Journal of Experimental Psychology:
Learning, Memory, and Cognition
2002, Vol. 28, No. 2, 293–302

0278-7393/02/$5.00 DOI: 10.1037//0278-7393.28.2.293

Visual Marking: Dissociating Effects of New and Old Set Size

Yuhong Jiang
Yale University

Marvin M. Chun
Vanderbilt University

Lawrence E. Marks
Yale University

Visual marking makes it possible to ignore old items during search. In a typical study, old items are previewed 1 s before adding an equal number of new items, one of which is the target. Previewing half of the items reduces the search slope relating response time (RT) to overall set size by half. However, this manipulation sometimes only reduces overall RT but not search slope (Experiment 1). By orthogonally varying the numbers of old and new items, Experiment 2 shows that old and new set sizes interactively affect visual marking. Given a constant new set size, the size of the old set has negligible effect on RT. However, increasing the new set size reduces the preview benefit in overall RT. Experiment 3 shows that this reduction may be restricted to paradigms that use temporal segregation cues. Studies should vary old and new set size orthogonally to avoid missing a visual marking effect where one may be present.

Selective attention to relevant aspects of the visual environment is crucial for human cognition, as we can attend to only a limited amount of information at any time (Pashler, 1998). Attention is guided by various cues, such as abrupt onset (Yantis & Jonides, 1984), location-based top-down attention (Paquet & Lortie, 1990), and familiar context (Chun & Jiang, 1998). The present study focuses on one mechanism that directs attention to target information: visual marking (Gibson & Jiang, 2001; Olivers, Watson, & Humphreys, 1999; Theeuwes, Kramer, & Atchley, 1998; Watson & Humphreys, 1997, 1998, 2000).

In studies of visual marking, a subset of distractors appears at least 400 ms ahead of other items, including the target. These distractors can be readily ignored or "marked," having barely any effect on time to search for the target. That is, old items are deprioritized relative to new items. Watson and Humphreys (1997, 1998) examined this benefit in previewing old items. In their conjunction search task, observers searched for the target—a blue *H*—among two sets of distractors of equal number—blue *A*s and green *H*s. In their simple feature search task, observers searched for a blue *H* among just the blue *A*s. The search was more efficient (search slope was shallower) in the simple feature condition than the conjunction search condition (Treisman & Gelade, 1980). In a third and critical condition, Watson and Humphreys (1997, 1998) presented a subset of the distractors—the green *H*s—1000 ms before the others. The search rate, defined as the slope of the function relating response time (RT) to overall set size, was much smaller in this preview condition than in the conjunction condition and indistinguishable from the search rate in simple feature search, suggesting that the previewed distractors were completely ignored. Similar preview benefits have been observed when the two sets of items differ only in time, not in color or other features (Olivers et al. 1999; Theeuwes et al., 1998). Watson and Humphreys (1997) suggested that the mechanism of visual marking in static displays is inhibition of previewed locations.

Our study examines the validity of two possible measures of visual marking: (a) slope and (b) intercept of the function relating RT to set size. In visual search tasks, target detection RT can be linearly regressed against the number of items in the display (set size). The resulting slope indicates how much additional time is needed to inspect each item, a good indicator of search rate. The intercept, by contrast, does not have quite a clear-cut explanation, because many factors may influence the duration between the onset of the stimuli and response, independent of set size.

Studies of visual marking have naturally relied on search slope as the main dependent variable. In most studies, previewing half of the distractors produced easily interpretable results. For instance, some studies obtained faster RT and a shallower slope in the valid preview than the invalid or no-preview condition, implying visual marking (Watson & Humphreys, 1997).[1] Others studies show

Yuhong Jiang and Lawrence E. Marks, Department of Psychology, Yale University, and the John B. Pierce Laboratory; Marvin M. Chun, Department of Psychology, Vanderbilt University.

This research formed part of a doctoral dissertation submitted by Yuhong Jiang to Yale University. The work was supported in part by a dissertation fellowship from Yale University to Yuhong Jiang, National Science Foundation Grant BCS-0096178 awarded to Marvin M. Chun, and National Institutes of Health Grant DC00271–15 awarded to Lawrence E. Marks. We thank Carol Fowler, Bradley Gibson, Glyn Humphreys, Frank Keil, Jeansok Kim, and Karen Wynn for their insightful comments and suggestions.

Correspondence concerning this article should be addressed to Yuhong Jiang, who is now at the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, NE 20-443B, Cambridge, Massachusetts 02139. E-mail: yuhong@mit.edu

[1] The baseline to assess visual marking is typically a no-preview condition, in which all items are presented simultaneously. A better baseline, more comparable to the valid preview condition, is an invalid preview condition. Here, old items are previewed but then move instantly to previously blank locations as the new items are added on the display. Thus, in terms of presentation sequence, the invalid preview condition is more comparable to the valid preview condition than the no-preview condition is. Pilot studies show that results in the invalid preview condition are comparable to results in the no-preview condition (Jiang, 2001).

comparable slopes and comparable overall RTs in the valid and invalid or no-preview conditions, implying no visual marking. However, measures of slope and overall RT may sometimes conflict, as when overall RT is faster in the valid preview condition compared with the invalid preview (or no-preview) condition, but the search slopes are equal (Watson & Humphreys, 1998). In the past, the lack of a difference in search rate in such a pattern of results has been taken as evidence for an absence of visual marking.

In this study, we provide evidence that marking may be present even in situations that do not reveal a difference in search rate. This methodological point has theoretical importance because it affects whether an experimental manipulation is interpreted to produce visual marking. Our argument is based on a new finding: Changes in search slope, and thus the likelihood of inferring the presence of visual marking, depend on both the number of old items and the number of new items. In some situations where visual marking is clearly present, search slope may fall only slightly, or not at all. This is likely to occur with the use of large set sizes containing equal numbers of old and new items. The following provides more background for this problem.

In a typical visual search, there is a single stimulus set that varies in size. But in studies of visual marking, there are three possible measures of set size. The experimental paradigm allows one to vary the number of new items (new set size), the number of old items (old set size), and, correspondingly, the total number of items (overall set size = old + new). It follows that there are three possible measures of the search slope, one for each of the three measures of set size. The slope of the function relating RT to new set size (i.e., new-set-size slope) is not directly relevant to understanding visual marking. Of central interest are search slopes for old set size and overall set size, two measures that, of course, are not independent. A mechanism that effectively rejects old items should lead to a reduction in both old-set-size slope and overall-set-size slope. In an ideal case in which all old items are rejected efficiently, the function relating RT to old set size should be flat (zero slope) in the valid preview condition, and, relative to the invalid preview condition, the overall-set-size slope should be reduced by the ratio of old set size to overall set size, a reduction of 50% when new and old set sizes are the same.

Interpreting the overall-set-size slope is problematic, however, if marking is somehow affected by variation in new set size. As described above, overall set size provides a good approximation of the effects of old set size (because the two covary), but only if new set size does not interact with old set size. In this study, we will demonstrate that new set size does, in fact, interact with old set size, and hence confounds the measure of overall set size. We propose that the measure typically used in the extant literature on visual marking—the slope of the function relating RT to overall set size—should be replaced by the slope of the function relating RT to old set size. To this end, researchers should overtly choose to measure old set size and design studies accordingly. Three experiments derive from these considerations.

Experiment 1 follows a design used by previous researchers (e.g., Watson & Humphreys, 1997) that measures the search slope in terms of overall set size. As in most previous studies, we kept new and old set sizes identical. Because the effect of new set size should be more obvious with larger numbers of items, new and old set sizes covered a large range. We examined whether visual marking may be revealed as a change in intercept, or overall RT, rather than a change in slope (of overall set size). Experiment 2 varies new and old set sizes orthogonally and determines the slopes of the function relating RT to both overall set size and old set size. We will show that the latter is a more sensitive measure of marking, and we delineate the source of the discrepancy between them. Finally, Experiment 3 asks whether a similar difference between overall-set-size slope and old-set-size slope is evident in a spatial analog of visual marking, or whether, as we suspect, it is unique to the temporal nature of the visual marking paradigm.

General Method

In the experiments reported here, observers performed a two-alternative-forced-choice search task, looking for a *T*, rotated to the left or right, among L-shaped objects. Two basic conditions were tested (Experiment 3 was an exception). In all the trials, half of the distractors were previewed for 1000 ms, after which the remaining distractors and the target were added. In the *valid preview* condition, the previewed items maintained their locations as the new items were added. In the *invalid preview* condition, the previewed items instantly moved at random to previously unoccupied locations. The difference in search efficiency between the two conditions reveals visual marking. Its strength can be indexed by the magnitude of slope reduction.

Participants

Participants were students from Yale University and Vanderbilt University who volunteered in the study for course credit or for pay. They were 18 to 34 years old; all had normal or corrected-to-normal visual acuity and normal color vision.

General Treatment of the Data

In this study, mean error rate was below 5% in all cells of the factorial design. An analysis of variance (ANOVA) calculated on accuracy revealed no significant effect. Analysis of accuracy is therefore not reported further. Median RT on correct trials within each cell of the factorial design was calculated for each individual. Means of the median RTs from different observers then underwent statistical analyses. In addition to RT, we report search rate and intercept. Whether overall set size or old set size is the predictor will be specified within the experiments. Statistical significance was based on an ANOVA using RT as the dependent measure.

Experiment 1

Visual Marking When Old and New Set Sizes Covary

Potential interactions among new set size, old set size, and preview condition would be most apparent when the set sizes vary widely. Thus, we tested overall set sizes of 17, 33, and 49, a range significantly greater than the typical range of 4 to 16. One of the items was the target. Half of the distractors on each trial were previewed for 1 s. The corresponding number of old distractors was 8, 16, and 24. Valid preview and invalid preview conditions will be contrasted. It should be noted that in this design, the new and old set sizes were positively correlated, similar to most extant studies of visual marking (Watson & Humphreys, 1997), and hence we will refer to this design as the “standard design.”

Large set sizes have been tested previously by Theeuwes et al. (1998) to measure the maximal number of old items ignored in visual marking. A change in slope, as measured by the function relating RT to overall set size, was apparent in that study, seemingly discounting the necessity to test large set sizes again. Yet, we are motivated to do so, for two reasons. First, nearly 10 experiments in our lab have failed to show a change in slope using conditions similar to those of Watson and Humphreys (1997) or Theeuwes et al. Second, observers in Theeuwes et al.'s study were first practiced and then tested in a fairly large number of valid preview trials (450). Practice may have enhanced the slope effect. We thus tested two groups of observers. One group was unpracticed, in that this was the first visual marking experiment that they ever performed. The other group was practiced, in that they had previously participated in three visual marking experiments, containing a total of more than 1,200 visual search trials, half of which were valid preview trials. Comparison of the two groups should reveal how practice affects visual marking. Eight practiced and eight unpracticed observers were tested.

Method

Equipment. The experiment was conducted on a Macintosh computer (PowerPC) with a 17-in. monitor, using MacProbe software (Hunt, 1994). Observers were tested individually in a room with normal interior lighting. They sat at an unrestricted distance from the computer screen of about 57 cm, at which distance, 1 cm corresponds to 1° visual angle.

Materials. Observers searched for a white *T* rotated left or right among white *L* distractors of four orientations presented on a gray background. Each item subtended 0.69 cm × 0.69 cm. The line segments forming the *L*s were offset by 1 pixel (0.03 cm) at their junctions. The width of each line segment was 0.09 cm. The locations of the items were randomly chosen from an invisible 8 × 8 matrix that subtended 20.00 cm × 20.00 cm. Each item was positioned at the center of a cell in the matrix.

Design. The experiment had a 2 × 3 factorial design. The *condition* factor had two levels: valid preview and invalid preview. In both conditions, half of the distractors were previewed for 1 s prior to the onset of the other distractors and the target. The conditions differed in whether the previewed items maintained their locations or changed to previously blank locations as the new items were added. The *set size* factor had three levels. Set size in this experiment refers to the total number of search items once all items were presented: 17, 33, or 49; one of these items was the target. The two preview conditions were intermixed within a session.

A target was present on every trial, rotated 90° to the left or right, with the constraint that an equal number of left and right *T*s were presented in each cell of the factorial design. Observers pressed one of two keys to identify the orientation.

Procedure. Each trial started with a blank screen for 1 s, followed by the *preview display* that lasted 1,000 ms. Then other items were added on the screen to form the *search display*, which stayed on until response. Correct response was rewarded by a beep. Approximately 500 ms later, the next trial started.

Each observer received 12 practice and 432 experimental trials. Before practice, observers were informed that when the old items stayed in place, the target would always be among the newly added items. Observers were instructed to respond as quickly as possible without sacrificing accuracy.

Results

First, we pooled results across all 16 observers. The means of median RT are shown in the left panel of Figure 1. Table 1 shows the search slopes and intercepts calculated on overall set size. A repeated measures ANOVA testing group (practiced vs. unpracticed observers), condition (valid vs. invalid preview), and set size (17, 33, or 49) showed that RT was comparable in practiced and unpracticed observers, $F(1, 14) < 1$. There was a significant main effect of preview condition, $F(1, 14) = 15.66$, $p < .01$, with faster RT in the valid preview condition. The overall preview benefit was comparable in the two groups of observers, reflected by a nonsignificant interaction between group and condition, $F(1, 14) = 1.81$, $p > .20$. RT also increased significantly as set size increased, $F(2, 28) = 85.75$, $p < .0001$. The interaction between group and set size was not significant, $F(2, 28) < 1$, indicating that search slope was comparable in practiced and unpracticed observers.

The interaction between condition and set size was not significant, $F(2, 28) = 2.40$, $p > .10$, indicating that the search slopes were comparable in the two preview conditions. However, the interaction among group, preview condition, and set size was significant, $F(2, 28) = 4.95$, $p < .02$. Follow-up tests showed that, for unpracticed observers, the search slopes in the two preview conditions did not differ, $F(2, 14) < 1$. The preview effect was reflected only in a benefit of the overall RT, $F(1, 7) = 5.86$, $p < .05$. In contrast, for practiced observers, the search slope in the valid preview condition (15 ms/item) was significantly shallower than that in the invalid preview condition (24 ms/item), $F(2, 14) = 8.03$, $p < .005$. The ratio of these two slopes is 0.62:1, a value significantly smaller than 1, $t(7) = 7.28$, $p < .0001$, but also significantly greater than 0.50, $t(7) = 2.43$, $p < .05$. Because there was no reduction in slope with naïve observers, the capacity of visual marking for these observers was less than 8. Practiced

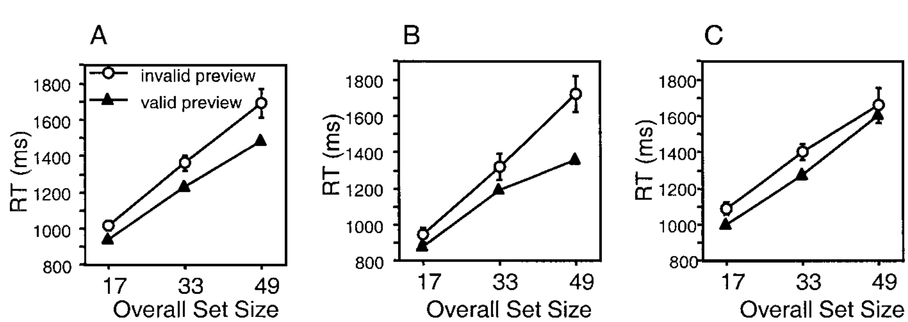

Figure 1. Results from Experiment 1: Visual marking measured when new and old set sizes positively covary on practiced and unpracticed observers. Panel A: All observers. Panel B: Practiced observers. Panel C: Unpracticed observers. RT = response time.

Table 1
Search Slopes (in ms/item) and Intercepts (in ms) of Response Time as a Function of Set Size in Experiment 1

Preview condition	Slope	Intercept
All 16 subjects		
Invalid preview	21 (2)	659 (67)
Valid preview	17 (2)	658 (61)
8 practiced search slopes		
Invalid preview	24 (3)	525 (67)
Valid preview	15 (2)	654 (61)
8 unpracticed search slopes		
Invalid preview	18 (3)	793 (97)
Valid preview	19 (3)	662 (110)

Note. Standard errors are shown in parentheses.

observers showed a reduction in overall-set-size slope on the order of 38%.

Discussion

Practiced and unpracticed observers were tested in Experiment 1 using large display set sizes. As in past studies of visual marking, Experiment 1 used a design that equated the number of new items and the number of old items on each display. The perplexing aspect of the data is that, for unpracticed observers, the benefit of previewing half of the distractors appears only in the overall RT, not in the slope of the search function relating RT to overall set size. We have failed to observe a change in search slope in nearly 10 experiments (Jiang, 2001), certainly nothing like the 50% reduction reported by Watson and Humphreys (1997).

The lack of a slope effect may partly be attributed to lack of sufficient statistical power. When data are pooled across all of Jiang's (2001) experiments that revealed a significant overall effect on RT, there are significant main effects of preview condition ($p < .0001$), overall set size ($p < .0001$), and their interaction ($p < .002$). The mean ratio between the search slopes of the valid and invalid preview conditions is 0.88, with values ranging from 0.45 to 1.29. The mean ratio is significantly greater than 0.50, $t(18) = 6.86$, $p < .0001$, the level predicted by an efficient visual marking system, but significantly smaller than 1.00, $t(18) = 2.26$, $p < .04$, the level predicted by a complete lack of visual marking. In short, there is a significant but small slope reduction when the data are pooled across approximately 100 observers. This modest and unstable slope effect is problematic. Nearly all experiments published so far by other researchers have reported a reduction in slope in the valid preview condition on the order of 50%; furthermore, most of these experiments tested only 12 unpracticed observers (e.g., Watson and Humphreys, 1997).

We have since ruled out several explanations of the lack of slope effect in our study (Jiang, 2001). First, there was no change in overall-set-size slope after introducing a color distinction between the old and the new items, a testing condition that more closely resembled the original study by Watson and Humphreys (1997). Second, changing the target discrimination task to a target detection task failed to induce a change in slope. Third, the discrepancy cannot be explained by the mixed-block design used in this study because a study in which the different preview conditions appeared in different blocks did not reveal a slope effect. Finally, the discrepancy was not due to the use of invalid preview condition as a baseline, because the slopes were similar in the no-preview and invalid preview conditions.

The significant change in slope found by other researchers using a design similar to ours can be explained partly by practice and partly by the small range of set sizes tested. Thus, Theeuwes et al. (1998) gave each observer 450 valid preview trials.[2] Although Watson and Humphreys (1997) tested naïve observers, they used smaller overall set sizes (4, 8, and 12), minimizing the effect of new items.

In sum, Experiment 1 showed a clear difference in intercept (or overall RT) between the valid and the invalid preview conditions. Problematic, however, is the lack of a consistent effect on the overall-set-size slope for unpracticed observers. This slope was calculated using overall set size as the predictor, and consequently was "contaminated" by the new set size. A better approach would be to calculate the slope of RT as a function of old set size, a cleaner indicator of the efficiency of marking the old items. Unfortunately, because the increase in old set size was always accompanied by an increase in new set size in Experiment 1, the effect of old set size cannot be isolated. Experiment 2 varied new and old set sizes orthogonally to find out whether the failure to find a reduction of slope in unpracticed observers was a true null effect or whether the covariation of new and old set sizes acted to mask the effect of old set size on visual marking.

Experiment 2

Decoupling the Size of the New and Old Sets

In the design of Experiment 1, where the number of items in the new and old sets was always the same, one would predict a reduction of search slope by 50% only if visual marking is equally effective at all new set sizes. This assumption may be too strong. Suppose that the visual marking mechanism becomes less effective as the number of new items increases, possibly because the marking signal dissipates over time. If so, then observers may be able to ignore a large number of old items if the new set is small but not when the new set is large. If this is true, then sampling the points at which new and old sets are both small and when they are both large would conceal evidence of visual marking, as explained in more detail in the *Results and Discussion* section of this experiment.

This problem applies to all studies that covary the number of new and old sets. Previous studies did not show this problem because they tested only small new set sizes (typically ranging from 2 to 8). Experiment 1 sampled a wider range of new set sizes,

[2] In addition to the amount of practice, which we knew to make a difference on the basis of the results of Experiment 1, there are other procedural differences between our study and Theeuwes et al.'s study (1998). These include baseline (invalid vs. no-preview), design (mixed vs. blocked conditions), task (two-alternative-forced-choice vs. detection), group of observers, and other differences. It is possible that when combined, all the procedural differences may contribute to the discrepancy in results.

from 8 to 24, exacerbating the possible effect of new set size and its interaction with old set size and preview condition.

To test the effect of the numbers of new and old items on the efficiency of visual marking, Experiment 2 varied new and old set sizes orthogonally. Old set size could be 3, 15, or 30, and new set size could be 3, 9, or 15. The efficiency of visual marking of old items is revealed by the slope of the function relating RT to the number of old items at each level of new set size.

If valid preview enhances only the overall RT of visual search across small and large new set sizes, as implied in Experiment 1, then the old-set-size slope should be comparable across the valid and the invalid preview conditions. This should be the case independent of new set size. In contrast, if valid preview enhances search rate, then the old-set-size slope should be shallower in the valid than the invalid preview condition. Finally, if the efficiency of marking the old items is affected by the number of new items, then the reduction in slope produced by previewing old items should itself decline as the new set increases in size.

In an ideal case, if visual marking is perfect in the valid preview condition and absent in the invalid preview condition, then the slope of the old set size function in the valid preview condition should be flat (near zero) across all levels of new set size. The RT in the old set size function in the invalid preview condition should all be linearly increasing across all levels of new set size. There should be a constant increase in the magnitude of the intercepts of the old set size function in both valid and invalid preview conditions as new set size increases, leading to a constant intercept difference between valid and invalid preview conditions across different levels of new set size. This ideal case predicts an absence of interaction between new set size and preview condition.

Method

This experiment closely duplicated Experiment 1, except that we manipulated the sizes of new and old sets independently. The three factors were: new set size (3, 9, or 15), old set size (3, 15, 30), and preview validity (valid or invalid). Seventeen unpracticed observers received 12 practice and 432 experimental trials.

Results and Discussion

Means of individual median RT are shown in Figure 2. The most striking finding is that, holding the number of new items constant, the search slope (RT vs. old set size) in the valid preview condition was almost flat. The search rate was 2, 5.5, and 10 ms/item when the new set had 3, 9, and 15 items, respectively. Across small and large new sets, RT was little affected by the number of old items, indicating that marking of old items was quite efficient. Using 10 ms/item as a rough criterion, the search slopes of the valid preview condition could be categorized as "efficient search" or "parallel search" (Treisman & Gelade, 1980; Wolfe, 1994, 1998). Table 2 shows the slopes and intercepts of linear functions relating RT to old set size.

We entered new set size, old set size, and preview condition as three within-subject factors in an ANOVA. All three main effects were significant: new set size, $F(2, 15) = 55.52$, $p < .0001$, with RT increasing as the number of new items increased; old set size, $F(2, 15) = 54.08$, $p < .0001$, with RT increasing as the number of old items increased; and preview condition, $F(1, 16) = 72.42$, $p < .0001$, with RT faster in the valid preview condition than the invalid preview condition.

The interaction between new set size and preview condition was significant, $F(2, 15) = 7.63$, $p < .005$. Specifically, when the new set had 3 items, the average difference in RT between the valid and the invalid preview conditions was 305 ms. When the new set increased to 9 items, the difference decreased to 186 ms, $F(1, 16) = 12.04$, $p < .003$. Increasing the new set size further, to 15 items, led to a slight but insignificant increase in the overall RT difference to 196 ms, $F(1, 16) < 1$. Thus, as new set size increased, the overall RT benefit of the preview diminished. This diminution in overall benefit cannot be attributed to a reduced effect on old-set-size slope, because the interaction among new set, condition, and old set was not significant, $F(4, 13) < 1$. Rather, the intercept effect, not the slope effect, decreased as new set size increased. It appears as if it took the observers longer to initiate the process of marking old items as new set size became larger. We call this finding the *differential intercept shift*.

The interaction between new set size and old set size was significant, $F(4, 13) = 3.83$, $p < .03$. The old-set-size slope was steeper when the new set had 15 items than when it had 9 or 3, $Fs(2, 15) > 4.89$, $ps < .03$. This was true in both the valid and the invalid preview conditions, as the three-way interaction was not

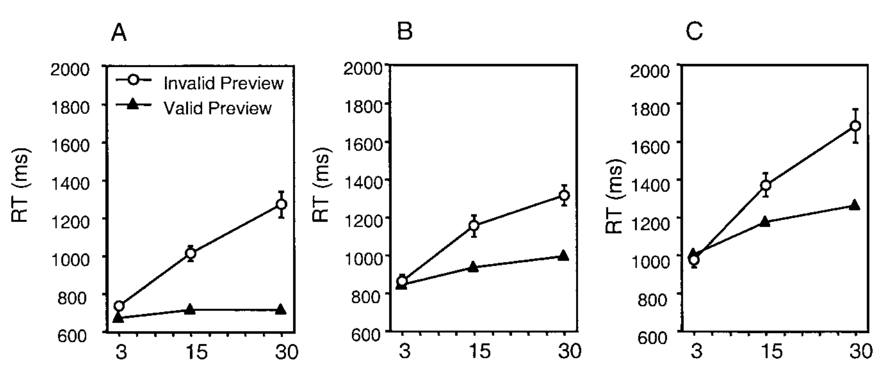

Figure 2. Results from Experiment 2: Visual marking measured when new and old set sizes were varied independently on unpracticed observers. Panel A: New set size = 3. Panel B: New set size = 9. Panel C: New set size = 15. RT = response time.

Table 2
Slopes (in ms/item) and Intercepts (in ms) of Response Time as a Function of Old Set in Experiment 2

New set	Preview condition	Slope	Intercept
3	Invalid preview	18 (2)	640 (50)
3	Valid preview	2 (1)	593 (58)
9	Invalid preview	16 (2)	772 (71)
9	Valid preview	6 (2)	737 (70)
15	Invalid preview	25 (4)	802 (78)
15	Valid preview	10 (2)	873 (90)

Note. Standard errors are shown in parentheses.

significant, $F(4, 13) < 1$. Visual search in general became less efficient as the size of the new set increased.

The interaction between preview condition and old set size was significant, $F(2, 15) = 38.43$, $p < .0001$, with search slope being much steeper in the invalid than the valid preview condition. Thus, there was a clear effect of preview evident in the old-set-size slope.

Finally, these data can be directly compared with the data collected in Experiment 1. That is, for a subset of the present data, the new set size equaled the old set size (both = 3 or 15). When analysis was restricted to these data, the slope of the function relating RT to overall set size was 26 ms/item for the invalid preview condition and 21 ms/item for the valid preview condition. That is, the reduction of slope in the valid preview compared with the invalid preview was about 19%. This ratio was significantly smaller than 1 ($p < .04$) but also far greater than the 50% reduction that characterizes an efficient marking process ($p < .001$).

The modest reduction in overall-set-size slope, as just calculated, sharply contrasts with the shallow old-set-size slope disclosed in Figure 2. The apparent paradox can be explained by the interaction between new set size and preview condition, as described earlier. The overall benefit in RT decreased as the new set increased in size, suggesting that the greater the number of new items, the smaller the preview benefit in overall RT. When we held the number of new items constant, this difference was incorporated into the intercept of the *RT–old-set function*. Consequently, the slope alone reflected the cost of old items. Conversely, when we selected the points at which new set size equaled old set size, the effect of new set size contributed to the slope of the *RT–overall-set-size function*. Consequently, this slope reflected not only the cost of old items, but also the effect of new items on the preview benefit.

Figure 3 illustrates two ways to calculate the search slopes. In the first design (which we refer to as the "preferred design"), new and old set sizes vary independently. Search slope is calculated using old set size as the predictor at each constant level of new set size, and each regression reveals the effect of old set size on RT, the proper measure of visual marking. This function can be written as

$$RT = b \times N_{old} + a; \quad (1)$$

where b is the search slope and a is a constant. If all of the old items can be effectively ignored, then the RT function in the valid preview condition will be flat. Not all studies vary new and old set sizes orthogonally. Extant studies typically covary new and old set sizes. Larger new sets are associated with larger old sets; each set comprises half of the distractors. We call this the "standard design." In this design, it is impossible to partial out the effect of new set size and, consequently, one is forced to measure the slope of RT as a function of overall set size. The RT function can be written as

$$RT = b \times (N_{new} + N_{old}) + a. \quad (2)$$

In the standard design, N_{new} always equals N_{old}. If all the old items can be effectively ignored, then the number of items searched in the valid preview condition will be half the number of items searched in the invalid preview condition. This leads to the prediction that the slope in the valid preview condition will be 50% as large as the slope in the invalid preview condition. Thus, an efficient visual marking process will lead to a flat old-set-size slope in the preferred design and a 50% reduction in overall-set-size slope in the standard design. This ideal case is shown in the left panel of Figure 3. Because new set size influences the preview

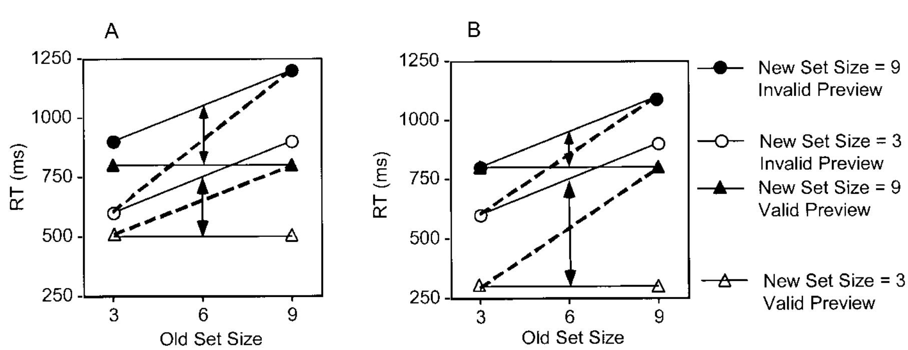

Figure 3. Illustration of how using two methods to calculate slope can lead to discrepancy. Panel A shows the ideal case in which the ratio of the invalid to valid slopes is 2:1. The intercept difference between valid and invalid preview conditions is consistent across the two levels of new set sizes. Panel B shows the actual finding in which the ratio of the invalid to valid slopes is close to 1:1. The intercept difference between valid and invalid preview conditions is smaller as the number of new items increases. The dashed lines show the slopes calculated using the standard design (Experiment 1). The ranges on the *x*- and *y*-axes are arbitrarily determined. RT = response time.

benefit, however, a 50% reduction in overall-set-size slope cannot always be achieved in the standard design. Experiment 2 is informative in this regard. Experiment 2 used the preferred design, which varied new and old set sizes independently. We found a near flat old-set-size slope in the valid preview condition across three levels of new set size. Thus, the new set size did not directly influence the efficiency of visual marking as measured by the old-set-size slope benefit. However, the size of the new set did affect the overall RT difference between the valid and the invalid preview conditions.

The right-hand panel of Figure 3 shows an exaggerated case where the new set size has no direct effect on the old-set-size slope of the valid preview condition but has a large effect on the intercept difference between the valid and the invalid preview conditions. As new set size increases, the average difference in overall RT between the valid and the invalid preview conditions is drastically reduced. This is the differential intercept shift. In this case, if we recalculate the slope of RT as a function of overall set size (the standard design), search in the valid preview condition may have a slope that is the same as or even larger than that in the invalid preview condition.

Thus, in the standard design, where the new set size equals the old set size, the old-set slope cannot be derived, and the overall-set slope is determined by the combined effects of new and old set sizes. Even when RT is completely independent of the old set size, the overall-set slope calculated in the standard design may not be reduced by the amount predicted by marking. If new set size correlates with old set size positively, as in most visual marking studies conducted in the past, one would underestimate the overall-set slope reduction. In many cases, the overall-set-size slope may be reduced by only a modest amount (about 12%). Conversely, if new set size correlates with old set size negatively, one would overestimate the overall-set-size slope reduction.

We note that we have simplified the discussion and attributed the effect of new set size almost entirely to a differential intercept shift. It is possible that increased new set size also produced other effects, such as the nonlinearity in the RT curve (see Figure 2) and a small modulation effect on the slope of visual marking. For instance, the slope of old set size was steeper as new set sizes increased, seemingly supporting the view that new set size may directly affect the slope effect of visual marking. However, the search curve was steeper for both valid and invalid preview trials as new set size increased; perhaps search in general became less efficient as the display became more crowded. We believe that the effect of new set size may not be restricted to the intercept shift. In any case, new and old set should be manipulated orthogonally as the benefit of preview is affected jointly by new and old set sizes.

In sum, the slopes of the search functions measured in Experiment 1 apparently reflected two effects: visual marking and a differential intercept shift. Old items were probably efficiently ignored, but the effect of marking on overall-set-size slope was masked by the differential intercept shift that varied across new set sizes. This interaction was not apparent in past studies because observers were typically tested with small numbers of new items (2–8) or were practiced. Results shown in Experiment 1 are thus not inconsistent with past studies. This discussion underscores the importance of manipulating the old and new set sizes orthogonally. Fifteen out of the 17 naïve observers tested in Experiment 2 gave a search slope smaller than 10 ms/item in the valid preview condition, indicating that the slope relating RT to old set size did not hinge on practice.

Experiment 3

Efficient and Inefficient Search

What produced the differential intercept shift? We have previously considered general task difficulty as the source (Jiang, 2001). We suspected that the preview benefit might always be smaller when the task is more difficult. To rule out this possibility, we varied search difficulty by manipulating the similarity between the target and distractors (Duncan & Humphreys, 1989; Jiang, 2001). The differential intercept shift was observed in both the easy and the difficult search conditions. Furthermore, the preview benefit was larger in the difficult search condition, indicating that the preview benefit sometimes may increase as the task becomes more difficult, and thus leading us to reject general task difficulty as the basis for the differential intercept shift.

Experiment 3 asks whether the differential intercept shift is observed only in the visual marking paradigm. The visual marking task is one case in which efficient and inefficient searches are combined in the same trial. Specifically, old items are rejected efficiently, whereas new items are searched inefficiently. The cue that separates the new and old items is temporal in nature (Jiang, Chun, & Marks, in press).

An analogous search task can be established by changing the temporal segregation cue to a spatial one. In Experiment 3, the target was a left or right rotated *T*. Its color was either white or black. The distractors were *L*s. Their colors were white, black, or red. Because the target was never in red, all the red distractors on the display could be rejected efficiently (Egeth, Virzi, & Garbart 1984; Kaptein, Theeuwes, & van der Heijden, 1995). The achromatic distractors, in contrast, had to be searched slowly. In a condition analogous to the invalid preview of visual marking, the target was presented among black and white items. Observers had to search through all items to detect the target. This is called the *exhaustive search* condition. In a condition analogous to the valid preview, the target was presented among achromatic and red items. Observers could reject the red items efficiently and search through the achromatic items slowly. This is called the *subset search* condition.

Similar to the visual marking task, we measured the slope relating RT to the set size of two different distractor sets. The *pseudo-old* set consisted of either red items that could be rejected efficiently leading to subset search (analogous to the valid preview condition in the visual marking paradigm) or achromatic items that could be rejected only by serial search (analogous to the invalid preview condition). The *pseudo-new* set consisted of achromatic items that could be rejected only by serial search (analogous to the new items). The question of interest is whether a differential intercept shift will appear in this paradigm.

If the differential intercept shift applies generally to paradigms in which efficient and inefficient searches are combined, the intercept shift should decrease as the size of the pseudo-new set (serial search) increases. Conversely, if the differential intercept shift is specific to visual marking, we should not find it in Experiment 3.

Method

Seven naïve observers were tested in 12 practice and 576 experimental trials. Three factors were manipulated. The first was search condition: The pseudo-old distractor set could be red or achromatic, producing "subset search" or "exhaustive search," respectively. The red pseudo-old set is analogous to the valid preview condition of the marking paradigm. The achromatic pseudo-old set is analogous to the invalid preview condition. When achromatic, the specific color was always chosen to be different from the target, which was also randomly chosen to be black or white on any given trial. The second factor was the number of items in the pseudo-old set: 3, 6, or 9. The last factor was the number of items in the pseudo-new set. The pseudo-new set was always achromatic, and its items always had a different color from that of the pseudo-old set. This set size was either 3 or 9. We have shown that using these levels of set size factors, a differential intercept shift can be observed in the visual marking paradigm (Jiang, 2001).

Each trial started with a fixation point that lasted 800 ms, after which the search display was presented. Observers were to press one of two keys to report the direction of the tilted *T*. They were instructed that the target would never be red, so they should ignore the red items. Acoustic feedback was given immediately after the response. The next trial commenced 1 s later. It should be noted that the two sets of distractors were presented simultaneously. A spatial segregation cue—color—separated them. Other aspects of the experiment were similar to Experiment 2.

Results

Means of median RT are shown in Figure 4. Using pseudo-old set, pseudo-new set, and search condition as the factors, an ANOVA showed significant main effects of the number of items in the pseudo-new set, $F(1, 6) = 38.57$, $p < .001$; the number of items in the pseudo-old set, $F(2, 12) = 12.14$, $p < .0001$; and search condition, $F(1, 6) = 34.56$, $p < .001$, with faster RT in the subset search condition than the exhaustive search condition.

The interaction between search condition and pseudo-old set was significant, $F(2, 12) = 7.00$, $p < .001$, showing shallower search slopes in the subset search condition. The interaction between pseudo-old and pseudo-new set was not significant, $F(2, 12) < 1$, nor was the three-way interaction, $F(2, 12) < 1$. Finally, the interaction between search condition and pseudo-new set was not significant, $F(1, 6) < 1$, *ns*. Thus, Experiment 3 gave no evidence of a differential intercept shift. The mean RT difference between the subset search and the exhaustive search conditions was 155 ms when the size of the pseudo-new items was 3, and the mean RT difference was 142 ms when the size of the pseudo-new items was 9. Thus, the differential intercept shift appears to be specific to the visual marking task; it does not appear in all tasks that combine efficient and inefficient searches. Table 3 shows the slopes and intercepts.

Discussion

Although the number of old items had little effect on search rate across small and large new set sizes in visual marking, the overall magnitude of the preview benefit was smaller when new set size was larger. Experiment 3 indicates that this differential intercept shift may be restricted to visual marking, which appears to rely on temporal segregation cues (Jiang et al., in press). In an analogous task that combined efficient and inefficient searches but used spatial segregation cues to differentiate the two sets of distractors, the overall difference in RT between the exhaustive search and the subset search conditions was not affected by the number of slowly rejected items. This implies that the differential intercept shift of visual marking is specific to new items segregated by temporal cues.

A complete understanding of the differential intercept shift requires further investigation. To be certain that this effect is worth studying, one needs to make sure that it is systematically induced by only a small number of factors. Two experiments serve as first steps toward this goal. Task difficulty does not account for the shift (Jiang, 2001). Furthermore, the shift is restricted to the visual marking paradigm, which relies on temporal grouping cues (Experiment 3). These experiments have narrowed the explanations of the intercept shift down to mechanisms related to the cue that separates new and old items. Here we provide some possible explanations.

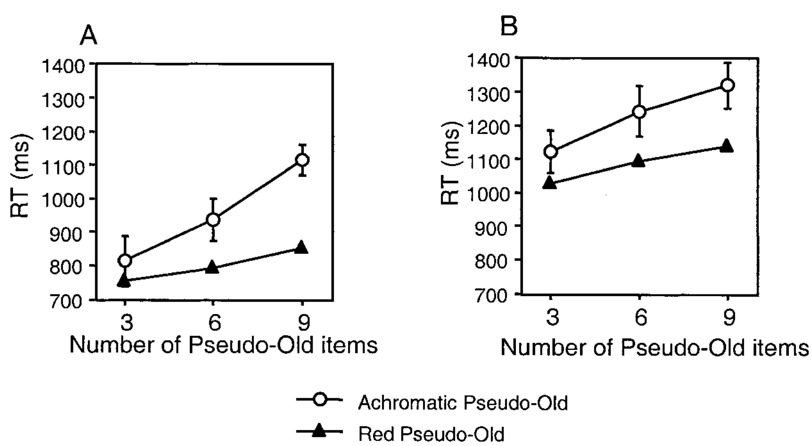

Figure 4. Results from Experiment 3: Subset search based on color—no differential intercept shift. Panel A: The number of pseudo-new items = 3. Panel B: The number of pseudo-new items = 9. The circles represent results from exhaustive search whereas the triangles represent results from subset search. RT = response time.

Table 3
Slopes (in ms/item) and Intercepts (in ms) of Response Time as a Function of Old Set in Experiment 3

Set 1 size	Search type	Slope	Intercept
3	Exhaustive	50 (9)	653 (51)
3	Subset	16 (4)	701 (51)
9	Exhaustive	33 (12)	1030 (140)
9	Subset	19 (11)	974 (107)

Note. Standard errors are shown in parentheses.

Our hypothesis for the differential intercept shift is based on our temporal grouping account of visual marking (Jiang et al., in press). This account states that visual marking is a two-step process. The first step is the segregation of the new and old items into two temporal groups. The segregation is accomplished by asynchronous presentation of the two sets of items in the time domain. Following the segregation, the second step is to deploy attention selectively to the group that contains the target—the new items. Evidence for the second step—attentional deprioritization of old items—has been obtained by previous studies, noticeably those of Watson and Humphreys (1997, 1998, 2000). Evidence for the first step—perceptual segregation of the new and old items through temporal asynchrony—has been advanced by our recent study (Jiang et al., in press). We have shown that the asynchronous presentation of the new and old items is necessary for visual marking to occur, as any cue that resets and synchronizes the two sets of items is detrimental.

Finally, we have discounted an important alternative explanation of the differential intercept shift. This alternative relates to the decay of the temporal segregation cue. As new set size increases, it takes longer for visual search to complete. Longer search requires a temporal segregation cue that lasts longer. If the segregation cue dissipates as search progresses, visual marking may become less efficient as the number of new items increase in number. Although it is plausible that the temporal segregation cue dissipates, so far we have not obtained convincing evidence to support it. In fact, at least two findings argue against it. First, if the temporal cue dissipates over time, then visual marking should be less efficient as search difficulty increases. This hypothesis was rejected in a previously mentioned experiment (Jiang, 2001). Second, if larger new set size is associated with less efficient visual marking, we should observe a reduction in old-set-size slope at larger new set sizes. Evidence for this hypothesis is ambiguous. Experiment 2 showed that the slope of the valid preview condition was steeper at larger new set sizes, apparently supporting the prediction. But the slope of the invalid preview condition was steeper at larger new set sizes as well, suggesting that the increased slope in the valid preview condition may stem from general inefficiency of visual search at larger set sizes. The hypothesis about dissipating visual marking with increasing new set size is therefore not supported by the data gathered so far. An understanding of the differential intercept shift is more likely to come from studies that test the properties of the temporal segregation cue. Regardless of the exact source of the differential intercept shift, its existence is clear. It is thus important for researchers to design studies that partial out contamination from new set sizes.

Conclusions

Visual marking should be tested by varying new and old set sizes independently, and marking should be measured by the function relating RT to old set size instead of overall set size. We have demonstrated that old-set-size slopes provide a more sensitive measure of the preview benefit in visual marking, which may be obscured when overall set size is used. Visual marking appears to be efficient across small (e.g., 3) and large (e.g., 9) sizes of the new set, because as many as 30 old items can be efficiently ignored. But the overall benefit in RT decreases as new set size increases, showing a differential intercept shift that can obscure differences in search efficiency at larger new set sizes. Because studies that measure overall set size typically covary new and old set sizes, the effect of old set size, and in turn, visual marking, cannot be isolated. The contamination of search by the size of new set may lead to premature rejection of visual marking if researchers wrongly accept the null effect in the overall-set-size slope. Future studies on visual marking should therefore independently vary new and old set sizes (as in Theeuwes et al., 1998). Additional studies are required to pinpoint the source of the differential intercept shift, especially because it appears to be restricted to visual marking, which uses a temporal cue to segregate two sets of distractors (Jiang et al., in press).

References

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

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

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 32–39.

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

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. (2001). Visual marking. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61* (10-B), 5589.

Jiang, Y., Chun, M. M., & Marks, L. E. (in press). Visual marking: Selective attention to asynchronous temporal groups. *Journal of Experimental Psychology: Human Perception and Performance.*

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 & Performance, 21,* 1053–1067.

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: Human Experimental Psychology, 52*(A), 679–715.

Paquet, L., & Lortie, C. (1990). Evidence for early selection: Precuing target location reduces interference from same-category distractors. *Perception & Psychophysics, 48,* 382–388.

Pashler, H. (1998). *Attention.* Hove, England: Psychology Press.

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

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

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. (1998). Visual marking of moving objects: A role for top-down feature-based inhibition in selection. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 946–962.

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

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

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9,* 33–39.

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 November 11, 2000
Revision received May 31, 2001
Accepted September 20, 2001 ■