Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 44 (2004) 1411–1426
Vision Research
www.elsevier.com/locate/visres

How fast can you change your mind? The speed of top-down guidance in visual search

Jeremy M. Wolfe [a,b,*], Todd S. Horowitz [a,b], Naomi Kenner [a], Megan Hyle [a], Nina Vasan [b]

[a] Brigham and Women's Hospital, USA
[b] Visual Attention Lab, Harvard Medical School, 64 Sidney St., Suite 170, Cambridge, MA 02139, USA

Received 24 June 2003; received in revised form 18 November 2003

Abstract

Most laboratory visual search tasks involve many searches for the same target, while in the real world we typically change our target with each search (e.g. find the coffee cup, then the sugar). How quickly can the visual system be reconfigured to search for a new target? Here observers searched for targets specified by cues presented at different SOAs relative to the search stimulus. Search for different targets on each trial was compared to search for the same target over a block of trials. Experiments 1 and 2 showed that an exact picture cue acts within 200 ms to make varied target conjunction search as fast and efficient as blocked conjunction search. Word cues were slower and never as effective. Experiment 3 replicated this result with a task that required top-down information about target identity. Experiment 4 showed that the effects of an exact picture cue were not mandatory. Experiments 5 and 6 used pictures of real objects to cue targets by category level.

1. Introduction

Sometimes we search the visual world for any item of interest. More commonly, however, we are looking for something specific: The coffee mug, the stamps, the stapler, etc. All of those objects might be in the same scene. The difference between search for one and search for another, therefore, is not governed by changes in the stimulus; it is governed by changes in the observer. When you are looking for the coffee mug, you somehow configure your visual system for coffee mugs (or, perhaps, for your specific mug). If you then need to search for the stapler, you change your mind in a manner that allows you to search the same visual input for a different target item. The purpose of this paper is to examine the time-course of that change of the mind.

This common sense notion of "top-down" control of search has been studied in the visual search literature. For example, Egeth, Virzi, and Garbart (1984) found that search for a red letter was twice as efficient if only half the letters were red. They concluded that search could be limited to the set of red letters, presumably by top-down control of the search process. The notion that top-down commands can guide the deployment of visual attention is central to search models like Guided Search (Wolfe, 1994, 2001; Wolfe, Cave, & Franzel, 1989). Much of the work in models of this sort has been devoted to determining the nature of these top-down commands. As discussed elsewhere, we believe that the commands are limited to some 12–18 types of feature (color, size, motion, depth, etc. Wolfe, 1998) and that the representation of those features that can be used for top-down guidance is coarse (e.g. red among green, not 650 nm red among 600 nm reddish, Nagy & Sanchez, 1990) and categorical (e.g. "steep" among "shallow", not angular difference between targets and distractors, Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). When we talk about the ability to "configure" the visual system to search for one target or another, we are talking about an ability to adjust the strength of internal connections in order to give more weight to a specific dimension (e.g. color) or a feature within a dimension

* Corresponding author. Address: Visual Attention Lab, Harvard Medical School, 64 Sidney St., Suite 170, Cambridge, MA 02139, USA. Tel.: +1-617-768-8818; fax: +1-617-768-8816.
E-mail address: wolfe@search.bwh.harvard.edu (J.M. Wolfe).
URL: http://www.search.bwh.harvard.edu.

0042-6989/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2003.11.024

(e.g. red). "Guidance" in Guided Search is presumed to be this sort of weighting operation (Wolfe, 1994). This is similar to the dimension weighting ideas of Mueller and colleagues (Found & Muller, 1996; Weidner, Pollmann, Muller, & von Cramon, 2002).

A parallel line of research, growing out of the signal detection approach to search, has measured the benefits of reducing uncertainty by providing top-down information about the location of potential targets (e.g. Davis, Kramer, & Graham, 1983; Lu, Lesmes, & Dosher, 2002; Palmer, Ames, & Lindsey, 1993; Solomon, Lavie, & Morgan, 1997) or about attributes of the stimulus (e.g. Eckstein, Whiting, & Thomas, 1996; Hubner, 1996; Monnier & Nagy, 2001).

While a mass of data indicates the existence of top-down guidance of search, less work has been devoted to the dynamics of how that guidance is implemented. In most search experiments, the issue is moot. An observer looks for a designated type of target for hundreds of trials. The establishment of the top-down "set" for that target occurs during a set of practice trials and presumably remains relatively stable over the course of a block of trials. Some information about the effects of changing the top-down set come from experiments where the target changes from trial to trial. A number of studies have shown that search—even search for salient feature singletons—is faster when the target identity remains the same from trial to trial than when it changes. This is true whether the change is within a feature dimension (trial K: red among green, $K+1$: green among red, Maljkovic & Nakayama, 1994) or across dimensions (trial K: red among green, $K+1$: Big among small, see for example Egeth, 1977; Treisman, 1988). The prior literature on this topic is reviewed much more extensively in Wolfe, Butcher, Lee, and Hyle (2003).

These priming experiments and the uncertainty experiments show that there is a top-down contribution to even the simplest of feature searches. However, they merely reveal the effects of different amounts of uncertainty in the top-down set. Observers are faster if they know that the target is red than if they know it is either red or green, and so forth. In the present work, we seek to uncover the time course of the transition from one top-down set (or a neutral set) to another.

The experiments reported here use cueing paradigms in which the target is specified just prior to the appearance of the search stimulus on each trial. There is an extensive literature on spatial cueing (e.g. Chastain, Cheal, & Lyon, 1996; Cheal, Lyon, & Gottlob, 1993; Posner, 1980; Posner, Nissen, & Ogden, 1978) (usefully reviewed in Chapter 4 of Pashler, 1997). In these studies, a cue indicates *where* the target will be, while the studies in this paper specify *what* the target will be. Luck and Vecera (2002) succinctly summarize the work on the time course of location cues. Exogenous cues (cues at the spatial location of the eventual target) are effective within about 100 ms. Endogenous cues are slower, taking on the order of 250–300 ms to reach full strength. If a salient exogenous cue is uninformative, it will still attract attention. However, the location will be inhibited shortly thereafter with the inhibitory effect reaching its maximum 300–400 ms after the onset of the cue.

Surprisingly, there do not seem to be comparable studies of the time course of cues to target identity in visual search. Blough (1989) showed that pigeons benefit from advance warning of the target identity but did not do a systematic study of time course. Hubner (1996) varied the cue-to-stimulus SOA (stimulus onset asynchrony) from 200 to 1000 but found little effect of this variation. Clearly, there must be a time before the cue is effective but Hubner did not find it in his work. In this paper, we report on a series of six experiments in which observers were given a cue to the identity of the target just before the onset of the search stimulus. The SOA between cue and search stimulus was varied. In most studies, results are compared to a "blocked" baseline in which target identity remained constant for a block of trials and/or an "uninformed" baseline in which target identity changed from trial to trial but the observer was not informed about target identity (Obviously, this only works for conditions where the target identity can be inferred from the stimulus and a general instruction to look for something like an odd-man-out).

Experiments 1 and 2 use standard conjunction search tasks. Target identity can be cued with either a picture or words. Results show that picture cues are fast. They are largely effective within 50 ms and fully effective (as good as the "blocked" conditions) within 200 ms. Unsurprisingly, word cues are slower. More surprisingly, even for well learned stimulus properties (e.g. "red", "vertical"), word cues are never as effective as picture cues, even at long SOAs. Experiment 3 shows that these results can be obtained using a search task where all of the items are different. In this case, use of the cue is mandatory. Given the special status of picture cues identical to the target, it is possible that they would capture attention in a mandatory manner (Folk, Remington, & Wright, 1994; Theeuwes, 1994). In Experiment 4, we show that this is not the case. A valid word cue can be used without interference from an invalid picture cue.

Experiments 5 and 6 use photographs of real objects as stimuli. The basic pattern of results is the same. Pictorial identity cues are superior to word cues. These studies allow us to differentiate between identity cues (this rabbit cues this same rabbit), type cues (this rabbit indicates that the target will be a rabbit, but possibly a different rabbit) and category cues (this rabbit indicates that the target will be an animal). As uncertainty increases the effectiveness of the cue decreases.

2. Experiment 1: Cueing in uncertain conjunction search

The basic design of these experiments is quite simple. Observers are searching for a target among a variable number of distractors. Target and distractors can change from trial to trial. In the uninformative control condition, observers are left to determine the identity of the target on their own. In the cued conditions, the target is specified just prior to the appearance of the search display. The SOA between the onset of the cue and the onset of the search array is varied in an effort to measure the time course of cue effectiveness. The basic sequence of events in a trial is shown in Fig. 1.

2.1. Methods

In the first experiment, each item in a search display could be big (3.3 deg by 0.9 deg at the 57 cm viewing distance) or small (2.2 deg by 0.3 deg), red (CIE: X. 48, Y. 34, 4.0 cd-m-sq) or green (CIE: X. 29, Y. 53, 14.2 cd-m-sq), and vertical or horizontal. This generated eight possible stimuli. On any given trial, observers searched for a conjunction of two of the features. For example, if the target was BIG and RED, two distractor types would be present: small RED and BIG green (note CAPS indicates a target feature). In this example, orientation would be irrelevant and would have been the same for all items. In a *mixed* set of trials, the next target item might be SMALL VERTICAL among big VERTICAL and SMALL horizontal with items being either all red or all green. In a *blocked* condition, the target would remain constant for the entire block of trials.

There were two types of control condition.

(1) In an *uninformative* mixed condition, observers searched for a unique item without knowing its identity. The target might be red vertical on one trial and big green on the next. This condition provided a ceiling, defining how bad performance could be. Note that the task can be done by identifying the unique item in the display. There were two blocks. In one, an uninformative picture cue (a white square) was presented before each trial. In the other, an uninformative word (“ready”) was presented.
(2) In the *blocked* conditions, the target was fixed. Each observer had a randomly assigned target type in this condition. This condition provided a floor, defining the best possible performance on the task. Two blocks were run with the same uninformative picture and word cues as the uninformative mixed conditions.

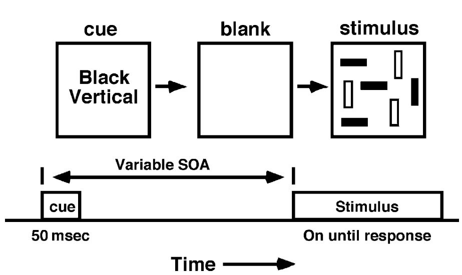

Fig. 1. The sequence of events in the experiments reported in this paper. A cue appears before each search stimulus. The cue could be a picture of the stimulus, words describing the stimulus, or an uninformative signal. The cue-stimulus SOA was varied. The designated target could change from trial to trial or it could remain constant for a block.

There were two types of experimental conditions.

(1) In the *picture cue* condition, observers saw an exact copy of the target as the cue prior to the appearance of the search stimulus. Observers were tested with four cue-to-stimulus SOAs: 50, 200, 400, and 800 ms. SOAs were blocked so that observers were tested for 300 trials with each SOA.
(2) The *word cue* conditions are identical to the picture cue conditions except that the cue was a pair of words defining the target rather than the image of the target (e.g. “BIG RED”).

Set sizes of 6, 12, and 18 items were displayed in a 19 by 19 deg region of the computer screen. Items were placed on a jittered 5 by 5 grid of locations. Twelve observers were tested for 3900 total trials. In the uninformative conditions, there were 150 trials in each of the two sessions. There were 300 trials for each of four SOAs in the informative mixed conditions. Finally, there were 600 trials in each of the two blocked condition sessions. Targets were present on 50% of trials. Set size was randomized across trials.

All observers had vision of at least 20/25 with appropriate correction. All passed the Ishihara color vision screen. Observers gave informed consent and were paid for their time. Studies were conducted on Macintosh computers running Matlab with the PsychToolbox (Brainard, 1997).

2.2. Results

RTs of less than 200 and greater than 4000 ms were excluded from the analysis (<1% of data). Mean RT as a function of SOA is shown for word and picture cues in Fig. 2.

The central question that motivated these studies was the speed of the implementation of top-down guidance. Fig. 2 makes it clear that the guidance is established very rapidly. A cue-stimulus SOA of just 50 ms provides the bulk of the guidance for a picture cue. For the picture cue, target-present RTs are significantly slower than the blocked RTs at SOA 50 ms ($t(11) = 4.6$, $p = 0.0007$). The difference is insignificant at 200 and 400 ms ($t(11) < 1.2, p > 0.25$). Curiously, it becomes marginally significant again at 800 ms ($t(11) = 2.6$, $p = 0.026$, uncorrected for multiple comparison). This pattern of

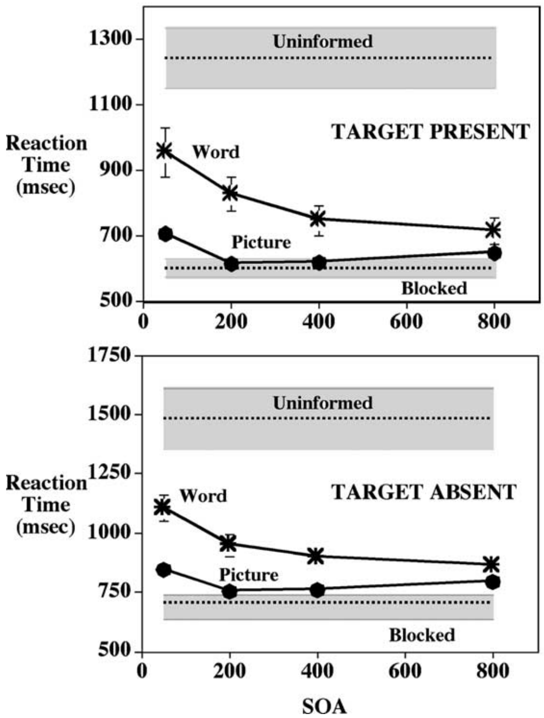

Fig. 2. Reaction Time as a function of SOA for picture and word cues in Experiment 1. Data are averaged across set size. Error bars and shaded areas show s.e.m. for the average RTs of the 12 observers.

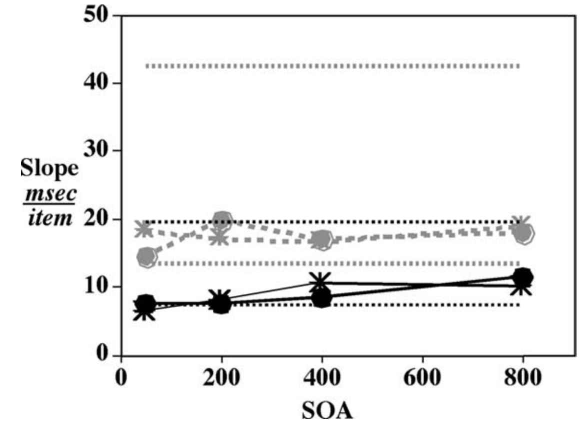

Fig. 3. Target present (black) and target absent (gray) slopes as a function of SOA. Solid circles show picture cue results. Stars show word cue results. Straight, horizontal, dashed lines show the results for the uninformed (upper line) and blocked (lower line) control conditions.

picture cues becoming less effective after SOAs of about 200 ms will recur in later experiments.

Unsurprisingly, the word cue takes longer to have its full effect. More interestingly, the word cue is never as effective as the picture cue ($p < 0.005$ for SOAs 50–400, $p < 0.05$ at 800 ms) and never reaches the level of the blocked condition ($p < 0.01$ at all SOAs).

A similar pattern is seen for target-absent trials. The only substantive difference is that blocked target-absent RTs are always somewhat faster than cued SOAs ($p < 0.05$ for picture cues and $p < 0.01$ for word cues). Error rates average 6% and did not show significant effects of cue type or SOA.

Fig. 3 shows the slope of the RT × set size function as a function of SOA.

The slope results show no significant change as a function of SOA. Unlike the RT data, the slope data show no advantage for picture over word cues. Slopes like these are typical of "guided" conjunction search (Wolfe et al., 1989). In some later experiments in this paper, we will see that this guidance can take time to develop. Short SOAs produce steep slopes. As information about the target becomes available, it can be used to guide attention toward some items and away from others. This produces an increase in efficiency. In this experiment, all SOAs show evidence of guidance. This suggests that observers waited for the information to become available and then performed a guided search. This shows up in the longer mean RTs at the short SOAs rather than in the slopes.

2.3. Priming effects

On some trials, the cue is the same as it was on the previous trial. In other search experiments with variable target identity, the repetition of a cue speeds RTs (Maljkovic & Nakayama, 1994; Olivers & Humphreys, 2003; Wolfe et al., 2003). In the present experiment, it is possible to distinguish four types of pairs of trials having the same target cue. On trial K, the target can be present (Y) or absent (N) and, on trial $K+1$, the target can be present or absent. These pairs of trials can be denoted as YY, YN, NY, and NN conditions. Fig. 4 shows average RT data for the cases where the target is present on trial $K+1$.

Comparing YY and NY conditions to the case where there is no cue repetition, it is clear that the priming

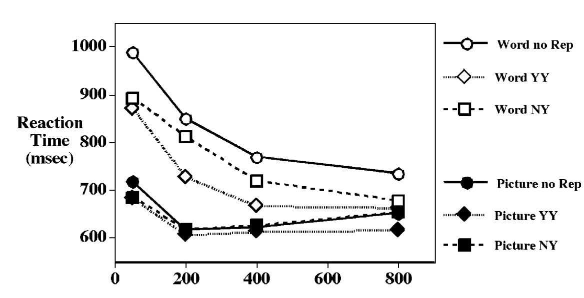

Fig. 4. Effects of cue repetition for correct, target present responses. "No Rep" conditions are those where the cue on trial K is different from that on $K+1$. The other conditions are those where the cue repeats. YY indicates target present on trial K and $K+1$. NY indicates target absent on trial K, present on $K+1$.

effect for picture cues is very weak (ANOVA: $F(2,11) = 3.3$, $p = 0.056$). Post-hoc tests reveal that the YY trials are somewhat faster than the no repetition baseline ($p = 0.04$). However, there is a much more substantial priming effect for word cues (ANOVA: $F(2,11) = 36.2$, $p < 0.001$). The effect does not interact with SOA ($F = 1$). Post-hoc tests reveal that NY trials are significantly faster than no repetition trials ($p < 0.001$) and YY trials are significantly faster than NY ($p < 0.002$).

For target absent trials, the priming effect is significant for both picture and word cues (ANOVA: $F(2,11) > 15, p < 0.001$ in both cases). Priming is about twice as great for the repeated word cue (102 vs 49 ms), a statistically reliable difference (ANOVA: $F(1,11) = 10.5$, $p = 0.0079$).

2.4. Discussion

Two salient facts emerge from this experiment. First, top-down guidance can be established very quickly. Indeed, the speed blurs the distinction between top-down and bottom-up. Picture cues have a very substantial effect within 50 ms of cue onset. This seems very fast for any "re-entrant" process (Di Lollo, Enns, & Rensink, 2000). More plausibly, observers are "set" to use the cue (Bacon & Egeth, 1994; Theeuwes, 1994) and this allows a picture cue to have its effect in a feed-forward manner—somehow setting weights for the appropriate features on its first pass through the system (Found & Muller, 1996; Weidner et al., 2002).

Secondly, it is interesting that word cues are not as effective as pictures. Note that observers are running hundreds of trials and, consequently, are very familiar with the appearance of these particular "big red" or "green vertical" stimuli. The priming data suggest that fully effective cueing requires that the observer must see the stimulus. A repeated word cue is more effective, most likely because the observer gets to see the specific target for trial $K + 1$ on trial K. Priming is greater when trial K is a target present trial perhaps because seeing the actual target is more effective that merely seeing the distractors. Cue repetition does not have an effect on target present trials when picture cues are used. The picture cue is its own prime and further priming is ineffective.

3. Experiment 2: Shorter SOAs

Experiment 2 is a variant of Experiment 1 using a finer grain of SOAs in order to explore the speed of the cueing effect in more detail.

3.1. Methods

There were two important differences between Experiments 1 and 2. First, in order to examine the rapid development of top-down guidance, nine SOAs were tested in Experiment 2: −50, 0, 25, 50, 75, 100, 150, 200, 400 ms. Negative SOA means that the search display appeared before the cue which was then presented at the center of the display. Second, cues were presented in the blocked condition. The target remained the same across all trials in a blocked condition (e.g. SMALL GREEN HORIZONTAL). However, the distractors could change. For example, on one trial the distractors might be big GREEN HORIZONTAL and SMALL red HORIZONTAL and on the next SMALL GREEN vertical and SMALL red HORIZONTAL. The cue for the first trial would be "small green" and, for the second, "green vertical". The target item is the same; the cue actually specifies the distractors. In addition, on half of the trials, the cue was an uninformative stimulus: in word cue conditions, the word "ready"; in picture cue conditions, a medium sized, white square.

Ten observers were tested. All had vision corrected to a minimum of 20/25, passed the Ishihara color screen, and gave informed consent. Each observer was tested with just one target-type in the blocked condition. Different targets were used for different observers. In the blocked condition, observers were tested in two blocks of 360 trials, one for word cues and the other for picture cues, for a total of 720 trials. Within each block, trials were evenly divided among target present and absent, nine SOAs, and two different cue types: informative (e.g. "red vertical" in picture or words) vs uninformative (white square or the word "ready") cues. Thus, there were 10 trials per cell.

In the mixed condition, target identity varied across trials. Observers were tested in four blocks of 450 trials with word cues and four blocks of 450 trials with picture cues, for a total of 3600 trials. Within cue condition, the 1800 trials were evenly divided among target present and absent and nine SOAs. Thus, there were 100 trials per cell.

3.2. Results

RTs less than 200 ms and greater than 4000 ms were excluded from the analysis (<1% of data). One subject was removed from analysis for excess errors (>10%). Errors otherwise averaged 6% and did not vary with SOA. Mean RT data are shown in Fig. 5.

The experiment replicates the main results of Experiment 1. Picture cues are effective rapidly—within about 100 ms in this experiment. Word cues are not as effective as picture cues (ANOVA: $F(1,8) = 193$, $p < 0.001$). The time course for word and picture cues is different as revealed by a significant interaction of SOA and cue type ($F(8,64) = 3.4$, $p = 0.0024$). Dashed lines in Fig. 5

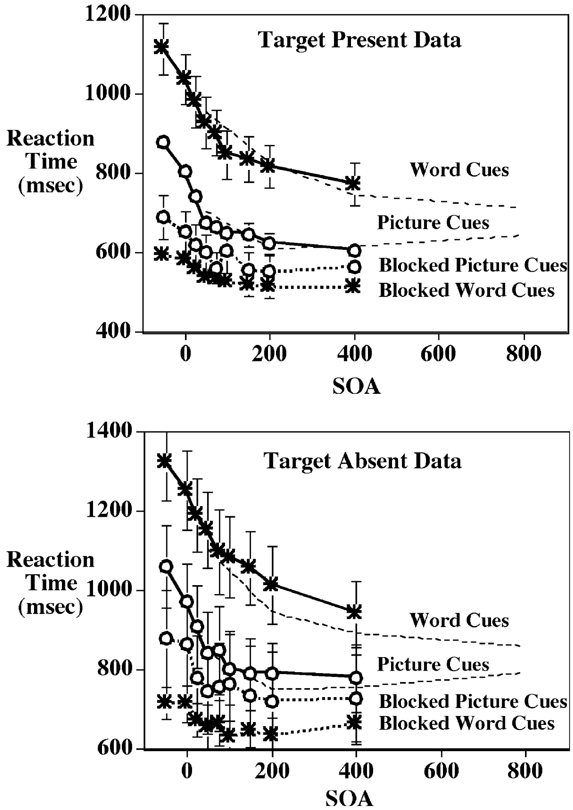

Fig. 5. Mean RT data for Experiment 2. Error bars are ±1 s.e.m. Dashed lines show data from Experiment 1 for comparison.

present the comparable word and picture cues data from Experiment 1, showing that the two experiments produce very similar results.

The difference between word and picture cues is made more dramatic if we look at the difference between blocked and mixed data for picture and word cues. First, there is no difference between informative and uninformative cues in the blocked conditions ($F(1,8) < 1$ for pictures and words) and, thus, the data are combined in Fig. 5. There is a significant effect of SOA ($F(8,64) > 8$, $p < 0.0001$ for both pictures and words). This suggests that the cue is serving as some sort of mask at short SOAs. Presumably, a similar masking effect influences the responses in the mixed conditions. The top-down effect can be seen by subtracting the blocked condition from the mixed. Thus, for picture cues there is little difference between mixed and blocked conditions, even at short SOAs. In the blocked conditions, the masking effect of picture cues appears to be somewhat greater than the effect of word cues. This effect is marginally reliable ($F(1,8) = 4.1$, $p = 0.07$).

Slopes are similar to those seen in Experiment 1. They do not vary systematically with SOA and are in the range of efficient "guided" search seen with conjunction stimuli of this sort. Average target present slopes range between 4.5 and 6.5 ms/item for blocked and mixed picture and word conditions. Average target absent slopes range between 11.4 and 16.4 ms/item.

3.3. Priming effects

In the mixed conditions, the main patterns from the preceding experiment recur: no significant benefit for priming with picture cues (ANOVA: $F(2,8) = 2.9$, $p = 0.08$) and a substantial benefit for priming with word cues (ANOVA: $F(2,8) = 6.7$, $p = 0.008$). On average, for word cues, the priming on trial $K+1$ is larger when the trial K was a target present trial (105 ms) than when it was a target absent trial (59 ms). Because of the relatively small numbers of trials, these results should be taken as suggestive rather than definitive.

Since this experiment was not specifically designed to investigate priming effects, there were not enough trials per observer to examine the priming effect of cue repetition as a function of SOA in the blocked conditions, so we have to pool over all SOAs. Recall that, in the blocked condition, the target is fixed and the cue specifies the distractors. We find that cue repetition, and therefore distractor repetition, does not have a significant benefit for either the word cues (1 ms difference for target present trials) or the picture cues (11 ms difference).

3.4. Discussion

The finer-grain analysis of SOA used in Experiment 2 confirms the main findings of Experiment 1. Picture cues are effective at short SOAs. Word cues are never as effective as picture cues even though observers were tested for thousands of trials with these simple stimuli. The lack of an effect of top-down information on slopes might appear to be a bit puzzling. After all, models like Guided Search (Wolfe, 1994, 2001) propose that conjunction searches are efficient because of the top-down guidance of attention to objects with the appropriate features. If guidance is absent at the start of the trial and develops during the course of the trial, why don't we see a slope effect? The answer may be that the appropriate guidance can be inferred without benefit of the cue. For example, given a display with several red horizontal and green vertical items, it is possible to infer that the target is either green horizontal or red vertical. Search can be guided on the basis of these inferences (Wolfe, 1992). This process of inferring the target is slow so the cue can speed search. However, it is still guided.

If this account is correct, then we might be able to see the effects of guidance on slope if we used a task that could not be accomplished without the cue. That is the purpose of Experiment 3.

4. Experiment 3: All unique items

4.1. Methods and stimuli

In Experiment 3, each item in the search display was unique. As a consequence, when the target changed from trial to trial, the task could not be done without the cue. Stimuli were colored, oriented bars, as in the previous experiments. These could be red or green, big or small, vertical, horizontal, or oblique; yielding $2\times2\times3=12$ stimuli. Set sizes were 3, 6, and 9 and items were chosen so that no item was ever duplicated within a search display.

Eleven observers were tested. All had vision corrected to a minimum of 20/25, passed the Ishihara color screen, and gave informed consent. Each observer was tested for 600 trials with a single target-type in the blocked condition: one block of 300 trials with picture cues and one block of 300 trials with word cues. Cues in the blocked condition were always presented at an SOA of 50 ms. Different targets were used for different observers. Observers were tested for 3000 trials in the mixed condition with SOAs of 0, 50, 200, and 400 ms. Cues were presented at the center of the screen. Picture cues were presented inside a white, outline box in order to distinguish them from the search array.

4.2. Results

RTs less than 200 ms and greater than 4000 ms were excluded from the data analysis(<1% of data). An unusual pattern appeared in the error data. False alarms are typically rare in search experiments (Chun & Wolfe, 1996). However, in the mixed trials with word cues, average false alarm (FA) rate was 17.8%. Three observers had FA rates of 24%, 40%, and 62%! When these observers are removed from the analysis, FA rate remains an unusually high 8.6%. Errors did not vary systematically with SOA. Word cues always contained two words. It seems possible that observers sometimes responded on the basis of the first word and an erroneous guess about the second (c.f. Spivey, Tyler, Eberhard, & Tanenhaus, 2001). Error rates for the eight remaining observers are shown in Fig. 6.

Of the eight remaining observers, one had very long RTs in the mixed word condition (averaging a second longer than the next slowest subject). This observer's data did not alter the basic pattern of average results, shown below. However, her data would increase the difference between word and picture cues in what might be a misleading manner. Accordingly, Fig. 7 shows RTs for seven observers.

The data show the now-familiar pattern of results. Picture cues have their effects very rapidly with RTs matching the blocked conditions within 200 ms. Word

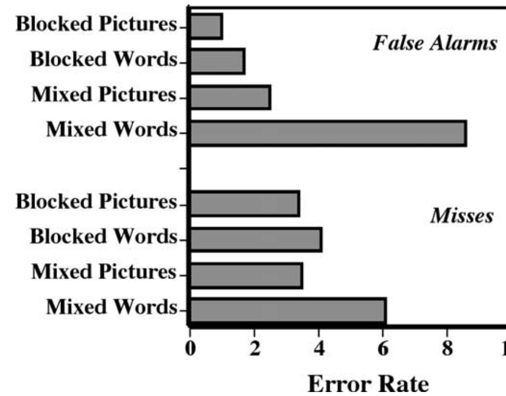

Fig. 6. Average error rates for the eight subjects in Experiment 3 with acceptable (<15%) error rates in the mixed words condition.

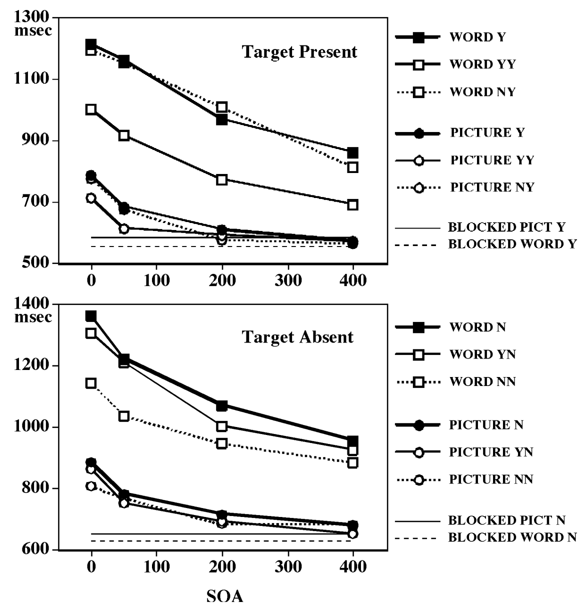

Fig. 7. Mean RT data for correct responses for seven observers in Experiment 3. Filled symbols indicate trials where the target cue on trial K was different from that on $K-1$. Open symbols show the data for trial K where $K-1$ had the same cue. Different curves indicate different pairs of trials. E.g. NY denotes target absent on the first trial and target present on the second.

cues are less effective. There is no priming effect for picture cues, presumably because the cue itself serves as an identity prime. The priming effect for word cues is large if and only if the target is present on trial $K-1$ as well as on the current trial K ("Word YY" line in Fig. 7). The word YY RTs are significantly shorter than unrepeated word cue RTs at all SOAs (all t-tests, corrected for multiple comparison, $p < 0.05$). The target on trial $K-1$ seems to serve as the prime for trial K. Unlike the previous experiment, no information about the distractors is useful. Accordingly, there is no priming at all in the word NY condition. Priming effects in the target absent conditions are not statistically reliable.

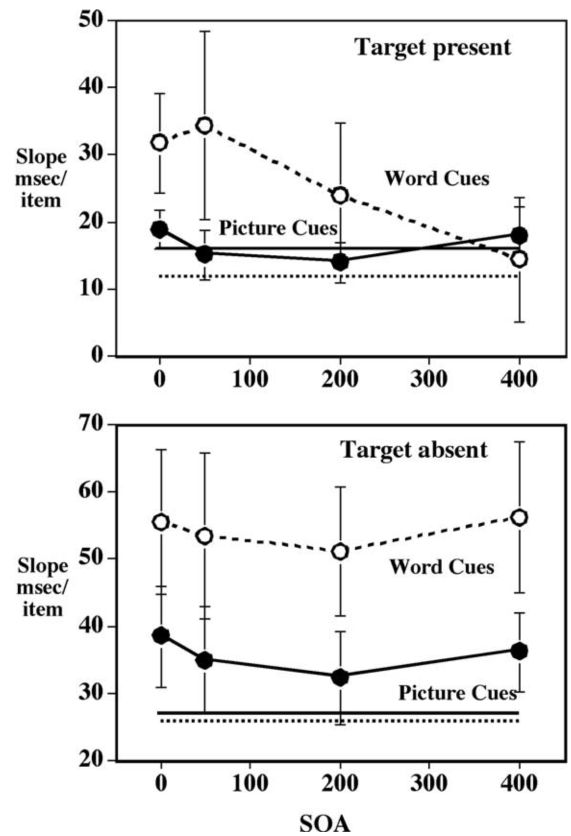

Fig. 8. Slope as a function of SOA for the picture and word cues in Experiment 3. Horizontal lines are the slopes for the blocked conditions (Solid = picture cue, Dashed = word cue).

Fig. 8 shows the slope data for Experiment 3. In this case, where guidance cannot be brought to bear in the absence of the cue, there is some evidence for an effect of cue on slope. The word cue, target present slopes are initially very inefficient and decrease as a function of SOA. However, while the main effects of SOA and set size are significant for word and picture cues (ANOVA: all $p < 0.01$), the slopes are too variable to achieve statistical reliability (ANOVA: set size x SOA interaction, $p > 0.1$ for both word and picture cues).

4.3. Discussion

Experiment 3 shows that the basic pattern of results, established in Experiments 1 and 2, does not change when the task cannot be done without a cue. The time course and the relative advantage for picture cues over words remains the same. The use of an SOA of zero ms makes it clear that the RT×SOA function does not represent the time that it takes to process the cue. Note that at zero ms a picture cue is already far more effective than a word cue. This can hardly reflect extensive processing of the cue in zero time. Rather, the RT×SOA function measures the lead time required to equate a cue and the full knowledge available when the target remains fixed across trials. The zero SOA data suggest that there is an initial "set up" time when the search stimulus appears before the search actually begins. During that time, the cue can start to "catch up" with the knowledge available in the blocked condition. This "set up" time presumably contributes to the intercepts of RT×set size functions in standard visual search tasks. Picture cues can catch up completely during this set up time. Word cues take longer. This is analogous to the concept of slack time in analyses of the psychological refractory period (Pashler, 1994).

5. Experiment 4: Is a picture a mandatory cue?

The clear advantage for picture cues over word cues could lead one to wonder if the effects of picture cues are mandatory. Would observers be able to ignore a fast but incorrect picture cue in favor of a slower, accurate word cue? In order to address this question, observers were presented with a word cue flanked by two identical picture cues. On two-thirds of the trials, the picture cue was invalid. On the remaining third, it matched the word cue. A schematic of an invalid picture cue trial is shown in Fig. 9.

On every trial, a valid word cue was presented at fixation. Invalid or valid picture cues flanked the word cue. After an SOA of 0, 50, 200, 400, or 800 ms, the search array was presented. Search stimuli were the same as in Experiments 1 and 2. The target was the odd items in a set of 10 or 18 items that were otherwise evenly divided between two distractor types. Thus, in the example in Fig. 9, the black vertical target is presented amidst black horizontal and white vertical distractors. In the actual experiment, items could be red or green, vertical or horizontal, big or small. The cues were presented on a gray background, slightly different from the black of the surround. Moreover, no search items could appear in the locations occupied by cues (words or pictures).

Eleven observers were tested for 300 trials at each SOA in the two-cue condition and for 100 trials at each SOA with words alone and with pictures alone.

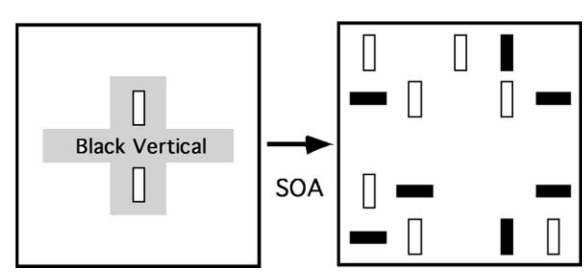

Fig. 9. A schematic invalid picture cue trial from Experiment 4.

5.1. Results

RTs less than 200 and over 4000 ms were removed from the data analysis. One observer had 10% of trials removed in this manner. For the remaining observers, only 0.3% of data were out of the acceptable range. Data from the one subject was excluded from analysis since this observer also had an excessively high error rate (>15%). For the remaining observers, the overall error rate was 6% with no systematic variation with SOA (ANOVA: $F(4,36) = 1.4$, $p = 0.25$) and a small (2%), but reliable effect of cue type (ANOVA: $F(4,36) = 4.8$, $p = 0.008$). This was due to a larger number of errors in the word alone cue condition and not to any tendency to make more errors with invalid picture cues.

Fig. 10 shows the mean RT data for the ten remaining observers.

It is clear that, when observers were gleaning information from the word cue, the performance was not much altered by the presence of invalid or supportive picture cues. There is a main effect of cue condition for target present trials ($F(3,9) = 3.2$, $p = 0.004$). Post-hoc analysis reveals what is obvious, that is due entirely to the differences between picture Only and other conditions. There are no significant differences among the other conditions. Target absent trials show the same pattern but with no significant effect of cue condition.

Slopes did not vary systematically with cue type, averaging 9 ms/item for target present trials and 25 ms/item for target absent.

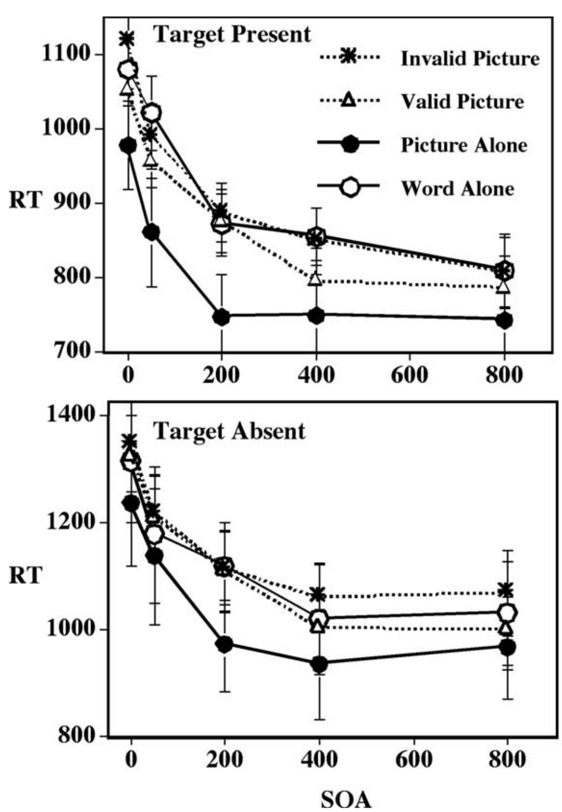

Fig. 10. Mean RTs (with s.e.m.) for Experiment 4. In the invalid and valid picture conditions, a pictorial cue was presented flanking a fully valid word cue. Only the flanking pictures were presented in the picture alone condition and only the word cue in the word alone condition.

5.2. Discussion

While picture cues can rapidly reconfigure the visual system for a search, this is not a mandatory effect. A reader of Lavie's work (Lavie, 1995; Lavie & Tsal, 1994) might argue that the picture cues were ignored because the word cue was so difficult that it forced a tight focus of attention on the words and excluded the pictures. This seems unlikely. However, it could be interesting to vary the demands imposed by the central cue in an effort to find a mandatory effect of invalid picture cues. In the absence of such data, it appears that observers can set themselves to read the word and ignore a picture cue. This is the case even though the effects of a picture cue may be related to the effects of an identity prime, which is considered to exert its effects automatically (Maljkovic & Nakayama, 1994).

Note that this result might be different if the picture cue had greater validity. For example, if the picture cue matched the word cue on three quarters of the trials, it might be hard to ignore on the one quarter of trials when it was invalid. The present data show that an unreliable picture cue does not have an automatic effect.

6. Experiment 5: Real objects

It could be that pictures were very effective cues in the preceding experiments because they were very simple stimuli. It might not take much time to reconfigure the visual system to look for a conjunction of a salient color and a salient orientation. However, we do not spend much time looking for Red Vertical stimuli in the real world. We look for chairs, lamps, cars, etc. In Experiment 5, observers search for photographs of such objects, isolated on a white background. There were 11 types of object. For each type (e.g. rabbit), there were two exemplars. Each observer was tested in two blocked conditions of 300 trials each. In the blocked conditions, the target item was fixed. In one block, the search stimulus was preceded by a picture cue. In the other, it was preceded by a word cue. Set sizes were 3, 6, and 9. Targets were present on 50% of the trials.

There were three mixed conditions, each run with SOAs of 50, 100, 200, 400 and 800 ms. In the mixed exact condition, the observers were tested for 200 trials per SOA with a picture cue that exactly matched the target, if present. In the mixed word condition, the observers were tested for 200 trials per SOA with a word cue that was a type match for the target, if present. That is, the word "rabbit" could apply to either instance of rabbit. In the mixed type condition, observers were tested for 400 trials per SOA with a picture cue that was an exact match to the target on 50% of target present trials (e.g. cue = rabbit1, target = rabbit1). On the other 50%, the target was the other instance of the cue type (e.g. cue = rabbit1, target = rabbit2). If rabbit1 was a target, rabbit2 was not a distractor (and similarly for other object types).

This experiment was designed so that, in mixed conditions, there was a one-third chance that the target and cue would not change from trial K to $K+1$. Methods were otherwise similar to previous experiments.

6.1. Results

RTs over 4000 ms and less than 200 ms were removed from analysis (<0.5% of data). Error rates were modest. Miss errors average 3.5% and False Alarms average 2.5%. There are significant effects of SOA and condition on error rate (ANOVA, all main effects, $p < 0.05$). These reflect small but reliable tendencies to make more errors at shorter SOAs and to make more errors when the cue is not an exact match to the target item.

Fig. 11 shows the mean RTs for the target present trials (target absent are less interesting, in this case, because, in the critical condition, there is no distinction between the absence of a type cue and the absence of an exact cue).

Fig. 11 reveals that the basic effect, seen in the previous experiments, extends to real objects. Solid circles show data from the block where different exact picture cues were intermixed. RTs rapidly approach the level of blocked trials. Word cues do not reach that level after 800 ms. Of most interest in this experiment are the data from the blocks where categorical and exact picture cues were intermixed. Here we see that exact cues (open circles) produce results similar to a block of exact cues (closed circles). Type picture cues (∗ with solid line) produce data similar to word cues (∗ with dashed line). These impressions are borne out by ANOVA. The main effect of condition is highly reliable ($F(3,9) = 34.7$, $p < 0.0001$). Post-hoc comparisons show that all pairs of mixed conditions differ significantly from each other except for the comparison between "exact mixed pictures and exact pictures (mixed with type)" and the comparison between type picture cues and word cues.

Slope data, shown in Fig. 12, suggest that exact cues provide better guidance than word or type cues in this case.

If guidance consists of giving more weight to some features and less to others, it is not surprising that exact pictures are superior to type picture cues or word cues. After all, the exact picture tells you the exact color (for example) of This Rabbit while the word or type cue merely informs you that something with a color appropriate to a rabbit will be the target. Again, ANOVA bears out the impression given by the figure. The main effect of condition is highly reliable ($F(3,9) = 13.7$, $p < 0.0001$). Post-hoc comparisons show that all pairs of mixed conditions differ significantly from each other except for the comparison between "exact mixed pictures and exact pictures (mixed with type)" and the comparison between type pictures and word cues.

6.2. Priming effects

As noted above, in mixed conditions there was a one-third chance that the target and cue would not change from trial K to $K+1$. This increased the number of repetitions in a manner that was not obvious to the observers though the higher frequency of repetition

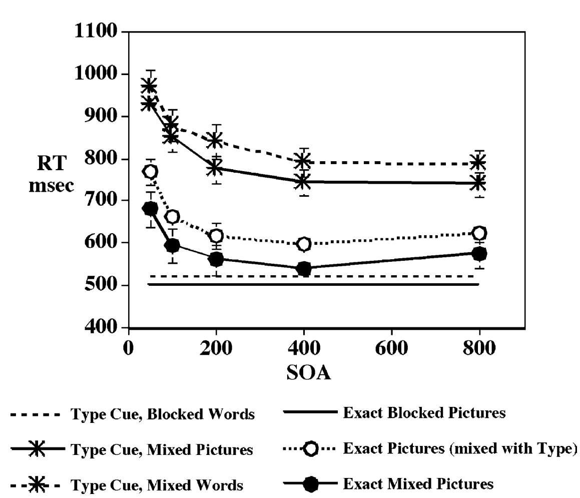

Fig. 11. Mean RTs (error bars = 1 s.e.m.) for Experiment 5. Note that "type cue, mixed pictures" were run in the same block as "exact pictures (mixed with type)". The type cue pictures behave like word cues while the exact pictures, mixed into the same block, behave more like exact pictures run by themselves.

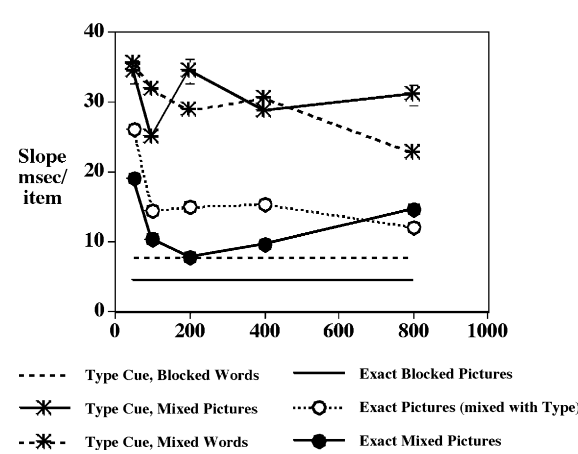

Fig. 12. Target present slopes for Experiment 5. Note that search is far more efficient when the cue is an exact match to the target.

might have had an implicit effect on behavior. Fig. 13 shows the target present data for trial K + 1 as a function of the status of trial *K*. For a given target-present trial, the prior target can have been target-present or target-absent and the cue and target can be either the repeated or not repeated. RTs for these four types of trial are plotted as a function of set size.

Looking at Fig. 13, it is clear that priming occurs when the same target is present on two successive trials. The main effect of priming condition is highly reliable ($F(1,9) = 30.4$, $p < 0.0004$). Looking at each of the four conditions separately (exact cues, exact cues mixed with type, type cues mixed with exact, and word cues), ANOVA shows that the four functions in each of the panels of Fig. 13 differ from each other ($F(3,9) > 11.0$, $p < 0.0001$, for all four conditions). As is obvious from the figure, post-hoc analysis reveals that this effect is due to the faster RTs in the Hit → Hit repeated cue case (all 9 relevant post-hoc comparisons, $p < 0.0024$). The only other significant post-hoc effect is that the repeated cue condition is *slower* than unrepeated for the Absent → Hit trials with type cues ($p = 0.025$). The size of the priming effect varies as a function of cue type as shown in Table 1. In the previous experiments, exact picture cues did not prime. Here they did prime but with a tendency to produce smaller priming effects than the word or type cues ($F(3,27) = 2.6$, $p = 0.07$). It is possible that this priming effect for exact pictures reflects a change in implicit strategy caused by the increased frequency of repetition or due to the use of different stimuli.

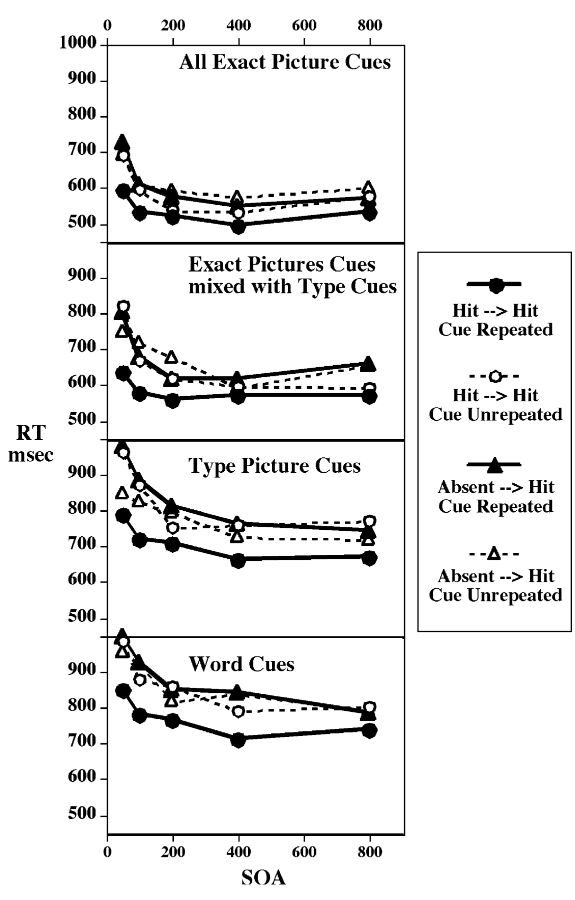

Fig. 13. Priming as a function of cue type and nature of the preceding trial. Note that priming occurs for all cue types when the preceding trial was a "Hit" trial. Priming is larger for word and type cues than for the exact cues.

Table 1
Average priming effect for different cue conditions (average RT for Hit → Hit priming trials subtracted from average RT for no repetition Hit → Hit trials)

Cue type	No repetition RTs—Hit → Hit RTs
Word cues	95
Type cues	114
Exact mixed with type	78
Exact cues	53
Absent → Absent (Exact)	64

The type cues behave like word cues in producing priming effects that are larger than the priming produced with an exact picture cue (*t*-tests comparing type and word cues to exact picture cues are significant, $p < 0.009$ in both cases).

6.3. Discussion

The central point to emerge from Experiment 5 is that type cues behave like word cues. This indicates that the difference between picture and word cues is not a problem with word reading. It is more likely that the difference reflects the privileged status of exact cues when it comes to setting up top-down guidance. Even if you have seen the "red" in the red vertical many times, seeing it again is more effective that recalling it. Even if you have seen both rabbits, seeing rabbit 1 does not set up the guidance for rabbit 2 the way that seeing rabbit 2 does. The presence of priming in the exact picture cue conditions may indicate that, with these stimuli, there are two components to the priming. As in the previous experiments, one of these would be a benefit from seeing the specific target item. The second would be a more general repetition priming effect where doing the same thing twice in a row (Yes, there is a rabbit; Yes, there is a rabbit) is faster than any change. In support of this notion, we can compare target absent trials that are exact copies of the previous trial (No, there is no rabbit, No, there still is no rabbit) with target absent trials where a different target was absent on the preceding trial (No, there is no rabbit, No, now there is no ball). The full repetition produces RTs that average 64 ms faster than the simple response repetition. This 64 ms is comparable to the 53 ms priming effect seen with picture cues in this experiment (see Table 1).

6.4. Experiment 6: Levels of categorical cueing

From the previous experiments, it might be proposed that there are two types of cues, exact picture cues and what could be called informational cues (words, categorical pictures, etc.). The purpose of Experiment 6 is to show that the specificity of that information is important. The logic of the experiment is shown in Fig. 14. A picture of an apple could be used to cue that EXACT apple, it could cue the TYPE "apple"—allowing for other apples, or it could cue a CATEGORY like "fruit". In Experiment 6, there were five categories of stimuli (fruit, tool, toy, bird, clothes). There were six types within each category (e.g. apple, cherry, grape, orange, peach, pear) and each type was represented by two examples. A control experiment was used to determine that all observers agreed on the names and categorical status of items. If a word cue was used, it could only specify a TYPE (apple) or CATEGORY (fruit). No effort was made to specify the EXACT item in this experiment (the apple sliced in half).

Fourteen observers were tested in each of eight conditions as specified in Table 2.

Note that, when a "type" picture cue was used, there could be an exact match or a type match between the target and the cue. For that reason, twice as many trials were run in the mixed-type-picture condition. In the categorical conditions, it would be possible, in principle, to have exact, type, and categorical matches of cue to target. However, in order to keep the experiment to a mere 6900 trials per observer, categorical picture cues always specified a member of the category that was not of the same type as the cue (e.g. apple could cue grape or peach but not the exact apple or another apple).

Stimuli were full color photographs of objects presented on a white background. Set sizes of 3, 6, and 9 were used. SOAs were 50, 100, 200, 400, and 800 ms. Stimuli were presented in a 22 by 22 deg field. In the blocked conditions, each observer was randomly assigned a cue (exact, type, or category as the situation demanded).

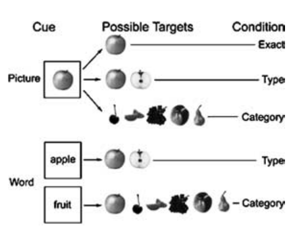

Fig. 14. Cue–target relationships in Experiment 6. A cue could exactly match the target, or it could specify its "type" or its "category". Words could only specify type and category.

Table 2
Conditions of Experiment 6

Blocked/ Mixed	Cue type	Picture/Word	No. of trials
Blocked	Category	Picture	300
Blocked	Exact	Picture	300
Blocked	Type	Picture	300
Mixed	Category	Picture	1000
Mixed	Exact	Picture	1000
Mixed	Type	Picture	2000
Mixed	Category	Word	1000
Mixed	Type	Word	1000

6.5. Results—blocked conditions

Three observers were removed from data analysis because of excessive errors in one or more conditions. The remaining eleven observers produced 1.5% Miss errors and 0.6% False Alarms in the blocked conditions. Errors decreased as SOA increased, as did RTs.

The eight conditions of the experiment produce too much data to present in a single figure. Fig. 15 shows the mean RT (±1 s.e.m.) data for the three blocked conditions.

These blocked conditions vary in the amount of top-down guidance that can be deployed. For example, if the cue is a green apple, in an exact cue condition, the cue specifies color. In the type cue condition, the cue restricts color (especially when the observer has learned that the two apples in this experiment are a green one

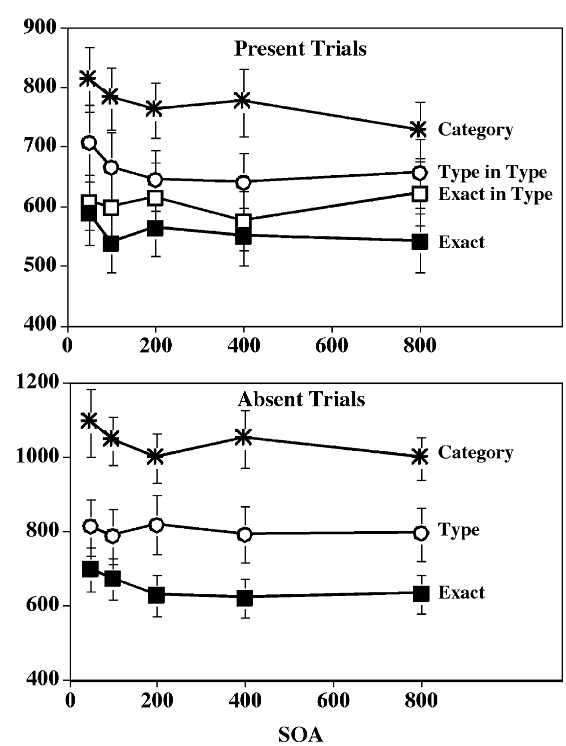

Fig. 15. Mean RT for the blocked conditions of Experiment 6. In the type cue condition, for target present trials, RTs are divided into those where the cue was an exact match to the target and those where the cue was a type match.

and a white, cut one). In the categorical condition, color guidance is nearly useless since fruit does not have a characteristic color. This is reflected in the significant effect of condition on mean target present RT (ANOVA: $F(3,39) = 14$, $p < 0.0001$). This being a blocked condition, there is little effect of SOA ($F(4,52) = 2.6$, $p = 0.048$). Target absent trials follow a similar pattern.

The amount of guidance should be reflected in slopes and it is. The target present slope for category cues, 21 ms/item, is significantly greater than the 7 ms/item slopes for the other conditions (*t*-tests, all $p < 0.01$). The difference between type and exact cue conditions is seen in the target absent trials where type cue slopes of 31 ms/item are significantly greater than exact cue slopes of 11 ms/item but significantly less than the 79 ms/item slope for the category condition (*t*-tests, all $p < 0.005$).

6.6. Results—mixed conditions

Miss error rates for the eleven observers whose data are analyzed were 5.3%. False Alarms were 3.7%. Error rates decreased as SOA increased. The same pattern is seen in the RTs, arguing against a speed-accuracy tradeoff explanation of the effects of SOA.

Fig. 16 shows RTs for the mixed conditions with the blocked conditions shown for comparison. As in the blocked conditions, observers are fastest when the cue is an exact match to the target and slowest with a categorical match. On blocks when the cue specifies the target type, Hit RTs are faster when the cue happens to be an exact match than when it is not.

The speed with which top-down guidance develops after a cue is seen more easily if we subtract the blocked

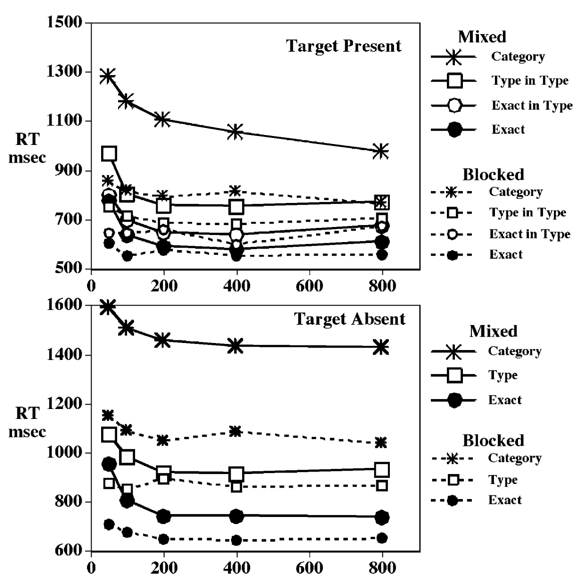

Fig. 16. RT as a function of SOA for the mixed conditions of Experiment 6. Blocked conditions, shown in Fig. 15, are replotted as dashed lines.

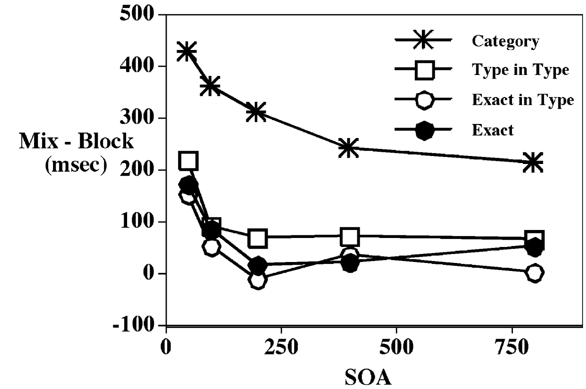

Fig. 17. Difference in RT between mixed and blocked conditions for the same type of cue (target present trials).

condition from the comparable mixed condition. This is shown in Fig. 17. Separate ANOVAs were performed for each of the four cueing conditions for the target-present data. Except for the exact cues in the type condition, there is a significant difference between mixed and blocked conditions ($F(1,10) > 5$, $p < 0.05$). For all cue types, there is a significant effect of SOA ($F(4,40) > 3$, $p < 0.03$). Finally, the interaction of SOA with the difference between mixed and blocked is always significant ($F(4,40) > 2.8$, $p < 0.05$).

These statistics confirm the message of Fig. 17. At short SOAs, notably SOA = 50 ms, mixed RTs are markedly slower than the relevant blocked baseline. After an SOA of 200 ms, exact picture cues produce RTs in the mixed conditions that are essentially the same as those in the blocked conditions. Type cues produce RTs that are somewhat slower (not statistically significant, in this case). Information from category cues is the least effective and the slowest.

The difference in top-down guidance between cue conditions is seen quite dramatically in the slope data (Fig. 18). Categorical cues provide little if any guidance. The slope does not change as a function of SOA and is consistently inefficient. The cue merely identifies the target. The exact cues work powerfully to make search efficient after a 200 ms SOA. There are two curious aspects to the slope data. First, the type cues seem to provide more guidance here than they did in a roughly equivalent condition in Experiment 5 (see “type cues, mixed picture” condition in Fig. 12). The difference between exact and type cues shows up in the target absent slope data (lower panel, Fig. 18). Specifying that a target was an “apple” or a “rabbit” produced guidance that could make search more efficient. Specifying categories like “animal” or “fruit” did not have this effect.

The second curious aspect of the slope data is the U-shaped nature of slope X SOA functions. The slopes reach a minimum around 100–200 ms and then appear to rise. This experiment lacks the power to determine if

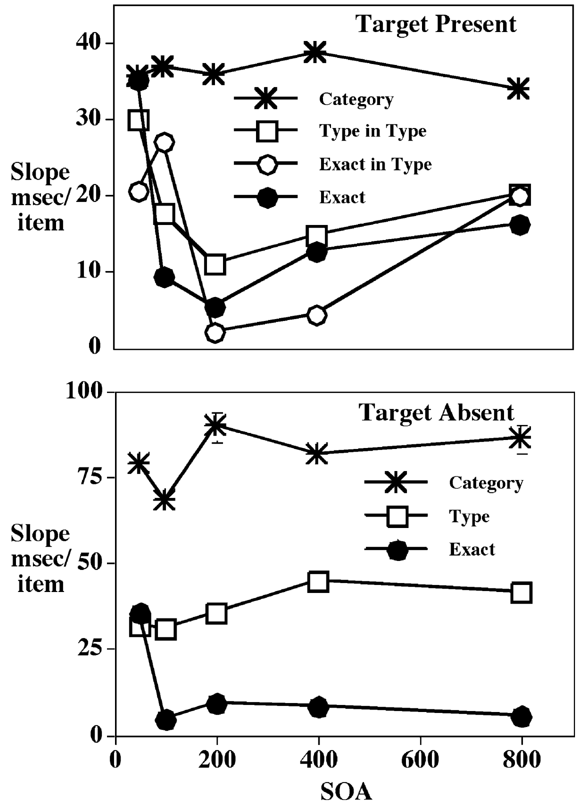

Fig. 18. Slopes of RT×set size functions as a function of SOA for mixed conditions of Experiment 6.

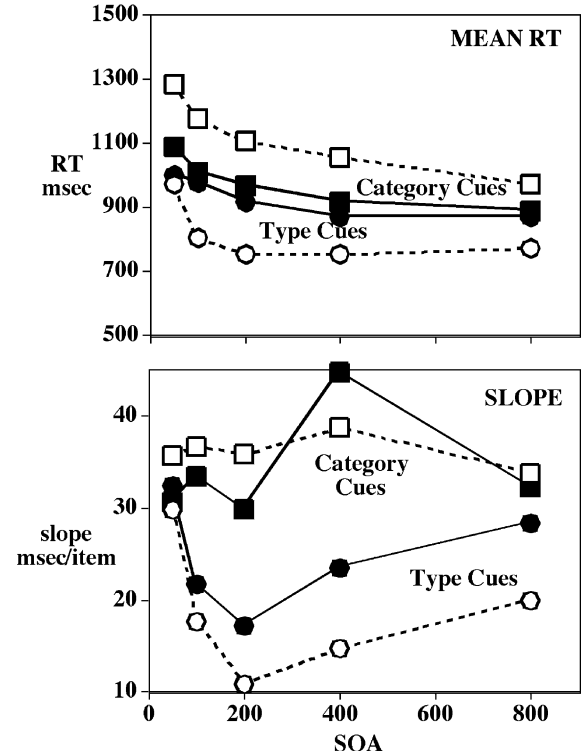

Fig. 19. Mean RTs and slopes for word cues (solid lines) in Experiment 6. Category cue data ("animal") are shown with squares. Type cues ("rabbit") are shown with circles. Equivalent picture cue data are shown with dashed lines.

this is reliable but it is worth noting that a similar, if less pronounced, shape can be seen in other experiments (see Figs. 8 and 12). This raises the possibility that, rather like the effects of an onset or singleton cue, the strong cueing effects of an exact picture cue might quickly rise to a peak and then decline. Sustained top-down guidance might be augmented by a transient component that fades during the course of a trial.

6.7. Results—word cues

Word cue results are similar to the picture cue results and are shown, for target present trials, in Fig. 19. When the word specified the target type (e.g. "rabbit"), mean RTs and slopes are similar to the "type in type" condition where a type picture cue was not an exact match to the target. The slope shows the same U-shape as in Fig. 18. When the word cue specified a category (e.g. "animal"), results are similar to the picture category cues. Word cues actually produce RTs that are somewhat faster than the picture cues, perhaps because a word can directly specify a category like animal where as a picture must use a specific animal to invoke the general class. Worse yet, in this experiment, the categorical picture cue was always the wrong animal, presumably adding to the word advantage. Slopes were similar in word and picture category cases. In both instances, there is no significant change in slope with SOA. Search is inefficient because the category specifies target identity but fails to produce guidance.

6.8. Discussion

The results of Experiment 6 indicate that the effectiveness of a cue depends on its ability to specify the information needed for the search. An exact picture cue provides all the information possible. These can swiftly act to reconfigure the visual system in order to guide attention toward the target. A type cue—word or picture—is less effective because it specifies less. For example, an exact picture of This Rabbit specifies the exact color of the target. A type cue is less precise, even when observers know that only two rabbits appear in the present experiment. This is akin to the uncertainty effects seen with simpler stimuli like spatial frequency gratings (Davis et al., 1983). Category cues are still less precise and produce little or no guidance in this case. Had we used other categories that had more obvious common features, we might have seen more guidance than we saw with the present set of categories.

7. General discussion

To summarize the results of these six experiments, information about the identity of a target acts quickly to configure the visual system to look for that target. A valid exact picture cue that precedes the search stimulus by less than 200 ms produces a search of the same speed and efficiency as blocked search. Other cues work more slowly and less completely, though perhaps we would have seen a "complete" effect if we had used SOAs longer than 800 ms. For example, one could look at Fig. 10 and conclude that the word cue effect would approach the exact picture cue effect by about 1600 ms SOA. In other experiments, this possibility is less clear (e.g. Fig. 11).

The results point to several cumulative aspects to the top-down information in a cue. The categorical cues of Experiment Six seem to provide information that is only useful once attention has selected an object. You know you are looking for an "animal". Once you select an item, it is clear enough whether or not it is an animal. However, the word "animal" does not provide information that allows you to guide attention toward some items and away from others in this experiment. As a result, the categorical cue fails to reduce the search slope. Search is similarly inefficient if observers look for animals throughout a block. More specific cues permit guidance because they specify some or all of the preattentively available features that define the target. Telling the observer to look for "red vertical" allows that observer to configure his visual system in a manner that allocates attention preferentially to red and vertical items. These components are under volitional control and can be considered to be "explicit" top-down guidance.

The advantage for exact picture cues and the evidence of priming effects point to what might be considered to be "implicit" top-down guidance (Wolfe et al., 2003). Here the border between top-down and bottom-up is blurred and depends on one's precise definition of the terms. An additional improvement in performance comes from seeing the actual target prior to searching for it. This sort of priming of search has been shown to be immune to the observer's explicit knowledge. Thus, knowing that the target will alternate between red and green from trial to trial does not help in simple pop-out search tasks (Maljkovic & Nakayama, 1994). If one considers this explicit knowledge to be part of the definition of "top-down", then the priming effect of exact picture cues is not a component of top-down control of search. However, if one defines "bottom-up" as attention-guiding information that is present in the search stimulus (e.g. salience based on local differences, Itti & Koch, 2000; Li, 2002) then "top-down" would be any information about target identity that is *not* present in the search stimulus. Under this definition, priming by exact cues can be considered to be an implicit form of top-down guidance.

Regardless of one's position on this essentially semantic issue, the added effect of an exact picture cue leads one to ask how exact an exact cue must be. If a picture of a rabbit guides attention to that same picture of that same rabbit, what happens if the cue is in a different orientation or size? What happens if the cue is black and white and the search stimuli are in color (or vice versa)? There is bound to be some tolerance for deviation from an exact match. The nature of that tolerance will inform us about the nature of the information that is abstracted from the cue.

Acknowledgements

This research was supported by the Transportation Security Administration and the National Eye Institute. We would like to thank David Burr and two anonymous reviewers for their contributions to the manuscript.

References

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception and Psychophysics, 55*(5), 485–496.

Blough, P. M. (1989). Attentional priming and visual search in pigeons. *Journal of Experimental Psychology—Animal Behaviour Process, 15*(4), 358–365.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443–446.

Chastain, G., Cheal, M., & Lyon, D. R. (1996). Attention and nontarget effects in the location-cuing paradigm. *Perception and Psychophysics, 58*(2), 300–309.

Cheal, M. L., Lyon, D. R., & Gottlob, L. R. (1993). A framework for understanding the allocation of attention in location-precued discrimination. *Quarterly Journal of Experimental Psychology*, MS.

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.

Davis, E., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. *Perception and Psychophysics, 33*, 20–28.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*(4), 481–507.

Eckstein, M. P., Whiting, J. S., & Thomas, J. P. (1996). Role of knowledge in human visual temporal integration in spatiotemporal noise. *Journal of Optical Society of America—A, 13*(10), 1960–1968.

Egeth, H. (1977). Attention and preattention. In G. H. Bower (Ed.), *The psychology of learning and motivation* (vol. 11, pp. 277–320). New York: Academic Press.

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

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*(2), 317–329.

Found, A., & Muller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a 'dimension weighting' account. *Perception and Psychophysics, 58*(1), 88–101.

Hubner, R. (1996). The efficiency of different cue types for reducing spatial-frequency uncertainty. *Vision Research, 36*(3), 401–408.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489–1506.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451–468.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics, 56*(2), 183–197.

Li, Z. (2002). A salience map in primary visual cortex. *Trends in Cognitive Sciences, 6*(1), 9–16.

Lu, Z.-L., Lesmes, L. A., & Dosher, B. A. (2002). Spatial attention excludes external noise at the target location. *Journal of Vision, 2*(4), 312–323.

Luck, S. J., & Vecera, S. P. (2002). Attention. In H. Pashler & S. Yantis (Eds.), *Stevens' Handbook of Experimental Psychology. 1: Sensation and Perception* (pp. 235–286). New York: Wiley and Sons.

Maljkovic, V., & Nakayama, K. (1994). Priming of popout: I. Role of features. *Memory and Cognition, 22*(6), 657–672.

Monnier, P., & Nagy, A. L. (2001). Set-size and chromatic uncertainty in an accuracy visual search task. *Vision Research, 41*(28), 3817–3827.

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of Optical Society of America—A, 7*(7), 1209–1217.

Olivers, C. N. L., & Humphreys, G. W. (2003). Attentional guidance by salient features depends on intertrial contingencies. *Journal of Experimental Psychology: Human Perception and Performance, 29*(3), 650–657.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 108–130.

Pashler, H. (1994). Dual task interference in simple tasks: Data and theory. *Psychological Bulletin, 16*, 220–244.

Pashler, H. (1997). *The psychology of attention*. Cambridge, MA: MIT Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Erlbaum.

Solomon, J. A., Lavie, N., & Morgan, M. J. (1997). The contrast discrimination function: spatial cuing effects. *Journal of the Optical Society of America, 14*, 2443–2448.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science, 12*(4), 282–286.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 799–806.

Treisman, A. (1988). Features and objects: The 14th Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40A*, 201–237.

Weidner, R., Pollmann, S., Muller, H. J., & von Cramon, D. Y. (2002). Top-down controlled visual dimension weighting: an event-related fMRI study. *Cerebral Cortex, 12*(3), 318–328.

Wolfe, J. M. (1992). "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research, 32*(4), 757–763.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*(2), 202–238.

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

Wolfe, J. M. (2001). Guided Search 4.0: A Guided Search model that does not require memory for rejected distractors. *Journal of Vision, Abstracts of the 2001 VSS Meeting*.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 483–502.

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.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 34–49.