Journal of Experimental Psychology:
Human Perception and Performance
2004, Vol. 30, No. 1, 79–91

0096-1523/04/$12.00 DOI: 10.1037/0096-1523.30.1.79

What Kind of Memory Supports Visual Marking?

Yuhong Jiang and Stephanie W. Wang
Massachusetts Institute of Technology

In visual search tasks, if a set of items is presented for 1 s before another set of new items (containing the target) is added, search can be restricted to the new set. The process that eliminates old items from search is visual marking. This study investigates the kind of memory that distinguishes the old items from the new items during search. Using an accuracy paradigm in which perfect marking results in 100% accuracy and lack of marking results in near chance performance, the authors show that search can be restricted to new items not by visual short-term memory (VSTM) of old locations but by a limited capacity and slow-decaying VSTM of new locations and a high capacity and fast-decaying memory for asynchrony.

Since the introduction of the phenomenon by Watson and Humphreys (1997), visual marking has received steady interest in the field of psychology due in part to its surprisingly large capacity (Theeuwes, Kramer, & Atchley, 1998). Unlike visual attention or short-term memory, which can monitor only 4 objects at a time, visual marking can "mark" more than 10 old objects (Jiang, Chun, & Marks, 2002a; Theeuwes et al., 1998).

Studies on visual marking typically rely on a visual search task with a temporal component (Donk & Theeuwes, 2001; Gibson & Jiang, 2001; Jiang, Chun, & Marks, 2002b; Kahneman, Treisman, & Burkell, 1983; Olivers & Humphreys, 2002; Olivers, Humphreys, Heinke, & Cooper, 2002; Olivers, Watson, & Humphreys, 1999; Watson & Humphreys, 1997, 2000, 2002). Two sets of items—old and new—are presented; the old items appear first, for approximately 1 s, and maintain all of their properties as the new items are added to the display. The task is to search for a target among distractors, and the target, when present, is always among the new items. Search time increases steadily as the number of new items increases, but search time is unaffected by the number of old items. The old items are marked so efficiently that they can be completely ignored, even when the new and old distractors are identical except for the temporal distinction (Jiang et al., 2002a; Theeuwes et al., 1998).

Theories of Visual Marking

What mechanism allows the search to be restricted to the new items? A few theories have been postulated. The inhibition account, proposed by Watson and Humphreys (1997), posits that an active mechanism keeps the locations of the old items inhibited. Inhibition lingers at these locations even after the addition of the new items. In turn, the old items are effectively ignored during the search for the target. The abrupt onset account, proposed by Donk and Theeuwes (2001), emphasizes the prioritization of new items because their presentation induces abrupt onset, which is known to capture visual attention (Jonides & Yantis, 1988). Finally, the temporal segregation account emphasizes the perceptual segregation of the two sets of items as a result of temporal asynchrony. Attention is then selectively allocated to the group that contains the target (Jiang et al., 2002a, 2002b). These accounts emphasize different aspects of the phenomenon and are not mutually exclusive. All imply some kind of memory that separates the old items from the new items. Such memory is necessarily a part, if not the essence, of visual marking.

Memory is involved in visual marking because once the new items are presented, the new and old items differ only in history. No perceptual distinctions in color or shape exist to keep the old items marked.[1] Consider the stimuli tested by Theeuwes et al. (1998). Both new and old distractors were drawn from the same pool of letters. After the presentation of the new items, the only distinction between new and old distractors was in memory, not on the display. To deprioritize old items, subjects had to rely on memory to differentiate new and old items. This is one of the important differences between visual marking and selective attention based on other grouping cues. For example, in a visual search task for a white target among white and black distractors (Jiang et al., 2002a), the perceptual distinction between the relevant set (white items) and the irrelevant set (black items) is present during search. Selective attention to the relevant set does not require an active memory about which locations contain white items and which locations contain black items. However, in visual marking, the difference between new and old items is in their history, and no cue is present to separate the two sets during search. Separate

Yuhong Jiang and Stephanie W. Wang, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

This research was supported by a postdoctoral research fellowship from the Helen Hay Whitney Foundation. We thank Derrick Watson and Hermann Müller for their comments.

Correspondence concerning this article should be addressed to Yuhong Jiang, who is now at the Department of Psychology, Harvard University, 33 Kirkland Street, Room 820, Cambridge, MA 02138. E-mail: yuhong@wjh.harvard.edu

[1] In some studies the new and old distractors differ in color or shape (e.g., Watson & Humphreys, 1997), but such perceptual differences are unnecessary for marking, which strongly manifests itself even when the new and old items differ only in time. Perceptual differences in color or shape may help separate the new and old items, but here we examine the more impressive aspect of visual marking—the ability to mark old items when there is no physical distinction between new and old items.

memory traces laid by new and old items are thus crucial for visual marking.

Theories that emphasize the preview period (Watson & Humphreys, 1997) can most naturally embrace the concept of a memory trace. The preview period provides abundant time to form and maintain a memory trace for the old items. After the addition of the new items, the two sets of items no longer differ by a perceptual property, but memory traces are associated with the old items. This in itself may be sufficient to separate the new and old items during search. Within this framework, one needs to address how old items are remembered and inhibited. Does inhibition of return keep the old items marked (Watson & Humphreys, 1997), does mnemonic search among old items leave a memory trace (Olivers et al., 2002), or are old locations stored in visual short-term memory (VSTM; Humphreys, Watson, & Jolicœur, 2002)?

Theories that focus on postpreview effects, such as the abrupt onset of new items (Donk & Theeuwes, 2001) or the temporal asynchrony of new and old items (Jiang et al., 2002b), cannot be easily extended to explain the memory traces. Both abrupt onset and temporal asynchrony are transient cues. Their effects need to be transformed into a sustained format to support extended visual search. One kind of transformation is to store the locations of abrupt onsets in VSTM. Visual search can then proceed within the VSTM store. This potential mechanism entails a capacity limit on the number of abrupt onsets that can be prioritized in visual search (Yantis & Johnson, 1990). Another kind of transformation is to store the distinction between the new and old items due to temporal asynchrony. This mechanism may also entail a capacity limit on how many groups can be distinguished, but it should be insensitive to the number of individual items within a group.

In this study, we explore the kind of memory involved in visual marking. Because the properties of such a memory—its content, capacity, and duration—are of interest no matter which theory one endorses, we characterize the kind of memory involved, but we remain largely agnostic about which of the three extant theories best accounts for visual marking. In the General Discussion, we discuss how the results fit within each theory.

Candidate Memory Processes

Candidate memory processes can be divided into memory for old items, memory for new items, and memory for the perceptual segregation between the new and old items.

Memory for Old Items

Mnemonic Search

Mnemonic search was initially proposed to explain how subjects avoid revisiting searched locations. Some researchers have argued that mnemonic search leaves inhibitory tags at visited locations in conjunction search (Klein, 1988; Klein & MacInnes, 1999). Others have argued that subjects do not keep any memory about which locations they have visited (Horowitz & Wolfe, 1998, 2001; but see Gibson, Li, Skow, Brown, & Cooke, 2000; Kristjansson, 2000). This controversy, when applied to visual marking, raises the question about whether subjects leave inhibitory tags at already searched new locations. But what about the old locations? Do subjects leave inhibitory tags after they have searched through the old items and failed to find the target?

Olivers et al. (2002) addressed this question by testing subjects in a double-search task. Their subjects searched for the target first among the old items and then among the new items. Mnemonic search would have left an inhibitory tag at the old locations and would have supported efficient rejection of old items. Results showed, however, that reaction time (RT) increased as the number of old (and new) items increased in the double-search procedure, suggesting that mnemonic search among old items is not what separates the new and old items in a typical visual marking task.

VSTM of Old Locations

A second possibility is that VSTM holds the locations of the old items. Such memory is formed during the preview and is actively used to eliminate these locations from search. Humphreys et al. (2002), when examining how visual marking is disrupted by secondary tasks, considered this a plausible mechanism. The idea of VSTM keeping track of old items, alternatively described by the visual index system[2] (Pylyshyn & Storm, 1988), receives some support because visual marking is disrupted by secondary tasks, which may tax the limited capacity of VSTM.

The dramatic differences in capacity,[3] however, suggest that VSTM of old locations is inadequate to account for visual marking. Whereas VSTM can hold only four to six locations[4] (Jiang et al., 2000; Phillips, 1974; Simons, 1996), visual marking can mark approximately a dozen old locations (Jiang et al., 2002a; Theeuwes et al., 1998).[5] However, the comparison between the capacity of VSTM and that of visual marking is hardly straightforward, because the capacity of VSTM is measured by using accuracy, whereas the capacity of visual marking is inferred by using RT. If each of a dozen old items is partially processed, then that may be sufficient for their complete deprioritization relative to the new items when assessed with RT. This does not mean that subjects can fully remember all of these locations. Thus, RT measures may overestimate the capacity of visual marking. In Experiment 1, we

[2] Here we loosely equate VSTM with the kind of visual attention used for tracking objects. There are several similarities between VSTM and object tracking to warrant such treatment. From an empirical standpoint, both have a capacity of approximately four, both rely on configuration-based representation, and both are object-based (Scholl & Pylyshyn, 1999); from an operational standpoint, both refer to the ability to hold online several items in working memory. We leave it open whether they are identical or just similar mechanisms.

[3] Subjects tend to represent the entire configuration (or layout) of items in VSTM (Jiang, Olson, & Chun, 2000; Yantis, 1992). This is why they can remember approximately six spatial locations (Simons, 1996) even though they may have a capacity limit of four units (Luck & Vogel, 1997). However, because configuration is hard to quantify, here we measure capacity in terms of the number of locations.

[4] This number assumes that subjects can use configural representation to hold locations in VSTM. When subjects are forced to remember individual locations, they perform poorly even when there are three locations to be memorized (Jiang et al., 2000).

[5] Separating locations in two temporal arrays can potentially lead to chunking of information into different temporal groups. This, however, does not dramatically increase the capacity of VSTM. A recent study on VSTM of two sequential arrays (Jiang & Kumar, in press) produced an estimate of VSTM capacity of approximately six spatial locations (for the two arrays combined), a value that is similar to estimates based on single arrays.

introduce a modification of the visual marking paradigm that allows an accuracy measure, and we directly test whether visual marking relies on VSTM of the old locations.

Memory for New Items

VSTM of New Locations

Alternatively, VSTM may hold the locations of the new items but not the old items. After all, working memory refers to the online manipulation of information (Baddeley, 1986; Logie, 1995). Inhibiting and eliminating old locations from search may be incongruous with actively keeping these locations in VSTM. If VSTM holds only the new locations but not the old locations, then the number of old items should be inconsequential to performance because old items do not impose load on VSTM. The number of new items influences how successfully old items are ignored. As new items increase in set size, the capacity of VSTM is exceeded, and performance suffers. Experiments 2 and 3 tested this hypothesis.

Memory for New and Old Perceptual Groups

Instead of remembering the old items or remembering the new items, the visual system may memorize the perceptual segregation between the two sets. We refer to this kind of memory as *memory for asynchrony.* By representing the distinction rather than the individual items within a set, this memory does not impose a capacity limit on the number of items per group. However, it may be tied to the transient signal that segregates the two groups, and thus, it may decay as the transient signal dissipates. Experiment 3 tested the presence of such a memory.

Experiment 1: VSTM of Old Locations

If visual marking is a result of holding old locations in VSTM, then the estimated capacity for marking old items should mirror the capacity of VSTM for spatial locations (Phillips, 1974). That is only four to six old locations can be effectively marked. In addition, because VSTM capacity is unaffected by shape or color changes of the placeholders (Jiang et al., 2000), visual marking should exhibit the same capacity limit regardless of whether the old items change features from the preview to the search display. Finally, if VSTM of old locations fully accounts for visual marking, then there is no need to remember which items are new. In turn, performance should not be sensitive to the number of new items. In the following section, we first introduce a revised paradigm that permits an accuracy measure in visual marking and then test these predictions.

An Accuracy Paradigm to Measure Visual Marking

The reason why past studies did not use accuracy to measure visual marking is because the target always differs from the previewed and new distractors in color or shape. When provided with sufficient time to view the displays, subjects can achieve perfect accuracy whether or not they have registered the temporal distinction between old and new items. The novel paradigm that we introduce here removes all cues distinguishing the target from the old items except for the temporal distinction. Thus, for example, subjects search for a *T* among *L*-shaped new items and *T*-shaped old items and report the direction of the target's rotation. The new distractors are rejected because of their shape, and the old distractors are rejected because of their temporal feature. Figure 1 shows a sample display for such trials (valid preview) as well as two control conditions (invalid preview and cross preview). For illustrative purposes, the dotted circles in Figure 1 indicate the locations of the old items. We varied the number of new items (new set size) and the number of old items (old set size). Because the target can rotate to one of four directions, by chance subjects can correctly identify the target on 25% of the trials.

Valid Preview

If subjects can separate the old items from the target on the basis of their temporal difference, and if they can do so independent of new and old set sizes, then accuracy should always be 100%. If there is a capacity limit for how many old items can be marked, let's say, only five old items can be fully marked, then if there are fewer than five old items, accuracy should be 100%. However, if there are more than five old items, then subjects would not know, among the unmarked old items and the target, which one is the target. They would have to guess. Their accuracy should be a combination of correctly marking five old items and randomly guessing from the remaining ones. Finally, if there is a capacity limit for how many new items can be distinguished from the old, then beyond that capacity subjects can no longer perceive the distinction between the target and the old items. They would have to guess from all the *T*s on the display. If there is one old item and the target, then subjects would be correct 50% of the time; if there are two old items, then they would be correct 33% of the time; and if there are three or more old items, they would be correct 25% of the time (note that accuracy cannot be lower than chance). Because the old set size was six or more in this experiment, ineffective marking should lead to near chance performance.

In sum, if the memory supporting visual marking has no capacity limit for new or old items, then performance should be 100% correct. If visual marking is primarily supported by VSTM of old locations with a capacity limit of 5, then as the old set size increases from 6, 9, to 12, performance should decrease dramati-

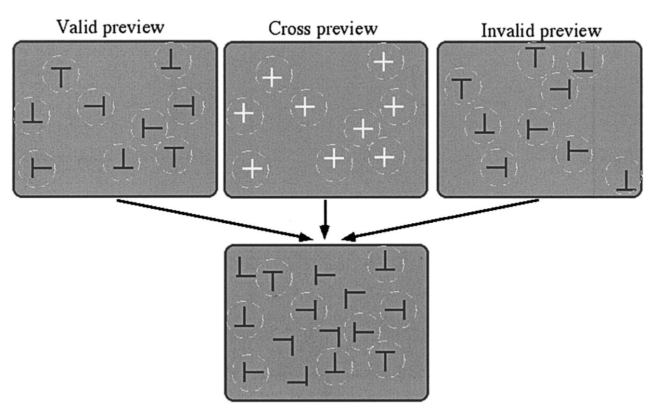

Figure 1. Sample displays tested in Experiment 1. The target was defined by the unique *T* among the new items, and the task was to determine its orientation (up, down, left, or right). Old items were all *T*s (or +s in the cross preview condition). The white dotted circles illustrate the locations of the old items for illustrative purposes. They were not actually in the experiment.

cally.[6] Alternatively, if visual marking is primarily supported by VSTM of new locations with a certain capacity limit, then accuracy should decrease as the new set size increases but should be insensitive to the old set size of 6 or more.

Invalid Preview

In this condition, the previewed old items move to new locations as the new items are added to the display. The temporal history established during the preview is no longer relevant. Therefore, the target cannot be distinguished from the old items by history. Subjects now have to guess which one among the *T*s is the target. Performance should be near but slightly above chance. If subjects try to remember the identity of the old *T*s and notice that there are two leftward *T*s during the preview and three leftward *T*s during the search display, then they can perform at a level that is above chance. Their ability to do so is limited because VSTM of shapes is poor. Yet this ability should be unaffected by the number of new items and should be insensitive to the range of old set sizes (6 to 12) that exceed the capacity of VSTM for shapes.

Cross Preview

In this condition, the shape and luminance of old items change at the onset of the new items. Such changes do not disrupt the abrupt onset of the new items (Enns, Austen, Di Lollo, Rauschenberger, & Yantis, 2001; Yantis & Jonides, 1984), nor do they influence the capacity of VSTM for old locations (Jiang et al., 2000). However, they disrupt the temporal asynchrony cues between the new and old items (Jiang et al., 2002b). Because both VSTM for old locations and attentional capture by abrupt onset are capacity limited, subjects may try to split their VSTM capacity between the old and new locations. The old items have priority because they are previewed for 1 s, during which they enter VSTM. Thus, performance should decrease more dramatically when the old set size increases than when the new set size increases.

Mixed Versus Pure Blocks

The three conditions listed above may be separately presented in pure blocks or randomly intermixed in mixed blocks. A mixed design is superior because any difference across conditions cannot be attributed to systematic differences in strategy. Yet marking is at least partly influenced by intention; when subjects ignore the search items and focus on a dot detection task, new and old locations are not distinguishable (Watson & Humphreys, 2000). Randomly intermixing invalid preview and cross preview with valid preview may reduce the efficiency of marking.[7] Experiment 1 tested half of the subjects in a mixed design and the other half in a blocked design.

Method

Subjects

Twelve college students volunteered to participate in this experiment. Their ages ranged from 19 to 27 years. All had normal or corrected-to-normal visual acuity.

Equipment

Subjects were tested individually in a room with normal interior lighting. The experiment was programmed using MacProbe (Hunt, 1994) and was displayed on a Macintosh computer with a 17-in. (43.18-cm) screen. Subjects viewed the display from an unrestrained distance of approximately 57 cm, at which distance 1 cm on the display subtended 1° visual angle.

Materials

We presented *T*, *L*, and + shaped items (each 0.8° × 0.8°) on randomly selected locations within an invisible 10 × 10 matrix (21° × 21°). *T*s and *L*s were in black, +s in white, and the background in gray. Each *T* or *L* was randomly rotated 0°, 90°, 180°, or 270°. The *L*s had a small offset at the junction between the two line segments (0.06°), whereas the *T*s had a large offset (0.40°), with one line bisecting the other symmetrically.

Procedure

Each subject was first tested in a standard visual search task to become familiar with the task. There were 2 practice and 96 experimental trials in this task. Each trial started with a fixation for 400 ms, followed by a search display (set size = 6 or 12) containing one *T* and several *L*s. Subjects were required to search for the *T* and press one of the four arrow keys to report the direction of its tail. Both accuracy and speed were emphasized. Once a response was made, the display was cleared, accuracy feedback was displayed for 400 ms, and 1 s later the next trial started.

Subjects were then tested in the visual marking task, constituting one of three conditions: valid preview, invalid preview, or cross preview. These conditions were tested in separate blocks for 6 subjects (the order of the blocks was counterbalanced across subjects) and were randomly intermixed for the other 6 subjects. There were 15 practice and 576 experimental trials. Each trial started with a fixation for 400 ms, followed by a preview display of 1 s and then a search display that lasted until response. Accuracy feedback was then displayed for 400 ms, and 1 s later the next trial started. Only accuracy was emphasized.

In the valid preview condition, several *T*s, each with a random orientation, were presented for 1 s during preview. Then several *L*s and one *T* were added to the search display. All of the old *T*s maintained their initial locations, shapes, and other properties. The task was to report the orientation (up, down, left, or right) of the unique *T* among the new items.

In the invalid preview condition, several *T*s were presented for 1 s during preview. As several *L*s and one *T* were added to the search display, all of the old *T*s moved instantly to previously blank locations. The task was to report the orientation of the new *T*.

In the cross preview condition, several +s were presented for 1 s during preview. As several *L*s and one *T* were added to new locations on the search display, all of the old +s changed to *T*s but remained at the old locations. The task was to report the orientation of the *T* at a previously unoccupied location.

In addition to the preview condition, two other factors were independently manipulated: new set size (6 or 12) and old set size (6, 9, or 12). Subjects were encouraged to memorize the old locations as a reliable method of spotting the target.

---

[6] The estimated accuracies are 91%, 67%, and 56% for old set sizes of 6, 9, and 12, respectively. This estimation is based on the following: Accuracy = $[5 \times 100\% + (N_{old} - 5) \times p\%]/N_{old}$, where $p\%$ is 50% for a set size of 6 and 25% for set sizes of 9 and 12.

[7] We thank Derrick Watson for this suggestion.

Results

Visual Search Practice

When searching for one *T* among many *L*s during the initial practice task, subjects were quite accurate (98.6% at Set Size 6 and 99.3% at Set Size 12), $t(11) = 1.83$, $p = .10$. RT was significantly shorter at Set Size 6 (median RT = 894 ms) than at Set Size 12 (median RT = 1,359 ms), $t(11) = 6.32$, $p < .01$.

Visual Marking

The factor of design (blocked vs. mixed) did not change the pattern of results during valid preview (all *p*s > .15). We thus pooled data across this factor and report data from all 12 subjects. Figure 2 plots accuracy in the marking tasks as a function of preview condition, new set size, and old set size. The pattern can be summarized as follows. Performance in the invalid preview condition was slightly better than chance and insensitive to new set size or old set size. Performance in the cross preview condition was better than that in the invalid preview condition. It declined as old set size increased and declined slightly as new set size increased. Performance in the valid preview condition was much higher than that in the cross preview condition. It was moderately influenced by old set size but declined by 17% when new set size increased from 6 to 12 and stayed at a much higher level than the cross preview condition.

An analysis of variance (ANOVA) on preview condition (valid, invalid, or cross preview), new set size (6 or 12), and old set size (6, 9, or 12) revealed significant main effects of preview condition, $F(2, 22) = 183.27$, $p < .01$; new set size, $F(1, 11) = 87.81$, $p < .01$; and old set size, $F(2, 22) = 14.17$, $p < .01$. In addition, there were significant interactions between preview condition and new set size, $F(2, 22) = 21.38$, $p < .01$, and between preview condition and old set size, $F(4, 44) = 3.68$, $p < .01$. The interaction between new and old set size and the three-way interaction were not significant (both *F*s < 1, *ns*).

The difference between the cross preview and invalid preview conditions reflects the operation of VSTM of previewed locations. VSTM of old locations assisted performance in the cross preview condition, as reflected by the significantly higher performance in this condition, $F(1, 11) = 44.34$, $p < .01$. The main effect of old set size was also significant, $F(2, 22) = 9.11$, $p < .01$, and the main effect of new set size approached significance, $F(1, 11) = 4.73$, $p < .06$. The interaction between condition and old set size was significant, $F(2, 22) = 6.40$, $p < .01$, and the interaction between condition and new set size approached significance, $F(1, 11) = 4.22$, $p < .07$, but the other interactions were not significant. Although follow-up tests showed that neither new nor old set size made any differences to performance in the invalid preview condition (both *p*s > .20), increases in both old set size and new set size significantly impaired accuracy for the cross preview condition, $F(2, 22) = 11.26$, $p < .01$, for old set size, and $F(1, 11) = 13.95$, $p < .01$, for new set size.

The difference between the valid preview and cross preview conditions reflects memory processes other than VSTM of previewed locations, because the effect of VSTM for old locations should be equivalent between the two conditions. These other processes must exist, because valid preview had significantly higher accuracy, $F(1, 11) = 124.31$, $p < .01$. The main effects of new set size and old set size were significant, $F(1, 11) = 343.65$, $p < .01$, for new set size, and $F(2, 22) = 18.93$, $p < .01$, for old set size. Also significant were the interactions between condition and new set size, $F(1, 11) = 21.67$, $p < .01$, and between condition and old set size, $F(2, 22) = 3.87$, $p < .05$. The other interaction effects were not significant. Follow-up tests showed that in the valid preview condition, performance declined by 17% (compared with a 5% decline in the cross preview condition) as new set size increased from 6 to 12 items, $F(1, 11) = 267.59$, $p < .01$, but was reduced mildly but significantly (5% compared with a 12% decline in the cross preview condition) as old set size increased from 6 to 12 items, $F(2, 22) = 5.68$, $p < .01$.

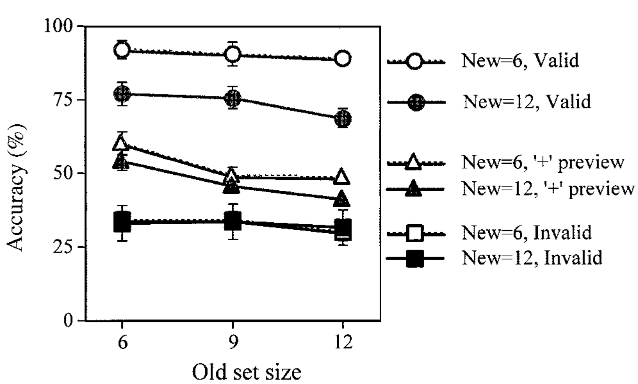

Figure 2. Results from Experiment 1: accuracy as a function of preview condition, new set size, and old set size. Chance performance was 25%. Error bars show standard errors.

Discussion

If VSTM of old locations is the primary means for visual marking, then accuracy should not be influenced by feature changes in the old items and should decline sharply as the old set size increases from 6 to 12 items. In addition, accuracy should not be affected by new set size. None of these predictions were supported by data from Experiment 1. First, feature changes mattered. Accuracy was much higher in the valid preview condition than in the cross preview condition. Second, performance declined only slightly as old set size increased. Furthermore, the number of new items significantly influenced accuracy. Thus, VSTM of old locations is inadequate as the mechanism to explain visual marking.

This conclusion implies that VSTM of old locations does not significantly contribute to visual marking. The large capacity for old locations in visual marking is perhaps not mysterious after all: Old locations are not stored in VSTM, so they do not impose load on VSTM. Imagine that someone has to remember the locations of 4 red squares presented among 18 black squares. The fact that the person can remember the red squares does not imply that he or she has a perfect capacity to hold 18 black squares. By this analogy, the term *marking* is perhaps a misnomer. Just as he or she does not mark the 18 black squares in order to process the 4 red squares, he or she probably does not mark the old items in order to search from the new items, at least not when marking is equated with VSTM. Performance is limited by new set size but not by old set size, perhaps because it is the new set, rather than the old set, that enters VSTM. The remaining experiments tested this hypothesis.

Whether the valid preview condition was intermixed with the invalid and cross preview conditions or was tested in a pure block did not influence the efficiency of visual marking. A similar result was previously obtained in the RT paradigm (Jiang et al., 2002b). It appears that subjects had paid as much attention to the valid preview trials whether or not they were mixed with other trials. These results do not reject the idea that efficient marking depends on attending to the search items (Watson & Humphreys, 2000). It simply means that attention was not significantly affected by mixed versus blocked design.

Accuracy Versus RT Paradigm

Because the accuracy paradigm introduced here is a novel method to assess visual marking, here we briefly compare results obtained from this paradigm and those obtained from a more typical RT paradigm. In previous studies, Jiang et al. (2002a, 2002b) manipulated similar factors (new set size, old set size, and preview condition) and used similar search tasks (*T* among *L*s) but measured visual marking in an RT paradigm. Figure 3 plots data from that paradigm (Jiang et al., 2002b, Experiment 1) as a comparison with those observed in Experiment 1. In that experiment, the old distractors were *L*s, so they differed from the target in shape and temporal features. Qualitatively similar results were obtained there: RT was relatively insensitive to old set size in the valid preview condition but was significantly affected by old set size in the cross preview condition. RT was sensitive to new set size in the valid preview and cross preview conditions, and the pattern of results was not influenced by blocked versus mixed design. These similarities suggest that the accuracy paradigm is a valid tool for assessing visual marking. The current study can now be connected with past studies of visual marking. The choice of the accuracy paradigm used in this study was motivated by the necessity to measure memory capacity, which is typically assessed on the basis of accuracy rather than RT.

The two paradigms do not always generate identical effects, however. For example, performance in the cross preview condition was indistinguishable from that in the invalid preview condition in an RT paradigm (Figure 3) but was significantly better than that in the invalid preview condition in the accuracy paradigm (Figure 2). This is likely due to the old locations not being effectively eliminated from search when held in VSTM. Therefore, subjects end up examining these locations—even though they know these locations do not contain the target—as well as the new locations in both the cross preview and invalid preview conditions. Thus, search speed should be similar. However, when provided with sufficient time to scrutinize their working memory to eliminate these locations, subjects can achieve good performance in the accuracy paradigm. This process of scrutiny may be too laborious to be used in the RT paradigm.

Another interesting difference between the RT and accuracy paradigms concerns the performance–old-set-size function when marking fails. In an accuracy paradigm in which the target is similar to the old items, any confusion of the temporal separation between new and old items decreases performance to chance level, whether there are 6, 9, or 12 old items. Thus, performance is largely independent of old set size but increasingly depends on guessing as new set size increases (Figure 2, cross preview condition). In an RT measure, confusion between new and old items leads to visual search among the old items, thus increasing the slope of the RT–old-set-size function (Figure 3).

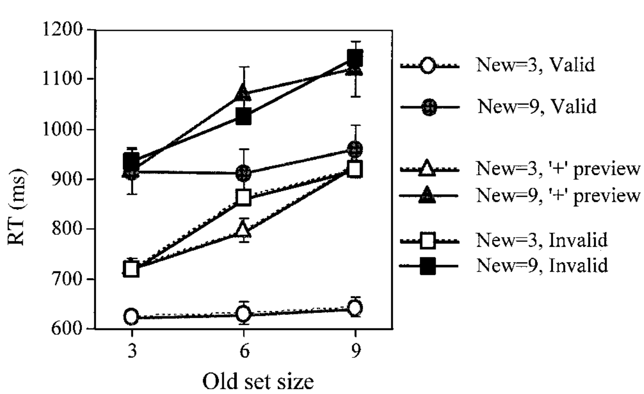

Figure 3. Typical results from standard reaction time (RT) paradigms. Note that the invalid preview condition was indistinguishable from the cross preview condition. Error bars show standard errors. Data are replotted from "Visual Marking: Selective Attention to Asynchronous Temporal Groups," by Y. Jiang, M. M. Chun, and L. E. Marks, 2002, *Journal of Experimental Psychology: Human Perception and Performance, 28,* p. 721. Copyright 2002 by the American Psychological Association.

Invalid Preview and Cross Preview

In Experiments 2–4, we examined only the valid preview condition. Here, we summarized results from the two control conditions tested in Experiment 1. The invalid preview condition was 32.7% correct on average (the range across subjects was 27% to 43%), significantly better than chance (25%), $t(11) = 5.19$, $p < .01$. This finding suggests that subjects were sensitive to the statistical information about the orientation of the *T*s.[8] It is interesting that no subjects reported trying to remember the identity of the old *T*s (e.g., 2 subjects left *T*s) so as to compare across the preview and the search displays. All subjects reported that they were randomly guessing. Thus, the visual system is sensitive to the statistical information without awareness (Stadler & Frensch, 1998). In Experiments 2–4, we used 32.7%, rather than 25%, as the baseline performance to calculate memory capacity.

Accuracies in the cross preview condition were 60%, 49%, and 49% for an old set size of 6, 9, and 12, respectively, when the new set size was 6. Accuracy dropped approximately 5% when the new set size was 12. As noted earlier (see Footnote 6), if subjects had a VSTM capacity of 5 locations, then the estimated accuracies would have been 91%, 67%, and 56% for a set size of 6, 9, and 12, respectively. The observed accuracy fell far short of these estimates. From the observed accuracy, we can estimate the capacity of VSTM to be approximately 2.5 to 3 locations, which is notably smaller than the typical estimate obtained from change detection tasks. Given the demand to remember these locations and to actively look for a target among other locations, it is perhaps not surprising that the capacity observed in the cross preview condition is lower than that seen in change detection tasks. The good performance seen in the valid preview condition is impressive in comparison.

[8] We thank Hermann Müller for this suggestion.

Experiments 2 and 3: VSTM of New Locations

An alternative to the VSTM of old locations mechanism is that one may rely on VSTM of new locations in visual marking. This hypothesis differs from the previous one not only in the content of VSTM (new or old set) but also in when VSTM is established. Whereas VSTM of old locations can be built up during the preview display, VSTM of new locations cannot start forming until the new items have been added to and segregated from the old items. VSTM of new locations relies on the assumption that new locations can be perceptually segregated from the old locations (Jiang et al., 2002b). Following the segregation, the new locations are transferred to VSTM and are maintained there. VSTM allows search to be restricted to new locations.

Consistent with this hypothesis, Experiment 1 revealed that as new set size increased, performance in the valid preview condition declined sharply. Prioritization of 6 new items over a group of old items was nearly perfect (90% correct), but prioritizing 12 new items was much harder (74% correct). This pattern is consistent with the idea that there is a capacity limit in prioritizing new items. Assuming that within the capacity of this memory accuracy is 100% and beyond the capacity accuracy is 32.7% (baseline performance estimated from Experiment 1), then capacity can be calculated by assuming measured accuracy as a weighted average of perfect memory and guesses:

$$\text{Accuracy} = [C \times 100\% + (N - C) \times 32.7\%]/N, \text{ if } N < C, \quad (1)$$

where C is the capacity and N is the number of items to be memorized. This simple formula calculates the capacity to be 7.3 new locations in Experiment 1.[9] This estimate is larger than the typical capacity limit of 4 to 6 in VSTM (Jiang et al., 2000; Phillips, 1974; Simons, 1996) and is larger than the limit of 2.5 to 3 observed in the cross preview condition of Experiment 1. Thus, the sensitivity to new set size supports the idea that new items may be transferred to and stored in VSTM, but the capacity difference suggests performance relies on other types of memory as well.

The additional mechanism may be a fast-decaying memory for the temporal distinction between the new and old items. A fast-decaying memory that does not have a capacity limit on the number of new or old items is consistent with results from Experiment 1. After all, in Experiment 1 it took approximately 400 ms longer to search from 12 new items than to search from 6. If there is a transient memory that decays rapidly within that period of time, then it should also lead to greater accuracy at a new set size of 6 than at 12. Because the estimated capacity from Experiment 1 exceeds the typical capacity limit of VSTM, it is possible that the transient memory for asynchrony and the sustained VSTM of new locations jointly support the performance.

To test whether VSTM of new locations is sufficient and necessary to account for visual marking, we compare VSTM and visual marking in the next two experiments in terms of their capacity and decay function. We first test how VSTM is affected by memory set size and retention interval and then test whether visual marking reveals similar capacity and decay function as VSTM.

Experiment 2: Direct Assessment of VSTM Capacity and Decay

In this experiment, we use change detection to assess how the number of to-be-remembered locations affects VSTM of spatial locations and how performance changes over a memory interval of 400 ms to 2 s. Experiment 2a was a standard change detection task: 6 or 12 locations were presented on the memory display; after a retention interval of 400 ms or 2 s, a probe display was presented. The task was to decide whether the two displays were identical or whether one item moved to a previously unoccupied location. Experiment 2b showed subjects a display of two groups of items, distinguished by color.[10] One group—the green squares—was designated as relevant. Subjects were asked to memorize the locations of the green squares (there were 6 or 12) and ignore the black squares (there were 12). After a retention interval of 400 ms or 2 s, 1 location was probed. The task was to decide whether this location was previously occupied by one of the green squares. The qualitative outcome—how set size and retention interval affect VSTM performance—serves as a basis for comparison with visual marking, tested in Experiment 3.

Method

Subjects

Twelve subjects were tested, 6 each in Experiments 2a and 2b.

Materials and Procedure

Each version of the experiment had 7 practice trials and 96 experimental trials. In Experiment 2a, 6 or 12 green squares (1° visual angle) were presented at randomly selected locations for 200 ms. After a blank interval of 0.4 s or 2 s, another display of green squares was presented. The probe display was either identical to the memory display (50% of trials) or changed in one location, and the task was to decide whether there was a change. To assist decision making, a red frame cued the critical item that could have changed in location. Subjects were asked to decide whether the cued item was at an old location or a new location. Error feedback was then displayed visually for 400 ms. The next trial started 1 s later.

In Experiment 2b, 6 or 12 green squares and 12 black squares were presented at randomly selected locations for 200 ms; after a blank interval of 0.4 s or 2 s, 1 white square was presented. The task was to decide whether the white square was at one of the locations previously occupied by the green squares. The probe landed on the location of 1 of the green squares (50%) or the black squares (50%). The same testing room and equipment were used as in Experiment 1.

Results and Discussion

Because the task was a same–different detection task, we calculated A' (Grier, 1971) to measure memory sensitivity. Figure 4 shows A' as a function of set size and retention interval. Table 1 shows the mean hit and false-alarm rates.

Because the two versions of the experiment produced nearly identical results with respect to set size and retention interval, we performed analysis on data pooled across the two data sets. An

[9] This estimate is based on 74% accuracy for a new set size of 12. The formula can be used only if the new set size is greater than the capacity.

[10] This task is designed to be a spatial analog of the visual marking task, the difference being what cue (color distinction or temporal asynchrony) separates the two sets of items (Jiang et al., 2002a).

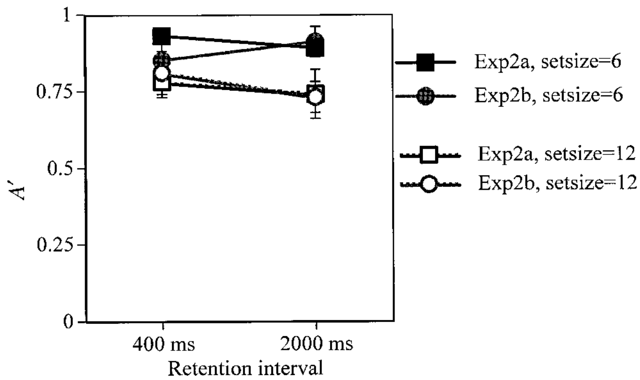

Figure 4. Results from Experiments (Exp) 2a and 2b: visual short-term memory as a function of set size and retention interval. Error bars show standard errors.

ANOVA on experiment (between-subjects), set size (within-subject), and retention interval (within-subject) revealed a significant main effect of set size, $F(1, 10) = 41.17$, $p < .01$. Neither the main effect of retention interval nor any other effect was significant (all $ps > .15$).

Thus, VSTM of spatial locations showed a severe capacity limit, but it was relatively insensitive to retention interval (400 ms vs. 2 s). Because overall accuracy was around 80% for a set size of six, the capacity of VSTM was approximately four.

Experiment 3: Visual Marking Capacity and Decay

If VSTM of new locations fully accounts for visual marking, then visual marking should also reveal a capacity limit of four to six new locations and show no decay between a retention interval of 400 ms to 2 s. Because VSTM of new locations starts at the onset of the new items and is no longer necessary once the target is found, the retention interval can be estimated by the target search time. For example, if it takes 800 ms to search for the target among 5 distractors, then the memory for new locations needs to be maintained for about 800 ms. If it takes 1,200 ms to search for the target among 11 other distractors, then the memory needs to be maintained for about 1,200 ms. The longer the search time, the longer the memory needs to be maintained. Thus, by varying search time, we can make an estimation of how fast visual marking decays.

We tested five conditions that varied in new set size and search difficulty. All conditions belonged to the valid preview condition. They served two different purposes: the first three conditions—easy-6, difficult-6, and easy-12—were used to infer the capacity of visual marking, whereas the last two conditions—feature-6 and feature-12—were included to test the existence of a memory for asynchrony. In the description below we first concentrate on the measurement of visual marking capacity.

Visual Marking Capacity and Decay

In the easy-6 condition, the new distractors (*L*s with small offset at the intersection) were relatively dissimilar to the target. In the difficult-6 condition, the new distractors (*L*s with large offset at the intersection) were similar to the target. Increasing the target–distractor similarity is known to increase search difficulty and hence RT (Duncan & Humphreys, 1989). Because VSTM of spatial locations is insensitive to the identity of the placeholders (Jiang, Chun, & Olson, in press), VSTM of new locations should be insensitive to this manipulation. Furthermore, because VSTM barely decays within 400 ms to 2 s, a range that covers the search RT for the easy-6 and difficult-6 conditions, VSTM of new items should not decay significantly for the difficult-6 condition. Thus, if VSTM of new locations fully accounts for visual marking, performance in the easy-6 and difficult-6 conditions should be equivalent in accuracy.

The third condition, easy-12, was designed to test the set size effect. We adjusted the level of difficulty such that RT was comparable between easy-12 and difficult-6 conditions. These two conditions, therefore, were equivalent in terms of memory decay. If visual marking depends on a limited-capacity memory, such as VSTM, then performance in easy-12 should be worse than that in difficult-6. Alternatively, if visual marking depends completely on an unlimited-capacity memory that decays rapidly, then response accuracy in easy-6 (which took the shortest amount of time) should be higher than that in both difficult-6 and easy-12, and the latter two should not differ from each other.

Feature Search and Memory for Asynchrony

Two other conditions, feature-6 and feature-12, were similar to easy-6 and easy-12, except that the target differed from the new distractors in color. Here, old items (black *T*s) were previewed for 1 s, and then new items (1 black *T* and several white *L*s) were added. As in the other conditions, subjects could succeed in the task only if they could maintain the temporal segregation between the target and the old distractors. Discussion about the feature conditions is deferred until after the presentation of the results.

Method

Subjects

Six new subjects were tested in this experiment.

Table 1
Hit and False-Alarm Rates Observed in Experiment 2

Rate	Set size = 6, delay = 400 ms		Set size = 12, delay = 400 ms		Set size = 6, delay = 2,000 ms		Set size = 12, delay = 2,000 ms	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Hit	.82	.04	.54	.06	.87	.03	.63	.05
False alarm	.14	.05	.12	.04	.17	.05	.30	.04

Procedure

Each subject was tested in a visual search task first and then in the visual marking task. Subjects searched for a *T* among *L*s in the search task and reported its orientation in a speeded response. Three types of *L*s were used to vary search difficulty. (a) In easy-6 and easy-12, the *L*s were dissimilar to the *T* because the offset at the intersection of the two segments was small (2 pixels). (b) In difficult-6, the *L*s were similar to the *T* because the offset was large (8 pixels; the offset in the *T*s was 11 pixels). These offset parameters were chosen so that easy-12 and difficult-6 conditions produced similar RTs. (c) Finally, the *L*s were white and the *T*s were black in feature-6 and feature-12 conditions. There were 20 trials in each condition.

Only the valid preview condition was tested in the visual marking task. The number of old distractors was 6, 9, or 12. There were five types of new items: easy-6, easy-12, difficult-6, feature-6, and feature-12. The presentation sequence was identical to that of the valid preview condition of Experiment 1: Old items were *T*s of random orientations and were previewed for 1 s before the addition of the new items, which contained one *T* and several *L*s. The five conditions were randomly intermixed in the experiment. Subjects completed a total of 240 trials. Figure 5 shows a schematic sample of the displays used in each condition.

Results and Discussion

Easy-6, Easy-12, and Difficult-6: Capacity and Decay Function

Visual search practice. RT measured in the visual search task is used to estimate the retention interval for the visual marking task. The longer it takes to search for a *T* among *L*s, the longer the memory for new locations needs to be maintained. Mean accuracies for easy-6, easy-12, and difficult-6 were 98%, 100%, and 95%, respectively, which were not significantly different from one another ($p > .10$). The means of the median RTs for easy-6, easy-12, and difficult-6 were 1,053 ms, 1,470 ms, and 1,620 ms, respectively. Easy-6 was significantly faster than both easy-12 ($p < .05$), and difficult-6 ($p < .03$); the latter two conditions were not significantly different from each other ($p > .30$). Combining speed and accuracy in a single measure of the ratio between RT and accuracy also failed to reveal any difference between difficult-6 and easy-12, $t(5) = 0.79$, $p > .45$. Thus, memory for new locations needed to be retained for approximately 1.0 s, 1.5 s, and 1.6 s for easy-6, easy-12, and difficult-6, respectively.

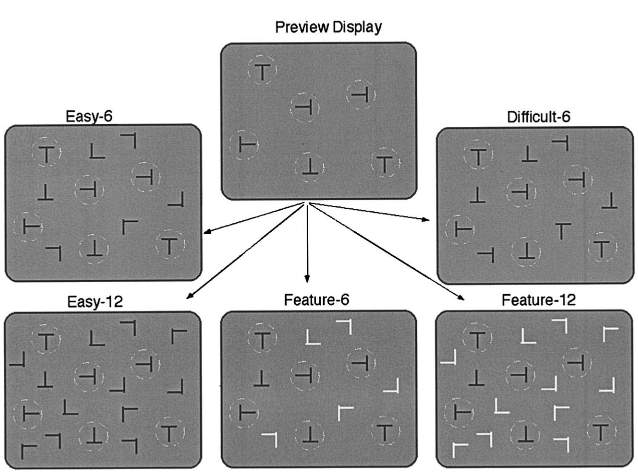

Figure 5. Sample displays from the five conditions of Experiment 3. The white dotted circles are for illustration only and were not actually presented in the experiment.

Because the observed RT also included the time necessary for selecting a response and executing the motor output, the above duration overestimated the duration that the memory for visual marking needed to be retained. Nonetheless, because response selection and motor output were comparable across conditions, the difference between the actual RTs is a valid estimate of the difference in the necessary retention intervals.

Visual marking. Because valid preview led to nearly perfect filtering of the old items, the search RTs measured in the visual search practice task were taken as a rough estimate of the duration that new locations needed to be maintained in memory. In addition, RT measured in the visual marking trials matched well with those measured in the visual search task. The means of individual subjects' median RTs (correct trials only) on the visual marking task were 1,013 ms for easy-6, 1,400 ms for easy-12, and 1,351 ms for difficult-6. Easy-6 was significantly faster than both easy-12 ($p < .03$) and difficult-6 ($p < .01$), and easy-12 was not significantly different from difficult-6 ($p > .50$). This pattern of data did not change when incorrect trials were included.

Thus, memory for new locations needed to be retained for approximately 1.0 s in easy-6 and 1.4 s in easy-12 and difficult-6. The difference in performance between easy-6 and difficult-6 thus measures how visual marking decays over a span of 1.0 to 1.4 s; and the difference between easy-12 and difficult-6 (similar in RT) measures how visual marking is affected by new set size. Figure 6 plots mean accuracy as a function of old set size and new set size.

Decay. We first compared easy-6 with difficult-6, which is a measure of how the memory of new locations (given consistent set size) decays over a span of approximately 1.0 s to 1.4 s. There was a significant main effect of difficulty, $F(1, 5) = 6.60$, $p < .05$, showing a significant decrease in accuracy as the search among new items was prolonged. The main effect of old set size was not significant ($F < 1$); neither was the interaction between old set size and new set condition significant ($F < 1$).

Set size. Because search RTs were comparable between easy-12 and difficult-6, the comparison is a relatively clean reflection of how memory of new locations is affected by set size. An ANOVA revealed a significant main effect of new set size, $F(1, 5) = 7.04$, $p < .05$, and no effect of old set size or interaction (Fs < 1). Accuracy was approximately 75% for difficult-6 and 60% for easy-12.

Neither VSTM of new locations nor an unlimited, fast-decaying memory alone can explain the observed results. VSTM of new locations, limited in capacity but slowly decaying, should be insensitive to differences in inferred retention interval between easy-6 and difficult-6. Yet performance was significantly worse for difficult-6. An unlimited, fast-decaying memory can account for the higher performance for easy-6, which required the shortest retention, but it cannot explain why easy-12 should have lower accuracy than difficult-6. It is thus necessary to propose that visual marking is supported by two kinds of memory: VSTM of new locations and a fast-decaying memory for asynchrony that does not have a capacity limit for the number of items per group. Memory for asynchrony is established by temporal segregation between the new and old items. This memory decays rapidly while a subset of new items (maybe four of them) are stabilized in VSTM.

Unlike VSTM, memory for asynchrony is a transient form of memory that shows significant decay between an inferred retention

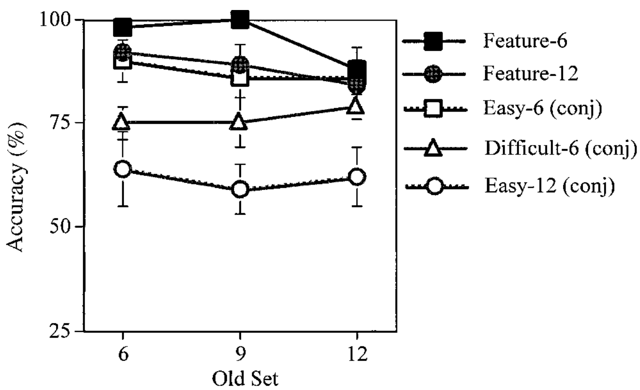

Figure 6. Results from Experiment 3: accuracy as a function of search type, new set size, and old set size. Error bars show standard errors. conj = conjunction search.

interval of 1.0 and 1.4 s. It decays too slowly to be confused with the kind of iconic memory that lasts for about 100 to 200 ms (Phillips, 1974; Sperling, 1960). It contains the distinction between two perceptual groups that are separated in the temporal domain. The specificity to temporal segregation is inferred from Experiment 2b, which tested memory for a perceptual group segregated by color (see also Driver & Baylis, 1989; Kaptein, Theeuwes, & van der Heijden, 1995). In Experiment 2b, VSTM was stable but limited in capacity. Thus, memory for asynchrony is a memory trace specific to perceptual segregation based on temporal cues. It may be the critical difference between visual marking and selective attention based on other kinds of perceptual segregation.

Feature-6 and Feature-12: Feature Search

To provide further evidence for the existence of a transient memory, we changed the search task among the new items into a single feature search.

Visual search practice. Accuracy was high, 99% for feature-6 and 98% for feature-12 ($F < 1$). RTs were 683 ms for feature-6 and 719 ms for feature-12, which were not significantly different, $F(1, 5) = 1.03$, $p > .30$. Thus, visual search for a color-distinguished target was not affected by the number of distractors.

Visual marking. Results are plotted in Figure 6. The striking result is that accuracy for feature-12 was as high as accuracy for easy-6, which required a conjunction search. This simply could not be accounted for by VSTM of new locations alone because VSTM was sharply limited by capacity beyond six locations. Furthermore, the difference in accuracy between feature-6 and feature-12 was much smaller than that between the corresponding conjunction search tasks: easy-6 and easy-12. That is, the memory that supported the two feature conditions showed a smaller dependence on a capacity-limited memory process (possibly VSTM).

Both observations were confirmed by an ANOVA on search condition, new set size, and old set size. There was a main effect of search type, $F(1, 5) = 17.40$, $p < .009$, with higher accuracy for feature search than conjunction search; a main effect of new set size, $F(1, 5) = 22.08$, $p < .01$, with higher accuracy at a new set size of 6 than 12; and a significant interaction between search type and new set size, $F(1, 5) = 28.11$, $p < .01$, showing a smaller set size effect for feature search. None of the other effects were significant (all ps $> .10$). The follow-up test showed that for feature search, accuracy was mildly affected by both new set size, $F(1, 5) = 4.31$, $p < .10$, and old set size, $F(2, 10) = 3.39$, $p < .08$; the interaction between new and old set size was not significant ($F < 1$). Applying data from feature-12 to Equation 1, the estimated combined memory capacity for new items was approximately 10, far exceeding the typical capacity of VSTM.

A transient memory for asynchrony is supported by two pieces of evidence: (a) There is a memory that decays relatively rapidly within the first 2 s, and (b) this memory has a large or unlimited capacity for both new and old items. When search among the new items was a single feature search, the number of new locations that could be simultaneously held was on the order of 10 or more. VSTM of new locations can contribute to a capacity of only about 4 to 6 locations, independent of the identity of the placeholders. In fact, whether the locations were occupied by color squares or by complex novel shapes had no effect on subjects' ability to remember these locations (Jiang et al., 2000). The excellent performance in the feature search task was in part a result of a transient memory that held the distinction between the new and old items. It decayed rapidly as the transient asynchrony cue dissipated. Such rapid decay may explain why performance in conjunction-12 (i.e., easy-12) was so much worse than that in feature-12.

It is important to note that although subjects could maintain 10 or more new items in memory in the feature search condition, this does not mean that their memory had a capacity of 10 units. The new items were most likely represented as a single group, separate from a different group of old distractors. In other words, the unit of representation in the memory for asynchrony is temporal group, and our data suggest that at least two groups could be transiently separated. Each group (or "chunk") may contain many items that share common temporal features. This characteristic is different from that of VSTM. When two arrays were presented with an interstimulus interval of 1 s, the locations on the two arrays competed for VSTM. A total of approximately 6 locations (about 4 from the second and 2 from the first) could be retained in VSTM for two sequential arrays (Jiang & Kumar, in press). Thus, VSTM does not effectively chunk information according to time, but the memory for asynchrony does.

Experiment 4: The Decay Function of VSTM Under Interference

In this experiment, we revisit the decay function of VSTM. The decay function measured in Experiments 2a and 2b was based on the change detection task in which a blank retention interval was inserted between the encoding and the retrieval displays. During the retention delay, there was no interference from other locations and no active process to search for a certain item. In contrast, the decay function in visual marking was measured while subjects were actively searching among the memorized locations and were ignoring items at other locations. The question that we address here is, Does VSTM decay rapidly when subjects are actively searching from the memory locations and are ignoring other interfering locations?[11]

To address this question, we designed a VSTM task that was as similar as possible to the visual marking tasks, including the requirement to search and to ignore distracting information while

[11] We thank Derrick Watson for raising this question.

holding information in memory. Subjects were asked to remember the spatial locations of four or six items (black +s), presented for 1 s. Then one of the +s changed to a *T*, and the others changed to *L*s. In addition, several *T*s were added to the display. Subjects were asked to report the unique *T* presented at the previewed locations. The important difference between this task and the cross preview condition tested in Experiment 1 was to which set the target belonged. In Experiment 1, the previewed locations changed to *T*s, and the target was the unique *T* among the new locations. In the current experiment, the previewed locations contained one unique *T*, and the new locations contained distractor *T*s.

For performance to be above chance in this task, subjects need to hold the previewed locations in VSTM and to search for the target within the VSTM store. They also have to ignore the new items. The time that it takes the subjects to find the target is the approximate amount of time that VSTM has to be maintained. This interval was varied by adjusting the similarity between the *T*s and the *L*s. In the easy-4 condition, the *L*s were quite different from the *T*s, so subjects needed to maintain VSTM only for about 1.2 s. In the difficult-4 condition, the *L*s were similar to the *T*s, so subjects needed to maintain VSTM for about 1.5 s. If VSTM decays rapidly when subjects have to search the VSTM store and to ignore distractor information, then we should observe a significant decrement in accuracy in difficult-4 as compared with easy-4.

Method

Subjects

Six subjects were tested in this experiment.

Design

Subjects were tested in 192 trials, with the three conditions, easy-4, easy-6, and difficult-4, randomly intermixed. In all conditions, after a fixation of 800 ms, a preview display containing several (four or six) black +s was presented for 1,000 ms. Then one of the +s changed to a white *T*, and the others changed to white *L*s. At the same time, nine white *T*s were added to previously unoccupied locations. Subjects were asked to remember the spatial locations of the previewed +s and report the direction of the unique *T* presented among the previewed locations. Once subjects made a response and the screen was cleared, an accuracy feedback was displayed for 400 ms. After an interval of 1 s, the next trial commenced.

Four +s were previewed in easy-4 and difficult-4, and six +s were previewed in easy-6. The *L*s at the previewed locations were dissimilar to the *T* (offset at the junction = 0 pixels) in the easy-4 and easy-6 conditions, and the *L*s were similar to the *T* (offset at the junction = 8 pixels) in the difficult-4 condition.

Results and Discussion

We used median RT for all trials as the estimate for how long the subjects needed to maintain the previewed locations in VSTM. The means of individuals' median RTs were 1,214 ms for easy-4, 1,245 ms for easy-6, and 1,483 ms for difficult-4, and the difference across conditions was significant, $F(2, 10) = 9.36$, $p < .01$. Pairwise comparisons revealed that subjects would need to hold the locations in VSTM longer during difficult-4 than easy-4, $t(5) = 3.33$, $p < .02$, or easy-6, $t(5) = 3.34$, $p < .02$. The two easy conditions did not differ significantly from each other, $t(5) < 1$.

Accuracy of spatial memory, however, was insensitive to the inferred retention interval. Accuracies were 74.5% in easy-4 and 74.7% in difficult-4; the difference was not significant, $t(5) < 1$. Both conditions had significantly higher accuracy than easy-6 (65.1%; both *p*s $< .02$). Thus, VSTM of spatial locations was sensitive to the number of locations to be remembered, but it was insensitive to an estimated retention interval between 1.2 and 1.5 s. VSTM did not show the same speed of decay as visual marking, even when VSTM was tested under conditions of interference and active search. Accuracy in this task was quite poor, corresponding to a VSTM capacity of approximately three locations. The low capacity reflected the interference from the additional new items that may have disrupted configural representation of the memorized locations. This experiment provides additional evidence that VSTM decays slowly but has a severe capacity limit.

General Discussion

Visual Marking and Memory

The memory processes involved in visual marking can be summarized as follows. The preview period establishes one set of items as old, which is temporally different from another set of new items. However, during the preview period subjects are not actively storing the old locations in VSTM, nor are they performing mnemonic search through the old locations to leave an inhibitory tag (Olivers et al., 2002). Once the new set is added to the display, the transient temporal asynchrony and the abrupt onset cues induce a transient perceptual segregation between the new and old items. The memory for asynchrony rapidly decays, during which a limited number of new locations—perhaps three or four of them—are transcribed into VSTM. If the target is found before the transient memory dissipates, the target information is also transcribed into VSTM and the search is terminated. If the target is not found before the transient memory completely decays, then the search continues within the new items in VSTM. This proposal is supported by data presented in this study.

One implication from this proposal is that the efficiency of deprioritizing old items decreases when new set size increases. In fact, if VSTM were the only memory involved in visual marking, one would have predicted that old items could be efficiently rejected only when the new set size was smaller than 4 to 6. This prediction receives some support from existing studies using the RT paradigm.[12] Two studies, one by Theeuwes et al. (1998) and one by Jiang et al. (2002a), tested new set sizes as large as 15. Only Jiang et al. (2002a) unconfounded new set size from old set size. There, new set size was 3, 9, or 15, and old set size was 3, 15, or 30. The slope of RT as a function of old set size was 2, 5.5, and 10 ms/item when new set size was 3, 9, or 15, respectively. Consistent with the prediction, the slope was completely flat (2 ms/item) when new set was 3 and was significantly above 0 when new set was 15. Inconsistent with the prediction, however, the slope was quite flat (5.5 ms/item) when new set size was 9, which had clearly exceeded VSTM capacity. However, this pattern of data becomes interpretable if performance is partly supported by the high-capacity, transient memory for asynchrony.

The necessity to include a transient memory for asynchrony is dictated by the fact that performance is better than that predicted if

[12] Most standard RT paradigms used a set size of 4 to 16; half were new (2 to 8) and half were old. This range of new set size may be too small to pick up the steep RT–old-set-size function.

visual marking is based on VSTM alone. In addition, visual marking efficiency decays rapidly within 1.0 to 1.5 s, whereas VSTM is maintained within such a range of retention interval. This memory can be considered as the trace laid by the transient temporal asynchrony cue and is thus specific to retaining perceptual segregation based on temporal differences.

Theories of Visual Marking Revisited

The three extant theories of visual marking emphasize different mechanisms: inhibition of the old items, attentional capture by new onset, and temporal asynchrony between the new and old sets. These three mechanisms are not mutually exclusive. In fact, Watson and Humphreys (2002) embraced abrupt onset in their inhibition hypothesis, and Jiang et al. (2002b) had considered temporal asynchrony as an additional factor to possible inhibition of old items. Thus, in this discussion we focus on the three complementary mechanisms rather than on the three theories that partially overlap.

We first consider the role of the preview period in visual marking. The processes that go on during the preview—for example, inhibition of old items—do not directly establish a memory trace for the old items in VSTM. That is, following the preview, the old items neither enter VSTM nor are tagged by mnemonic search. Its role is perhaps to assist in the isolation of new items, so that old items do not compete with new items for entry into the limited-capacity VSTM. Active search through old items makes them a more powerful competitor to VSTM and leads to failed visual marking (Olivers et al., 2002). Secondary tasks draw away necessary resources from the executive control that monitors which set enters VSTM and also leads to failed visual marking (Humphreys et al., 2002; Watson & Humphreys, 1997).

Another important role of the preview is to establish different time courses for the new and old items. Attentional capture produced by the abrupt onset of the new items (Donk & Theeuwes, 2001) allows a subset of new items to enter VSTM. However, the number of abrupt onsets that can simultaneously capture attention is limited to about four (Yantis & Johnson, 1990). When there are more than four new items, a transient memory for asynchrony allows the new and old items to be continuously separated. As this memory dissipates, performance will rely completely on VSTM. The faster subjects can find the target among the new items, the less dependent they will be on VSTM. Practice in visual search, performing feature rather than conjunction search, and other processes that enhance search speed all lead to effective visual marking that is less affected by the number of new items.

It remains to be seen whether inhibition of old items during preview leaves a memory trace that is not accounted for by the memory for asynchrony. So far, evidence for a memory trace of old items has come from the probe detection task, in which RT to detect a probe dot at the old locations was found to be longer than that at the new locations or at a location not previously previewed (Watson & Humphreys, 2000). Because memory for asynchrony registers the difference between new and old items, and the new items are subsequently enhanced and the old items are inhibited, delayed probe detection at old locations after the asynchrony cue can be explained without postulating an active template (or memory) for old items. Whether probe detection is delayed before the asynchrony cue remains to be tested in the future. If inhibition of old locations exists independent of the presentation of new items, then the current list of memory for marking would need to be expanded to include it.

Because the transient segregation cue is disrupted by changes, memory for asynchrony does not apply to visual marking of moving displays. When new and old items move, the two are segregated primarily by differences in motion grouping. This case is somewhat analogous to the cross preview condition tested in Experiment 1, where "marking" relies heavily on transcribing new or old items (or both) to VSTM. Because VSTM represents the pattern (or configuration) of items, configural effects may be observed in dynamic visual marking (Watson, 2001).

In summary, in this study we have shown that VSTM of old locations does not play a significant role in visual marking. Both VSTM of new locations, limited in capacity but slow decaying, and a transient memory for asynchrony that has a high capacity but decays rapidly contribute to visual marking. Future studies should further characterize how inhibition during the preview, abrupt onset, and temporal segregation assist memory processes involved in visual marking.

References

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

Donk, M., & Theeuwes, J. (2001). Visual marking beside the mark: Prioritizing selection by abrupt onsets. *Perception & Psychophysics, 63,* 891–900.

Driver, J., & Baylis, G. C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 448–456.

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

Enns, J. T., Austen, E. L., Di Lollo, V., Rauschenberger, R., & Yantis, S. (2001). New objects dominate luminance transients in setting attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1287–1302.

Gibson, B. S., & Jiang, Y. (2001). Visual marking and the perception of salience in visual search. *Perception & Psychophysics, 63,* 59–73.

Gibson, B. S., Li, L., Skow, E., Brown, K., & Cooke, L. (2000). Search for one versus two identical targets: When visual search has a memory. *Psychological Science, 11,* 324–327.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75,* 424–429.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394,* 575–577.

Horowitz, T. S., & Wolfe, J. M. (2001). Search for multiple targets: Remember the targets, forget the search. *Perception & Psychophysics, 63,* 272–285.

Humphreys, G. W., Watson, D. G., & Jolicœur, P. (2002). Fractionating the preview benefit in search: Dual-task decomposition of visual marking by timing and modality. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 640–660.

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

Jiang, Y., Chun, M. M., & Marks, L. E. (2002a). Visual marking: Dissociating effects of new and old set size. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 293–302.

Jiang, Y., Chun, M. M., & Marks, L. E. (2002b). Visual marking: Selective attention to asynchronous temporal groups. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 717–730.

Jiang, Y., Chun, M. M., & Olson, I. R. (in press). Perceptual grouping in change detection. *Perception & Psychophysics.*

Jiang, Y., & Kumar, A. (in press). Visual short-term memory for two sequential arrays: One integrated representation or two separate representations? *Psychonomic Bulletin & Review.*

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

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43,* 346–354.

Kahneman, D., Treisman, A., & Burkell, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 510–522.

Kaptein, N. A., Theeuwes, J., & van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1053–1067.

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334,* 430–431.

Klein, R., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science, 10,* 346–352.

Kristjansson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science, 11,* 328–332.

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, 390,* 279–281.

Olivers, C. N. L., & Humphreys, G. W. (2002). Visual marking under attentional blink conditions: More evidence for top-down limited capacity inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 22–42.

Olivers, C. N. L., Humphreys, G. W., Heinke, D., & Cooper, A. (2002). Prioritization in visual search: Visual marking is not dependent on a mnemonic search. *Perception & Psychophysics, 64,* 540–560.

Olivers, C. N. L., Watson, D. G., & Humphreys, G. W. (1999). Visual marking of locations and feature maps: Evidence from within-dimension defined conjunctions. *Quarterly Journal of Experimental Psychology, 52*(A), 679–715.

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

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38,* 259–290.

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

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

Stadler, M. A., & Frensch, P. A. (Eds.). (1998). *Handbook of implicit learning.* Thousand Oaks, CA: Sage.

Theeuwes, J., Kramer, A. F., & Atchley, P. (1998). Visual marking of old objects. *Psychonomic Bulletin & Review, 5,* 130–134.

Watson, D. G. (2001). Visual marking in moving displays: Feature-based inhibition is not necessary. *Perception & Psychophysics, 63,* 74–84.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review, 104,* 90–122.

Watson, D. G., & Humphreys, G. W. (2000). Visual marking: Evidence for inhibition using a probe-dot detection paradigm. *Perception & Psychophysics, 62,* 471–481.

Watson, D. G., & Humphreys, G. W. (2002). Visual marking and visual change. *Journal of Experimental Psychology, 28,* 379–395.

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

Yantis, S., & Johnson, D. N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 812–825.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601–621.

Received September 4, 2002
Revision received July 23, 2003
Accepted July 23, 2003 ■