also matched. In change trials with whole display probes, two of the three stimuli in the initial display exchanged either their shapes or their colors, thus changing the binding of two of the colored shapes and leaving the third unchanged. With single probes, a shape and a color from the original display were recombined. The probes were presented either in the old locations (keeping the location of the color on half the binding change trials and of the shape on the other half), or in new, previously unoccupied locations.

We were interested also in possible changes over time in the representations stored in VWM. If location is initially important to the binding process, it might become less important as the delay between the initial presentation and the recognition probe increases, allowing the consolidation of feature–feature bindings. We therefore tested two groups of participants, one with a brief interval (100 msec) between initial display and probe and one with a longer interval (900 msec). Since we did not want to introduce differential interference from verbal articulation in one case and not in the other, neither group used articulatory suppression. (The effects of articulatory suppression are tested in the same conditions in Experiment 4.)

#### Method

**Stimuli.** The stimuli were drawn from the same shapes and colors as in the previous experiments. On each trial, three items (all differing in color and shape) were presented in randomly chosen locations

in a  $3 \times 3$  invisible grid subtending  $16^{\circ}8' \times 12^{\circ}7'$ . The probe displays either matched the originals or exchanged the binding of two features from the original display. The probes were either presented in the same locations as the originals or in new, previously unfilled locations (see Figure 5).

**Design and Procedure.** Four factors were varied in this experiment, the first three within participants and the last between participants.

1. Matching or changed binding. On match trials, the probe items were identical to the original items. On changed-binding trials, the same features were present in the probe as in the original display, but with whole display probes, two bindings were changed by exchanging the shapes between two objects on half of the trials and the colors on the other half of the trials. With single probes, the probe had the shape of one of the original items and the color of another. When the location was an old one, the probe was presented in the location of the matching shape on half the trials and in the location of the matching color on the other half. Again there was an equal number of trials of each type and they were randomly mixed within blocks.

2. Old or new locations: In whole display trials, the test display occupied either the same three locations as the original display or three previously empty locations in the same grid. In single probe trials, the single probe occupied either the same location as in the original display or a new location that had previously been unfilled. In changed binding trials with old locations, the locations matched in shape on half the trials and color on the other half. The location factor was blocked, and the order was counterbalanced between participants.

3. Whole display or single probe. With whole display probes, there were three colored shapes in the probe as well as in the initial display. With single probes, there was only one item in the probe display. There was an equal number of trials of each type and they were randomly mixed within blocks. Figure 5 shows examples of each combination of same and different locations and same and different bindings for the whole display condition.

4. Delay of 900 msec or 100 msec: One group of 13 participants (5 men and 8 women) was tested with a delay of 900 msec between the initial display and the probe, and another group of 13 participants (5 men and 8 women) was tested with a delay of 100 msec.

Each participant did 32 trials in each combination of binding, location, and probe type. On each trial the sequence of events was as follows: A small black cross was presented at center screen for 1 sec, then a blank screen appeared for 900 msec, then an initial display of three items flashed for 150 msec, followed by a 900-msec blank interval for the long delay group and a 100-msec blank interval for the short delay group; finally, a test display appeared and remained until a response was made or for 4 sec, whichever came first. Participants were told that the same features would be present in the original display and in the probe. They were told to indicate by a keypress whether there was a change of binding in the probe display or whether the shapes and colors were paired as in the original display. In the new locations condition, they were told that the test items would appear in new locations, but that this was not relevant to their task: They were to respond solely to the bindings of color and shape. The ratio of match to changed binding trials was 1:1, and this was balanced across all levels of the other two factors. No articulatory suppression was used, since the 100-msec delay allowed no time for it, and we did not wish delay to be confounded with possible articulatory interference. Participants were instructed to aim for accuracy; if they made an error, they received a tone as feedback. At the beginning of the experiment, they practiced 16 trials of each block type and were given a chance to ask questions. At the conclusion, they were given an opportunity to comment and ask questions and then were debriefed.

#### Results and Discussion

The mean accuracy in each condition is shown in Figure 6. An ANOVA showed a significant effect of location [ $F(1,24) = 52.34, p < .0001$ ], with better perfor-

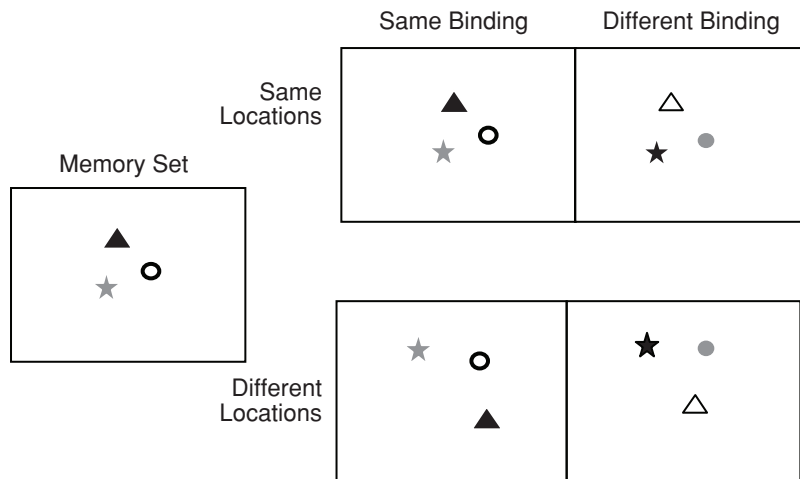


Figure 5. Schematic description of design of Experiment 3.

mance in the old locations (83.9% vs. 75.3% accuracy). There was also a significant interaction of probe type and location [ $F(1,24) = 82.28, p < .0001$ ], showing that the location difference was much reduced in the single probe trials (3% difference vs. 14%). Neither delay nor probe type nor their interaction had significant effects, but the three-way interaction between delay, location, and probe type was highly significant [ $F(1,24) = 14.80, p < .001$ ]. Relative locations were available only in the whole display conditions and these seemed to be mediating the effects of location. At the short delay, changing the locations in the whole display probes led to a drop of 18%, whereas at the long delay the drop was only 9%. This result suggests that the object files were initially closely dependent on location and that the direct link between the features was gradually consolidated over the first second of delay, reducing without eliminating the dependence on location.

Although some apparent motion occurred on the single probe new location displays with the short delay, the motion does not seem to have had much effect on performance. The results are almost identical to those obtained with the 900-msec delay. With the whole display probes, the apparent motion was not present or was very much weaker, probably because the three new locations were randomly related to the three original locations so that the three items could not be seen to move coherently. It is unlikely therefore that apparent motion is responsible for the large effect of location with the short delay.

Finally, in the present experiment, match versus changed binding gave highly significant effects [ $F(1,24) = 40.27, p < .0001$ ], with an overall bias of 84% to respond “match” versus 75% to respond “change.” There was also a shift in this bias across the different probe types: The single probe favored the response “change” by 3%, whereas the whole display probes favored the response “match” by 14%, giving a significant interaction between probe type and response [ $F(1,24) = 73.44, p < .0001$ ].

The default assumptions seemed to be that the whole display probes were the same as the originals and that the single probes were changed. Thus there were strong biases in opposite directions, favoring change blindness with the whole display and match blindness with the single probe. Participants were particularly likely to miss the change of binding in the new locations of the whole displays.

The results support a number of conclusions: First, locations are clearly more important in memory for binding than they are in memory for features, where they had no main effect on recognition. With memory for features, even when the binding was unchanged, probes in old locations were recognized only about 3% more than those in new locations, whereas with memory for binding, the drop in recognition due to new locations with whole display probes averaged 9% at the 900-msec delay. The greater importance of location in memory for binding is consistent with the claim from FIT that spatial attention is involved in binding features and establishing object files. Second, the effect of location is considerably stronger immediately after presentation than after a 900-msec delay, indicating that the bindings initially depend heavily on location but become less dependent with time. The bindings may start as feature–location bindings, whereas the feature–feature bindings take time to consolidate, reducing the dependence on location as time passes and the object file becomes more firmly established. Third, the location effects are almost entirely restricted to the whole display probes, implying that the locations are coded as a spatial pattern of relative rather than absolute locations.

#### EXPERIMENT 4

##### Interference With Memory for Binding

In this experiment, we ran some variations on Experiment 3 to look at different kinds of interference that might reveal more about the nature of the VWM representations.



The method was identical to that used in Experiment 3 except that in Experiment 4A, participants were required to use articulatory suppression (they repeated “Coca-Cola” over and over throughout each trial), and in Experiment 4B, they saw a single word presented between the initial display and the probe. The word was different on each trial; we used a set of 570 words selected from the MRC psycholinguistic database (Coltheart, 1981) with the following constraints: two or more syllables, a frequency of 10 or more, and a concreteness between 500 and 700. They were told to read the word to themselves, but no other task was required. The goal was to see whether simply attending to another visual stimulus with different properties and semantic characteristics would disrupt the memory for binding. Two groups of 12 (6 men and 6 women in Experiment 4A and 4 men and 8 women in Experiment 4B) participated in each of these two additional experiments.

Figure 7 shows the results, together with the control condition from the 900-msec delay in Experiment 3. The pattern is very similar across the three experiments with just a slight and nonsignificant reduction in recognition memory with articulatory suppression and with word interference. We ran separate between-groups ANOVAs, pairing the 900-msec baseline condition from Experiment 3 with both the articulatory suppression and the intervening word experimental variations. Table 2 shows the significance levels in these ANOVAs.

There were no significant main effects of experiment, although the impairment due to articulatory suppression approached significance, suggesting some minimal benefit from verbal coding. Either participants did not attempt to store verbal labels or the labels were not particularly helpful in this recognition memory task. Performance in Experiment 4B was not impaired by the presentation of an extraneous word during the delay period. VWM clearly differs from iconic memory. The material in VWM has been encoded into a format that is no longer vulnerable to intervening visual events, at least when

they are different enough in category (word vs. colored shapes).

Old versus new locations had a highly significant effect in both ANOVAs. As in Experiment 3, performance dropped significantly when the probe items were presented in new locations rather than when they reappeared in the original locations. Again these results are in clear contrast with the findings for feature memory in Experiment 1, suggesting that location plays an important role specifically in the retention of binding as well as in its encoding. Location change also interacted strongly with probe type, with large effects on whole display trials and little effect on the single probe trials, probably because the pattern of relative locations is preserved only with whole display probes.

Again there was an overall bias to respond “same” rather than “different” (80% vs. 73%) and also a highly significant two-way interaction between probe type and match versus changed binding. The mean correct responses were 90% “match” and 63% “changed” for the whole display, but the bias was reversed to 71% “match” and 83% “changed” for single probes. Finally, the three-way interactions between probe type, location, and match versus changed binding were also highly significant. These three-way interactions reflect the participants’ very poor performance in detecting the change of binding in a new location with whole display probes.

### EXPERIMENT 5 The Effects of Longer Delays on VWM for Binding

We had two goals for this experiment: First, to observe what happened to the role of location in memory for binding after somewhat longer delays and second, to ascertain whether the previous results reflected specific strategies that might have been absent when the participants could

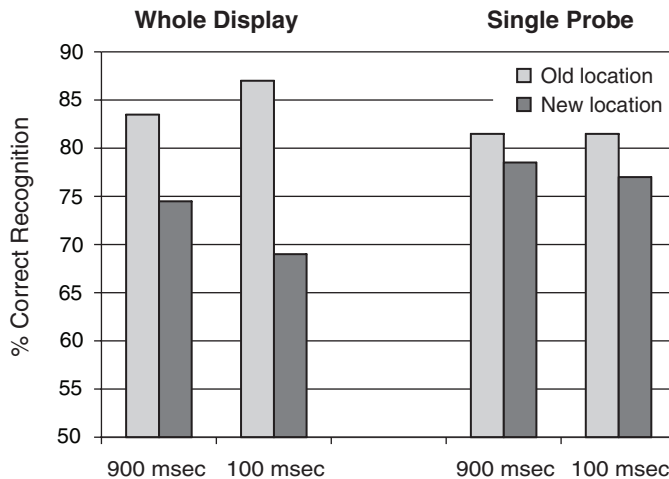


Figure 6. Accuracy in recognition of binding in VWM as a function of delay and of old or new locations.

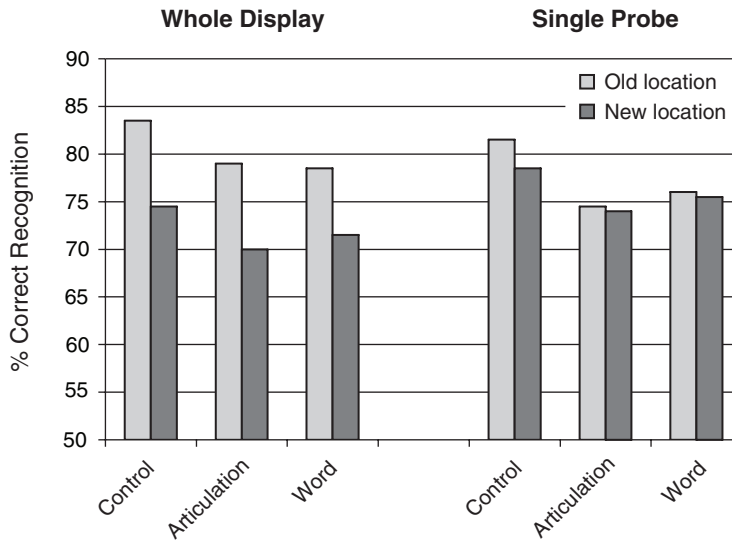


Figure 7. Accuracy in recognition of binding with articulatory suppression and with intervening visual presentation of a word.

no longer predict either the delay interval or the probe locations. In the present experiment, we randomly mixed both the locations and the delays within blocks, rather than testing them separately. It seemed possible that participants might encode the display differently when they knew whether it would be probed immediately or after a delay and whether the probe items would appear in the same or different locations.

**Method**

**Apparatus and Stimuli.** These were the same as in Experiment 3.

**Procedure.** The procedure was the same as for the whole display condition in Experiment 3 with the following exceptions: Two additional delays, 3 sec and 6 sec, between the initial display and the probe were added to the 100-msec and 900-msec delays tested before, and the four delays were presented randomly within the blocks. The old and new location trials were also randomly mixed within blocks rather than separated as they had been in the earlier experiments. Participants repeated “Coca-Cola” throughout each trial in order to prevent verbal coding of the displays.

**Results**

An ANOVA on Experiment 5 again showed a significant overall benefit to memory for binding when it was

tested in old rather than in new locations [ $F(1,14) = 38.70, p < .001$ ]. There was no overall effect of delay [ $F(3,14) = 2.27, p = .094$ ], suggesting that VWM for binding retains its strength for at least a few seconds after the stimuli have disappeared, provided that there is no additional interference. However, a more detailed analysis shows marked changes in the nature of the memory trace over those first few seconds (see Figure 8). There was a significant interaction of location with delay [ $F(3,42) = 9.61, p < .001$ ], reflecting the fact that items in old locations were much better recognized than items in new ones after the very short intervals but that this advantage dropped smoothly across delay from 0.1 to 6 sec (see Figure 8). Memory for items in new locations steadily increased over time up to 3 sec, then began to drop in parallel with memory for items in the old locations. After 3 sec have passed, recognition of the bindings in a new location is almost as good as recognition of them in the original location. Before that, there is substantial benefit from a matching location and also substantial interference when the probes appear in new locations. The results suggest that immediately after presentation, location is an integral part of the representation of the objects, and/or is actively used in the retrieval

Table 2  
ANOVAs on Effects of Articulatory Suppression and Word Interference

	Articulation	Word interference
Experiment (baseline vs. interference)	n.s. ( $p = .075$ )	n.s.
Location (old vs. new)	35.38***	30.89***
Probe type (single vs. whole display)	n.s.	n.s.
Location × probe type	34.17***	11.64***
Match versus changed binding	35.39***	23.33***
Match × probe type	170.33***	163.24***
Match × probe type × experiment	6.67*	n.s.
Match × probe type × location	32.53***	18.86***
Match × probe type × location × experiment	n.s.	6.95*

Note—Both relative to 900-msec delay with no articulation. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

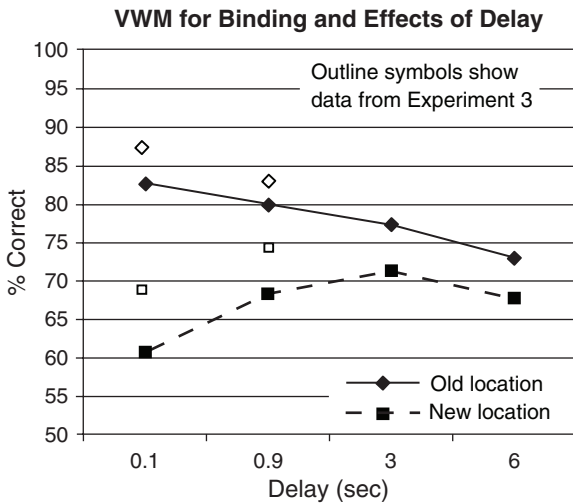


Figure 8. Effects of longer delays on VWM for binding in Experiment 5.

process. As time passes, however, the binding between the features within objects is consolidated and need no longer be mediated primarily through the features' shared location.

There is a significant bias to respond "same" to binding overall—in other words, to miss the changes in binding [ $F(1,14) = 63.60, p < .001$ ], which suggests a propensity to form illusory conjunctions in memory to match those that have just been seen. There is also a significant interaction of response bias with delay [ $F(3,42) = 12.35, p < .001$ ], as shown in Figure 9. The bias peaks around 0.9 sec and has disappeared at 6 sec. There is no three-way interaction with location. As the original binding between the features is consolidated over time, the propensity to form illusory conjunctions may also be reduced.

The participants in Experiment 5 were less accurate overall than those in Experiment 3. The drop in performance may be explained by the increased uncertainty about where the probe would appear and how soon and also by the fact that there was articulatory suppression throughout the trials in Experiment 5 but not in Experiment 3. However, the pattern of performance on old and new location trials with the two short delays was strikingly similar across the experiments, suggesting that these effects are inherent to VWM for binding and are not induced by different strategies adapted to the particular conditions being tested.

## GENERAL DISCUSSION

Our research was directed at two related questions: (1) What is the relation between memory for locations and memory for their contents, both features and bound objects? (2) Is binding automatic when attention is directed to each item in a display, or do we have the option of storing only the separate features? Data from Wheeler and Treisman (2002) suggest that VWM can hold separate fea-

tures, perhaps in the form of traces activated in separate feature maps during perceptual processing, as well as the bound objects that are formed when attentional resources are available. This raises the question of whether we have strategic control over the format used in VWM.

## Location

One of the striking findings was recognition's greater dependence on location when memory for the bindings was tested than when memory for the features was tested. When bound objects were explicitly retrieved, the original locations provided strong support. If locations and objects were stored independently, it is hard to see how these strong dependencies could occur. At the least, there must be strong connections facilitating retrieval of the object through its location and perhaps of the location through the object. Evidence consistent with the idea that object retrieval can be facilitated by location information comes from a study of memory for objects in natural scenes by Hollingworth and Henderson (2002). They showed that fixating the location in which an object had appeared facilitated its retrieval.

Features too are clearly stored and retrieved preferentially through their locations, as shown by the strong interactions of location with binding and with match versus new feature. However, when features are tested in new locations, they can be accessed independently of their original locations without loss of accuracy. Retrieval from VWM may initially and naturally occur at the object level, where features have been bound to their locations and to each other; however, if this route is ruled out because the locations are new, unbound features can still be retrieved, perhaps from separate feature modules, by matching to those in the probe. These conclusions are consistent with those drawn by Wheeler and Treisman (2002).

Another finding from our tests of VWM is that locations seem to be defined as relative locations in a spatial configuration rather than as absolute locations, either in

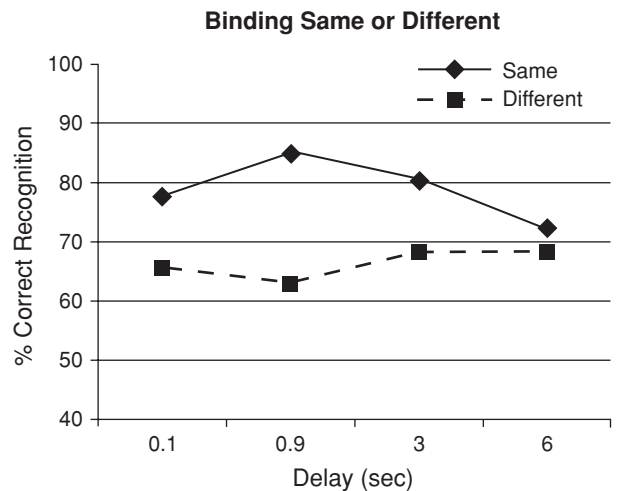


Figure 9. Response biases across delays in Experiment 5.

the real world or on the retina. Evidence for this comes from the fact that with single probes, where no pattern of spatial relations was available, location matches had no effect on feature memory and little effect on memory for binding. A similar conclusion was reached by Jiang et al. (2000), who found that changes in spatial configuration disrupted visual short term memory for color or shape even though the configuration was irrelevant to the task. It is possible that absolute locations can also be stored when they are relevant to the task, but we find little evidence that they contribute to memory when they are not explicitly needed.

### Binding

Features are also better retrieved when probed in their original conjunctions rather than in new ones. A new binding presented in a location that had contained one of the rebound features often caused the other rebound feature to be interpreted as new, yielding only 51% correct recognition on match trials. A critical question is whether the changed bindings also damage recognition when probed in new locations. The fact that they do (participants' scores declined 8% on changed vs. same bindings in new locations) implies that features are bound to each other as well as to their locations. Thus it seems that features are automatically bound in memory, even when this is not required and even when it actually interferes with feature recognition. **Binding is not optional; it occurs as a direct consequence of attending to an object and storing it for later recognition.** This is a new finding to add to the model proposed by Wheeler and Treisman (2002).

Note that in order for changes in location and/or binding to affect performance even when they are irrelevant to the task, the changes must be mediated by changes in the perceptual and/or memory representations, not simply by objective differences in similarity between the original displays and the probes. Thus changes in location and/or binding provide evidence that (1) **when only the features are relevant, they are nevertheless at least to some degree bound to each other and to their locations, and (2) when the bindings are relevant, locations play an important role in their early retention and retrieval.**

### Verbal Coding

There was little decrement in accuracy of recognition for the group that was required to repeat "Coca-Cola" during the presentation, delay, and response. If articulatory suppression is assumed to remove or reduce the possibility of verbal coding, we can conclude that participants did not rely on verbal coding with these brief displays of colored shapes. The interaction between articulation, probe type, and same versus different binding suggests that if there is any interference from articulation, it occurs primarily in the detection of a match in the single probe condition.

### Visual Interference

The presentation of an extraneous visual pattern (the word) in the same general area as the display had no ef-

fect on recognition of color–shape bindings. Thus pulling visual attention to an irrelevant intervening stimulus is not sufficient to disrupt visual memory for binding. It is possible, of course, that more similar visual stimuli or a more attention-demanding task would interfere. Our data are consistent with those of Woodman, Vogel, and Luck (2001), who tested feature memory rather than binding. They asked whether an intervening search task interfered with VWM and found that the estimated memory capacity was 3.2 items with memory alone and 2.7 with search added. Using the same type of items (Landolt Cs) in memory and in search did not increase the interference, ruling out a similarity-based account. Finally, in another experiment, the search stimuli were shown, but participants could ignore them. Again memory dropped from 3.3 to 2.6 items. Thus the requirement to search had no additional effect beyond some small, shared visual disruption. On the other hand, when one of the search stimuli had to be remembered along with the colored squares, memory was considerably impaired. Woodman et al. concluded that "objects can be attended at a perceptual level without automatically being entered into working memory."

Overall, the absence of effect from these two manipulations suggests that the characteristics of visual memory that we have probed in this paradigm are fairly robust. VWM is much less affected by articulatory suppression or by visual interference from an extraneous visual word than it is by the factors that we varied within experiments—the new or old locations and the nature of the probe.

### Delay

Memory for binding, although limited, seems neither to decay very fast nor to improve much across brief intervals after presentation. It is possible, of course, that if we had tested longer delays, we would have seen more decrements. The main change over time in the first few seconds after presentation is seen in the effect of location. There is a marked improvement in performance for new locations, accompanied by a decrease in performance for old locations. Apparently, the traces become less location specific over the first few seconds of delay, dropping to an asymptotic advantage for old locations of only 6% after 3 sec (Figure 8). There was also a marked bias to respond "same" (or to rely on feature matches and miss changes of binding) at the shorter delays. This bias disappears once the delay reaches 6 sec.

How does our framework relate to earlier theories of visual short-term memory, such as those proposed by Phillips (1974) and Irwin (1992)? Phillips suggested that short-term visual memory goes through two stages—an initial, brief, high-capacity, topographic sensory store, subject to masking, and a more limited but longer lasting schematic memory that survives for a few seconds. Irwin (following Coltheart, 1980) distinguished three stages: precategorical visible persistence and nonvisible information persistence, both of which contribute to traditional measures of iconic memory, and postcategorical visual short-term memory. Visible persistence depends

on exposure duration whereas information persistence does not; however, information persistence shares with visible persistence properties such as a high capacity for detailed form and structure information, and information persistence remains maskable, whereas Coltheart's post-categorical visual memory is more abstract and is limited to about five items. Where would our studies fit in relation to these hypothesized stores? At 100-msec delay, visible persistence has ended, since the stimulus itself lasted 150 msec and visible persistence is measured from stimulus onset rather than offset (di Lollo, 1980). Could performance at our 100-msec delay be the result of informational persistence? Informational persistence lasts about 150–300 msec after stimulus offset (Irwin & Yeomans, 1986; Phillips, 1974; Yeomans & Irwin, 1985). However, there is no evidence that masking by the probe display played any role in the old location conditions, although informational persistence, like sensory persistence, should be subject to visual masking. If anything, accuracy was better at the 100-msec than at the 900-msec delay and better in the old locations than in the new. Nevertheless, the dependence on location fits better with the sensory than with the schematic store.

An alternative account that seems more consistent with our data emphasizes the distinction between different types of information: the separate features and locations on the one hand and the object files on the other, all of which may coexist and contribute to performance at early intervals after presentation. Features and locations are registered automatically in separate maps and their traces may last for brief durations after the stimulus is removed. Meanwhile, object files are formed as attention scans the display to integrate its features. The object files are initially tied to locations as attention is focused on them. As time passes, there is a gradual consolidation in which the object features are bound more firmly to each other and become less dependent on their locations. We conceive of the change from early to late object files as more of a continuum than a dichotomous translation. In the early stages, object files, together with the persisting traces of features and locations, may function most like the sensory and informational persistence described by Irwin and others. In the later stages, attended object files become gradually abstracted from sensory details and from their locations and are consolidated into the more lasting representations attributed to VWM.

### Whole Display Versus Single Probe

We hoped that the comparison of whole display with single probe cues for retrieval might throw light on the retrieval process in VWM. Would the reinstatement of the complete context make the matching process easier, or would it add to the decision load? In the Wheeler and Treisman experiments (2002), memory for binding was better with a single probe than with the whole display, whereas memory for features showed no effect of probe type. Wheeler and Treisman proposed that when three items were presented in the probe and the bindings had to

be checked, attention had to focus on each item in turn to ensure accurate binding. This pulled attention away from the stored items and led to some loss of binding in memory. When the probe was a single item, there was less need for attention, since the probe's features could be bound by default. When only the separate features were tested, differential demands on attention were less important.

In the present experiments, we see the reverse pattern: Memory for features was better with the single probes, whereas memory for bindings was, if anything, worse with the single probe than with the whole display, averaging 75% and 79%, respectively (see Figure 10). Does this inconsistency cast doubt on the Wheeler and Treisman (2002) model? We think not. The discrepancy can be explained by some important differences of method in the two experiments. First, in the present feature-memory experiments, we were not directly interested in the relative difficulty of recognizing changes of features versus changes of binding, so we changed only one feature on new feature trials, whereas Wheeler and Treisman changed two features on whole display new feature trials in order to match the load across features and bindings. This made their feature recognition tests much easier than the present ones. Second, our feature recognition tests were also disrupted by changes of binding, which never occurred in the feature memory tests of Wheeler and Treisman. Again, this would damage the whole display tests more than the single probes, since more items differed on changed binding trials.

Finally, in both the binding and the feature tests, the relation between the locations in the original display and in the probes was different in the present experiments from those in Wheeler and Treisman (2002). In the whole display trials of their color–shape experiment, Wheeler and Treisman reused the same locations but exchanged items between those locations, whereas in their single probe trials, the probe was always presented at the center of the display in a location that was previously empty. Thus the interference on their whole display trials could have come partly from overwriting the previous memory traces, whereas the single probe would not have interfered with any location-specific traces. In the present experiments, we used new locations in half the trials and old locations with the original allocation of items in the other half (apart from the change in binding on new binding trials). Thus the present participants may have benefited on old location trials from the preserved location–feature bindings, actually performing better on whole display than on single probe trials. If we look only at trials with new locations, we replicate the Wheeler and Treisman finding of better performance with the single probe (78% vs. 72%), consistent with the attention overload hypothesis suggested in the earlier article.

Our findings suggest that the comparison between these two probe types is complicated by a number of different factors: Whole display probes load attention and decision processes more than single probes, and they may overwrite the earlier memory traces if new items appear in the old lo-

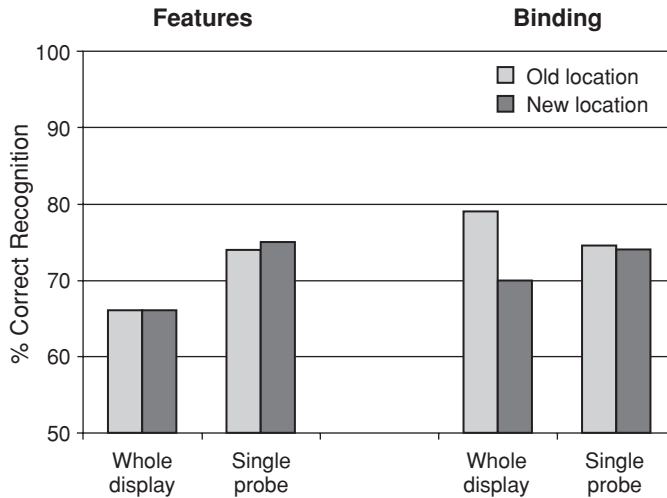


Figure 10. Comparison of whole display and single probe in tests of memory for features and memory for binding.

cations. On the other hand, whole display probes may help retrieval by reinstating the complete context including the relative locations if these are preserved in the probes.

### Bias

In looking at the relative frequencies of responses of “same” and “different,” we made an interesting observation. The overall bias was to respond “same” with the whole display and “different” with the single probe, particularly with probes in the old locations. Thus task differences seemed to induce different strategies for performing the tasks: With the whole display probes, most items from the initial display matched items in the probe, so it made sense to look for a change and to have “no change” as a default assumption. On the other hand, in the single probe condition, most of the initial display items did not match the single probe, so it made more sense to look for a match. Note that these overall biases cannot explain differences in performance *within* the two probe types. In the old locations, participants were better not only on match trials but also on changed binding trials for which the correct response was “different,” and in new locations, they were better not only on changed binding trials but also on match trials for which the correct response was “same.”

### Capacity

Our three-item displays were correctly recognized on about 80% of trials. Using the formula devised by Pashler (1988) for estimating the capacity of VWM, capacity  $k = [S \times (H - F)] / (1 - F)$ , where  $S$  items are presented,  $H$  is the hit rate, and  $F$  is the false alarm rate, the capacity in our experiments averaged only about two objects, substantially fewer than the estimate of three to four objects reported in a feature memory experiment by Luck and Vogel (1997). The estimate given by Luck and Vogel may be higher in part because they tested displays with more than three items presented, giving more scope for higher

recognition scores, and also in part because they tested feature memory without changing the binding.

However, the change of bias with the different probe types described in the previous paragraph raises a question about this measure of capacity. We assumed in making the calculations that detecting a change counted as a hit. But a hit could instead have been the correct detection of a match between probe and memory set. The estimates change with the assumption selected. For example, with the single probes in new locations, the estimates of capacity ( $k$ ) are 2.35 if participants were looking for a change of binding and 2.02 if they were looking for a match. If we follow the response biases and infer change blindness in whole display probes and match blindness in single probes, we should take responses of “different” as hits with the whole display probes and responses of “same” as hits with the single probes. The capacity would then be 1.57 for the whole display and 2.02 for the single probe, making performance in the whole display condition a little worse than in the single probe condition. If we reverse these assumptions, the whole display estimate would be 2.45 and the single probe estimate would be 2.35, slightly worse than the whole display estimate. The former assumptions seem more plausible here, but in general we would argue that the ambiguity needs to be resolved in each experiment before VWM capacity can be estimated.

### Conclusion

The view of VWM that emerges from these experiments is of a representation that is still close to the perceptual encoding. Attended objects are automatically bound to their locations within a spatial configuration (not to absolute locations), and their features are integrated with each other. There is in addition some memory for unbound features, perhaps for items that did not receive attention in the brief presentations we used or for items that lost their binding during the memory delay. The locations be-

come less salient, or the objects become detached from their locations, over the first seconds after presentation, but at a 900-msec delay, locations still play a substantial role in feature retrieval and are still critical in retrieving the bindings. The implications are consistent with the FIT view that features are initially registered preattentively in separate feature maps, while at the same time the filled locations activate areas in a map of locations. If focused attention is directed to each filled location, it binds the features to their locations and to each other perceptually, forming object files that are, at least initially, addressed by their current location. Once the object files have been set up and the features bound, the objects may be identified, verbally labeled, and stored independently of their locations, but the initial representations integrate the locations with the features in bound object tokens. These form the main content of VWM.

Finally, how do we reconcile this framework for our experiments with the behavioral and neural evidence that objects and locations occupy separate stores? Our experiments directly tested only memory for “what,” providing indirect evidence about memory for “where.” Experiments 3–5 cast doubt on the possibility of storing bound objects without their locations, at least in the initial phases of memory encoding. However, the inference need not be symmetrical. We suggest that the pattern of filled locations may be stored without the features or objects that occupy them if the task requires it, and in this context the features or objects would act simply as place markers.

## REFERENCES

- BADDELEY, A. (1992). Is working memory working? The fifteenth Bartlett lecture. *Quarterly Journal of Experimental Psychology*, **44B**, 1-31.
- BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.
- CARLESIMO, G. A., & OSCAR-BERMAN, M. (1992). Memory deficits in Alzheimer's patients: A comprehensive review. *Neuropsychology Review*, **3**, 119-169.
- CARLESIMO, G. A., PERRI, R., TURRIZIANI, P., TOMAIUOLO, F., & CALTAGIRONE, C. (2001). Remembering what but not where: Independence of spatial and visual working memory in the human brain. *Cortex*, **37**, 519-534.
- COLTHEART, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, **27**, 183-228.
- COLTHEART, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, **33A**, 497-505.
- 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.
- DI LOLLO, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, **109**, 75-97.
- HANLEY, J. R., YOUNG, A. W., & PEARSON, N. A. (1991). Impairment of the visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology*, **43A**, 101-125.
- 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**, 145-156.
- HECKER, R., & MAPPERSON, B. (1997). Dissociation of visual and spatial processing in working memory. *Neuropsychologia*, **35**, 599-603.
- HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 113-136.
- IRWIN, D. E. (1992). Visual memory within and across fixations. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 146-165). New York: Springer-Verlag.
- IRWIN, D. E., & ANDREWS, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press.
- IRWIN, D. E., & YEOMANS, J. M. (1986). Sensory registration and informational persistence. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 343-360.
- 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.
- KAHNEMAN, D., TREISMAN, A., & GIBBS, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, **24**, 175-219.
- LOGIE, R. H., & MARCHETTI, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 105-115). Amsterdam: North-Holland.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- MAGNUSSEN, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences*, **23**, 247-251.
- MOTTAGHY, F. M., GANGITANO, M., SPARING, R., KRAUSE, B. J., & PASCUAL-LEONE, A. (2002). Segregation of areas related to visual working memory in the prefrontal cortex revealed by rTMS. *Cerebral Cortex*, **12**, 369-375.
- MUNK, M. H. J., LINDEN, D. E. J., MUCKLI, L., LANFERMANN, H., ZANELLA, F. E., SINGER, W., & GOEBEL, R. (2002). Distributed cortical systems in visual short-term memory revealed by event-related functional magnetic resonance imaging. *Cerebral Cortex*, **12**, 866-876.
- NUNN, J. A., POLKEY, C. E., & MORRIS, R. G. (1998). Selective spatial memory impairment after right unilateral temporal lobectomy. *Neuropsychologia*, **36**, 837-848.
- PASHLER, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378.
- PELLI, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.
- PHILLIPS, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, **16**, 283-290.
- POSTLE, B. R., & D'ESPOSITO, M. (1999). “What”—then—“where” in visual working memory: An event-related fMRI study. *Journal of Cognitive Neuroscience*, **11**, 585-597.
- POSTLE, B. R., STERN, C., ROSEN, B., & CORKIN, S. (2000). An fMRI investigation of cortical contributions to spatial and nonspatial visual working memory. *NeuroImage*, **11**, 409-423.
- PRABHAKARAN, V., NARAYANAN, K., ZHAO, Z., & GABRIELI, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, **3**, 85-90.
- RAO, S. C., RAINER, G., & MILLER, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, **276**, 821-824.
- RIZZUTO, D. S., MAMELAK, A. N., SUTHERLING, W. W., FINEMAN, I., & ANDERSEN, R. A. (2005). Spatial selectivity in human ventrolateral prefrontal cortex. *Nature Neuroscience*, **8**, 415-417.
- SMITH, E. E., JONIDES, J., KOEPE, R. A., AWH, E., SCHUMACHER, E. H., & MINOSHIMA, S. (1995). Spatial versus object working memory: PET investigations. *Journal of Cognitive Neuroscience*, **7**, 337-356.
- STEFURAK, D. L., & BOYNTON, R. M. (1986). Independence of memory for categorically different colors and shapes. *Perception & Psychophysics*, **39**, 164-174.
- TREISMAN, A. [M.] (1977). Focused attention in the perception and retrieval of multidimensional stimuli. *Perception & Psychophysics*, **22**, 1-11.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- 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.

- VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 1436-1451.
- VUONTELA, V., RÄMÄ, P., RANINEN, A., ARONEN, H. J., & CARLSON, S. (1999). Selective interference reveals dissociation between memory for location and colour. *NeuroReport*, **10**, 2235-2240.
- WHEELER, M. E., & TREISMAN, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.
- WILSON, F. A., O'SCALAIDHE, S. P., & GOLDMAN-RAKIC, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, **260**, 1955-1958.
- WOODMAN, G. F., & LUCK, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, **11**, 269-274.
- WOODMAN, G. F., VOGEL, E. K., & LUCK, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, **12**, 219-224.
- WOODMAN, G. F., VOGEL, E. K., LUCK, S. J., & HOLLINGWORTH, A. (2007). *Is the same visual working memory store used to represent all visual objects?* Manuscript in preparation.
- YEOMANS, J. M., & IRWIN, D. E. (1985). Stimulus duration and partial report performance. *Perception & Psychophysics*, **37**, 163-169.
- ZIMMER, H. D., SPEISER, H. R., & SEIDLER, B. (2003). Spatio-temporal working-memory and short-term object-location tasks use different memory mechanisms. *Acta Psychologica*, **114**, 41-65.

(Manuscript received March 29, 2005;  
revision accepted for publication August 8, 2005.)