Journal of Experimental Psychology:
Human Perception and Performance
2012, Vol. 38, No. 2, 350–366

0096-1523/12/$12.00 DOI: 10.1037/a0025794

Determining the Capacity of Time-Based Selection

Derrick G. Watson and Melina A. Kunar
University of Warwick

In visual search, a set of distractor items can be suppressed from future selection if they are presented (previewed) before a second set of search items arrive. This *visual marking* mechanism provides a top-down way of prioritizing the selection of new stimuli, at the expense of old stimuli already in the field (Watson & Humphreys, 1997). Typically, this *preview benefit* has been examined by measuring the efficiency of detecting a target item contained within a newly presented set. Such work has led to the suggestion that the capacity of the visual marking mechanism is extremely high. Here we present five experiments which measured performance for selecting and responding to all the new stimuli rather than a single target item within the new set. The findings illustrate that when selecting and responding to all new items intentionally trying to prioritize new stimuli has a capacity limit of approximately six to seven items and that this limit depends partly on properties of the stimuli. The findings are discussed in terms of mechanisms of time-based selection, attentional capacity limits, and the task demands of multiple item selection.

Keywords: visual marking, visual search, attentional capacity, onset capture

The visual world presents us with vast amounts of information at any particular point in time. For some tasks, certain pieces of information may be more relevant to our goals than others—for other tasks the reverse may be true. Given that we do not have the capacity to process everything that is thrown at us, our attentional system needs to discount information that is of little current value and enhance the processing of important information required for our present behavior. There are a number of ways this can be achieved. At one end of the scale lay relatively automatic processes. For example, the appearance of a new perceptual object or a sudden luminance change can attract attention automatically to enhance processing at that location (Posner, 1980; Yantis & Jonides, 1984; Davoli, Suszko, & Abrams, 2007; Rauschenberger, 2003). In addition, we appear to be biased against reattending to a recently processed and rejected location (Posner & Cohen, 1984; Klein, 1988, 2000) in favor of searching new and previously unexamined locations.

In contrast, some selection processes appear to be under more voluntary control. For example, we have the ability to intentionally attend to a region of space on the basis of symbolic information (Jonides, 1981; Posner, Snyder, & Davidson, 1980). We can also influence which types of stimuli attract our attention by adopting an "attentional set." This serves to bias attention toward stimuli containing particular properties such as color, abrupt onsets, or motion (contingent orienting, Folk, Remington, & Johnston, 1992; Folk, Remington, & Wright, 1994; Folk & Anderson, 2010). Furthermore, the visual system is also able to "hold on to" and track a subset of target items among other items even if the nonselected items are identical to the attended ones (Pylyshyn & Storm, 1988).

Given that our cognitive resources are finite, these selection processes are subject to certain limitations and constraints. For example, the number of new objects that can attract attention automatically is, in some circumstances, limited to approximately three or four items (Yantis & Johnson, 1990; Yantis & Jones, 1991; but see Donk & Theeuwes, 2003 for an example where this limit is exceeded). Similarly, the number of items that can be tracked (Pylyshyn & Storm, 1988; Pylyshyn, 2004; Horowitz et al., 2007), or enumerated efficiently—subitized (Kaufman, Lord, Reese, & Volkman, 1949; Trick & Pylyshyn, 1993, 1994; Watson, Maylor, & Bruce, 2005, 2007) is also limited to approximately three to four objects. Other evidence shows that the precision with which attention can be biased to certain classes of stimuli might be relatively coarse (i.e., top-down limits can be set to which type of stimuli capture attention, Folk et al., 1992). Such capacity limits are also highlighted in difficult visual search tasks wherein reaction times (RTs) to find a predefined target increase linearly with the number of items in the display, as participants search each item in turn (see Wolfe, 1998, for a review).

Time-Based Selection

The work outlined above deals with selection of stimuli across space. However other work has examined selection across both space and time (Atchley, Jones, & Hoffman, 2003; Donk & Theeuwes, 2001, 2003; Gibson & Jiang, 2001; Jiang, Chun, & Marks, 2002a,b; Osugi, Kumada, & Kawahara, 2009, 2010; Theeuwes, Kramer, & Atchley, 1998; Watson & Humphreys, 1997, 1998, 2000; for a summary see Watson, Humphreys, & Olivers, 2003). In order to study whether time of appearance could be used to help

This article was published Online First October 17, 2011.

Derrick G. Watson and Melina A. Kunar, Department of Psychology, University of Warwick, United Kingdom.

We are grateful to Naz Al-Aidroos and Mieke Donk for valuable comments and suggestions regarding this study and thank Craig Scott and Edward Nottingham for their assistance with data collection.

Correspondence concerning this article should be addressed to Derrick G. Watson, Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom. E-mail: d.g.watson@warwick.ac.uk

guide selection to new items, Watson and Humphreys (1997) developed a modified visual search task. Typically in visual search, participants are asked to indicate the presence or absence of a prespecified target among a varying number of distractors (e.g., Treisman & Gelade, 1980). The time taken to respond as a function of the number of distractors (i.e., the search slope) indicates the difficulty of the search, with steep slopes indicating difficult/inefficient search and shallow slopes indicating easy/efficient search (Wolfe, 1998).

Watson and Humphreys (1997) divided a standard color–form conjunction search task (Treisman & Gelade, 1980) into two temporally separated parts—a method known as the preview paradigm. One set of old distractors (green *H*s) was first presented (previewed) for 1000 ms. Following this, a set of new items (blue *A*s) which would contain a blue *H* target (when present) was added to the first display. The task was to detect whether the blue *H* target was present or absent among the new items. Performance in this preview condition was compared with a full element baseline (FEB) in which all the display items appeared simultaneously and a half element baseline (HEB) in which only the new items from the preview condition were presented. Search efficiency, in terms of the RT × Set size function, was found to be more efficient in the preview condition than in the FEB and statistically equivalent to the HEB. That is, search in the preview condition was the same as if the old items had not been presented. This indicates that participants were able to prioritize the new items and ignore the old stimuli, an effect termed the *preview benefit*. Watson and Humphreys (1997) proposed that this prioritization for new items occurred as the result of participants suppressing the old (previewed) items via a top-down capacity-limited inhibitory mechanism, which they termed *visual marking*.

However, this proposal has not gone without challenge. For example, Donk and colleagues (e.g., Donk, 2005, 2006; Donk & Theeuwes, 2001, 2003; Donk & Verburg, 2004) proposed that the preview benefit occurs because the abrupt luminance onsets associated with the new items capture attention automatically. When they eliminated luminance onsets by making the new objects isoluminant with the background, a preview benefit did not occur (Donk & Theeuwes, 2001, although see Braithwaite, Hulleman, Watson, & Humphreys, 2006). A second alternative by Jiang and colleagues (Jiang, Chun, & Marks, 2002ab; Jiang & Wang, 2004) proposes that the old and new items form asynchronous temporal groups. Following this grouping process, attention can then be directed and selectively allocated to either the old or to the new group. Although there is evidence that both onsets and grouping contribute to obtaining a preview benefit, these accounts alone have trouble explaining some other results from preview work (e.g., Kunar, Humphreys, & Smith, 2003; Humphreys, Watson, & Jolicoeur, 2002; Olivers & Humphreys, 2003; Braithwaite, Humphreys, & Hodsoll, 2003, 2004; Osugi, Kumada, & Kawahara, 2010; Watson, Braithwaite, & Humphreys, 2008). For example, abrupt luminance onset signals are likely to play a role in prioritizing new elements when available. However, it is also possible to obtain a preview benefit when stimuli are isoluminant with their background (e.g., Braithwaite, Hulleman, Watson, & Humphreys, 2005). Currently, the data provide evidence for the involvement of inhibition of old items, grouping and luminance onsets in driving the preview effect—although the exact contribution of each might depend on stimulus characteristics and task demands.

The Capacity of Time-Based Selection

Irrespective of the underlying mechanisms involved in prioritizing new visual information, the existing literature suggests that time-based visual selection appears to have a relatively large capacity. There are two strands of evidence supporting this. First, previous work has demonstrated that up to 30 old items (Jiang, Chun, & Marks, 2002a) and 15 new items (Theeuwes, Kramer, & Atchley, 1998) can be prioritized with no upper boundary yet established of the number of items that can be ignored. Second, one might expect that if the ability to ignore old items was limited to a small number of objects then search in the preview condition would be efficient at small set sizes but would then become less efficient at larger set sizes. As compared with the baseline conditions, this would produce a Bilinear RT × Set size search function in the preview condition. However, the majority of previous work has shown that preview search slopes are typically linear across the range of set sizes previously examined (but see Al-Aidroos, Emrich, Ferber & Pratt, in press, for recent evidence that bilinear preview search slopes can be obtained when a sufficient range of set sizes are used).

In contrast, other work has suggested that there might be limits to how many items can be selected with priority. Examining eye movements in preview search (see also Watson & Inglis, 2007), Emrich, Ruppel, Al-Aidroos, Pratt, and Ferber (2008) presented participants with 0 to 10 preview items to which eight new items were added and measured fixation frequency to old or new stimuli. The results showed that the probability of preferentially fixating a new item gradually decreased as the number of fixations increased. By about the fifth fixation, there was no difference in the probability of fixating either a new or an old, previewed item. Emrich et al. concluded that the ability to prioritize new stimuli appeared to decay with each fixation so that time-based selection might be limited to, at the most, approximately four new items. Note however, that in these experiments search efficiency (in terms of Manual RT × Set size slope measures) still remained more efficient than in their FEB condition, showing the usual preview benefit.

Comparing data from manual RT responses and that from eye movement studies seems to highlight a disparity in the capacity limit of the number of old items that can be ignored. It appears that attentional search and the observer's manual responses can remain efficient across high set sizes in time-based selection but eye movement behavior shows a limitation in the number of old items that can be discounted. One difference between the eye movement measure and the manual RT measure is that, with eye movements, an overt response is made to multiple items as part of the search process. In contrast, in the previous work on preview search, only a single manual response to one specific target has had to be made (even though multiple eye movements might still be required to locate the target). It is possible that if multiple manual overt responses have to be made to the new items then a capacity limit to process new items might also be revealed. This might arise because making multiple responses might interfere with the processes/representations needed to maintain the old from the new or might require the distinction between old and new information to be maintained for a greater length of time. We investigate this here.

Purpose of the Present Study

The current work examined capacity limits in preview conditions in which all the new objects have to be responded to manually. This contrasts with previous work where only a single target had to be detected and responded to among multiple new distractors.

Similar to previous studies, participants were first presented with a preview display. Following this, an additional set of items was added and participants were asked to respond to each new item by making a mouse click or touch screen response to them. This process departs from that typically used because participants had to respond to every new item rather than making a single response to a target presented among a set of distractors. We also used displays in which there was no simple feature difference (such as color or shape) between the old and the new stimuli (i.e., all stimuli were the same). This methodology allowed us to stress the mechanisms of time-based selection because the only way that participants were able to differentiate the old items from the new was by their temporal appearance.

In order to determine the capacity of time-based selection we employed a similar methodology as that used in Multiple Object Tracking (MOT) studies (Pylyshyn & Storm, 1988; Pylyshyn, 1994; Horowitz et al., 2007). In MOT tasks participants are asked to track a number of items among a set of distractors. Various measures within this literature have been developed to determine the tracking capacity of the visual attentional system. In the present work, the previewed items are treated as equivalent to the non-tracked items in MOT studies and the new items are equivalent to the tracked items.

We also examined the effect of display heterogeneity. Previous visual search studies have shown that increasing the similarity of distractors can allow them to be grouped and rejected more efficiently (Duncan & Humphreys, 1989, 1992). Thus we might expect that homogeneous preview displays might be grouped and suppressed more easily than previews which contain heterogeneous items, allowing better prioritization of newly appearing items. Accordingly, Experiments 1 to 3 examined color and shape heterogeneity within the displays. To summarize the results, when responding to all new items we found that the capacity of time-based selection was six to seven items. Furthermore, the capacity limit was greater with homogenous displays than with heterogeneous displays. In Experiment 4 we introduced a delay between the onset of the new items and when participants could start to respond. This delay reduced the new item capacity estimate by about one item. In Experiment 5 participants responded to each new item directly using a touch screen in one condition and a mouse click in another. Despite the change in response type a similar capacity of six to seven items was obtained. Overall the findings suggest that when all the new items have to be manually responded to there is a capacity limit of approximately six to seven items. This limit is reduced when stimuli are heterogeneous but remains relatively robust over different response types.

Experiment 1: Small Set Sizes and Color and Shape Heterogeneity

In Experiment 1 we established the basic methodology and examined capacity limits for ignoring up to eight old items and prioritizing up to eight new items. In addition to set size we also examined the effect of stimulus heterogeneity. In the mixed color condition the displays consisted of red, green, blue, and yellow randomly oriented letter *L*s. In the single color condition all the stimuli within a trial were of the same color (red, green, blue, or yellow).

Method

Participants

Twelve participants (3 male), ages 18 to 20 years (mean [*M*] = 18.8 years) took part for payment or course credits. All were undergraduate students at the University of Warwick and reported normal or corrected-to-normal visual acuity.

Stimuli and Apparatus

Stimuli were generated and presented by a custom program running on a Pentium-based PC computer attached to a Sony 19" CRT monitor. Displays were presented at a resolution of 1024 × 768 pixels running at refresh rate of 100Hz. The stimuli consisted of red (RGB = 200,0,0), green (RGB = 0,200,0), blue (RGB = 0,0,200), and yellow (RGB = 200,200,0) letter *L*s (15 mm × 15 mm) rotated at 0°, 90°, 180°, or 270°. Displays were generated by randomly placing the stimuli into the cells of a centrally placed 6 × 6 invisible grid. Center-to-center interstimulus spacing was 80 pixels (28 mm) and individual stimulus positions were varied by up to ± 8 pixels (3 mm). In the single color condition the stimuli in a single trial were either all red, all green, all blue, or all yellow. In the mixed color condition the preview items consisted of an equal number of each of the four colors, and the new items consisted of an equal number of all four colors. Thus all four colors were present in the same quantities in both the preview and search displays. This meant that participants were not able to determine which were the new items on the basis of color information contained in either the preview or the new item display. The orientation of each stimulus was selected randomly.

Design and Procedure

A 2 (color: single or mixed) × 3 (display size: 8, 12, or 16 items) within-participants design was used. A single trial consisted of a black screen for 500 ms, followed by a gray (RGB = 150, 150, 150) central fixation dot, followed by a preview display of four, six, or eight randomly oriented *L*s. After 1000 ms an additional four, six, or eight *L*s were added respectively (i.e., the number of new items was equal to the number of old items). Participants were instructed to ignore the preview items and to select the new ones when they appeared. Stimuli were selected by moving a mouse cursor over the item and left-clicking on it. Selecting an item immediately reduced its luminance (the transparency alpha level was changed from 1 to 0.3) After completing their selection, participants moved the mouse cursor to the right side of the screen until it disappeared and pressed the space key to initiate the start of the next trial (see Figure 1). This procedure allowed participants to take a break within the block by delaying when they pressed the space key. The number of items that participants could select was unrestricted. The experimental condition was comprised of 240

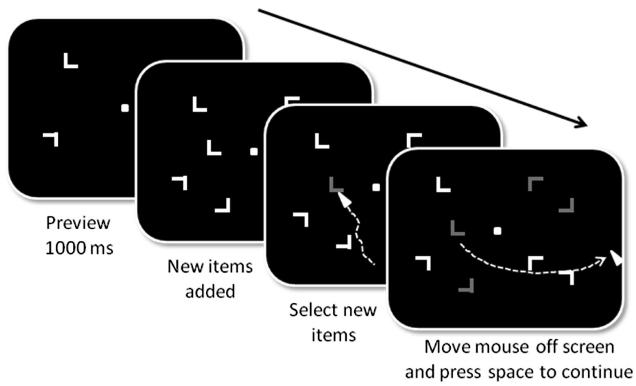

Figure 1. An example trial sequence from Experiment 1. A single trial sequence consisted of: a blank screen (500 ms), a fixation dot (1000 ms), a preview of old distractors (1000 ms), the presentation of new (target) items. Participants selected the new elements by clicking on them. When finished they moved the mouse cursor off screen to the right of the display and pressed the space bar to start the next trial.

trials (40 trials per color × display size combination) and participants completed a short practice block of practice trials directly before the experiment proper.

The main dependent variable of interest was the capacity limit for prioritizing new items. However, we also examined the total number of items that participants selected, the proportion of correct responses, the proportion of correct responses as a function of how many items had been previously selected, and response times. All of these measures were used to characterize observers' performance.

Results

Prior to these analyses, two trials (<0.07%) were discarded as outliers because the RT of the first response was less than 100 ms.

Number of Items Selected

The mean total number of items selected (both new and old) as a function of new item set size (four, six, eight) and color (single/mixed) is shown in Figure 2A (note that in all conditions, the number of new items was equal to the number of old items). As can be seen in both conditions, participants selected fewer items than the number of new items presented, particularly at the larger set sizes. A 2 (Color: mixed/single) × 3 (New item set size: four, six, or eight) within-subjects analysis of variance (ANOVA) revealed that participants selected more items in the single-color condition (M = 5.73) than in the mixed-color condition (M = 5.67), $F(1, 11) = 7.84$, mean standard error [MSE] = .009, $p < .05$, $\eta_p^2 = .416$, and, as one might expect, selected more items as the new set size increased, $F(2, 22) = 377.05$, $MSE = .174$, $p < .001$, $\eta_p^2 = .972$. The increase in number of items selected with set size was greater in the single color condition than in the mixed color condition, $F(2, 22) = 3.52$, $MSE = .007$, $p < .05$, $\eta_p^2 = .242$. As shown in Figure 2A, participants appeared to select proportionally fewer items as the new set size increased. To confirm this we calculated the slope of the number of items selected as a function of the new set size (this slope should be less than one if participants tended to select fewer items as set size increased). This revealed mean slopes of 0.84 responses/new set size (single color) and 0.81 responses/new item set size (mixed color) and both slopes were significantly less than one, $t(11) = 3.70$, $p < .005$, and $t(11) = 4.59$, $p = .001$, respectively. In both conditions participants selected proportionally fewer items overall as the number of new items increased.

Proportion Correct

To take into account that participants were free to select as few or as many items as they wanted, equation (1) was used to calculate the proportion of correct responses on each trial. Overall means are shown in Figure 2B.

$$\textit{Proportion correct} = \frac{\textit{new items selected} + \textit{old items not selected}}{\textit{total number of items}} \tag{1}$$

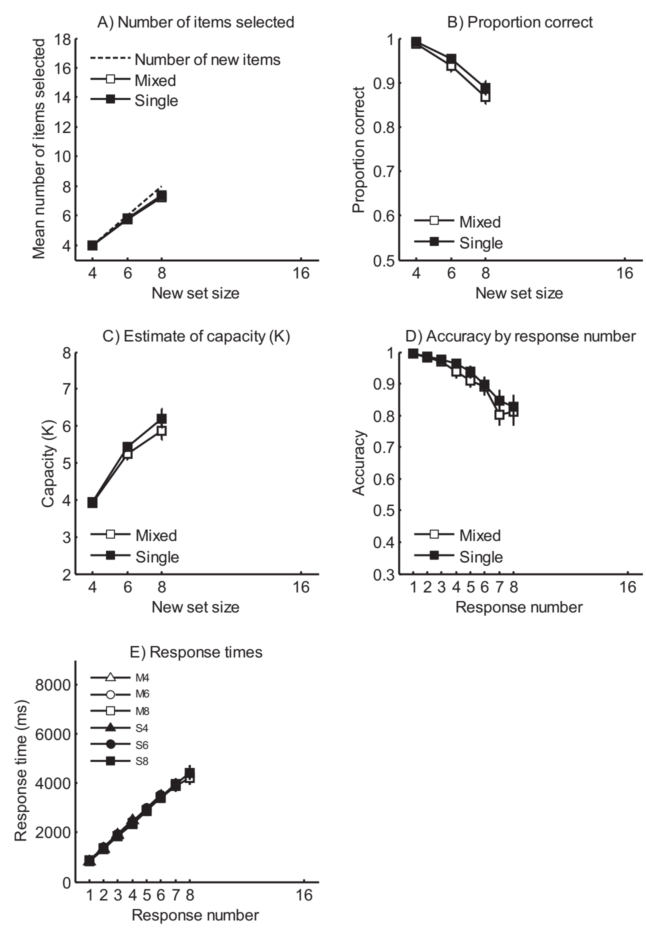

Figure 2. Results for Experiment 1. Panels A, B, and C show the total number of items selected, the proportion correct, and the estimate of new item capacity as a function of stimulus color (mixed or single) and new set size. Panels D and E show accuracy (mean probability of selecting a new item) and mean RTs as a function of stimulus color and the number of previous responses within the trial. Error bars indicate ±1 standard error [SE].

The proportion correct was lower overall for the mixed than for the single color displays, $F(1, 11) = 19.21$, $MSE = .000$, $p = .001$, $\eta_p^2 = .636$. Accuracy also decreased as the number of new items increased, $F(2, 22) = 67.25$, $MSE = .001$, $p < .001$, $\eta_p^2 = .859$ and this decrease was marginally greater in the mixed color displays than in the single color displays, $F(2, 22) = 3.01$, $MSE = .000$, $p = .07$, $\eta_p^2 = .215$.

Estimate of Capacity (K)

Because the number of new items was equal to the number of old items, following Wolfe, Place, and Horowitz (2007; see also Hulleman, 2005), we calculated the capacity of new item selection according to equation (2), where K = capacity, T = the number of new items, and p = the proportion correct response. The overall mean capacity estimates are shown in Figure 1C.

$$K = T(2p - 1) \tag{2}$$

A 2 (Color: single/mixed color) × 3 (Set size, four, six, eight) ANOVA revealed that capacity was greater for single color displays than for mixed color displays, $F(1, 11) = 26.28$, $MSE = .022$, $p < .001$, $\eta_p^2 = .705$, and increased as the number of new items increased, $F(2, 22) = 83.62$, $MSE = .335$, $p < .001$, $\eta_p^2 = .884$, reaching approximately six items at a new set size of eight. The capacity increase with new set size is not surprising and most likely reflects that at the smaller set sizes performance was at ceiling. However, the capacity increase with set size was smaller for the mixed color condition than for the single color condition, $F(2, 22) = 5.69$, $MSE = .024$, $p = .01$, $\eta_p^2 = .341$, indicating an influence of stimulus heterogeneity on capacity.

Accuracy as a Function of Response Number

It is possible that participants become worse at selecting the new items over time, or as a function of how many items they have already selected. This might occur if, for example, the ability to separate the old from the new deteriorates with time, perhaps because of weakening suppression of the old. To examine this possibility we calculated the mean accuracy in selecting a new item as a function of the number of previous responses (e.g., the accuracy of the first response made, the accuracy of the second response made etc.). The resulting overall mean proportion accuracy scores are shown in Figure 2D. The figure plots the mean accuracy for eight responses, however, because one participant never clicked on more than seven items, our statistical analysis was restricted to the first seven selections. A 2 (color) × 7 (response number) ANOVA showed that accuracy was overall greater with the single color displays than with the mixed color displays, $F(1, 11) = 11.20$, $MSE = .001$, $p < .01$, $\eta_p^2 = .504$. The probability of correctly clicking on a new item also decreased as a function of the number of previously clicked items, $F(6, 66) = 32.37$, $MSE = .003$, $p < .001$, $\eta_p^2 = 0.746$, but this decrease did not differ between the mixed and single color displays, $F < 1$. Thus, the ability to accurately select the new stimuli appeared to decrease over time suggesting a weakening representation.

Response Times

We calculated mean response times for each item selected as the time taken to respond from the onset of the display. This allowed us to determine the overall rate at which observers selected each item. In order to avoid including means based on small sample sizes (because there were few trials on which participants made more responses than the number of new items), the analysis was restricted to the first n responses, where n was equal to the new set size. Overall, mean RTs as a function of color (mixed or single), new set size, and response number are shown in Figure 2E. The response functions were highly linear suggesting that the rate of responding remained relatively constant throughout the trial.

It is interesting that although the rate of selection of each new item within a trial remained constant, the particular rate of search varied between trials, depending on set size. A 2 (Color: single/mixed) × 3 (New set size, four, six, eight) ANOVA showed there to be an increase in response rate for displays with a higher set size (553.3, to 530.4, to 507.6 ms/item for new set sizes of four, six, and eight, respectively), $F(2, 22) = 11.28$, $MSE = 1109.46$, $p < .001$, $\eta_p^2 = .506$. However, neither the main effect of display color, $F < 1$, nor the Color × Set size interaction, $F(2, 22) = 1.30$, $MSE = 823.10$, $p = .292$, $\eta_p^2 = .106$, reliably influenced the response rate.

Discussion

Experiment 1 examined the capacity limits in time-based selection with up to eight old and eight new items of a single or mixed color. Our analyses revealed that the overall capacity for selecting new stimuli was approximately six items and increased as a function of the new item set size (most likely reflecting ceiling effects at the smaller set sizes). Of note, this capacity is larger than four, the capacity limit of new items observed using eye movement measures (Emrich et al., 2008), but is less than 15 new (Theeuwes et al., 1998) and 30 old (Jiang et al., 2002a) items, suggested on the basis of previous studies in which only a single target had to be selected and responded to.

The results also found that display heterogeneity influenced the capacity, where capacity was larger for displays containing items of the same color, as compared with mixed color displays. This might reflect the ease of grouping and rejecting homogeneous items (Duncan & Humphreys, 1989, 1992). If the old stimuli could be grouped and rejected more efficiently when they were homogenous, this might lead to more new items being detected because the difference between groups would be more pronounced. This has implications for the different accounts of preview search. From an inhibitory marking perspective (Watson & Humphreys, 1997) being able to group the old distractors more efficiently might aid the ability to generate (or maintain over time) an inhibitory template toward them. From an onset account (Donk & Verburg, 2004) additional grouping within the new elements might also help to separate them from the old. Predictions from the temporal asynchrony account (Jiang et al., 2002b) are less straightforward. On the one hand increasing homogeneity within the old and new items might allow the old items to be grouped more easily and distinguished from the new. However, it could also predict that when all items were identical the new items would also group more easily with the old, based for example, on their common color.

There were also a number of secondary findings. In terms of the number of items selected, participants tended to underestimate the number of new items that appeared on the screen, making proportionally fewer selections at the larger set sizes than at the smaller set sizes. The reduction in the number of responses as set size increased was also greater for the mixed color displays than for the single color displays. This suggests that the efficiency of mechanisms prioritizing the new items decreases over time/number of responses and is worse with mixed color displays. This finding also suggests that maintaining a suppressive mechanism against a group of mixed color items is more difficult than against single color group.

We also examined response times as each new item was selected. Within a trial the response rate was constant; however, the overall rate of responding increased as a function of set size. One possible reason for this is that participants modulated their rate of responding depending on the perceived difficulty of the selection task. When participants were confronted with a large number of new items they might have perceived the need to select them more quickly than when fewer new items appeared. If so, this suggests that observers were aware that their ability to select new stimuli declined over time and thus modulated their response rate accordingly to compensate.

Experiment 2: Large Set Sizes and Color Heterogeneity

Experiment 1 suggested that depending on color heterogeneity, the capacity of time-based selection was approximately six items. However, the range of set sizes used did not allow us to find an upper bound, because the capacity estimate had not reached asymptote by the largest set size. Accordingly, in Experiment 2 we increased the maximum set size up to 32 items (16 old and 16 new). We also changed the stimuli to filled circles. This change was made for two main reasons. First, we wanted to generalize the findings to a different stimulus type/shape. Second, in Experiment 1, the displays consisted of randomly rotated letter *L*s and so all the displays were heterogeneous in terms of stimulus shape/orientation (regardless of color heterogeneity). This shape heterogeneity might have reduced the observed capacity of time-based selection in a similar way to that of color heterogeneity. Using filled circles ensured that all the displays were homogeneous with respect to stimulus shape.

Method

Participants

Twelve participants (six male), ages 19 to 22 years ($M = 19.9$ years) took part for payment or course credits. All were undergraduate students at the University of Warwick and reported normal or corrected-to-normal visual acuity.

Stimuli and Apparatus

The stimuli consisted of filled circles (12 mm in diameter) and the total display sizes were 8, 16, 24, and 32 items. Thus the preview display consisted of 4, 8, 12, or 16 items, and after 1000 ms, an additional 4, 8, 12, or 16 items were added, respectively.

Design and Procedure

A 2 (color: mixed or single) × 4 (set size: 8, 16, 24 or 32 items) within-participants design was used. Participants completed 320 experimental trials with an equal number of trials for each Set size × Display color combination (40 replications per cell). The remaining aspects were the same as Experiment 1.

Results

The results were analyzed in the same way as in Experiment 1. Fourteen trials (<0.4%) were discarded as outliers because the first selected item had an RT of less than 100 ms.

Number of Items Selected

The mean total number of items selected as a function of new item set size (4, 8, 12, 16) and color (single/mixed) are shown in Figure 3A. As in Experiment 1, participants selected fewer items in the mixed color ($M = 7.89$) condition than in the single color condition ($M = 8.16$), $F(1, 11) = 15.39$, $MSE = .113$, $p < .005$, $\eta_p^2 = .583$, selected more items as the number of new items increased, $F(3, 33) = 142.25$, $MSE = 1.76$, $p < .001$, $\eta_p^2 = .928$, and this increase with new item set size was greater for the single

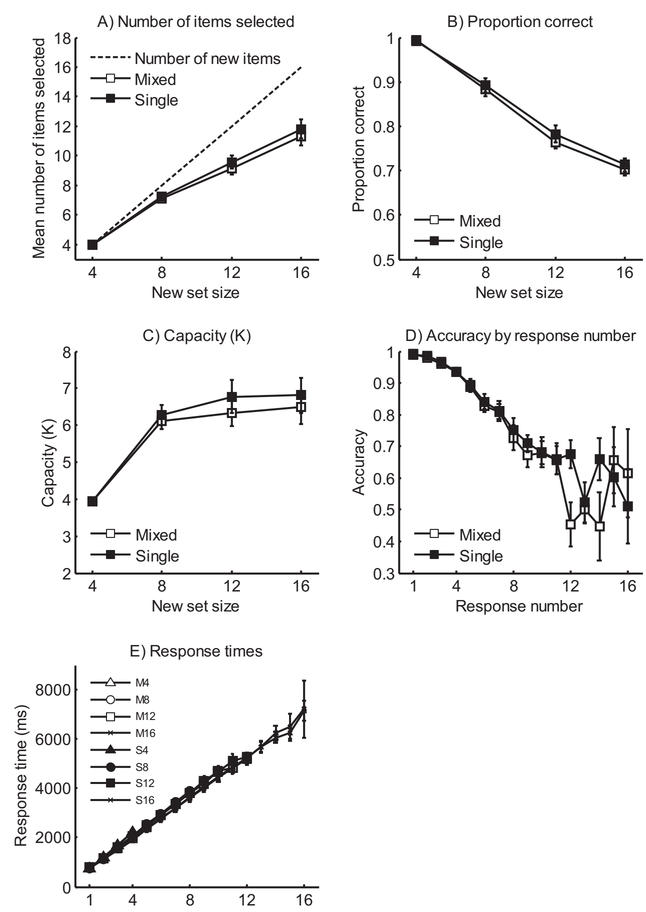

Figure 3. Results for Experiment 2. Panels A, B, and C show the total number of items selected, the proportion correct, and the estimate of new item capacity as a function of stimulus color (mixed or single) and new set size. Panels D and E show accuracy and mean RTs as a function of stimulus color and the number of previous responses within the trial. Error bars indicate ±1 *SE*.

color displays than for the mixed color displays, $F(3, 33) = 5.30$, $MSE = .070$, $p < .005$, $\eta_p^2 = .325$. The slope of the number of items selected as a function of new set size was 0.64 (single color) and 0.60 (mixed color) and was significantly less than one in both cases, $t(11) = 6.59$, $p < .001$, and $t(11) = 8.21$, $p < .001$, respectively. As in Experiment 1, participants selected fewer items than the number of new items that had appeared in the display.

Proportion Correct

The proportion correct responses showed a similar pattern to those of Experiment 1 and are shown in Figure 3B. Fewer correct responses were made in the mixed color condition than in the single color condition, $F(1, 11) = 7.33$, $MSE = .000$, $p < .05$, $\eta_p^2 = .400$, accuracy decreased with set size, $F(3, 33) = 277.09$, $MSE = .001$, $p < .001$, $\eta_p^2 = .962$, and this decrease was marginally greater for the mixed color displays than for the single color displays, $F(3, 33) = 2.47$, $MSE = .000$, $p = .079$, $\eta_p^2 = .184$.

Estimate of Capacity (K)

Using the same equation as in Experiment 1, mean estimates of capacity leveled off at approximately six to seven items (see Figure 3C). As in Experiment 1, the capacity was greater for the single color displays than for the mixed color displays, $F(1, 11) = 8.10$, $MSE = .151$, $p < .05$, $\eta_p^2 = .424$. The capacity also increased with set size, $F(3, 33) = 37.02$, $MSE = 1.06$, $p < .001$, $\eta_p^2 = .771$, and this increase was marginally greater with the single color displays than with the mixed color displays, $F(3, 33) = 2.59$, $MSE = .092$, $p = .07$, $\eta_p^2 = .190$.

Accuracy as a Function of Response Number

The mean accuracy for selecting new items as a function of the number of previously selected items is shown in Figure 3D. Because some participants never clicked on more than 12 items we restricted our statistical analysis to the first 12 responses. A 2 (color: single/mixed) × 12 (number of items already selected) ANOVA revealed a significant main effect of color, $F(1, 11) = 12.68$, $MSE = .004$, $p < .005$, $\eta_p^2 = .535$, response number, $F(11, 121) = 48.82$, $MSE = .010$, $p < .001$, $\eta_p^2 = .816$, and a significant Color × Response number interaction, $F(11, 121) = 4.28$, $MSE = .006$, $p < .001$, $\eta_p^2 = .280$. Responses were overall more accurate for the single color condition than the mixed color condition, decreased with number of items selected, and this decrease was larger for the mixed color displays. As in Experiment 1, this suggests that the mechanisms driving new item prioritization decreased in efficiency over time.

Response Times

Mean RTs as a function of color (mixed or single) and set size are shown in Figure 3E. Mean RT × Set size slopes were calculated individually for each participant and were analyzed with a 2 (color: single/mixed) × 4 (set size) ANOVA. As in Experiment 1, the rate of responding increased as the number of items in the display increased, $F(3, 33) = 8.49$, $MSE = 2435.60$, $p = .001$, $\eta_p^2 = .436$, with response rates of 484.8, 447.4, 426.2, 419.6 ms/item with 4, 8, 12, and 16 new items respectively. However, neither the main effect of display color, nor the Color × Set size interaction approached significance, both Fs < 1.

Discussion

Experiment 2 replicated Experiment 1, however, here we increased the maximum total set size to 32 items (16 old and 16 new) and used homogeneous shape stimuli (filled circles). As shown, even with these changes the capacity limits leveled off at approximately 6.4 and 6.8 items for the mixed and single color displays, respectively. The use of homogeneously shaped stimuli (filled circles instead of randomly rotated letter *L*s) in Experiment 2 did not substantially increase the capacity estimate over that indicated by Experiment 1.

Of note, the capacity difference between mixed and single color displays was greater at the larger sets sizes than at the smaller set sizes. This might indicate the influence of differing combinations of processes underlying the prioritization of new objects. For example, it is known that when searching for a single target item, a small number of onsets (approximately four) can capture attention automatically (Yantis & Johnson, 1990; Yantis & Jones, 1991).[1] In the current displays, the selection of a small number of new items might be driven by a combination of automatic capture of attention by new onsets and suppression of the old items.

If automatic capture is insensitive to display heterogeneity then there would be little difference between the mixed and single color displays at the small set sizes. In contrast, at the larger set sizes, given the capped limit of onsets that can capture attention, the prioritization of the new items may rely more on old item suppression. As suggested earlier, suppressing an old set of items might be influenced by item heterogeneity because we would expect grouping the old items to be stronger in single color displays than in mixed color displays (Duncan & Humphreys, 1989, 1992; Humphreys & Müller, 1993). This may make it easier to reject the old items, and thus separate them from the new stimuli. In addition, a single color preview display might allow feature-based (color) inhibition to contribute to their suppression. According to this account, item heterogeneity should have a greater influence at the larger set sizes where inhibitory suppression of the old items contributes more to prioritizing the new stimuli.

The remaining findings were similar to and strengthen those of Experiment 1: i) more new items were missed as new set size increased, ii) response times increased linearly with each item selected, iii) response rates were greater with the larger set sizes. We consider these findings further in the General Discussion.

Experiment 3: Fixed Number of Selections and the Influence of Shape Heterogeneity

In Experiments 1 and 2, participants were free to select as many or as few items as they deemed appropriate before proceeding to the next trial. As the results showed, observers selected fewer items than had actually been presented especially at the larger set

[1] Note however, that this limit was determined based on search for and response to a single target item among onset elements. The limit might be considerably smaller if all the new items had to be localized or responded to.

sizes. Although this provides a legitimate measure of capacity limits, it is also possible that it might have underestimated the capacity of time-based selection. If participants adopted too conservative a criterion for ending the trial, their willingness to make additional selections might have declined over time. For example, if participants were wary of making mistakes then they might have selected only the new items that they were absolutely certain of, leading to a reduction in potential performance (see Chun & Wolfe, 1996 for a related discussion of how people decide to end target absent trials in visual search tasks). To address this issue, Experiment 3 required participants to make a number of responses equal to the number of new stimuli that had been presented (i.e., if there were eight new items then participants had to select eight items in total).

Experiments 1 and 2 also showed that color heterogeneity influences capacity, with capacity being higher in single compared with mixed color displays. Homogeneous shape displays (filled circles, Experiment 2) did not appear to provide any benefit compared with heterogeneous shape displays (randomly oriented letter *L*s, Experiment 1), however this was not formally tested. Accordingly the second aim of Experiment 3 was to examine the effect shape heterogeneity; all displays were now a single color but consisted of filled circles or randomly oriented letter *L*s. If shape heterogeneity acts in a similar way to color heterogeneity then we would expect capacity to be greater with filled circles than with randomly oriented letter *L*s.

Method

Participants

Thirteen participants (three male), ages 18 to 20 years (M = 18.9 years) took part for payment or course credits. All were undergraduate students at the University of Warwick and reported normal or corrected-to-normal visual acuity.

Stimuli and Apparatus

The stimuli and apparatus were the same as in Experiment 2 except that all the displays were a single color (they were all green) and the stimuli within a display were either all *L*s or all filled circles (as in Experiment 1 and 2, respectively).

Design and Procedure

A 2 (shape: circle or L) × 4 (display size: 8, 16, 24, or 32 items) within-participants design was used. The procedure was basically the same as in Experiments 1 and 2 except that on each trial participants had to select a total number of items equal to the number of new stimuli that had been added. When the final item had been selected the screen went blank to signal the end of the trial and, as in the previous experiments, participants moved the mouse cursor to the right side of the screen to initiate the next trial. The total number of trials in a single block was 320 (40 replications per cell). The remaining aspects were the same as Experiment 2.

Results

Five trials (<0.2%) were discarded as outliers because the first selected item had an RT of less than 100 ms. One participant was removed from analysis due to initial RTs of 0 ms on 29% of their trials. The data were analyzed in the same way as Experiment 2 except that we did not consider the total number of items selected because participants were required to make a fixed number of selections equal to the number of new items that had been presented.

Proportion Correct

Proportion correct responses were calculated as a function of the number of new items presented and overall means are shown in Figure 4A. Participants showed a greater proportion of correct responses in the circle-stimulus (homogeneous) condition than in the *L*- stimulus shape (heterogeneous) condition, $F(1, 11) = 34.49$, $MSE = .001$, $p < .001$, $\eta_p^2 = .758$, and the proportion correct decreased as the number of new items increased, $F(3, 33) = 295.10$, $MSE = .001$, $p < .001$, $\eta_p^2 = .964$. The decrease in accuracy was greater with the *L* stimuli than for the *C* stimuli as set size increased, $F(3, 33) = 6.20$, $MSE = .000$, $p < .005$, $\eta_p^2 = .360$. Thus shape heterogeneity had a similar effect, in terms of decreasing accuracy as color heterogeneity did.

Estimate of Capacity (K)

Mean estimates of capacity are shown in Figure 4B. Capacity was greater with the circle-stimulus displays than with the *L*-stimulus displays, $F(1, 11) = 30.64$, $MSE = .470$, $p < .001$, $\eta_p^2 = .736$. Capacity also increased with set size, $F(3, 33) = 35.18$, $MSE = 1.06$, $p < .001$, $\eta_p^2 = .765$, leveling off at approximately 5.9 items in the *L* displays and 7.1 items in the circle displays. The increase with set size was greater for the circle displays than the *L*

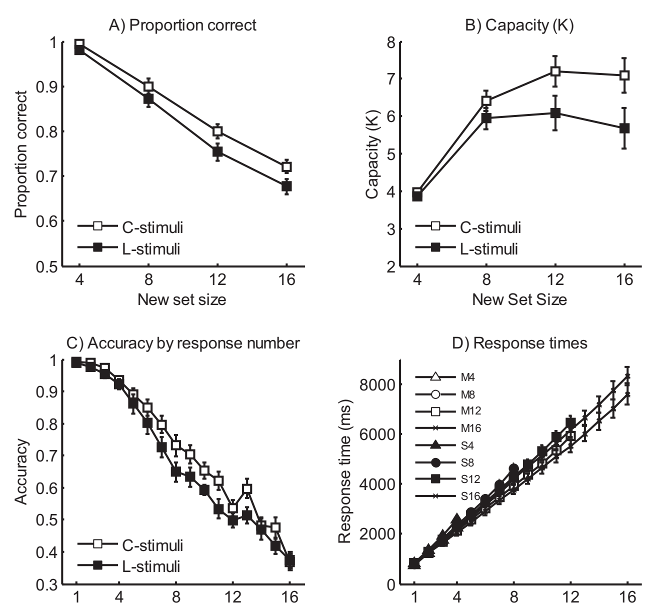

Figure 4. Results for Experiment 3. Panels A and B show the proportion correct, and the estimate of new item capacity as a function of stimulus shape and new set size. Panels C and D show accuracy and mean RTs as a function of stimulus shape and the number of previous responses within the trial. Error bars indicate ±1 *SE*.

displays, $F(3, 33) = 13.77$, $MSE = .153$, $p < .001$, $\eta_p^2 = .556$. An independent t test comparing capacity limits for Experiments 2 and 3 (collapsed across all conditions) revealed that there was no significant difference in capacity, $t(22) = .148$, $p = .884$.

Accuracy as a Function of Response Number

The mean accuracy for selecting new items as a function of the number of previously selected items is shown in Figure 4C. Accuracy was overall greater with the circle stimuli than with the *L* stimuli, $F(1, 11) = 23.00$, $MSE = .008$, $p = .001$, $\eta_p^2 = .676$, decreased as the number of items selected increased, $F(15, 165) = 152.79$, $MSE = .007$, $p < .001$, $\eta_p^2 = .933$, and this decrease was larger for *L*-stimulus displays than for circle-stimulus displays, $F(15, 165) = 1.85$, $MSE = .003$, $p < .05$, $\eta_p^2 = .144$. As in Experiments 1 and 2, accuracy decreased as new set size increased and was greater with homogeneous stimuli than with heterogeneous stimuli.

Response Times

Mean correct RTs as a function of stimulus shape and set size are shown in Figure 4D. Mean RT × Set size slopes were calculated individually for each participant and were analyzed with a 2 (stimulus shape) × 4 (set size) ANOVA. The rate of responding increased as the number of new items increased, $F(3, 33) = 20.50$, $MSE = 1281.19$, $p < .001$, $\eta_p^2 = .651$, with response rates of 545.7, 509.4, 483.5, and 470.8 ms/item for new set sizes of 4, 8, 12, and 16 new items respectively. Response rate was also faster overall for the circle stimuli (470.4 ms/click) than for the *L* stimuli (534.4 ms/click), $F(1, 11) = 32.15$, $MSE = 3056.85$, $p < .001$, $\eta_p^2 = .745$. The Stimulus type × Set size interaction was not significant, $F(3, 33) = 2.36$, $MSE = 659.29$, $p = .09$, $\eta_p^2 = .176$. As in Experiments 1 and 2 participants increased their rate of responding as the number of new items increased.

Discussion

One aim of Experiment 3 was to examine the capacity limit observed when participants were required to make as many selections as there were new items. In Experiments 1 and 2, participants could end the trial when they had selected as many items as they liked. This might have led to an underestimation of the true capacity if participants had set a relatively conservative criterion for selecting the new stimuli. This possibility was addressed in Experiment 3 by requiring the number of responses to be equal to the number of new items added. Despite this change, the estimated capacity was again six to seven items depending on stimulus type and did not differ from that obtained in Experiment 2, suggesting that participants had set a reasonable quitting threshold in the earlier experiments.

The second aim of Experiment 3 was to examine the influence of stimulus shape heterogeneity. Experiments 1 and 2 showed that color heterogeneity reduced the capacity of new item selection. In Experiment 3 we also found a clear effect of stimulus shape. The capacity for selecting new items was larger with homogeneous displays of filled circles than with heterogeneous displays of randomly oriented letter *L*s. The findings of secondary interest also confirmed the previous experiment's results: proportion correct decreased as a function of the number of previously selected items, and response rates were greater for larger new set sizes.

A common finding across Experiments 1 to 3 was that the probability of correctly selecting a new item decreased with the number of responses made. There could be two reasons for this. First, the decrease in accuracy could occur if participants selected the new items in order of how confident they were that they were new. In other words, they selected the ones that they were most certain of first. This would lead to an increase in errors as more items were selected. Alternatively, or in addition, the representations supporting the prioritization of the new elements might decay over time (and thus the ability to separate the old from the new) once the new items have been presented. Related to this possibility, Watson, Compton, and Bailey (in press) showed a decreasing preview benefit over a 3-s preview period when stimuli changed shape over time. In contrast, Braithwaite et al. (2006) showed that a robust preview benefit can be obtained with relatively long preview periods (e.g., 3 s, see also Blagrove & Watson, 2010). However, neither of these studies tested the effect of time on separating old from new after the new items had been presented. Accordingly, in Experiment 4, to determine whether time-based representations decay over time once the new items have appeared, we introduced a delay between the presentation of the new stimuli and when participants were allowed to start making their selections.

Experiment 4: Delayed Responding

In Experiment 4 we examined whether introducing a delay between the onset of the new stimuli and when participants could start responding would have an influence on the capacity of new item selection. There were two conditions. In the no delay (0 ms) condition, participants could start to respond as soon as the new items had been presented (as in the previous experiments). In the delay condition, participants were not allowed to start responding until 3200 ms had elapsed. We chose 3200 ms as the delay time because this was the approximate average time that it took participants to select eight of the new items in the previous experiments and by which time a substantial reduction in the selection accuracy of new items had occurred.

Participants

Twelve participants (five male), ages 19 to 33 years ($M = 22.8$ years) took part for payment (£5). All participants were recruited from the University of Warwick's participant pool and reported normal or corrected-to-normal visual acuity.

Stimuli and Apparatus

The stimuli consisted of filled green circles as used in Experiment 3. The apparatus was the same as used in Experiments 1 to 3.

Design and Procedure

The procedure was similar to Experiment 3, except where stated otherwise. A 2 (delay: 0 ms or 3200 ms) × 3 (display size: 8, 16, or 32 items) within-participants design was used. In the 0-ms delay condition, following a fixation dot, a set of 4, 8, or 16 filled green circles appeared. After 1000 ms, 4, 8 or 16 new circles were displayed along with a 50-ms 1000Hz tone which indicated that participants could start to click on the new items. The 3200-ms delay condition was similar except that the tone indicating that participants could start to select the new items was presented 3200-ms after the onset of the new items. Participants were instructed that they could not select any items until the tone had been presented. Each participant completed one block of 120 trials (20 replications per cell) for the 0-ms delay and 3200-ms delay conditions. Participants were given a practice session before each block.

Results

The data were analyzed in the same way as Experiment 3. Two trials ($<$0.1%) were discarded as outliers because the first selected item had an RT of less than 100 ms.

Proportion Correct

Proportion correct responses were calculated as a function of the number of new items presented and overall means are shown in Figure 5A. Participants showed a greater proportion of correct responses in the 0ms delay condition than in the 3200 ms delay condition, $F(1, 11) = 17.22$, $MSE = .001$, $p < .005$, $\eta_p^2 = .610$, and the proportion correct decreased as the number of new items increased, $F(2, 22) = 403.89$, $MSE = .001$, $p < .001$, $\eta_p^2 = .973$. The decrease in accuracy with set size was greater in the delay condition than in the no delay condition, $F(2, 22) = 4.03$, $MSE = .000$, $p < .05$, $\eta_p^2 = .268$.

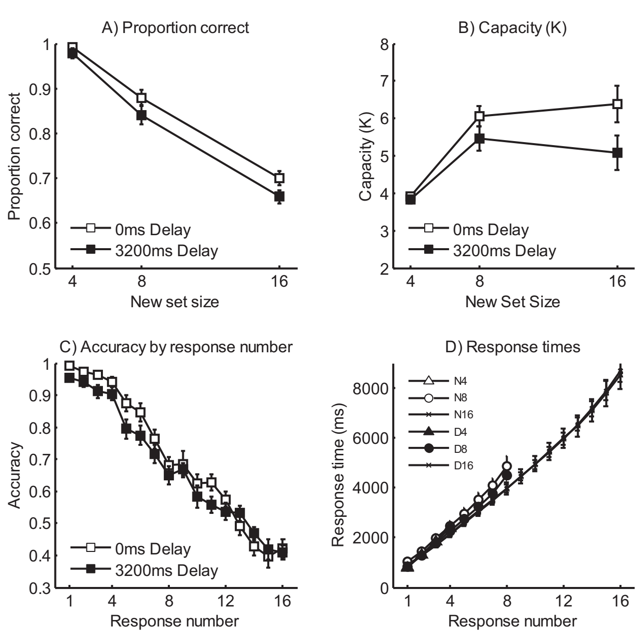

Figure 5. Results for Experiment 4. Panels A and B show the proportion correct, and the estimate of new item capacity as a function delay (0 or 3200 ms) and new set size. Panels C and D show accuracy and mean RTs as a function of delay and the number of previous responses within the trial. Error bars indicate ± 1 *SE*.

Estimate of Capacity (K)

Mean estimates of capacity are shown in Figure 5B. Capacity was greater overall in the no delay condition than when a delay was present, $F(1, 11) = 20.19$, $MSE = .404$, $p = .001$, $\eta_p^2 = .647$. Capacity also increased with set size, $F(2, 22) = 26.22$, $MSE = 1.07$, $p < .001$, $\eta_p^2 = .704$, leveling off at approximately 6.2 items in the no delay and 5.3 items in the delay condition. The increase in capacity with set size was greater for the no delay condition than in the delay condition, $F(2, 22) = 13.44$, $MSE = .164$, $p < .001$, $\eta_p^2 = .550$.

Accuracy as a Function of Response Number

The mean accuracy for selecting new items as a function of the number of previously selected items is shown in Figure 5C. Accuracy was overall greater in the no delay condition than in the delay condition, $F(1, 11) = 14.07$, $MSE = .006$, $p < .005$, $\eta_p^2 = .561$, and decreased as the number of items selected increased, $F(15, 165) = 151.34$, $MSE = .006$, $p < .001$, $\eta_p^2 = .932$. The Delay $\times$ Response number interaction was not significant, $F(15, 165) = 1.57$, $MSE = .005$, $p = .088$, $\eta_p^2 = .125$. Of note, accuracy was poorer in the delay condition from the first response made, $t(11) = 4.01$, $p = .002$.

Response Times

Mean correct RTs as a function of delay condition and set size are shown in Figure 5D. Mean RT $\times$ Set size slopes were calculated individually for each participant and were analyzed with a 2 (delay condition) $\times$ 3 (set size) ANOVA. The overall average rate of responding was 518.0 ms/item and there was a trend for response rates to be greater for the larger, new set-size displays. However, neither main effect of delay condition, set size, or their interaction approached significance, $F < 1$, $F(2, 22) = 1.61$, $MSE = 2977.34$, $p = .223$, $\eta_p^2 = .128$, $F(2, 22) = 2.15$, $MSE = 1372.22$, $p = .140$, $\eta_p^2 = .164$, respectively.

Discussion

Experiment 4 examined why accuracy in the previous experiments decreased with the number of responses made. One possible reason was that participants simply selected the items that they were most sure of first and that each subsequent response led to a reduction in maintaining the old/new distinction. Alternatively, it could be that the representations underlying the prioritization of new elements simply decayed over time, leading to the less accurate rejection of old items and the less accurate selection of new elements. This was tested in Experiment 4 by introducing a delay between the onset of the new items and the time when participants were allowed to start responding. The results suggest that introducing a delay before participants could start to respond reduced new item capacity but only by approximately one item. This shows that the representations allowing the old and new to be separated decays, slightly, over time after the new items have appeared.

Of note, accuracy was reduced even for the very first response selection made showing that the decay is not contingent on any responses having being made. This finding contrasts with those of previous work showing that a robust preview benefit can be obtained even with relatively long preview durations (e.g., Braithwaite et al., 2006). The difference in results may be because in this experiment we examined the influence of time after the new items had been presented (which was not the case in previous work). It would seem that maintaining the old from the new becomes more difficult with time after the new items have appeared.

However, note that although we found a robust time-based decay, decay alone is unlikely to fully account for the reduced accuracy with each selection seen in Experiments 1–3. In previous experiments, after approximately 3200 ms (approximately seven or eight selections) accuracy had reduced to approximately 60–70%. In contrast, in Experiment 4, after 3200 ms (i.e., response No. 1) accuracy was reduced but remained relatively high (approximately 95%). It was again only after seven or eight responses that in this experiment, accuracy was reduced to 60–70%. This suggests that both time-based decay and the actual responses made interfered with accuracy in time-based selection.

In the experiments so far, participants have selected the new stimuli using mouse clicks. In Experiment 5, we examine the influence of an alternative, more direct response method on new item capacity.

Experiment 5: Mouse Versus Touch Screen Response

The findings so far have indicated a capacity limit for intentionally prioritizing and responding to all new stimuli of approximately six to seven items. In the experiments participants indicated which were the new items by an indirect method of making a mouse click on them. It is possible that different response methods could lead to different capacity limits. Experiment 4 demonstrated that delaying the onset of new item selection caused a reduction in new item capacity. Thus more efficient selection methods might increase the capacity for selecting new items if it takes less time to select each stimulus. Also the additional complexity of mapping mouse to screen movements might impact on the maintenance of time-based representations, as compared with more direct selection methods. In Experiment 5, in order to test the stability of the observed six-to-seven–item limit, we compared responding via mouse clicks with responding via a direct touch-screen response. Examining the influence of such direct versus indirect responding is also important given the increasing trend in technology to use touch screen devices.

Method

Participants

Twelve participants (eight male), ages 19 to 32 years ($M = 23.5$ years) took part for payment (£5). All participants were recruited from the University of Warwick's participant pool and reported normal or corrected-to-normal visual acuity.

Stimuli and Apparatus

The stimuli were the same as those presented in the filled circle condition of Experiment 3 with the following changes. Each display consisted of 4, 8, or 16 green-filled circles (10 mm diameter) for 1000 ms followed by an additional 4, 8, or 16 items. In all conditions, the stimuli were presented on a Viewsonic VX2258wm 22" LCD touch screen monitor with a resolution of 1920 × 1080 pixels (the native resolution of the display panel). Stimulus spacing was 120 pixels (30 mm) and stimuli were jittered by ± 12 pixels (3 mm) in the vertical and horizontal directions.

Design and Procedure

The procedure was similar to Experiment 3, except where stated otherwise. A 2 (response: touch screen or mouse) × 3 (display size: 8, 16, or 32 items) within-participants design was used. The mouse response condition was essentially the same as in Experiment 3. When half the display items had been selected the stimuli were removed and participants initiated the next trial by moving the mouse of the screen to the right of the display. In the touch-screen response condition participants selected the new stimuli directly by touching them with a finger from their right hand. The next trial was initiated by touching the far right side of the screen within a 120-pixel band (6% of the display width) of the display edge. Each participant completed one block of 120 trials (20 replications per cell) for each of the response conditions. Participants were given a practice session before each block.

Results

The data were analyzed in the same way as Experiment 3. One trial (<0.04%) was discarded as an outlier because the first selected item had an RT of less than 100 ms.

Proportion Correct

Proportion correct responses were calculated as a function of response type and the number of new items presented and overall means are shown in Figure 6A. Proportion correct decreased as the number of new items increased, $F(2, 22) = 198.84$, $MSE = .002$, $p < .001$, $\eta_p^2 = .948$. However, neither the main effect of response type, $F(1, 11) = 1.39$, $MSE = .000$, $p = .264$, $\eta_p^2 = .112$, nor the Response type × Set size interaction, $F < 1$, were significant.

Estimate of Capacity (K)

Mean estimates of capacity are shown in Figure 6B. Capacity increased with set size, $F(2, 22) = 26.06$, $MSE = 1.86$, $p < .001$, $\eta_p^2 = .703$, leveling off at approximately 6.4 items. Neither the main effect of response type, nor the Response type × Set size interaction approached significance, both Fs < 1.

Accuracy as a Function of Response Number

The mean accuracy for selecting new items as a function of the number of previously selected items is shown in Figure 6C. Accuracy decreased as the number of items selected increased, $F(15, 165) = 101.86$, $MSE = .009$, $p < .001$, $\eta_p^2 = .903$. However, neither the main effect of response number, $F < 1$, nor the Response number × Response type interaction were significant, $F(15, 165) = 1.32$, $MSE = .003$, $p = .198$, $\eta_p^2 = .107$.

Response Times

Mean correct RTs as a function of stimulus shape and set size are shown in Figure 6D. Mean RT × Set size slopes were calculated individually for each participant and were analyzed with a 2

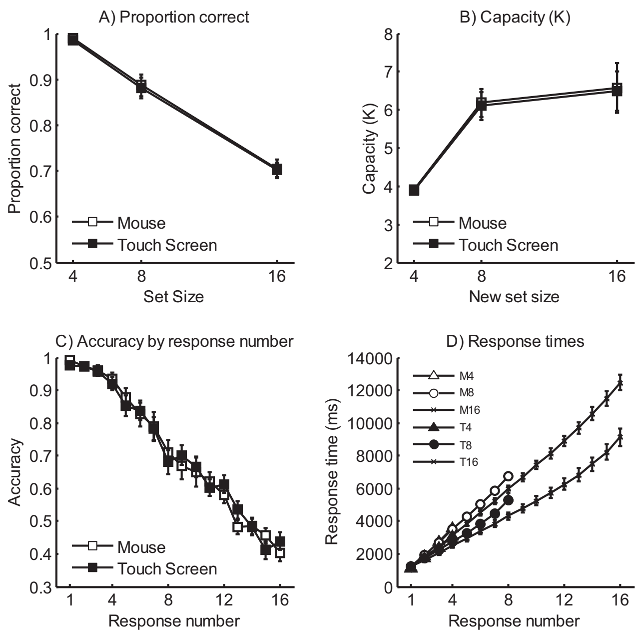

Figure 6. Results for Experiment 5. Panels A and B show the proportion correct, and the estimate of new item capacity as a function of response method (touch screen or mouse) and new set size. Panels C and D show accuracy and mean RTs as a function of response method and the number of previous responses within the trial. Error bars indicate ±1 *SE*.

(response type) × 3 (set size) ANOVA. This revealed that the rate of responding was faster in the touch screen condition (570.8 ms/selection) than in the mouse response condition (785.6 ms/selection), $F(1, 11) = 38.06$, $MSE = 21816.88$, $p < .001$, $\eta_p^2 = .776$, and increased as set size increased, $F(2, 22) = 31.49$, $MSE = 2403.87$, $p < .001$, $\eta_p^2 = .741$. The Response type × Set size interaction was not significant, $F(2, 22) = 2.46$, $MSE = 1970.54$, $p = .108$, $\eta_p^2 = .183$.

Discussion

The main aim of Experiment 5 was to examine the stability of the new item capacity limit across different response methods. In one condition participants responded via indirect mouse responses and in the other they selected the new stimuli directly via a touch screen. As might be expected, participants responded at a higher rate (approximately 200 ms per item faster) when using the less complex and more direct touch screen way of responding. However, response method had no other statistically reliable influence. Of importance the capacity limit of six to seven items remained relatively stable across the different response modes.

The fact that capacity remained unchanged is interesting considering that the two different response methods produced different rates of responding. From the results of Experiment 4 we might have expected a reduced capacity limit for the slower responses because that would add additional time to each selection. However, this did not occur suggesting that relatively modest differences (here approximately 200 ms/selection) have little influence. Consistent with this, even a 3200-ms delay imposed before responses could be made only resulted in a relatively modest (one item) drop in capacity. We consider this further in the General Discussion. All remaining findings relating to proportion correct, accuracy by response number and rate of responding as a function of new item set size mirrored those of the earlier experiments.

General Discussion

Previous work has examined the ability of people to ignore old information and prioritize the selection of new stimuli, using the preview paradigm. In a typical preview condition a set of distractor stimuli are displayed (previewed) for some period of time, and then a second set of stimuli are added. Although there are some variations, participants usually have to indicate whether or not a prespecified target is present or absent from the newly presented stimuli. This preview condition is compared with baselines in which all the items are presented simultaneously (the FEB) or only the new items from the preview condition are presented (HEB). Such studies have shown that search in the preview condition is more efficient than in the FEB and matches that of the HEB, consistent with observers being able to ignore “old” stimuli currently in the field of view and prioritize “new” stimuli when they appear.

Based on this method, previous work has suggested that up to 30 old items can be suppressed and up to 15 new items can be prioritized with no upper bound yet established (Jiang, Chun, & Marks, 2002a; Theeuwes et al., 1998). This implies that the capacity of time-based selection is relatively large. However, in contrast to this approach, Emrich et al., (2008) have shown that with displays containing up to 10 old and 8 new items, eye movements appear to be biased to approximately four or fewer new objects. Beyond four fixations, eye movements were equally likely to land on an old or a new item. This suggests that there might be some upper limit to how many new items can be selected with priority. Clearly there appears to be a discrepancy between the results obtained with behavioral manual RT measures and overt eye movement behavior in terms of new item selection.

The work in the present study addressed this incongruity. Here, participants were asked to make responses to all the new stimuli. Under these conditions a consistent capacity limit for selecting new items of approximately six to seven items emerged. This is considerably less than the minimum capacity limit of 30 old and 15 new items obtained when responding to a single target (Jiang et al., 2002a; Theeuwes et al., 1998). One explanation for this difference might lie in the different task demands here, as compared with previous manual RT studies. In the current study, participants had to localize and overtly respond to all the new items. That is, every new item was a target.

In contrast, in previous work, although participants had to prioritize multiple new items, they only had to select and make an overt response to a single target. It is possible that the process of localizing and responding to multiple targets interferes with the maintenance of the representations used to suppress the old elements. This seems plausible because localizing a target would require the involvement of spatial processing mechanisms which might also be needed in maintaining a spatial representation of the old items. Consistent with this idea of shared resources, Humphreys, Watson, and Jolicoeur, (2002) found that a visual modality load task presented halfway through the preview period disrupted the preview benefit. However, an auditory load task did not (a similar argument can be constructed in terms of prioritizing the onsets associated with new elements). In contrast to responding to all the new items, in previous work, only a single target had to be selected and responded to. Hence any interference following the initial target selection was not measured on the ability to select additional new items for response (see also Watson & Humphreys, 1999).

The observed capacity in these experiments was also greater than the capacity limit suggested from the eye movement study of Emrich et al. (2008). Based on the number of fixations to each new item Emrich et al. (2008) proposed that the capacity limit in preview search was approximately four or less. There are several potential explanations for this difference. First, there might be differences in the effects of eye movements as compared with manual responses in terms of disruption to the preview benefit. For example, generating saccades might interfere more with the visual representations used to prioritize the new items than manual responses because of the close links between eye movement control and visual attention (Shepherd, Findlay, & Hockey, 1986; Deubel & Schneider, 1996). Against this possibility, Watson and Inglis (2007) found little effect of eye movements made during the preview period. However, it remains possible that eye movements made after the new items are presented and which are involved in visual search through the items have a more disruptive effect.

A second explanation for this difference could relate to the type of information that is available for deliberate compared with incidental eye movements. In Emrich et al., the eye movements were an integral part of the attentional search process or a byproduct of it. In contrast, here the manual selection movements were deliberately made to each item in turn, and were perhaps guided by representations which were not available to incidental eye movements.[2] Finally, it is possible that in Emrich et al.'s study participants could have localized and processed more than one item within a single fixation. If so, the capacity limit for new items from their study would have been underestimated. Consistent with this, in visual search tasks, the number of fixations is typically much smaller than the number of search items (Zelinsky & Sheinberg, 1997) and search can proceed even when no eye movements are permitted (Klein & Farrell, 1989). Thus although the number of fixations might be limited to four or fewer new elements, the number of new items able to be responded to or searched through might be higher, consistent with the work presented here.

Decaying Representations and Visual Short Term Memory

One reason why time-based selection might be limited in this study to six to seven items is that the suppression of the old items could not be maintained for relatively long periods of time. This might seem unlikely given that previous research has obtained robust preview benefits with extended durations of up to 3 s (Braithwaite et al., 2006; see also Blagrove & Watson, 2010). However, these studies only examined the effect of the preview duration before the new items were added, and not after (but see also Watson & Inglis, 2007). In addition, as compared with RTs in previous work, in which a single target had to be responded to, the response times in the present study were much longer. For example, in Watson and Humphreys' (1997) work, average RTs were typically less than 1.5 s whereas here, at the largest set sizes, the average RTs ranged up to approximately 8 s. It is possible that a relatively long delay after the new items have been presented results in a decay of the signals associated with the new items and/or a reduction in old item inhibitory processing.

Addressing this possibility, Experiment 4 showed that delaying the response by approximately 3 s after the new items had been added led to a robust, although modest, reduction in new item capacity of approximately one item. This reduction in accuracy was detectable even for the first response made suggesting that time alone can result in a small decay of the necessary representations. This could be due to either a difficulty in maintaining an inhibitory template of the old items, a decay of onset signals associated with the new items (Jiang & Wang, 2004) or a combination of both.

Consistent with this idea, Jiang and Wang (2004) suggested that new items are associated with an asynchronous signal, depicting their unique onset from other items in the field. This asynchrony signal enables the new items to selectively enter the limited capacity visual short-term memory (VSTM Phillips, 1974; Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001) in order to be processed as a target or not.[3] If the target was not present then the remaining new items entered VSTM for examination provided that their asynchrony signals had not yet decayed to baseline. Such a rapid decaying signal might be equivalent to the high-capacity but fragile VSTM component suggested by Sligte, Scholte, and Lamme (2008). It follows that when all items need to be selected/responded to there might not be sufficient time to process them before such signals have decayed. This would then lead to a capacity limit similar to that normally associated with VSTM. In terms of the present work, our estimate of new item capacity of six to seven items is larger than the four-item VSTM object limit proposed by Luck and Vogel (1997), but similar to the six locations VSTM limit suggested by Jiang, Olson, and Chun (2000).

Linked to the decaying signals idea, we found that participants increased their rate of responding as the new set size became larger. One reason for this could be that observers were aware that their ability to select all the new items decreased over time. To compensate for this they might then have adjusted their rate of responding with the belief that this would allow them to select more items at the larger set sizes. Note however, that as described earlier, although a robust time-based decay was found, the reduction due to time alone was relatively modest. Thus time alone is unlikely to account for the full drop in performance as each item was selected. Instead, it seems that localizing and responding to new items also caused a reduction in maintaining the old from the new. This might occur if localizing a response to each item interferes with an inhibitory representation set up and maintained against the old items (Humphreys et al., 2002).

Our results also mesh with other work on VSTM. In particular, VSTM appears to store information at the *object level*. That is, even though the capacity is limited to four objects, it is relatively unlimited by the complexity of those items. For example, 16 features which are *chunked* into four objects (four features per

[2] We thank Naz Al-Aidroos for this suggestion.

[3] Note that the relative strength of such signals could be further increased by inhibition of the old items.

object) can be stored as accurately as just four features distributed across four different objects (one feature per object, Luck & Vogel, 1997). With this chunking ability in mind, Kumar and Jiang (2005) asked whether time of appearance could also be used to chunk information and hence effectively increase the capacity of VSTM. They presented participants with two brief (27 ms) arrays of five dots separated by stimulus onset asynchronies (SOAs) of 27 ms to 1227 ms. Participants were then given a two-alternative forced choice task to indicate which of two probe dots fell at the location of a previous item presented in one of the two arrays. If time of appearance could be used to chunk the stimuli in each array then performance should improve as the time between the two arrays increased. However, unlike other grouping cues, separating each array in time did not increase the number of locations that could be stored in VSTM, as compared with when both arrays were (in effect) presented simultaneously (a similar result was found for color memory). Kumar and Jiang (2005) also found that performance for items from the second array was better than for items from the first array. This might be because items from the second array acted as a mask (at the short SOAs) or because items in the second array competed with and displaced items from the first array (at the longer SOAs, Kumar & Jiang, 2005).[4]

There are two notable differences between the present method and those of Kumar and Jiang (2005). First, in Kumar and Jiang (2005) the stimuli from the two arrays were presented separately (i.e., stimuli from the first array never appeared with stimuli from the second array). Second, participants had to process and memorize both arrays. In contrast, in our work, i) the first set of stimuli were presented alone and the second set were added to the first set, and ii) participants had to try to ignore the first set and selectively process the second set of new items (rather than integrate both sets). Jiang et al., found that overall VSTM capacity was shared between both sets of items. In contrast, in terms of VSTM, our work shows that the old items could be effectively ignored and the full capacity of VSTM (approx six to seven locations) could be dedicated to storing just the new items.[5]

The Influence of Stimulus Heterogeneity and Stability Across Different Response Types

A consistent finding across the first three experiments was that stimulus heterogeneity, either color (Experiment 1 and 2) or shape (Experiment 3), decreased the capacity of time-based selection. One explanation for this is that stimuli of the same color or shape could be suppressed/rejected more efficiently because they could be more easily processed as a single group (Duncan & Humphreys, 1989, 1992). This has implications for understanding the possible mechanisms responsible for the preview effect. From an inhibitory visual marking perspective (Watson & Humphreys, 1997), increased grouping between the preview items might allow for the more efficient development and a stronger/more robust inhibitory template to be maintained against the old items. Similarly, from an abrupt onset account (Donk & Verburg, 2004) common features could allow the new items to be grouped more efficiently. However, from a temporal asynchrony view (Jiang et al., 2002ab; Jiang & Wang, 2004), on the one hand increased grouping of the old items would predict better performance because the “old group” could be established more efficiently. On the other hand, in the single color condition the old and new groups would also be of the same color, leading to stronger grouping of *all* items (both old and new) and a reduction in preview efficiency.

In terms of VSTM, Kumar and Jiang (2005; Experiment 4) compared memory for conjoined versus individual features across two array presentations. In their three-feature condition, each array consisted of three colors or three orientations. In their six-feature conjoined condition, each array consisted of three items each consisting of two features (color and orientation) to give a total of six features per array. Overall performance between the two conditions did not differ (and both were better than a six-feature baseline condition), consistent with object-based chunking of visual information. Of most interest, the influence of array separation was also the same for the two conditions. Thus stimulus heterogeneity had little influence when items from both arrays had to be memorized (see also Awh, Barton, & Vogel, 2007). In contrast, when trying to ignore one set of stimuli, we found that heterogeneous stimuli led to a reduced new item capacity. This suggests that feature-based influences play a larger role when trying to separate stimuli in time than when trying to integrate them.

It is possible that even our heterogeneous stimuli still contained some small grouping cues. If this is the case then our new item capacity limit of six to seven items might be reduced if these remaining grouping cues were removed (in other words our capacity estimate may not be a lower bound). However, we think that any change in new-item capacity based on a possible further reduction of grouping cues would be modest at best. This is particularly relevant when we consider the heterogeneous, rotated *L* displays. In visual search these stimuli are known to produce only inefficient search as they show minimal evidence of interitem grouping (e.g., Duncan & Humphreys, 1989). Knowing this, we believe that any grouping in these displays is already negligible, at best.

Interestingly, the influence of stimulus heterogeneity was most pronounced at the largest set sizes. This could point to the fact that a number of processes contribute to driving the preview benefit. For example, abrupt luminance onset signals may well play a part in the selection for new elements even when the old items are also being intentionally suppressed (Atchley et al., 2003; Donk & Theeuwes, 2001; Herrero, Crawley, van Leeuwen, & Raffone, 2007; Watson & Humphreys, 2002). The lack of an effect of stimulus heterogeneity at the smallest set sizes might reflect the contribution of noninhibitory mechanisms (such as abrupt onset capture for a small number of items) that might not be susceptible to the ease with which old items could be suppressed.

[4] When display durations of 350 ms were used, Ihssen, Linden, and Shapiro (2010) found a modest increase in VSTM capacity of approximately 0.5–0.75 items for detecting an object feature change in sequentially presented arrays. This increase may have been the result of a performance boost in detecting a change in the second display.

[5] An alternative possibility is that participants actively tried to remember a combination of old and new locations in VSTM in order to perform the task. However, we think that this is unlikely because prior work in which participants actively process (search) through a previewed set of stimuli abolishes the preview benefit (Olivers et al., 2002).

Influence of Response Type

Experiment 5 found that regardless of a change in response type, the capacity estimates for the selection of new items were indistinguishable. Of course it remains possible that other, more different types of response methods could produce different capacity estimates. However, our work suggests that the difference in response type might have to be quite extreme. For example, in terms of visual processing time scales a difference of 200 ms/item between response methods (found between our mouse and touch screen manipulation) is substantial and yet the capacity estimates remained stable. Determining the influence of more extreme response methods could therefore be a useful goal for future research.

The work here suggests that a number of factors, related to both stimulus properties and the response task are likely to affect the capacity of time-based selection. In the present work we adopted a new preview methodology, requiring a response to each and every new item. Given the multiple responses, elongated response times, and the fact that old and new items were only physically distinguishable by their temporal onset, one could argue that the search task here was more difficult than previous preview search tasks. If so, the limit of six to seven items most likely represents what is probably a conservative estimate. For example, localizing and selecting the position of each item new might be especially difficult if prioritization processes also involve location-based representations (Watson & Humphreys, 1997). It follows that if the task did not require a location based response (e.g., naming the colors of the new items) we might obtain a higher capacity limit.

Conclusion

The present findings show that when people have to select and respond to all new stimuli then humans' ability to process such new information is limited to a relatively stable capacity of approximatley six to seven items. Furthermore, the data suggest that the similarity between display items can have an impact on this limit. Perhaps surprisingly, having display items in different colors appeared to hinder the ability to detect new stimuli. Thus although separating items based on their color *can* be used to help guide our search (Wolfe, Cave, & Franzel, 1989; Wolfe, 1994), it appears that at least under some conditions single color displays might be better.

References

Al-Aidroos, N., Emrich, S. M., Ferber, S., & Pratt, J. (in press). Visual working memory supports the inhibition of previously processed information: Evidence from preview search. *Journal of Experimental Psychology: Human Perception and Performance.*

Atchley, P., Jones, S. E., & Hoffman, L. (2003). Visual marking: A convergence of goal- and stimulus-driven processes during visual search. *Perception & Psychophysics, 65,* 667–677. doi:10.3758/BF03194805

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18,* 622–628. doi:10.1111/j.1467-9280.2007.01949.x

Blagrove, E., & Watson, D. G. (2010). Visual marking and facial affect: Can an emotional face be ignored? *Emotion, 10,* 147–168. doi:10.1037/a0017743

Braithwaite, J. J., Hulleman, J., Watson, D. G., & Humphreys, G. W. (2006). Is it impossible to inhibit isoluminant items or does it simply take longer? Evidence from preview search. *Perception & Psychophysics, 68,* 290–300. doi:10.3758/BF03193676

Braithwaite, J. J., Humphreys, G. W., & Hodsoll, J. (2003). Color grouping in space and time: Evidence from negative color-based carryover effects in preview search. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 758–778. doi:10.1037/0096-1523.29.4.758

Braithwaite, J. J., Humphreys, G. W., & Hodsoll, J. (2004). Effects of colour on preview search: Anticipatory and inhibitory biases for colour. *Spatial Vision, 17,* 389–415. doi:10.1163/1568568041920096

Braithwaite, J. J., Humphreys, G. W., Watson, D. G., & Hulleman, J. (2005). Revisiting preview search benefits at isoluminance: New onsets are not necessary for the preview advantage. *Perception & Psychophysics, 67,* 1214–1228. doi:10.3758/BF03193554

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30,* 39–78. doi:10.1006/cogp.1996.0002

Davoli, C. C., Suszko, J. W., & Abrams, R. A. (2007). New objects can capture attention without a unique luminance transient. *Psychonomic Bulletin & Review, 14,* 338–343. doi:10.3758/BF03194074

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827–1837. doi:10.1016/0042-6989(95)00294-4

Donk, M. (2005). Prioritizing selection of new elements: On the time course of the preview effect. *Visual Cognition, 12,* 1373–1385. doi:10.1080/13506280444000742

Donk, M. (2006). The preview benefit: Visual marking, feature-based inhibition, temporal segregation, or onset capture? *Visual Cognition, 14,* 736–748. doi:10.1080/13506280500193230

Donk, M., & Theeuwes, J. (2001). Visual marking beside the mark: Prioritizing selection by abrupt onsets. *Perception & Psychophysics, 63,* 891–900. doi:10.3758/BF03194445

Donk, M., & Theeuwes, J. (2003). Prioritizing selection of new elements: Bottom–up versus top–down control. *Perception & Psychophysics, 65,* 1231–1242. doi:10.3758/BF03194848

Donk, M., & Verburg, R. C. (2004). Prioritizing new elements with a brief preview period: Evidence against visual marking. *Psychonomic Bulletin & Review, 11,* 282–288. doi:10.3758/BF03196571

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. doi:10.1037/0033-295X.96.3.433

Duncan, J., & Humphreys, G. W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 578–588. doi:10.1037/0096-1523.18.2.578

Emrich, S. M., Ruppel, J. D. N., Al-Aidroos, N., Pratt, J., & Ferber, S. (2008). Out with the old: Inhibition of old items in a preview search is limited. *Perception & Psychophysics, 70,* 1552–1557. doi:10.3758/PP.70.8.1552

Folk, C. L., & Anderson, B. A. (2010). Target-uncertainty effects in attentional capture: Color-singleton set or multiple attentional control setting? *Psychonomic Bulletin & Review, 17,* 421–426. doi:10.3758/PBR.17.3.421

Folk, C. L., Remington, R., & Johnson, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044. doi:10.1037/0096-1523.18.4.1030

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 317–329. doi:10.1037/0096-1523.20.2.317

Gibson, B. S., & Jiang, Y. H. (2001). Visual marking and the perception of salience in visual search. *Perception & Psychophysics, 63,* 59–73. doi:10.3758/BF03200503

Herrero, J. L., Crawley, R., van Leeuwen, C., & Raffone, A. (2007). Visual marking and change detection, *Cognitive Processing, 8,* 233–244. doi:10.1007/s10339-007-0180-0

Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception & Psychophysics, 69,* 172–184. doi:10.3758/BF03193740

Hulleman, J. (2005). The mathematics of multiple object tracking: From proportions correct to number of objects tracked. *Vision Research, 45,* 2298–2309. doi:10.1016/j.visres.2005.02.016

Humphreys, G. W., & Müller, H. J. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25,* 43–110.

Humphreys, G. W., Watson, D. G., & Jolicoeur, P. (2002). Fractionating the preview benefit: Dual-task decomposition by timing and modality. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 640–660. doi:10.1037/0096-1523.28.3.640

Ihssen, N., Linden, D. E. J., & Shapiro, K. L. (2010). Improving visual short-term memory by sequencing the stimulus array. *Psychonomic Bulletin & Review, 17,* 680–686. doi:10.3758/PBR.17.5.680

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. doi:10.1037/0278-7393.28.2.293

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. doi:10.1037/0096-1523.28.3.717

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.

Jiang, Y., & Wang, S. W. (2004). What kind of memory supports visual marking? *Journal of Experimental Psychology: Human Perception and Performance, 30,* 79–91. doi:10.1037/0096-1523.30.1.79

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye movement. In J. B. Long & A. D. Baddeley (Eds), *Attention and performance* (9th ed., pp. 187–203). Hillsdale, NJ: Erlbaum.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *The American Journal of Psychology, 62,* 498–525. doi:10.2307/1418556

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334,* 430–431. doi:10.1038/334430a0

Klein, R., & Farrell, M. (1989). Search performance without eye movements. *Perception & Psychophysics, 46,* 476–482. doi:10.3758/BF03210863

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4,* 138–147. doi:10.1016/S1364-6613(00)01452-2

Kumar, A., & Jiang, Y. (2005). Visual short-term memory for sequential arrays. *Memory & Cognition, 33,* 488–498. doi:10.3758/BF03193066

Kunar, M. A., Humphreys, G. W., & Smith, K. J. (2003). History matters: The preview benefit in search is not onset capture. *Psychological Science, 14,* 181–185. doi:10.1111/1467-9280.t01-1-01439

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions, *Nature, 390,* 279–281. doi:10.1038/36846

Olivers, C. N. L., & Humphreys, G. W. (2003). Visual marking inhibits singleton capture. *Cognitive Psychology, 47,* 1–42. doi:10.1016/S0010-0285(03)00003-3

Olivers, C. N. L., Humphreys, G. W., Heinke, D., & Cooper, A. C. G. (2002). Prioritization in visual search: Visual marking is not dependent on a mnemonic search. *Perception & Psychophysics, 64,* 540–560. doi:10.3758/BF03194725

Osugi, T., Kumada, T., & Kawahara, J. (2009). The spatial distribution of inhibition in preview search. *Vision Research, 49,* 851–861. doi:10.1016/j.visres.2009.03.002

Osugi, T., Kumada, T., & Kawahara, J. (2010). Visual marking survives graphical change if meaning is retained. *Attention, Perception & Psychophysics, 72,* 2144–2156.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16,* 283–290. doi:10.3758/BF03203943

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32,* 3–25. doi:10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160–174. doi:10.1037/0096-3445.109.2.160

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition, 11,* 801–822. doi:10.1080/13506280344000518

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197. doi:10.1163/156856888X00122

Rauschenberger, R. (2003). When something old becomes something new: Spatiotemporal object continuity and attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 600–615. doi:10.1037/0096-1523.29.3.600

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 38,* 475–491.

Sligte, I. G., Schole, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLoS ONE, 3,* e1699. doi:10.1371/journal.pone.0001699

Theeuwes, J., Kramer, A. F., & Atchley, P. (1998). Visual marking of old objects. *Psychonomic Bulletin & Review, 5,* 130–134. doi:10.3758/BF03209468

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. doi:10.1016/0010-0285(80)90005-5

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 331–351. doi:10.1037/0096-1523.19.2.331

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102. doi:10.1037/0033-295X.101.1.80

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92–114. doi:10.1037/0096-1523.27.1.92

Watson, D. G., Braithwaite, J. J., & Humphreys, G. W. (2008). Resisting change: The influence of luminance changes on visual marking and the preview benefit. *Perception & Psychophysics, 70,* 1526–1539. doi:10.3758/PP.70.8.1526

Watson, D. G., Compton, S., & Bailey, H. (in press). Visual Marking: The influence of temporary changes on time-based visual selection. *Journal of Experimental Psychology: Human Perception and Performance.*

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. doi:10.1037/0033-295X.104.1.90

Watson, D. G., & Humphreys, G. W. (1998). Visual marking of moving objects: A role for top-down feature based inhibition in selection. *Jour-*

nal of Experimental Psychology: Human Perception and Performance, 24, 946–962. doi:10.1037/0096-1523.24.3.946

Watson, D. G., & Humphreys, G. W. (1999). The magic number four and temporo-parietal damage: Neurological impairments in counting targets amongst distractors. *Cognitive Neuropsychology, 16,* 609–629.

Watson, D. G., & Humphreys, G. W. (2000). Visual marking: Evidence for inhibition using a probe-dot paradigm. *Perception & Psychophysics, 62,* 471–481. doi:10.3758/BF03212099

Watson, D. G., & Humphreys, G. W. (2002). Visual marking and visual change. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 379–395. doi:10.1037/0096-1523.28.2.379

Watson, D. G., Humphreys, G. W., & Olivers, C. N. L. (2003). Visual marking: Using time in visual selection. *Trends in Cognitive Sciences, 7,* 180–186. doi:10.1016/S1364-6613(03)00033-0

Watson, D. G., & Inglis, M. (2007). Eye movements and time-based selection: Where do the eyes go in preview search? *Psychonomic Bulletin & Review, 14,* 852–857. doi:10.3758/BF03194111

Watson, D. G., Maylor, E. A., & Bruce, L. A. M. (2005). Search, enumeration and aging: Eye movement requirements cause age-equivalent performance in enumeration but not in search tasks. *Psychology and Aging, 20,* 226–240. doi:10.1037/0882-7974.20.2.226

Watson, D. G., Maylor, E. A., & Bruce, L. A. M. (2007). The role of eye movements in subitizing and counting. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1389–1399. doi:10.1037/0096-1523.33.6.1389

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238. doi:10.3758/BF03200774

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–74). Hove, East Sussex, England, U. K.: Psychology Press.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433. doi:10.1037/0096-1523.15.3.419

Wolfe, J. M., Place, S. S., & Horowitz, T. S. (2007). Multiple object juggling: Changing what is tracked during extended multiple object tracking. *Psychonomic Bulletin & Review, 14,* 344–349. doi:10.3758/BF03194075

Yantis, S., & Johnson, D. N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 812–825. doi:10.1037/0096-1523.16.4.812

Yantis, S., & Jones, E. (1991). Mechanisms of attentional selection: Temporally modulated priority tags. *Perception & Psychophysics, 50,* 166–178. doi:10.3758/BF03212217

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. doi:10.1037/0096-1523.10.5.601

Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 244–262. doi:10.1037/0096-1523.23.1.244

Received January 14, 2011
Revision received July 27, 2011
Accepted September 2, 2011 ■