# Organization of Visual Short-Term Memory

Yuhong Jiang and Ingrid R. Olson
Yale University

Marvin M. Chun
Vanderbilt University

The authors examined the organization of visual short-term memory (VSTM). Using a change-detection task, they reported that VSTM stores relational information between individual items. This relational processing is mediated by the organization of items into spatial configurations. The spatial configuration of visual objects is important for VSTM of spatial locations, colors, and shapes. When color VSTM is compared with location VSTM, spatial configuration plays an integral role because configuration is important for color VSTM, whereas color is not important for location VSTM. The authors also examined the role of attention and found that the formation of configuration is modulated by both top-down and bottom-up attentional factors. In summary, the authors proposed that VSTM stores the relational information of individual visual items on the basis of global spatial configuration.

Ever since the pioneering study by Phillips (1974) on the visual short-term memory (VSTM) of patterns, the issue of representation has been a central topic in the study of VSTM (Baddeley, 1986; Logie, 1995). This issue can be parsed into several interrelated questions. The most intensively studied question is the capacity of VSTM. How much visual information can humans retain from one moment to the next? The answer to this question is fairly clear: The capacity of VSTM is severely limited. For example, people cannot integrate information gathered from successive fixations on spatial-based coordinates, which suggests that very little information can be retained from previous fixations (Irwin, 1991; Irwin, Brown, & Sun, 1988; Irwin, Yantis, & Jonides, 1983; Rayner & Pollatsek, 1983; but see Davidson, Fox, & Dick, 1973). Similarly, researchers have shown that VSTM is very poor for unattended information in scene perception and in social interactions (Levin & Simons, 1997; Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1998).

Specifically, the capacity of VSTM is often quoted to be approximately four items (Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974; Simons, 1996). However, different stimuli produce different estimates of VSTM capacity. When letter stimuli (Pashler, 1988) or simple features (Luck & Vogel, 1997) are used, the capacity hovers around four to

five items. In contrast, memory for spatial locations produces larger variability. Simons (1996) found that people performed almost perfectly with 5 locations (the largest set size tested), but the capacity for spatial locations was clearly taxed when confronted with more locations in other studies. For example, when 8 to 32 locations were presented, participants in Phillips's (1974) study performed well below ceiling when relying only on VSTM. However, their performance was nearly perfect if they were allowed to rely on visual sensory memory. Visual sensory memory, also known as iconic memory (Neisser, 1967), has higher capacity but shorter durability than VSTM (Sperling, 1960). Thus, iconic memory is distinguishable from VSTM and is not considered further in this article.

The second question concerns the resolution of VSTM. How detailed are VSTM representations? Research on picture memory has found that visual memory representations are not detailed. Although people can store hundreds of pictures in a short period of time (Nickerson, 1965), the visual details of the pictures seem to be poorly retained (Intraub, 1997). O'Regan and his colleagues have argued for a very coarse visual representation (O'Regan, 1992; O'Regan, Rensink, & Clark, 1999; Rensink et al., 1997). They suggested that detailed visual representations are unnecessary because people can always rely on the outside world as their visual memory. Although implicit measures may reveal more detailed visual representations (see Chun & Nakayama, in press; Moore & Egeth, 1997), such measures have yet to be developed to quantify the upper limit of VSTM. So to summarize, the literature suggests that the capacity of VSTM is around four to five items and that the representation of items in VSTM is not detailed.

Although there is ample evidence that VSTM capacity is limited, the nature of representations in VSTM is not well understood. Specifically, what is the unit of VSTM capacity (feature, object, or location)? How are items organized and related to one another? The unit question has been examined recently by Luck and Vogel (1997), and past research on the unit question is briefly described below. The VSTM organization issue has not gained much attention, however, and

this will be the main focus of our article. We will turn to the organization issue after a brief review of the unit question.

As we mentioned above, people can retain about four to five letters or single features in VSTM (Luck & Vogel, 1997; Pashler, 1988). What is the unit of this capacity? It can be features, objects, or locations. Luck and Vogel (1997) sought to determine whether the unit was at the feature level or the object level. To investigate this, they presented participants with objects containing conjunctions of features. For example, they presented four objects, each of which was a conjunction of color, orientation, size, and the presence or absence of a gap. Participants were asked to hold these features in VSTM. If the unit of VSTM is feature, participants should only be able to retain 4 features. If the unit is object, participants should be able to retain all 16 features because they conjoin to form four objects. The results showed that participants were able to retain all 16 features, suggesting that the unit of representation in VSTM was at the object level rather than the feature level. Luck and Vogel (1998) further proposed that the underlying neural mechanism for object-based VSTM could be synchronization of firing across populations of cells (Singer & Gray, 1995). Neurons that encode different features of the same object fire in synchrony. Information can thus be encoded together as one unit.

Thus, investigations of the unit of VSTM representation have shown that the VSTM capacity for features can be greatly enlarged by grouping features into a single object. This confers significant benefits to the visual system. The human visual system efficiently extracts regularities in visual perception and learning (Chun & Jiang, 1998, 1999), and visual memory is no exception.

In this investigation, we examined the representation issue in VSTM, but at a different level than the unit-of-capacity question. In particular, we asked how items in VSTM are organized and related to each other. That is, does VSTM represent individual items independently, or does it also represent relational information between items? If relational information is represented, how can we characterize the nature of that relational representation?

## Relational Processing in VSTM

Items held in VSTM may be represented independently such that memory for one item is not affected by other items. Conversely, individual items in VSTM may influence each other, and researchers interpret this to reflect *relational processing* in VSTM.

There are at least two types of relational encoding in visual perception: (a) semantic schemata formed from long-term learning and (b) perceptual grouping based on gestalt principles. Neither of these was the focus of our investigation, for reasons described below. The influence of semantic schemata on object identification has been extensively studied (Biederman, Mezzanotte, & Rabinowitz, 1982; Intraub, 1997; Intraub, Bender, & Mangels, 1992). For example, knowing that you are viewing a kitchen scene facilitates identification of probable objects such as a refrigerator or an oven (Biederman et al., 1982). However,

semantic context can be used only if relevant schemata have been stored in long-term memory and if schemata can be abstracted from the current display. For simple visual stimuli devoid of long-term schematic cues, perceptual grouping provides more immediate organization (Kanizsa, 1979).

The effect of perceptual grouping on VSTM is demonstrated when the similarity of to-be-remembered items is manipulated. For example, when eight colored objects are presented, this typically exceeds VSTM capacity for colors. However, people can remember all eight colors if all the objects have the same color, or if half of them are one color (e.g., green) and half of them are another color (e.g., red). When items can be grouped on the basis of similarity or other grouping cues, such as spatial proximity or closure (Kanizsa, 1979), the visual representations are simplified and the capacity of VSTM increases concomitantly.

Perceptual grouping may be limited, however, as it only applies to situations in which grouping cues are salient. What if grouping cues are not salient? The central goal of our study was to establish and characterize relational processing in the absence of grouping cues. For example, when several different colors are scattered randomly on the display, how are they represented in VSTM?

## Experimental Paradigm

To study the above question, we adopted the change-detection task used in several studies of VSTM (Lee & Chun, 1999; Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974; Rensink et al., 1997; Simons, 1996). In this task, a visual display is briefly presented to the observers (e.g., 400 ms). After a short delay interval (e.g., 500–1,000 ms), another visual display is presented. The first display is the *memory image*, which contains visual stimuli that the observers have to memorize. The second display is the *probe image*, which is either identical to the memory image or slightly different. The difference can be an addition of a new stimulus, a deletion of an old stimulus, or a switch of two old stimuli, et cetera. The task is to detect whether a change has occurred. On the probe image, the item that is changed is called the *critical probe*. Other items that do not change are called *distractor probes*. One important feature of this task is that the change is usually readily detectable if attention is directed to the critical probe in advance. Therefore, failure to detect a change cannot be attributed to perceptual failure to detect a weak signal.

In this task, the number of to-be-remembered items is called *memory set size*. The number of probe items that need to be compared with the memory image is called *probe set size*. In most studies, the memory set size and probe set size are the same. A general finding is that performance deteriorates as memory set size increases, reflecting capacity limitations in VSTM.

Our paradigm is different from previous change detection studies in one important way. In previous studies, the only item that could change was the critical probe; distractor probes stayed the same across memory and probe images. In our paradigm, however, distractor probes can change in various ways. They may stay the same as in the memory

image, they may change identity or location, or they may simply disappear. The critical probe, however, is clearly demarcated either by a unique feature or by a box cue in most experiments. Therefore, participants always know which items should be compared with the memory image.

In our task, distractor probes serve a critical role. Our assumption is that memory items may be encoded in relation to each other. When distractor probes stay the same, relational information in the memory image is preserved in the probe image. If relational information is processed by VSTM, the match between memory and probe images should facilitate memory retrieval and comparison (Tulving, 1974; Tulving & Thomson, 1973). However, distractor probes may also change in our experiments, producing a change in relational cues for VSTM retrieval. If relational information is important for VSTM, consistent relational cues should facilitate performance and inconsistent cues should hurt performance.

## Outline of Experiments

The logic discussed above is used in our experiments. We conducted eight experiments to investigate how multiple items are represented in VSTM. Our studies show that visual memory for a single target item is affected by various changes to surrounding distractor items, even when the location of the target is fully specified during memory retrieval. Our experiments are grouped into five sections. Section 1 examines VSTM for color and demonstrates that the representation of a given color is not independent of the color of other items in the display. Section 2 examines VSTM for spatial locations and shows that the representation of a given location is not independent of the location of other items, even when there are explicit instructions to ignore the configuration. Section 3 examines the interaction between VSTM for identities (color and shape) and VSTM for spatial configurations. Here we show that spatial configuration is important for accurate memory of colors and shapes. Section 4 delineates configuration in more detail and shows that (a) the configuration effect can be parsed into a purely spatial component and an identity component in color VSTM and (b) the formation of a configuration can be based on a subset of items in the visual display. Section 5 shows that the influence of color pattern on spatial location VSTM is negligible, suggesting that spatial configuration plays a more primitive role in our VSTM paradigm.

## Section 1: Is VSTM of One Color Independent of Other Colors?

When various colors are randomly scattered on the display with no salient grouping cues, does relational information between items influence VSTM? Or is the short-term memory of each item independent, subject only to overall general capacity?

Our pilot study provided preliminary evidence that VSTM of a given color was affected by the fate of other colors simultaneously presented on the display, suggesting that

relational information between items may be processed in VSTM. In the pilot study, we briefly presented several colored squares on the memory image. After a blank interval of 1 s, a probe display with the same number of colors was presented. On the probe display, one color square was enclosed by an outline box, and the observers' task was to decide whether this color was the same as the memory color located there earlier. Colors not cued by the outline box either maintained their previous colors (*minimal color change condition*) or they all changed colors (*maximal color change condition*). Change detection of the critical probe was contrasted between the minimal and the maximal color change conditions. Results showed that performance was significantly poorer when accompanied by color changes of the distractor probes. This result is consistent with the hypothesis that relational information between color items is processed in VSTM. As distractor probes changed colors on the probe image, the retrieval context of the critical probe became inconsistent with the encoding context, resulting in impaired performance.

However, results from the pilot study may also be explained without referring to relational processing. Specifically, change noise was increased in the maximal color change condition when distractor probes changed their colors. Observers had to differentiate $N$ changes from $N - 1$ changes, where $N$ is the number of colors presented. In the minimal color change condition, observers needed to differentiate one change from zero changes. The former task was more difficult because the signal had to be differentiated from a larger noise baseline (Stevens, 1975). Therefore, the pilot study did not provide unequivocal support to the relational processing hypothesis.

To demonstrate relational processing between items in the color VSTM task, we carried out Experiment 1. In this experiment, we presented eight target colors on the memory display. The critical probe was cued by an outline box on the probe display. There were two types of probes (see Figure 1). In the minimal change condition, the items not cued by the box maintained their previous colors and locations on the probe display. In the *single probe* condition, only the critical probe was presented in the probe image. Relational information was thus absent in the probe display. How would observers represent the colors in this task?

Observers could memorize each color independently of other colors. Because both probe tasks required observers to discriminate one change from zero changes, the two probe conditions were equivalent in the amount of signal strength and baseline noise. No performance difference should be expected if colors were represented independent from one another. However, observers might memorize each color in relation to other colors on the display. For example, they might encode a given color as "a red square sitting next to a yellow and a blue square." Of course, such encoding is visual in nature, not verbal. During the retrieval process, these relational cues were not present in the single probe condition. The absence of a retrieval cue should produce performance impairment in the single probe condition. Poorer performance in the single probe condition would
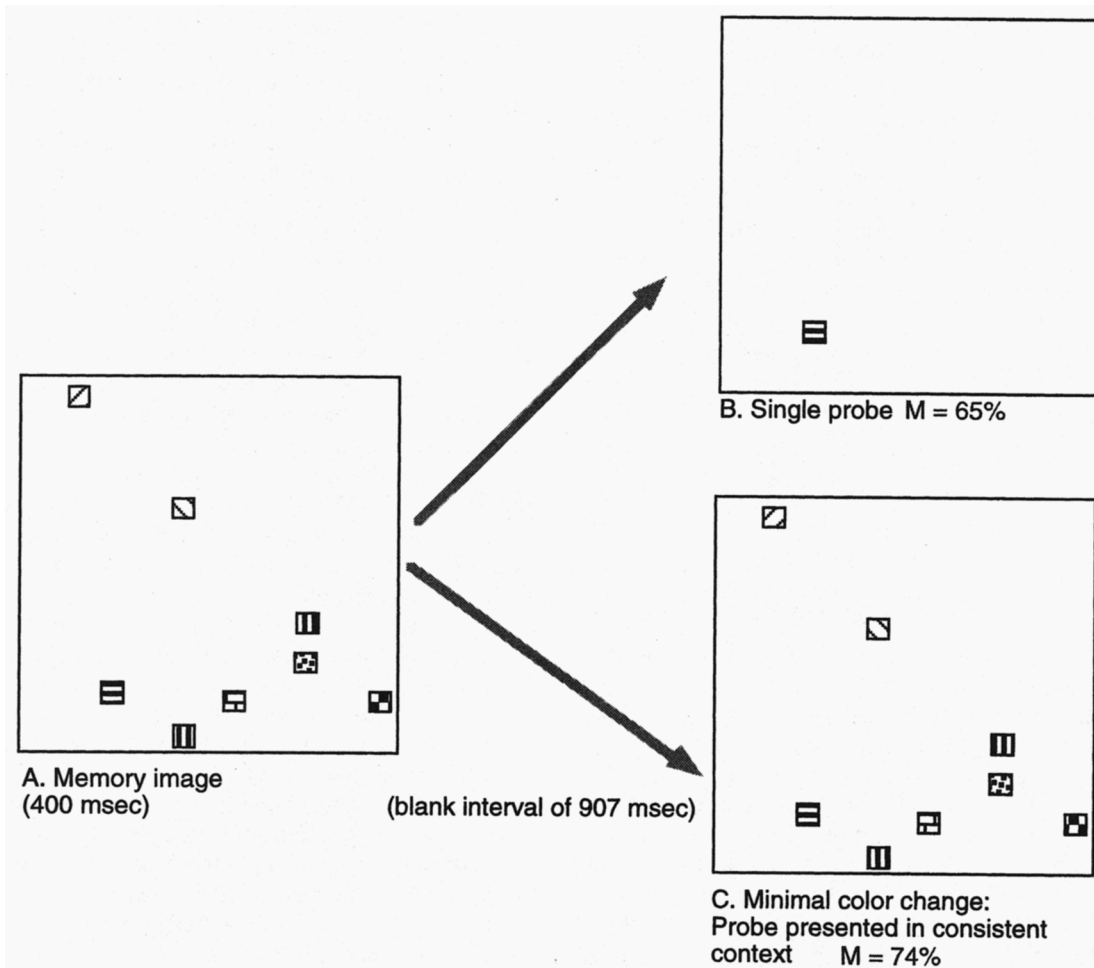
*Figure 1.* Samples of displays used in Experiment 1. Different textures on different squares represent the different colors used in the experiment. In the actual experiment, the texture of every item was smooth; items only differed in color.

provide unambiguous support for the relational processing hypothesis.

### Experiment 1

#### Method

*Participants.* In all of the experiments reported in this article, participants were recruited from a Yale University graduate and undergraduate participant pool. Their ages ranged from 18 to 30 years. All had normal or corrected-to-normal visual acuity and normal color vision. Some naive observers participated in more than one experiment, but none of them were aware of the experimental design or of our hypothesis. Yuhong Jiang participated in all but the first experiment. Her performance was not noticeably different from the performances of naïve observers.

In Experiment 1, 8 naive observers were tested.

*Materials and procedure.* On each trial, two images were presented, separated by a blank interval. The memory image contained eight colors; the probe image contained either eight colors (minimal color change condition) or a single probe. The critical probe was cued by a white outline box (1.0 cm × 1.0 cm) in both conditions. Each color square (0.8 cm × 0.8 cm) was

randomly positioned in an invisible 10 cm × 10 cm matrix that subtended 15.6 cm × 15.6 cm. The position of each square inside a cell was slightly jittered. The eight colors were produced by permutation of the red (absent = 0 or present = 255), green (0 or 255) and blue (0 or 255) phosphors. The background was gray (127 on red, green, and blue phosphors). On a particular display, repetition of colors (e.g., two instances of red) was allowed, with the restriction that the same color could not repeat more than three times on each display.[1]

---

[1] The reason for allowing some repetition in this experiment was to discourage a "repetition detection" strategy. Because the colors came from eight different colors, a nonrepeated display on the memory image contained all eight different colors. When a change happened for such displays, it necessarily contained a repetition of a color. Participants could just count whether a repetition had occurred on the probe display to perform perfect change detection even if they did not remember any memory colors. To prevent such strategies, every trial contained at least one color repetition in the memory image. But the amount of repetition of the same color was small (at most three times). Grouping cues were not salient for the displays.

Participants initiated each trial by pressing the space bar. Each trial started with a white fixation cross (0.2 cm × 0.2 cm) lasting 507 ms. The memory image was then presented for 400 ms and erased. After a blank interval of 907 ms with the gray background, the probe image was presented until the participant responded. These time parameters were chosen to be similar to those used by Luck and Vogel (1997).[2] Sound feedback concerning the accuracy of a trial was provided immediately after response. We instructed the observers that they should try to respond as accurately as possible without worrying about their speed. Nevertheless, we recorded both accuracy and response time (RT).

Four participants carried out 16 practice trials and one block of 80 experimental trials (80 = 2 probe conditions × 2 change conditions × 20 cases). Another 4 participants carried out 16 practice and two experimental blocks (each block with 80 trials). Trials were presented in random order.

*Apparatus.* The experiment was conducted on a Macintosh computer with a 17 in. (43 cm) screen. The experiment was programmed with MacProbe software (Hunt, 1994). Participants were tested individually in a room with normal interior lighting. The unrestricted viewing distance was about 57 cm, at which 1 cm corresponds to a 1° visual angle.

### Results

In this and later experiments, we calculated several measures for each condition: (a) overall accuracy, (b) memory sensitivity ($d'$), (c) response bias ($\beta$), (d) hit rate, (e) false alarm rate, and (f) RT. We report statistical analyses performed only on the first three measures (accuracy, $d'$, and $\beta$). Mean values of hit and false alarm rates are presented in the Appendix.[3] RT measures are not reported because our tasks were unspeeded. RT will be mentioned only when other measures showed null results in the manipulation or when a speed–accuracy trade-off occurred, which only happened in one session of Experiment 3.

The mean accuracy in the minimal color change condition was 74.4% ($SE = 3.4\%$). The mean accuracy in the single probe condition was 64.5% ($SE = 3.7\%$). The difference between these two conditions was significant, $t(7) = 3.11$, $p < .017$. Because the observers might adopt different response criteria for the two conditions, we calculated their memory sensitivity and response bias for each condition, on the basis of signal detection theory (Gescheider, 1976). Sensitivity was significantly higher in the minimal color change condition ($d' = 2.13$, $SE = .38$) than in the single probe condition ($d' = 0.85$, $SE = .21$), $t(7) = 2.54$, $p < .039$. In addition, observers adopted a significantly more liberal response criterion in the single probe condition ($\beta = 0.72$, $SE = .37$) than in the minimal color change condition ($\beta = 2.12$, $SE = .32$), $t(7) = 3.93$, $p < .006$. The sensitivity measure indicated that memory was more accurate in the minimal color change condition than in the single probe condition. The response bias measure showed that observers were more liberal in reporting a change in the single probe condition.

### Discussion

We manipulated distractor probes in this experiment. VSTM for the critical probe was clearly affected by the fate of other probe items in the context. Both the pilot study and Experiment I showed that VSTM performance was best when the distractor probes remained the same as in the memory image. Performance was severely impaired as the distractor probes changed colors from the memory image to the probe image. Such impairment may be partly attributed to increased noise baseline (Weber's law; Gescheider, 1976). However, unless some relational encoding and retrieval process is also assumed, the results from Experiment 1 cannot be completely explained by increases in noise baseline. Performance was impaired when only a single probe devoid of any change was presented. Thus, some relational information is processed in color VSTM.

How should we characterize such relational information? Similarity grouping is certainly not the mechanism at work, although it may have an additional effect when present. In our experiments, we tried to minimize repetitions of colors in the displays to reduce color grouping. In the absence of grouping cues, the only cue available would be the spatial configuration of the items. We now turn to experiments that examined the role of spatial configuration in VSTM.

### Section 2: Location VSTM Is Based on Spatial Configurations

Before examining whether global spatial configuration has any effect on color VSTM, we need to establish that spatial configuration is important for VSTM of locations. After all, both locations and configurations are inherently spatial. Location and configuration may be more closely linked than color and configuration.

Past research on VSTM and transsaccadic memory has shown that location short-term memory may be closely tied to configuration (Irwin, 1991; Phillips, 1974). For example, Phillips asked participants to memorize a spatial pattern produced by random placement of multiple dots in a square matrix. On the probe image, the whole pattern could shift its center position. Memory performance was contrasted between the conditions in which the pattern did or did not shift its center position. Results showed that participants were able to ignore the movement of the pattern unless the interval between the two images was too short (e.g., 50 ms). When iconic memory did not contribute to performance, the movement of the pattern did not affect memory accuracy.

---

[2] Unpublished experiments from our lab showed a similar pattern of results using shorter memory presentation (e.g., 100 ms) or longer retention period (e.g., 1,500 ms), suggesting that the particular time parameters used here were not critical to the findings. Luck and Vogel (1997) have also shown that several properties of VSTM (e.g., capacity, object benefit) are not affected by variation of the temporal parameters within a certain range.

[3] There are two reasons why we opted not to report detailed analysis on hit rate and false alarm rate. First, information contained in hit and false alarm is redundant with sensitivity and response bias results; and second, when observers adopt different response biases, hit rate (and false alarm rate) in different conditions cannot be directly compared. However, because hit and false alarm rates provided more direct sense of the results than $d'$ and beta, we attached the mean values in the Appendix.

This suggests that spatial locations can be encoded independently of their absolute spatial position in VSTM. Similar conclusions were obtained in a saccadic eye movement study, demonstrating that transsaccadic memory and VSTM share similar properties (Carlson-Radvansky & Irwin, 1995; Irwin, 1991).

Clearly, these studies suggest that spatial locations can be memorized in a relational manner when observers are instructed to memorize the whole pattern of locations. However, it is not clear whether observers can ignore the configuration information that does not form the memory target in a task or when configuration encoding is actually detrimental to performance. If we show that spatial configuration is utilized even when it may be detrimental to a task, this would suggest that spatial-configuration encoding is obligatory for VSTM.

We carried out two experiments to examine the role of spatial configuration in VSTM of individual spatial locations. In both experiments, observers were instructed to memorize individual spatial locations. They were told that the configuration produced by the items should be ignored because it may impair memory accuracy of individual locations. On each trial, participants saw several green squares presented at random locations. This image was quickly erased, and, after a short blank interval, another image of green squares was presented. On the probe image, one square was cued by a red outline box. The task was to decide whether this target square was positioned at a previously empty new location, or a previously occupied old one.

Experiment 2A contrasted three probe types: minimal location change, maximal location change and single probe. Figure 2 shows a sample display of these probe conditions. The global spatial configuration was preserved across the two displays in the minimal location change condition but was absent in the probe display of the single probe condition. Also, a new (potentially disrupting) configuration was presented in the maximal location change condition. If location VSTM is based on configuration even when the task does not involve direct configuration comparison, performance should be disrupted in the maximal location change condition. In addition, the absence of a consistent configuration should lead to poorer performance in the single probe condition compared with the minimal location change condition.

Experiment 2B tested three probe types, two of which were the same as in Experiment 2A: minimal location change and single probe. The third type was a *partial probe*, in which, on the probe display, half of the original memory items simply disappeared while the other half stayed on the screen. This condition was added to reduce the location uncertainty of the single target probe. The remaining items in the partial probe served as a reference frame for the critical probe. However, because half of the items disappeared, the global configuration of the partial probe display differed markedly from the original configuration. Thus, if configuration is key for VSTM of spatial locations, we should find poorer performance in the partial probe condition than in the minimal location change condition.

## Experiment 2A

We manipulated two factors in this experiment: probe condition and memory set size. As described above, there were three probe conditions: minimal location change, single probe, and maximal location change (see Figure 2). The memory set size was varied from 3 to 12 items. Trials in different conditions were randomly mixed within a single test session.

### Method

*Participants.* Six naive observers and Yuhong Jiang participated in this experiment.

*Materials and procedures.* Materials and procedures similar to those in Experiment 1 were adopted here, unless otherwise noted. The memory image contained 3, 6, 9, or 12 green squares (0.8 cm × 0.8 cm) presented at random locations. The probe image had three conditions (See Figure 2). In the single probe condition, one green square enclosed by a red outline box (1.0 cm × 1.0 cm) was presented. In the other two conditions, an equal number of squares as in the memory image were presented. One of the squares was enclosed by a red outline box. The others either remained at previously presented locations (minimal location change) or were positioned at randomly selected, previously blank locations (maximal location change). The task was to decide whether the square with a red outline box was at a new location (50% of the trials) or a previously occupied location.

Participants performed 16 practice trials and 480 experimental trials (480 = 4 memory set size × 3 probe types × 2 change present or absent × 20 cases). Locations were randomly chosen on each trial. Trial order was randomly mixed. Other aspects of the experiment were the same as in Experiment 1.

### Results

Figure 3 shows the mean accuracy as a function of memory set size and probe type. We found a significant main effect of probe type (maximal location change, minimal location change, or single probe), $F(2, 12) = 40.72$, $p < .001$, with the best performance in the minimal location change condition and worst performance in the maximal location change condition; as well as a significant main effect of memory set size, $F(3, 18) = 17.37$, $p < .001$, with lower performance at larger set sizes. The interaction between probe and set size was not significant, $F(6, 36) = 0.67$, *ns*.

To ensure that the above accuracy difference was not just an artifact of different response criteria in the different probe conditions, we calculated sensitivity and response criterion for each probe condition. Because set size factor did not interact with probe condition, we pooled across the set size factors when calculating false alarm rate (reporting a change when the critical probe did not change) and hit rate (correct detection of the critical probe change). Sensitivity was the worst in the maximal location change condition, $d' = 0.60$, $SE = .11$. It improved in the single probe condition, $d' = 1.50$, $SE = .23$, and was best in the minimal location change condition, $d' = 2.06$, $SE = .27$. These sensitivity values were significantly different from each other, $ts(6) > 2.83$, $ps < .03$. In the response criterion measure, we found a trend
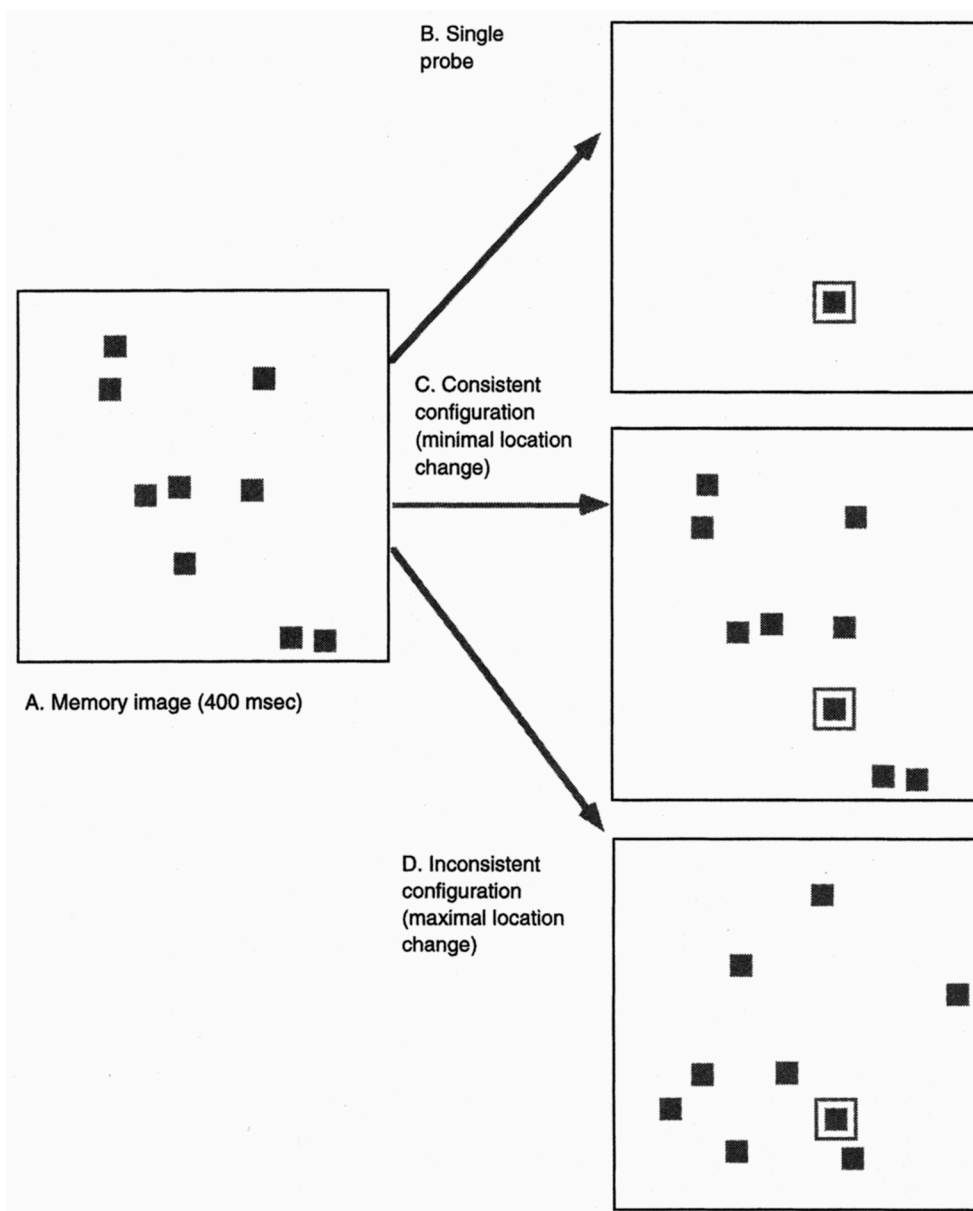
*Figure 2.* Samples of displays used in Experiment 2A.

toward more stringent change detection responses, from the maximal location change to the single probe to the minimal location condition. The beta values were 0.82 ($SE$ = .07), 1.58 ($SE$ = .42), and 3.41 ($SE$ = 1.19) for the maximal location change, single probe, and minimal location change conditions, respectively. However, the differences in the beta values failed to reach statistical significance, $ts(6) < 2.14$, $ps > .07$.

## Discussion

Even when observers were not required to memorize the spatial configuration of items, their VSTM of individual locations was apparently affected by the consistency of global configuration between encoding and retrieval. An inconsistency of configuration between the memory and probe images produced significant impairment in VSTM of an individual location. The degree of impairment was surprising. Performance was below 70% even when the memory set size was as small as three items in the maximal location change condition. The severe drop in performance may be partly attributable to increased baseline noise in that condition. However, the whole pattern of results cannot be explained solely in terms of increased noise, because performance was also impaired in the single probe condition in which no additional noise was introduced.

Results from this experiment suggest that the exact location of a single item cannot be encoded in absolute spatial coordinates efficiently; its encoding is wedded to spatial context information. The memory for the configura-
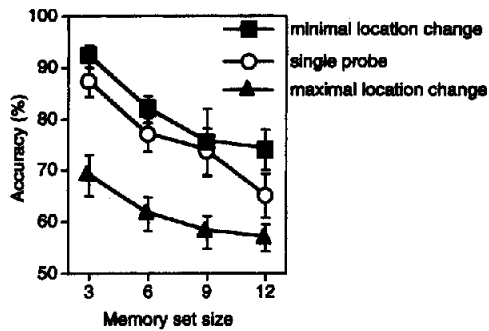
*Figure 3.* Spatial visual short-term memory location: Mean accuracy as a function of probe type and memory set size in Experiment 2A.

tion can be maintained with consistent cues in the minimal location change condition. It is also largely preserved in the single probe condition where no cues are presented, but is severely disrupted by a new configuration.

## Experiment 2B

A potential criticism for the single probe versus minimal location change condition comparison is that targets in the single probe condition lacked spatial reference cues. This may result in poorer localization. Other items may have served not as configuration cues, but as anchoring reference points. It follows that if we had provided anchoring reference points, performance for the single item might not be impaired at all. Although we believe that at least one function of configuration is to provide an accurate reference frame, we nevertheless decided to test whether any anchoring points are sufficient for accurate VSTM retrieval.

To test this idea, we tested three probe conditions: minimal location change, single probe, and partial probe. The condition of central interest was the partial probe condition (see Figure 4 for an example). In this condition, half of the items originally presented on the memory image disappeared on the probe image. The other half preserved their locations. Therefore, this condition contained no irrel-

evant location changes. Moreover, the old locations should serve as anchoring points to accurately localize the critical probe. However, if global configuration is the key, then performance in the partial probe condition should suffer. Because of the absence of half of the items, the configuration of a partial probe is phenomenally different from the original configuration.

The configuration hypothesis thus predicts that performance should be poorer in the partial probe condition than in the minimal location change condition. In addition, performance may also be poorer in the partial probe condition than in the single probe condition, because the single probe did not present a new configuration, which may disrupt performance. However, if any anchoring is sufficient, performance in the partial probe condition should be better than in the single probe condition and could even be as good as in the minimal location change condition.

## Method

All aspects of the method were identical to Experiment 1 except where noted. Eight observers were tested (7 naive observers and Yuhong Jiang). The first image contained eight locations. The second image contained one (single probe), four (partial probe), or 8 (minimal location change) locations. The critical probe was always cued by a red outline box on the probe image. Distractor probes always remained in the same locations as in the memory image. In the minimal location change condition, all seven distractor probes were presented on the probe image. In the partial probe condition, only three randomly chosen distractor probes were presented at their previous locations. The other four were not presented. In the single probe condition, all seven distractor probes disappeared. Each observer was tested in a single session of 16 practice trials and 120 experimental trials (120 = 3 probe types × 2 change present and change absent × 20 cases).

## Results

Results supported the configuration hypothesis rather than the anchoring hypothesis. Mean accuracy was highest in the minimal location change condition ($M = 84\%$, $SE = 2\%$). It declined in the single probe condition ($M = 72\%$, $SE = 4\%$). Accuracy was lowest in the partial probe condition
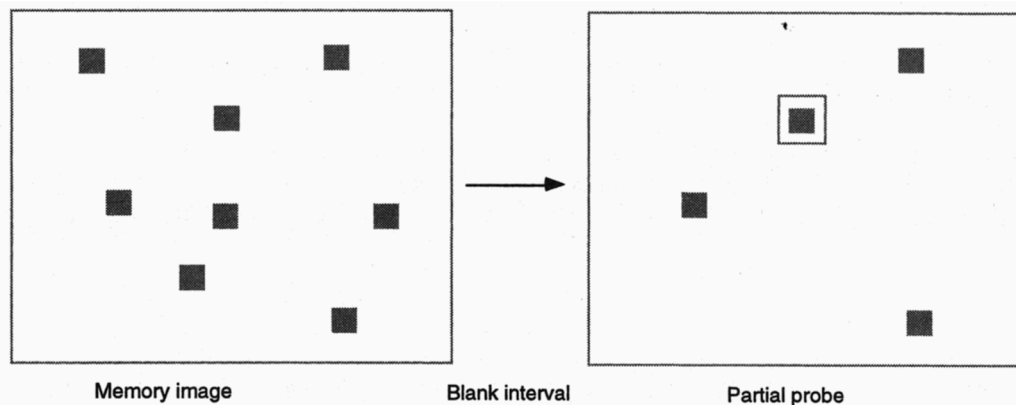


*Figure 4.* Samples of the partial probe condition used in Experiment 2B.

($M = 65\%$, $SE = 4\%$). The main effect of probe type was significant, $F(2, 14) = 6.21$, $p < .012$. In post hoc comparisons, we found that the minimal location change condition had significantly higher accuracy than both the single probe, $t(7) = 3.27$, $p < .014$, and the partial probe, $t(7) = 3.73$, $p < .007$, conditions. The 7% accuracy difference between the single probe and partial probe conditions was not significant, $t(7) = 1.03$, $p > .33$.

To examine whether the accuracy impairment in the partial probe condition reflected real memory-sensitivity impairment or response bias, we calculated sensitivity and response bias for each of the three probe conditions. Sensitivity decreased in the following order: minimal change ($d' = 2.69$, $SE = .50$), single probe ($d' = 1.81$, $SE = .63$), and partial probe ($d' = 1.05$, $SE = .21$). Pairwise comparisons between the three probe conditions showed a significant difference in sensitivity between partial probe and minimal change, $t(7) = 3.08$, $p < .018$. The sensitivity of the single probe condition failed to differ from the other two conditions significantly, $ts(7) < 1.05$, $ps > .30$. In the response bias measure, although observers became more likely to report that a change was detected as they went from the minimal change condition ($\beta = 2.24$, $SE = .39$) to the single probe condition ($\beta = 1.74$, $SE = .36$) to the partial probe condition ($\beta = 1.40$, $SE = .22$), these beta-rate differences failed to reach statistical significance, all $ts(7) < 1.58$, $p > .16$.

*Discussion*

It is clear that the VSTM of locations is configuration based. Although configuration may provide a reference frame to localize the critical probe, Experiment 2B indicates that the presence of a few reference-anchoring points is not sufficient for VSTM of individual locations. In addition, short-term memory of a single location is disrupted when the input configuration is different from the memory configuration. This suggests that global configuration information is encoded in VSTM for locations. This finding is consistent with that from Phillips's (1974) and Irwin's (1991) studies. Nevertheless, our results provide a stronger test of the effect of spatial configuration in VSTM of locations. Specifically, our participants were explicitly instructed to memorize each individual location and ignore the global configuration change. Although observers in our task knew that only a single location was to be compared, they were not able to register each location independently of the others.

Why did observers utilize global configuration even though they were instructed to represent individual locations? We believe there may be two reasons. First, items provide rich reference points for each other. By encoding the global configuration, item localization within that configuration becomes more accurate. Second, spatial configuration allows "chunking" of locations. Instead of encoding each location separately as an independent unit, several locations can be grouped together by forming a global pattern. Because VSTM has severe capacity limitations, the configuration-chunking strategy may greatly enhance processing efficiency within VSTM capacity limits.

The importance of configuration information in the change-detection paradigm provides us with a tool to examine how VSTM organizes other visual features. We now return to colors and examine whether the VSTM of features other than location is also configuration based.

## Section 3: Color and Shape VSTM Are Based on Spatial Configurations

Section 1 demonstrated that relational information was encoded in color VSTM. How should we characterize such relational information? Color grouping is one type of cue that can be used to process the relations between items. When several items have the same colors, they may be grouped together. Such redundancy can improve color VSTM. In our experiments, however, grouping information was absent because color redundancy was minimized by reducing repetitions of colors. In the absence of grouping cues, the only cues available are spatial cues.

In the traditional color change detection paradigm (e.g., Luck & Vogel, 1997), color and location must be accurately conjoined. To detect a color change correctly, observers have to remember not only what the colors are but also where they are. However, it is not clear how color–location conjunctions are formed. There are two possibilities. First, people may just bind individual colors with their corresponding absolute locations. Second, people may form a spatial configuration of all items and bind colors to different parts (points) of this configuration. Both methods are sufficient for performing the task.

On the basis of the results from Section 2, it seems that configuration binding is more likely to be used in the color-change-detection task. As we have seen in Section 2, relational configuration processing greatly enhanced location memory. In the color change detection task, because people have to remember location information, it may be beneficial for them to form a spatial configuration of the items. Colors are then "glued" to different parts of this configuration.

To investigate the effect of configuration in VSTM of colors or shapes, we designed a task in which the exact location of each individual item was not a task requirement. In the color session, for example, we presented four colors located in four different quadrants in both the memory and the probe images. Participants' task was to decide whether colors in the same quadrant had changed or not. In this task, participants were required only to memorize in which quadrant each color was presented (e.g., green in the upper left quadrant). They were explicitly instructed that the exact location of the color inside a quadrant was not important. There were three probe types. In the same location condition, locations across the two images were exactly the same. In the same configuration condition, the configuration of the four items remained the same but the exact location of each item changed because the global layout was expanded. In the different configuration condition, items moved locations in such a way that the whole configuration was changed. A sample of the displays is shown in Figure 5. If VSTM of colors is based on configuration, we should see disrupted
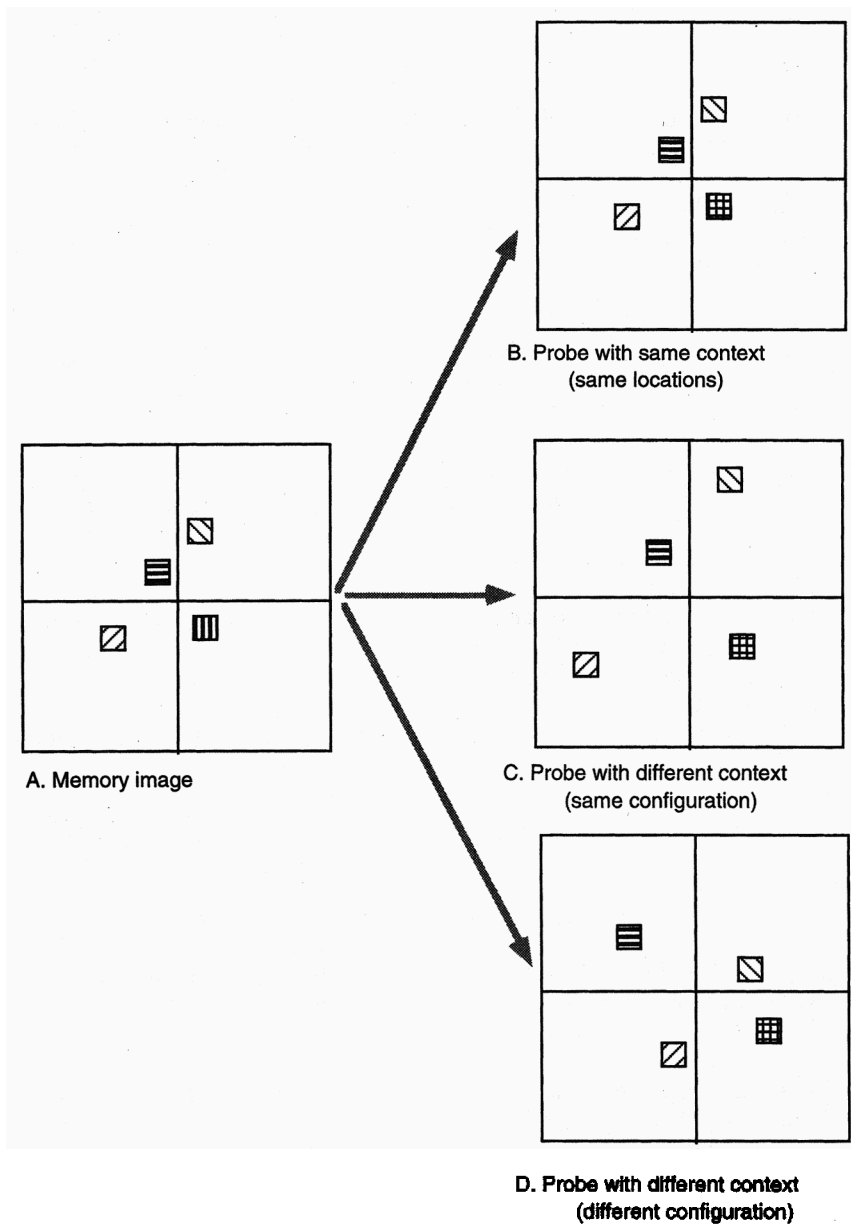
*Figure 5.* Samples of displays used in Experiment 3. Different textures on different squares represent the different colors used in the experiment.

performance in the different-configuration condition compared with the other two. If color–location binding is carried out for each location separately, we should find equivalent performance in the same-configuration and the different-configuration conditions. The same design was used in the shape session, except that colors were replaced by geometric shapes and observers performed a shape change detection task.

The effect of spatial configuration on the short-term memory of pictures has been tested in previous studies (Santa, 1977; Zimmer, 1998). In Santa's study, for example, the spatial configuration of three geometric figures was transformed from a triangular configuration in the memory image to a linear array in the probe image. This transformation resulted in significantly slower RTs to detect a shape change. Zimmer replicated this result using pictures and extended the results to word stimuli. These earlier studies suggested that we should find a difference between the same-location and the different-configuration conditions.

Nevertheless, our experimental design has several improvements over the earlier studies. First, in Santa's (1977; see also Zimmer, 1998) study, spatial transformation had no effect on accuracy because accuracy was very high in general. Our study used a larger memory set size (four items instead of three). We expected this larger set size to push performance below ceiling, thereby allowing us to observe

the effect of configuration on memory accuracy. Second, the configurations used in earlier studies were simple and familiar (e.g., triangular or linear arrays). In addition, there was little variation of configuration across trials; the same triangular and linear configurations were used again and again. In our experiments, configurations were more complex and had much greater variation. It is important to test whether people can utilize such random and complex configurations to assist VSTM of identities.

Lastly, the addition of the same-configuration condition in our design allowed us to tease apart the effect of individual-item location change and global configuration change. In earlier studies, spatial transformation always resulted in both configuration change (from triangular to linear or vice versa) and item location change (one object moved from the upper left location to the middle horizontal location, for instance). Thus, the influence of spatial transformation could not be attributed unambiguously to a configuration effect. In our study, the same-configuration condition involved individual location change but not configuration change. By comparing performance in this condition with the same-location and the different-configuration condition, we could separate the influence of configuration change and individual location change.

## Experiment 3

### Method

The materials and procedure were the same as Experiment 1 except where noted. Five naive observers and Yuhong Jiang performed the task. There were two sessions: a color VSTM session and a geometric shape VSTM session. The set of geometric shapes used in this experiment is shown in Figure 6A. In each session, four items (color squares in one session and white geometric shapes in another session) were presented on both the memory and the probe images. They were randomly located in an invisible 8 by 8 matrix with the constraint that each item occupied a unique quadrant. The size of the memory image was 14.1 cm × 14.1 cm. A dim cross (gray scale 150; 26.25 cm × 26.25 cm) was always presented at the center of the display to serve as a reference frame. Instead of detecting whether a color change occurred at each exact location, participants were required to detect a color change in each quadrant.[4] They were specifically instructed that the colors might migrate within a quadrant from the two images and that they should ignore such location change.

In the same location condition that comprised one third of the trials, items were located in exactly the same locations across the two images. The remaining trials had different locations. For half of the different location trials, the configuration of the memory image was expanded by a factor of 1.70. The size of each individual item was constant. This was the same-configuration condition. The rest of the trials had different configurations produced by changing item locations randomly within each quadrant. The magnification factor (1.70) and search matrix size (8 × 8 cm) were chosen to equate the expected amount of absolute location change between the same-configuration and different-configuration conditions. The algorithm to calculate expected location change in the two conditions can be obtained from us. Observers performed 16 practice and 120 experimental trials in each session (120 = 3 probe types × 2 change present or absent × 20 cases). The order of sessions was counterbalanced across participants.
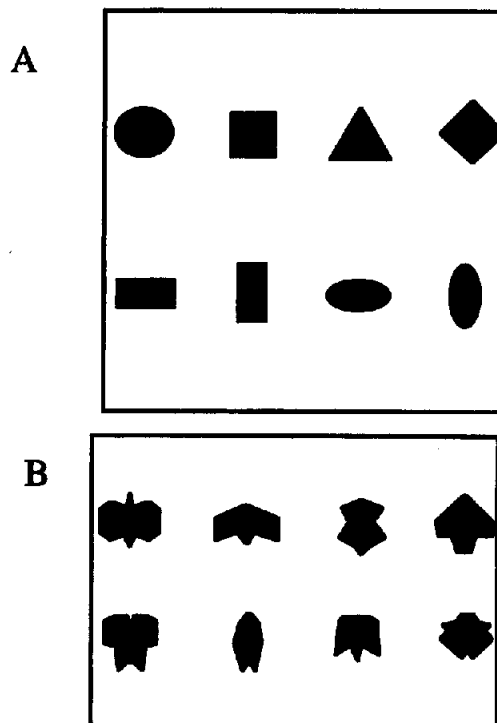


*Figure 6.* Figure 6A shows geometric shapes used in Experiment 3. Results from Experiment 3 can be generalized to novel shapes shown in Figure 6B.

### Results

Figure 7 shows the mean accuracy data. The pattern of results supports the configuration-binding hypothesis over the individual binding hypothesis.

In the color session, accuracy was significantly affected by probe type, $F(1, 5) = 16.36$, $p < .001$. Accuracy was lower in the different-configuration condition than in the same-location, $t(5) = 2.93$, $p < .033$, and the same-configuration, $t(5) = 5.86$, $p < .002$, conditions, suggesting a configuration effect. This drop in accuracy could not be attributed to individual location movement per se. Although the same amount of individual location movement occurred in the same-configuration condition, accuracy in that condition was not impaired compared with the same-location condition. In fact, accuracy was significantly higher in the same-configuration than in the same-location condition, $t(5) = 2.74$, $p < .041$. This result was probably produced by a speed-accuracy trade off between these two conditions. For the same-location and same-configuration conditions, RT was 845 ms and 918 ms, respectively. This difference was significant, $t(5) = 3.08$, $p < .028$, suggesting that a speed-accuracy trade off had indeed occurred for these two conditions.

---

[4] In a follow-up study, we found the same pattern of results when the critical probe was indicated by an outline box cue. Thus, the findings obtained in this experiment were not produced by any differences across conditions in the decision stage.
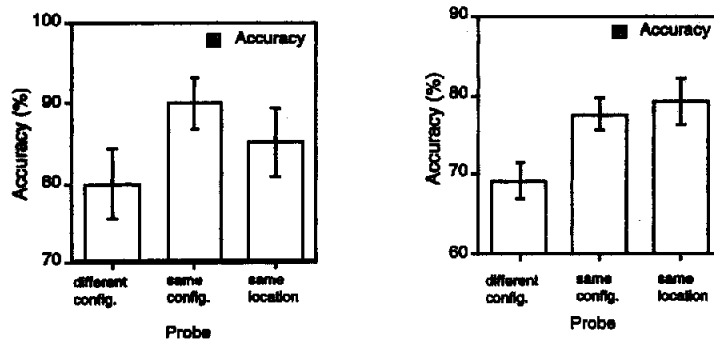
*Figure 7.* Color (left panel) and shape (right panel) visual short-term memory: Mean accuracy as a function of probe type in Experiment 3. config = configuration.

We computed sensitivity and response bias for each probe condition. These measures are shown in Table 1. Pairwise comparisons show that sensitivity was lower in the different configuration condition than in the other two conditions, $ts(5) > 2.90$, $ps < .034$. Sensitivity in the same-configuration and same-location conditions did not differ, $t(5) = 2.27$, $p > .07$. In terms of response bias, observers were more liberal to report a color change in the different-configuration condition than in the same-location condition, $t(5) = 4.18$, $p < .009$. The beta value of the same configuration condition did not differ from the beta values of the other two conditions significantly, both $ts(5) < 1.85$, $ps > .10$.

In the shape session, accuracy was also significantly affected by probe type, $F(1, 5) = 12.04$, $p < .002$. Pairwise comparisons showed that accuracy was significantly poorer in the different-configuration condition than in the other two conditions, $ts(5) > 3.65$, $ps < .015$. The difference between the same-location and the same-configuration conditions was not significant, $t(5) = 1.00$. No speed–accuracy trade offs were present.

Table 1 shows the sensitivity and response bias measures. We found that memory sensitivity was significantly lower in the different-configuration condition than in the other two conditions, $ts(5) > 3.35$, $ps < .02$. Memory sensitivity was comparable in the same-configuration and same-location conditions, $t(5) = .49$, $p > .30$. In the response bias measure, the beta values of all three conditions were comparable, all $ts(5) < 1.87$, $ps > .10$.

## Discussion

When participants had to detect whether the color or the shape of an item had changed within a quadrant, changes in

global configuration between memory and probe images disrupted their performance. This result clearly demonstrates that the VSTM of colors and shapes is based on the spatial configuration of memory items even when participants are not required to encode the exact location of each item. In fact, performance was not affected by location change of individual items, as long as the global configuration of all of the items was preserved. This second finding suggests that VSTM of colors and shapes is not based on individual locations. What is important for VSTM of color and shape is the spatial configuration of items. Similar results have been obtained in our lab with novel shapes as the memory attribute (see Figure 6B).

These three experiments not only confirm the finding from Section 1 that relational information is encoded in VSTM of random colors, they also point to the nature of such relational cues: spatial configuration. This is an interesting organizational cue given that spatial configuration is an emergent property of individual items. Encoding of configuration is immediate. For example, using visual search tasks, Chun and Jiang (1998) found that spatial configuration can be extracted and matched to memory within the first few hundred milliseconds of visual presentation. Configuration encoding does not require any Gestalt grouping cues or long term schemata, although these cues may have an additional effect when present.

These results fit well with the object-recognition literature that shows configural effect in recognition of novel objects and faces (Carlson-Radvansky, 1999; Gauthier & Tarr, 1997; Palmer, 1977; Tanaka & Farah, 1993). These researchers presented objects composed of several connected parts. When the configuration of the global object changed,

Table 1

*Sensitivity and Response Bias in the Identity VSTM Task of Experiment 3*

| Condition | Color session: sensitivity ($d'$) | Color session: response bias ($\beta$) | Shape session: sensitivity ($d'$) | Shape session: response bias ($\beta$) |
|---|---|---|---|---|
| Different configuration | 1.85 (.4) | 1.22 (.1) | 0.88 (.2) | 0.82 (.1) |
| Same configuration | 3.78 (.8) | 2.04 (.7) | 1.69 (.1) | 1.28 (.2) |
| Same location | 3.42 (.8) | 3.19 (.5) | 1.75 (.2) | 1.41 (.3) |

*Note.* Standard errors are in parentheses. VSTM = visual short-term memory.

observers became less accurate at identifying parts of the object. This decreased accuracy suggests that configuration information is important for object identification. In addition, global shape exerts an effect on local-element identification (Navon, 1977). When observers have to identify the local letters that give rise to a global letter, their reaction time is slowed down if the global letter matches a different response than the local letter. Our results support these past findings. Nevertheless, our data go beyond a redemonstration of configural effects in part and local element identification. Unlike these other studies, the "parts" of the configurations in our experiments were not physically connected, nor were there other explicit grouping cues such as colinearity. Furthermore, our global configurations were randomly generated and did not share response mappings with local elements. Thus, our task may be more analogous to processes involved in global scene organization rather than individual object recognition. In both cases, however, configuration information is important.

In the next section, we examined configuration representations in more detail. Specifically, what information is contained in a configuration? Does the formation of configurations necessarily involve all items in a display?

## Section 4: Characteristics of Configuration: Components and Formation

### Experiment 4A: Components

The purpose of this experiment was twofold. First, it was designed to replicate the configuration effect found in Section 3 using a different paradigm. In Section 3, change detection was based on quadrants. This task requirement is different from most change detection tasks. Experiment 4A went back to the typical location-based change detection task. Second, we investigated whether relational information used in color VSTM is completely spatial. Suppose that a spatial configuration devoid of color was presented on the probe display. Would it support VSTM of colors just as well as when the original colors were attached to appropriate parts of the configuration?

To answer this question, we utilized three probe types. Both colors and locations were preserved across the memory and the probe images in the minimal-change condition. In the *faded, same-configuration* condition, locations were preserved across the two images, but color information was lost in the probe image because all distractors turned colors to dim gray. Spatial configuration was preserved in this condition, but the appropriate color cues were absent. Lastly, in the *faded, different-configuration* condition, distractor locations changed and distractors faded to dim gray. On the basis of results from Section 3, we expected that performance should be better in the minimal-change condition than in the faded, different-configuration condition. In addition, any difference between the two fade conditions would reveal pure effects of spatial-configuration changes. Lastly, any difference between the minimal-change and the faded, same-configuration conditions would reveal the effect of additional color cues.

### Method

Methods were similar to Experiment 1 except where noted. We tested 9 observers (8 naive observers and Yuhong Jiang) in this experiment. The first image contained four, six, or eight colors randomly positioned on the screen. The second image contained the same number of items. The critical probe item was cued by a white outline box. Participants' task was to detect whether a color change occurred at this cued location. In the minimal-change condition, the other, noncued items were exactly the same in color and location as in the first image. In the faded, same-configuration condition, all of the noncued items also remained at previously occupied locations, but their colors were all changed to gray (gray value was 140, the background was 127). Phenomenally, the colors of these items just seemed to fade away into dim gray. In the faded, different-configuration condition, all of the noncued items turned gray and were positioned at previously blank locations. Participants performed 16 practice and 360 experimental trials (360 = 3 set size × 3 probe conditions × 2 change present or absent × 20 cases).

### Results and Discussion

Figure 8 shows mean accuracy as a function of set size and probe type. The main effect of probe type was significant, $F(2, 16) = 15.68$, $p < .001$, with best performance in the minimal-change condition and worst performance in the faded, different-configuration condition. The main effect of set size was significant, $F(2, 16) = 23.09$, $p < .001$, with better performance at smaller set sizes. The interaction between these two factors was not significant, $F(4, 32) = 1.12$, $p > .36$.

In follow-up planned comparisons, we found a significant accuracy difference between the minimal-change condition and the faded, different-configuration condition, $F(1, 8) = 42.87$, $p < .001$. This result supports the claim that color VSTM is based on bindings to spatial configuration. When the configuration was changed, color VSTM was disrupted.

This configuration effect can be parsed into two components: a purely spatial configuration devoid of color cues and a configuration effect that encodes color cues. The comparison between the faded, same-configuration and the faded, different-configuration gives us a measure of the effect of spatial configuration in the absence of color cues. The
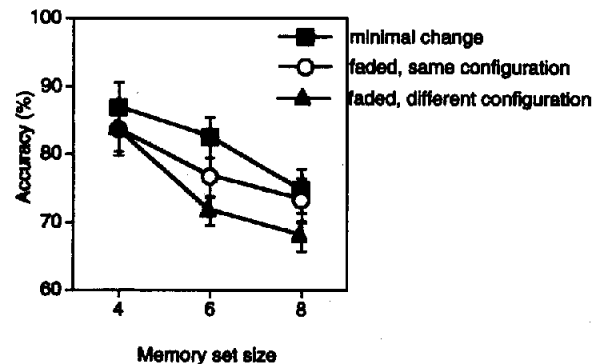


*Figure 8.* Color visual short-term memory: Mean accuracy as a function of probe type in Experiment 4A.

difference between these two conditions was significant in accuracy, $F(1, 8) = 6.00, p < .040$. In another words, when the spatial configuration was preserved, accuracy was higher. The effect of color cues in configuration encoding was measured by the difference between the minimal-change condition and the faded, same-configuration condition. Accuracy in the faded, same-configuration condition was significantly lower, $F(1, 8) = 8.15, p < .021$.

The difference between the three conditions was less clear when sensitivity and response bias were calculated, pooled across the set size factor. In the sensitivity measure, $d'$ values were 1.53 ($SE = 0.18$), 1.96 ($SE = 0.37$), and 1.94 ($SE = 0.20$) for the faded, different-configuration, faded, same-configuration, and minimal-change conditions, respectively. Only the difference between the faded, different-configuration and minimal-change conditions was significant for the $d'$ measure, $t(8) = 3.44, p < .009$. The large mean $d'$ value in the faded, same-configuration condition was boosted by an outlier ($d' = 4.42$) and should be treated with caution. The $d'$ in this condition was not consistently higher than in the faded, different-configuration condition, $t(8) = .50, p > .60$. In terms of response bias, observers were significantly more likely to report a change in the faded, different-configuration condition ($\beta = 1.66, SE = .27$) than in the minimal-change condition ($\beta = 0.68, SE = .15$), $t(8) = 4.08, p < .004$. The beta value in the faded, same-configuration condition was intermediate ($\beta = 0.98, SE = .61$) between the other two conditions, and it failed to differ significantly from either of the other two conditions.

Thus, the results in this experiment were somewhat weak in the sensitivity measure. However, accuracy showed a partition of the configuration effect into spatial and color identity components; thus, we tentatively concluded that both contribute to configuration effects in color VSTM.

### Experiment 4B: Formation

What constitutes a configuration? Is a configuration formed by all items presented on the display or can it consist of just a subset of items?

In this experiment, we provided both top-down and bottom-up cues for target items. The question of interest is whether distractor items influence VSTM of the target items. On both memory and probe images, four colored squares (excluding white) and 12 white squares were presented. The task was to detect a color change. Top-down attentional cues of the targets were provided by instruction. Participants were informed that only the four nonwhite items were relevant to the task. Bottom-up cues were provided by the physical difference between targets (color squares) and distractors (white squares). The color squares popped out from the uniformly white distractors. There were two types of probe displays. In the same-distractor-location condition, the locations of the 12 white squares were preserved across the two images. In the different-distractor-location condition, the 12 white squares changed locations across the two images. All of them were presented at previously blank locations on the probe image.

There were two critical differences between this experiment and the earlier experiments. First, distractors were physically similar to the target in the earlier experiments because of the high degree of color heterogeneity. In this experiment, color targets popped out from white, segregating the targets from distractors. Second, in earlier experiments participants were instructed to memorize all of the items on the memory image. The cue that indicated which item was the target and which were distractors was presented only on the probe image. In this experiment, however, distractors were marked starting from the memory image. The joint benefit of a clear physical distinction and a head-start attentional cue may help observers extract the relevant configuration. If the configuration of all items is obligatorily encoded, distractor location changes should disrupt performance. However, if configuration formation can be restricted to attended items that are also physically distinctive from distractors, performance should not be affected by distractor location changes.

### Method

Six observers (5 naive observers and Yuhong Jiang) were tested in 16 practice trials and 80 experimental trials (80 = 2 probe types × 2 change present or absent × 20 cases). Both memory and probe images had four target colors (excluding white) and 12 white distractor squares. The locations of the target colors were always the same across the two images. In the probe image, the white distractors could all remain at previous locations (same-distractor-location condition) or all change to new (previously) blank locations (different distractor location). The task was to memorize the four target colors and detect whether a color change occurred for one of the targets. All other aspects of the method were similar to Experiment 1.

### Results

In contrast to Experiments 3 and 4A, which showed a significant effect of distractor location change on color VSTM, we found no effect of distractor configuration changes in this experiment. Accuracy was similar between the same-distractor-location condition ($M = 85\%, SE = 4\%$) and the different-distractor-location condition ($M = 87\%, SE = 3\%$). The $d'$ for the same-distractor-location condition ($d' = 2.26, SE = 0.30$) and that for the different-distractor-location condition ($d' = 2.81, SE = 0.50$) were similar in magnitude. RT was comparable as well ($M = 805$ ms for same location and 809 ms for different locations). No statistical significance was observed for the accuracy, sensitivity, or RT measures ($ps > .15$). The only measure that approached significance was response bias. In this experiment, observers adopted a more stringent criterion to report a change in the different location condition ($\beta = 1.77, SE = 0.50$) than in the same location condition ($\beta = 0.91, SE = 0.19$), $t(5) = 2.37, p < .064$, suggesting that the change in distractor locations was registered. However, VSTM sensitivity was not affected by such changes, indicating that a subset of items can be utilized to form the relevant configuration.

## Discussion

These results suggest that configuration encoding is flexible because it can be based on the subset of items that are attended and distinctive from distractors. We should note, however, that such flexibility may be limited. Distractors in this experiment were not only unattended, but also had a homogeneous white color that was distinct from attended items. The physical distinction was preserved in both memory and probe images. Performance may suffer if this physical distinction is obscured. In fact, in a follow-up experiment in which we manipulated top-down attention (through precue) and bottom-up attention (through physical cue), we found that both top-down and bottom-up cues contributed to observers' ability to extract the relevant configuration.

In summary, configuration encoding was important for color VSTM (Experiments 3 and 4A). Colors were bound not to their individual locations, but to the configuration. Such configuration cues have two components: pure spatial configuration information and another component with color information tagged. Formation of the configuration does not have to be based on all items in the display. Rather, configuration can be restricted to attended items if these items are also physically distinctive from distractors in both the memory and the probe images. Both attentional precues and bottom-up cues are important for formation of targets' configuration.

## Section 5: How Does Color Affect VSTM of Locations?

Spatial configuration has an enormous effect on color VSTM. What about the reverse case? Does color configuration have any effect on location VSTM? The answer to this question has implications for the organization of VSTM. On one hand, VSTM may be completely object based. Processing one dimension of an object may result in automatic processing and organization of another dimension. For example, in color VSTM, location is processed and a configuration is formed. Conversely, in spatial location VSTM, color may also be processed and a color configuration may be formed. On the other hand, spatial configurations may have more central status in change detection tasks, producing an asymmetry in the interaction between space and color. Color VSTM is affected by spatial configuration, but spatial location VSTM may not be affected by color pattern.

There is an additional reason to examine how colors may affect VSTM of locations. It is not completely clear whether the configuration effect we found in the above experiments was simply a demonstration of the general encoding specificity effect (Tulving & Thompson, 1973). The encoding specificity theory predicts a decrement in memory performance when there is an inconsistency between encoding and retrieval. However, some inconsistencies may be more critical than others in VSTM. For example, configuration information is critical, and performance is impaired when there is configural inconsistency. But surface color may not

be critical when it is not the to-be-remembered attribute, so color inconsistency across encoding and retrieval may not produce interference. However, if the configuration effect is produced by general encoding specificity, we may find a memory impairment regardless of the type of inconsistency between encoding and retrieval.

We carried out two experiments to test the effect of color pattern change on memory of spatial locations. Random color change was tested in Experiment 5A and color grouping change was tested in Experiment 5B.

### Experiment 5A

In this experiment, both memory and probe images contained eight items at random locations. The critical probe on the probe image was cued by a white box. The task was to detect a location change of this item. Other items maintained their previous locations. Items on the memory image were all of different colors. On the probe image, items either maintained their previous colors or they switched colors with each other. As we have seen in Experiment 1, such random color change led to impaired color VSTM. But would it have an effect on VSTM of spatial locations?

### Method

Five naive observers and Yuhong Jiang were tested in this experiment. They carried out 16 practice trials and 80 experimental trials (80 = 2 probe type × 2 change present or absent × 20 cases). Eight colored squares were presented on each trial in both the memory and the probe images. Participants' task was to detect location change of the critical probe, which was cued by a white box on the probe display. On half of the trials, the color of each item stayed the same across memory and probe images. This was the minimal color change condition. On the other half of the trials, items switched colors with each other on the probe display so that the color of each item changed across the two images. This was the maximal color change condition. Other procedures were the same as in Experiment 1.

### Results

Observers' performance was not affected by random color change. Accuracy was 84.6% ($SE = 3\%$) in the minimal-color-change condition and 86.7% ($SE = 2\%$) in the maximal color change condition. This difference was not significant, $F(1, 5) < 1$. For the minimal- and maximal-change conditions, $d'$ values were 2.90 ($SE = 0.64$) and 2.40 ($SE = 0.11$), respectively, and these values did not differ significantly from each other, $t(5) = .89, p > .40$, suggesting that memory sensitivity was not affected by color change. Response bias was also comparable between the two conditions ($\beta = 1.49$, $SE = 0.40$ for the minimal color change condition; $\beta = 1.37$, $SE = 0.41$ for the maximal color change condition), $t(5) = 0.79, p > .40$. The same pattern of results was evident in RT measures, which were 958 ms ($SE = 132$ ms) in the minimal color change condition and 995 ms ($SE = 139$) in the maximal color change condition, $F(1, 5) = 1.17, p > .30$. These results stand in sharp contrast

to Experiment 1B, in which random color change produced severe impairment on color VSTM.

## Experiment 5B

In Experiment 5A, probe color change produced no effect on VSTM of locations, indicating that color pattern information was not used for VSTM of locations. However, color might still have an effect on location VSTM in appropriate situations. In this experiment we tested the idea that color grouping cues may affect location VSTM. In Experiment 5A, colors could not form groupings because each item had a unique color. In this experiment, we used only two colors. Half of the items in both memory and probe images were colored red, the other half green. In the *same-grouping* condition, the same set of items was colored green (or red) across the two images. In the *different-grouping* condition, half of the green (or red) items on the memory image turned red (or green) on the probe image. The latter case gave a strong percept of change in color grouping. The critical probe item was always cued by a white box, and it never changed color. Observers' task was to detect location change of this item. We also varied memory set size. The set size was 4, 8, or 12.

If color pattern has any influence on spatial location VSTM, such a grouping change is probably the most sensitive form of disruption that one can use, because color grouping exerts a spatial organization on the elements. As we have seen in Experiment 4, bottom-up cues can support selective configuration encoding when instructions encourage people to do so. However, under the instruction to attend to all items and ignore color changes, color grouping may not influence configuration encoding at all.

### Method

We tested 6 observers (5 naive observers and Yuhong Jiang) in this experiment on 240 experimental trials (240 = 3 set size × 2 probe types × 2 change present or absent × 20 cases). The memory set size had three levels: 4, 8, or 12. Half of the items were colored red; the other half were colored green. In the same-grouping

condition, the colors were preserved across memory and probe images. In the different-grouping condition, half of the red (or green) items on the memory image changed to green (or red) on the probe image. The critical probe was cued by a red box, and it never changed color. The task was to detect location change of the cued item. All other aspects of the experiment were identical to Experiment 1.

### Results

Results are shown in Figure 9. Color-grouping change had no effect on accuracy at all, $F(1, 5) = .01$, $ns$. It did slow down RT, but the difference was not statistically significant, $F(1, 5) = 2.05, p > .20$. The set size effect was significant in accuracy, $F(2, 10) = 25.36, p < .001$. The interaction between probe type and set size was not significant $(F < 1)$. When pooled across the set size factor, memory sensitivity in the same grouping condition $(d' = 2.64, SE = 0.53)$ did not differ significantly from that in the different grouping condition $(d' = 2.09, SE = 0.33)$, $t(5) = 1.17, p > .25$. Response bias was comparable between the two conditions as well ($\beta = 2.50$, $SE = 0.35$ in the same grouping condition; $\beta = 1.51$, $SE = 0.48$ in the different grouping condition), $t(5) = 1.67, p > .15$. Apparently, if irrelevant color grouping had an effect on spatial VSTM, it was too small to produce a statistically significant behavioral change. This finding does not mean that color grouping cannot be used to facilitate spatial VSTM. Under conditions in which color grouping is relevant, it may exert a strong influence on encoding of relevant spatial configuration. For example, if participants are instructed to attend to a subset of items, the color distinction between the attended and ignored sets should play a major role in VSTM processing, as we saw in Experiment 4B. However, under the current condition, in which all items had to be attended, irrelevant color grouping had negligible effect on VSTM of spatial locations.

### Discussion

In this section, we found that neither random color changes nor color grouping changes had an influence on VSTM of spatial locations. This finding stands in contrast to
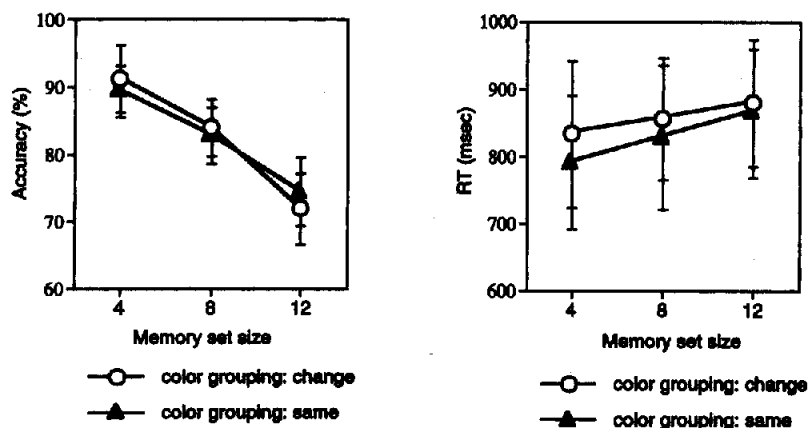


*Figure 9.* Location VSTM: Mean accuracy and response time as a function of memory set size and color grouping condition in Experiment 5B.

findings that random location change resulted in severe impairment of color VSTM. Thus, the influence between spatial location and color is asymmetric. In a follow-up study, we also observed similar asymmetric influence between spatial configuration and novel shapes, as shown in Figure 6B. That is, VSTM of novel shapes was affected by spatial configuration changes, but VSTM of spatial locations was not affected by element shape changes. This latter result suggests that the asymmetric influence between element identity and configuration was not restricted to color per se. Although both object shape and spatial configuration have spatial properties, they do not influence each other in a symmetric manner.

Why is there an asymmetric interaction between configuration and identities? Change detection tasks require correct localization of the target item. Although global spatial-configuration information is not necessary for any of the tasks here, configuration information may be registered in VSTM because it assists localization of each individual element. Configuration information also helped color and shape memory in Experiment 3, possibly because a configuration allows individual colors to be chunked together as a big unit. However, color or shape contour of an element does not aid element localization. Thus, the asymmetry may be attributable to the unequal amount of information provided by each dimension. We have not developed a method to quantify how much information a configuration or element identities or colors provide, but on intuitive grounds it seems plausible that spatial configuration provides more useful cues for encoding information into VSTM. At least in the paradigm used in our experiments (and that of most other VSTM studies), spatial configurations may play a more primitive role than other spatial attributes, such as color grouping.

## General Discussion

The experiments presented in this paper allow us to draw two general conclusions about VSTM. First, VSTM is labile and can be easily disrupted. Discrepancies between the memory image and the probe image impair VSTM retrieval. Section 1 demonstrated that irrelevant color changes disrupted color VSTM. The results indicate that relational information between items is encoded in VSTM. Supporting this view, memory performance was disrupted by a single probe display, which minimized changes in noise but impoverished the global organization of the display.

Our second conclusion is that spatial configurations form the basis of relational encoding in VSTM. To examine the nature of relational information, we first tested the effect of spatial configuration on VSTM of locations in Section 2. Section 2 established that location VSTM is based on spatial configuration. Disruption of the original memory configuration or the absence of configuration cues produced impaired retrieval of a target location. This conclusion from Section 2 is not surprising alone. It is intuitively appealing that VSTM of locations relies heavily on spatial configuration. The significance of these results is that they provided a basis for

exploring the nature of VSTM representations for other dimensions, such as color and shape.

In Section 3, we found a large impact of spatial configuration on color and shape VSTM. Section 4 demonstrated that representations of color configurations in VSTM are based on both a purely spatial component and another component that has color information bound to the configuration. Section 4 also showed that both top-down and bottom-up attentional cues affect the encoding of configurations. Lastly, Section 5 showed that the influence between spatial configuration and color is asymmetric. Neither random color change nor color-grouping change disrupts VSTM of spatial locations, suggesting spatial configuration is more primitive in our paradigm.

These findings support the second main conclusion of our study: VSTM organization is based on spatial configurations that specify not only the location of the target but also its relationship to other items in the display. This is true whether the to-be-remembered stimulus dimension is a spatial location or an identity attribute, such as color or shape. We offer the following conjectures on how VSTM is organized.

First, our findings demonstrated that organization of VSTM is configuration-based. When a visual image is presented, humans immediately form a spatial configuration of the items (see also Chun & Jiang, 1998). Items in VSTM are more closely bound to locations within a spatial configuration, rather than to absolute locations in space (Experiment 2B). Thus, information is not represented independently in VSTM, but rather in relation to other items appearing in the same spatial configuration. Items are represented within this context, and memory is impaired when such spatial organization cues are distorted or absent.

Context effects have a rich tradition in memory research, but most of this work has been limited to verbal memory (Riccio, Richardson, & Ebner, 1984; Tulving, 1974). However, it is instructive to see that the verbal and semantic context of a word plays an important role in verbal memory retrieval, as spatial configuration does for VSTM encoding and retrieval. For example, memory of a target word (e.g., *grade* in *good grade*) was affected by whether it was retrieved in the same context as the original encoding sentence (e.g., *good grade*) or in a different context (e.g., *steep grade;* Light & Carter-Sobell, 1970). The match or mismatch of environmental cues such as the place of learning also influences retrieval of the target (Godden & Baddeley, 1975). Thus, past work indicates that context plays a central role in verbal memory, as it does for visual memory in our study. Our new findings demonstrate that context in VSTM is specified by the global spatial configuration. This can be distinguished from verbal context information because our spatial configuration manipulations are devoid of semantic schemata.

Second, the formation of a critical configuration is somewhat flexible. Attention and stimulus-driven cues help the visual system extract the relevant aspects of a configuration. However, top-down attention is not sufficient for discarding distractors. Rather, it works in tandem with bottom-up grouping cues. Distractors that are physically distinctive from targets are excluded from the configuration,

as shown in Section 4. In Experiment 4B, we found that when participants were asked to ignore distinctive distractors, they were successful, whereas unpublished data from our lab (Jiang, Olson, & Chun, 1999) indicated that observers were not successful in their attempt to ignore nondistinctive distractors. In Experiment 5B, we found that distinctive but attended items all entered into configuration formation. This interaction between top-down and bottom-up factors has been very important in visual-attention literature (e.g. Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe, 1994), and such principles of attentional selection are relevant for understanding what information gets represented in VSTM.

Finally, we speculate that the organization of VSTM is hierarchical. In our study, we focused on the organization of item relations—that is, how items are linked together in VSTM. Configuration coding is the key so far, and other Gestalt grouping cues may be important as well. At an elemental level of organization, one may ask how each individual item is represented. This is the question addressed by Luck and Vogel (1997), who proposed that features were organized on the basis of the object. For example, color and orientation can be stored together as long as they form a single object. Our experiments are directed at a higher level—how items are related to each other. It is possible that the organization of VSTM is carried out at multiple levels. At the global level, items are linked through configuration. At the item level, individual items may be stored as integrated objects (Luck & Vogel, 1997).

Both levels of organization can contribute to chunking in VSTM. Object chunking can increase the capacity of VSTM from 4 individual features to 16 (Luck & Vogel, 1997). It is more difficult to quantify how much benefit configuration encoding confers to VSTM capacity. However, the benefit of having a constant spatial configuration across encoding and retrieval can be clearly seen in accuracy improvements. In our various experiments, accuracy ranged from 10% to 20% better when configuration cues were preserved than when they were not.

More broadly, configuration formation may be a basic and general organization rule in visual processing. The formation of configuration is rapid (see Chun & Jiang, 1998). Spatial configurations can be implicitly learned within 5 to 10 repetitions and used as contextual information to guide attentional deployment in visual search (contextual cueing; Chun & Jiang, 1998). Spatial configurations can also be encoded incidentally to facilitate other behaviors, such as counting (Palmeri, 1997) and tracking (Yantis, 1992). However, configuration encoding as an organizational rule has been ignored. When discussing organization principles in vision, researchers quickly point to various Gestalt principles (Kanizsa, 1979) or contextual cues that rely on overlearned and highly familiar objects and scenes (Biederman et al., 1982). Even when configural processing is emphasized in object identification, the parts that form the configuration are usually connected (Palmer, 1977). Our study demonstrates that, in the domain of VSTM, configural organization can be exerted on much more arbitrary visual input at a very short time scale (namely, within a few hundred milliseconds in a single trial). Incidental configurations are formed on random visual input devoid of other grouping cues. How fundamental configuration formation is in other domains is an important question to be addressed in the future.

In conclusion, we proposed that VSTM is organized around spatial configurations of the visual input. We demonstrated that VSTM is easily disrupted by changes in the distractors between the memory and probe images in a change detection task. Configuration change impaired both location VSTM and identity VSTM (e.g., color and shape). Configuration formation is affected by both top-down instructions and bottom-up grouping cues. Configural organization is proposed to increase the capacity and fidelity of VSTM.

## References

Baddeley, A. (1986). *Working memory.* Oxford, England: Oxford University Press.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143–177.

Carlson-Radvansky, L. A. (1999). Memory for relational information across eye movements. *Perception & Psychophysics, 61,* 919–934.

Carlson-Radvansky, L. A., & Irwin, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1441–1458.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10,* 360–365.

Chun, M. M., & Nakayama, K. (in press). On the functional role of implicit visual memory for the adaptive deployment of attention across views. *Visual Cognition.*

Davidson, M. L., Fox, M. J., & Dick, A. O. (1973). Effect of eye movements on backward masking and perceived location. *Perception & Psychophysics, 14,* 110–116.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Gauthier, I., & Tarr, M. J. (1997). Becoming a Greeble expert: Exploring mechanisms for face recognition. *Vision Research, 37,* 1673–1682.

Gescheider, G. A. (1976). *Psychophysics: Method and theory.* Hillside, NJ: Erlbaum.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66,* 325–331.

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

Intraub, H. (1997). The representation of visual scenes. *Trends in Cognitive Sciences, 1,* 217–222.

Intraub, H., Bender, R. S., & Mangels, J. A. (1992). Looking at pictures but remembering scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 180–191.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23,* 420–456.

Irwin, D. E., Brown, J. S., & Sun, J.-S. (1988). Integrating visual information across saccadic eye movements. *Journal of Experimental Psychology: General, 117,* 276–287.

Irwin, D.E, Yantis, S., & Jonides, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception & Psychophysics, 34,* 49–57.

Jiang, Y., Olson, I. R., & Marvin, M. M. (1999). [Selective attention and visual short-term memory]. Unpublished raw data.

Kanizsa, G. (1979). *Organization in vision: Essays on gestalt perception.* New York: Praeger.

Lee, D., & Chun, M. M. (1999). *What are the units of visual short-term memory: Features, objects, or spatial locations?* Manuscript submitted for publication.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review, 4,* 501–506.

Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 1–11.

Logie, R. H. (1995). *Visual–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, 309,* 279–281.

Luck, S. J., & Vogel, E. K. (1998). Visual and auditory working memory-capacity: Response from Luck and Vogel. *Trends in Cognitive Sciences, 2,* 78–80.

Moore, C., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 339–352.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Nickerson, R. S. (1965). Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology, 19,* 155–160.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461–488.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change blindness as a result of "mudsplashes." *Nature, 398,* 34.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441–474.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 324–354.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44,* 369–378.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16,* 283–290.

Rayner, K., & Pollatsek, A. (1983). Is visual information integrated across saccades? *Perception & Psychophysics, 34,* 39–48.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373.

Riccio, D. C., Richardson, R., & Ebner, D. L. (1984). Memory retrieval deficits based upon altered contextual cues: A paradox. *Psychological Bulletin, 96,* 152–165.

Santa, J. L. (1977). Spatial transformations of words and pictures. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 418–427.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science, 7,* 301–305.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5,* 644–649.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18,* 555–586.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74,* 1–29.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* New York: John Wiley & Sons.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 46*(A), 225–245.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Treisman, A. M., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 459–478.

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62,* 74–82.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24,* 295–340.

Zimmer, H. D. (1998). Spatial information with pictures and words in visual short-term memory. *Psychological Research, 61,* 277–284.

*(Appendix follows)*

# Appendix

## False Alarm Rate and Hit Rate Found in This Study (Percentages Given; Standard Error in Parentheses)

| Experiment | Task | Condition | False alarm | Hit |
|---|---|---|---|---|
| 1 | Color change detection | Minimal color change | 11 (4) | 60 (9) |
| | | Single probe | 48 (4) | 77 (5) |
| 2A | Location change detection | Minimal location change | 10 (3) | 71 (5) |
| | | Single probe | 22 (5) | 73 (2) |
| | | Maximal location change | 50 (4) | 73 (5) |
| 2B | Location change detection | Minimal location change | 9 (3) | 76 (4) |
| | | Single probe | 24 (8) | 64 (6) |
| | | Partial probe | 23 (5) | 58 (8) |
| 3—color | Color change detection based on quadrant | Same location | 5 (3) | 75 (7) |
| | | Same configuration | 8 (4) | 88 (4) |
| | | Different configuration | 16 (3) | 76 (6) |
| 3—shape | Shape change detection based on quadrant | Same location | 18 (3) | 76 (4) |
| | | Same configuration | 19 (5) | 74 (7) |
| | | Different configuration | 44 (5) | 68 (12) |
| 4A | Color change detection | Minimal change | 14 (4) | 77 (3) |
| | | Faded, same configuration | 31 (5) | 86 (3) |
| | | Faded, different configuration | 36 (5) | 85 (3) |
| 4B | Color change detection of the nonwhite targets | Same distractor location | 18 (5) | 88 (3) |
| | | Different distractor location | 10 (3) | 84 (4) |
| 5A | Location change detection | Same colors | 13 (4) | 82 (5) |
| | | Different colors | 14 (5) | 88 (3) |
| 5B | Location change detection | Same color grouping | 6 (2) | 72 (6) |
| | | Different color grouping | 13 (3) | 78 (7) |