On the limits of top-down control of visual selection

Jan Theeuwes · Erik Van der Burg

Abstract In the present study, observers viewed displays in which two equally salient color singletons were simultaneously present. Before each trial, observers received a word cue (e.g., the word *red,* or *green*) or a symbolic cue (a circle colored red or green) telling them which color singleton to select on the upcoming trial. Even though many theories of visual search predict that observers should be able to selectively attend the target color singleton, the results of the present study show that observers could not select the target singleton without interference from the irrelevant color singleton. The results indicate that the irrelevant color singleton captured attention. Only when the color of the target singleton remained the same from one trial to the next was selection perfect—an effect that is thought to be the result of passive automatic intertrial priming. The results of the present study demonstrate the limits of top-down attentional control.

Keywords Attention · Attention: Selective · Attentional capture

Imagine a situation in which two uniquely colored and highly distinguishable objects are present in the visual field. Each time before you start searching, you are told which of the two objects you need to select. For example, on one trial, you need to select the red object, and on the next trial, you need to select the green one. On the face of it, this should be no problem: Everyone expects that people can

J. Theeuwes (✉) · E. Van der Burg
Department of Cognitive Psychology, Vrije Universiteit,
Van der Boechorststraat 1,
1081 BT Amsterdam, The Netherlands
e-mail: J.Theeuwes@psy.vu.nl

select the object they are told to select. This intuitive assumption is reinforced by most theories on visual search that predict that people can select the object needed for their task. For example, the contingent capture hypothesis predicts that observers can set themselves on each trial for a particular color (Folk, Remington, & Jonhnston, 1992). The guided search theory argues that top-down set can increase the salience of the relevant feature dimension (in this example: the feature "red" or "green") so that attention is (mostly) guided to relevant features only (Wolfe, 1994; Wolfe, Butcher, Lee, & Hyle, 2003). Other theories argue that visual selection is biased toward "some kind short-term description of information" that describes what is needed for the task at hand (Duncan & Humphreys, 1989). This attentional template ensures that stimuli that match that description are selected over those that do not match (Bundesen, 1990). Even though most theories predict efficient top-down selection, here we show that selection in a top-down manner is inefficient; people cannot flexibly select the object needed for their task.

Recently, in a series of experiments, Theeuwes and colleagues (Theeuwes, Reimann, & Mortier, 2006; Theeuwes & van der Burg, 2007, 2008) tested the boundary conditions of top-down selection. For example, in Theeuwes et al. (2006), observers were confronted with pop-out displays in which one element either had a unique shape (one green diamond among five green circles) or a unique color (a red circle among five green circles). Observers had to respond to the orientation of a line segment inside the unique popping-out element. Before each trial, observers received a cue telling them with 80% validity the likely feature property of the upcoming target singleton. For example, observers received as a cue the word *color* (or the word *red*) and knew with 80% validity that the line segment they were looking for would be

presented within the red-colored circle. In the 20% invalid trial, the target line segment would appear in the shape singleton (the green diamond). In the neutral condition, no information was provided about the properties of the target singleton. The results showed that preparing for the upcoming target feature had no effect on reaction time (RT; Theeuwes et al., 2006) or on perceptual sensitivity (Theeuwes & van der Burg, 2007). In other words, knowing that the popping out target is going to be a color singleton instead of a shape singleton did not affect the speed of selection or perceptual sensitivity.

These previous results were considered to be controversial: Many theories of visual search (Bundesen, 1990; Duncan & Humphreys, 1989; Folk et al., 1992; Müller, Reimann, & Krummenacher, 2003; Treisman, 1988) would predict that endogenously preparing for the target feature should improve visual search and should enhance perceptual selectivity. For example, according to Müller and colleagues (e.g., Müller et al., 2003), a verbal cue telling the upcoming feature dimension should enable observers to allocate more weight to the relevant visual dimension. Because these previous experiments were inconsistent with their theory, Müller and Krummenacher (2006) tried to replicate Theeuwes et al.'s (2006) findings. Interestingly, they found numerically exactly the same cue benefit (9 ms) as did Theeuwes et al. (2006). However, unlike in Theeuwes et al. (2006) and Müller and Krummenacher, this 9-ms effect was statistically reliable presumably because they made sure that observers had an extra incentive to use the cue. Therefore, it seems that under the right circumstances, these very small verbal cuing effects may become statistically reliable.

In addition to the concern that observers may not have used the cue in Theeuwes et al. (2006), there was the concern that the task may have been too easy to see an effect. Indeed, in each display, there was one popping-out element with either a unique color or a unique shape. Importantly, in each display, the only popping-out element was always the target. Because the task was so easy, knowing the likely dimension of the upcoming target may not have helped much. This is consistent with the biased competition model of attention (Desimone & Duncan, 1995). According to this view, competition among elements is needed to see an attention effect (for a recent review, see Beck & Kastner, 2009). Because in our display there was only one popping-out element, there was basically no competition. However, if two salient elements are simultaneously present in the visual field, visual attention is needed to resolve the neural ambiguity sp that only the relevant object continues to be represented in the system (Luck, Girelli, McDermott, & Ford, 1997). Related to the previous point: Maybe, because the task was so easy, observers may have chosen not to use the verbal cue to

prepare for the upcoming target dimension. This may explain why we did not see an effect of the cue.

To circumvent these concerns, in the present study, we addressed the question of whether top-down knowledge can improve visual search in a task in which there was competition between two popping-out elements. Moreover, we ensured that observers had to use the cue in order to be able to respond. We employed a visual search task in which there were always two salient elements (e.g., a red and green target circle) popping out from the background of five gray circles. Within all circles of the display, there were vertical or horizontal line segments, and observers responded to the orientation of the line segment located within the colored target singleton. Observers were verbally cued before display onset (e.g., the word *red* or *green* appeared) what the target color would be on the current trial. They responded to the orientation of the line segment within the target singleton. For example, when the word *red* would be presented as a cue, observers needed to respond to the line segment inside the red singleton and to ignore the green singleton distractor. Clearly, in contrast with our previous studies (Theeuwes and van der Burg, 2007, 2008), observers had to process and use the cue; otherwise, they were not able to do the task.

The present task was simple, and the intuitive prediction was that people should be perfectly able to do this task. The interval between cue and search display was long enough for observers to set their attentional bias toward the relevant color. For example, Theeuwes and van der Burg (2008) used a rather complicated location cue that indicated that the target would appear, for example, "at the 10 o'clock" position on an imaginary clock. In this experiment, a 1.5-s cue to target interval was very efficient in generating strong location cuing effects even though observers had to translate the complicated verbal cue (e.g., the 10 o'clock position) into an effective attentional set. It is therefore reasonable to assume that a fairly simple cue (e.g., such as "attend to red") as used in the present study, is also effective when observers have 1.5 s to prepare.

As a measure of the efficiency of selection, we used the identity intrusion technique first introduced in Theeuwes and Burger (1998). The idea is that if observers are able to perfectly select the target singleton and respond to the line segment inside of it, then the identity of the line segment positioned in the distractor singleton (the other popping-out element) should have no effect on responding. However, if selection is not perfect, and on a few trials attention goes to the distractor singleton before going to the target singleton, then a congruency effect is expected. In other words, if the line segment in the distractor singleton is identical to that inside the target singleton (i.e., congruent), observers should respond faster than when the line segments are incongruent (e.g., a horizontal line segment inside the target

singleton and a vertical line segment in the distractor singleton) observers should be relatively slow (see Gibson, & Bryant, 2008; Schreij, Owens, & Theeuwes, 2008; Theeuwes, 1995; Theeuwes & Burger, 1998). Therefore, the presence or absence of a congruency effect isolates the efficiency of selection. Note that if attention does not go to the distractor singleton, then one does not expect a congruency effect because focal attention is needed at the location of the distractor singleton to reveal the identity of the line segment inside of it (see Theeuwes, van der Burg, & Belopolsky, 2008). In this case, if the identity of the line segment inside the distractor singleton is not available, it cannot affect the speed of responding toward the line segment inside the target singleton.

Experiment 1

To isolate the pure top-down attentional set effect, in Experiment 1, we ensured that intertrial bottom-up priming would be minimum. Previous studies investigating this so-called "priming of pop out" have demonstrated that in these type of tasks, priming is basically bottom-up in origin (Maljkovic & Nakayama, 1994; Olivers & Hickey, 2010) and cannot be counteracted by top-down control (Pinto, Olivers, & Theeuwes, 2005; Theeuwes et al., 2006). Note, however, that some studies have provided evidence that "priming of pop out" is top-down modulable (Geyer & Müller, 2009). To avoid any influence of intertrial priming, there was never a repetition of the target or distractor color from one trial to the next.

Method

Observers Nine students (six females; mean age = 21.1 years, ranging from 18 to 25 years) participated in Experiment 1 as paid volunteers. Each observer received €7 an hour. All observers were naive as to the purpose of the experiment. The data from one observer was excluded from further analyses because of an overall high error rate (>10%).

Stimuli and Apparatus Experiments were run in a dimly lit, air-conditioned cubicle. Observers were seated at approximately 80 cm from the monitor (120-Hz refresh rate). Figure 1 shows an example of an experimental trial.

A trial started with a presentation of a light-gray fixation dot (radius 0.4; 29.3 cd/m^2) in the middle of the screen for a fixed period of 900 ms. The luminance of the dark-gray background was kept constant during the experiment (0.58 cd/m^2). Immediately following the presentation of the fixation dot, a word cue presented in white was displayed for a period of 850 ms at the center of the
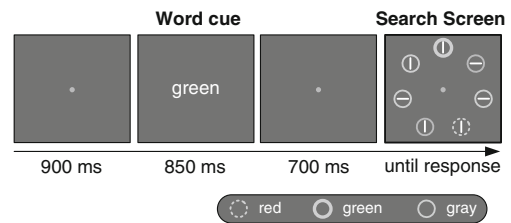

Fig. 1 Example trial of Experiment 1. Observers were asked to search for one of two color singletons and to respond to the line segment inside the one that was indicated by the cue. The orientation of the line segment inside the (cued) target singleton could be congruent (e.g., both horizontal) or incongruent (e.g., one horizontal and one vertical) with the orientation of the line segment inside the other color distractor singleton. In this example, it is congruent since the line segment inside the target singleton (which is green) is similar to the line segment inside the distractor singleton (which is red)

screen. This cue indicated the color of the circle they had to select. The characters of the word cue were presented in 18-point Courier new font (0.46° height; 0.43° width; 29.3 cd/m^2), and was either the word "rood," "groen," "geel," "blauw," "paars," or "oranje" (i.e., red, green, yellow, blue, purple and orange in Dutch, respectively).

After the cue was presented, the fixation dot reappeared for a fixed period of 700 ms followed by a search display with seven items equally spaced around the fixation dot on an imaginary circle with a radius of 5.3° until observers made a response. The display consisted of five gray outline circles (radius 1.0°; 15.6 cd/m^2), and two uniquely colored circles, each with a unique color (red, 15.6 cd/m^2; yellow, 86.0 cd/m^2; blue, 7.53 cd/m^2; purple, 14.2 cd/m^2; green, 15.6 cd/m^2 ; and orange, 34.9 cd/m^2). To prevent intertrial priming, none of the colors were repeated on the next trial. Within the center of each outline circle, a horizontally or vertically oriented white line segment (length 1.2°; 29.3 cd/m^2) was presented.

Procedure Observers responded to the orientation of the line segment (i.e., horizontal or vertical) presented inside the colored target singleton, which was indicated by the word cue. Note that it is impossible to do the task properly without attending to the word cue. Observers used the "z" and "m" keys to respond to the horizontal or vertical line segment, respectively. A tone was presented when observers committed an error. Observers were instructed to remain fixated on the central fixation dot during the course of a trial. The next trial started after 1,500 ms.

Experimental design The independent variables were all varied within subjects and within blocks of trials. The orientation of the line segment inside the target singleton was randomized with an equal probability of a horizontal or a vertical target line on each trial. The orientation of the line segment in the distractor singleton was also randomized

with equal probability of a horizontal or vertical line in each trial. The orientation of each line in the gray circles was either horizontal or vertical and randomly determined. The distance between the target and distractor singleton was balanced. Observers practiced one block of 48 trials. Then, they completed six experimental blocks of 48 trials each. After each block, they were informed on the screen of their mean RT and number of errors of that block, which they copied on a sheet of paper to help them maintain a high level of performance.

Results

We discarded the data from both the practice blocks and the first three trials of each block. RT outliers (RTs > 2,500 ms) were also discarded (0.4%). We determined whether there were effects of congruency on RTs and errors.

RT data A *t* test indicated that observers were faster to respond to the line segment inside the target singleton when this line segment was congruent with the line segment inside the distractor singleton (703 ms) than when the line segment inside the target singleton was incongruent with the line segment inside the distractor singleton (729 ms), $t(7) = 4.2$; $p < .005$. The presence of this congruency effect indicates that observers were unable to exclusively select the target singleton indicated by the cue. The congruency effect indicates that, at least on a subset of trials, observers selected the distractor singleton before selecting the target singleton, giving rise to the congruency effect.

Error data Overall mean error rate was 4.2%. In the congruent condition, the error rate was 3.7%; in the incongruent condition, it was 4.5%. This difference was not reliable.

Furthermore, a separate ANOVA with color (red, green, yellow, blue, purple, and orange) as a within-subjects variable was included to investigate the difference between the physical colors of the targets. Even though there were RT differences between the various colors (ranging from 693 ms for yellow to 751 ms for purple), there was no reliable statistical effect of the color used, which suggests that these colors were more or less equally salient.

Discussion

The present results showing a clear congruency effect indicate that observers were not able to selectively direct their attention to the color of the target singleton, even though they had about 1.5 s to prepare for the color of the upcoming target singleton. The present task presents a situation that should allow for optimal top-down control of selection: Observers were instructed to select an item with one of two clearly distinguishable colors. Even though this should pose no problem to the visual system, our results show otherwise: The clear congruency effect indicates that at least on a subset of trials, attentional selection was not perfect; attention must have gone first to the distractor before it went to the target.

Note that in this task, observers must have used the verbal cue to decide to which color singleton to respond. If they would have ignored the cue, their responses would have been at chance level. Clearly, our 4.2% error rate indicates that observers used the cue.

The present findings are inconsistent with those theories of visual search that assume a large role for top-down control (Duncan & Humphreys, 1989; Folk et al., 1992; Lien, Ruthruff, & Johnston, 2010). Because in the present experiment the color singletons used were about equally salient (there was no difference in RTs between the various colors used), an increase in the top-down weight for either one or the other color singleton should have facilitated top-down selection of the target singleton. The presence of the congruency effect suggests that the top-down weight toward the target color singleton could not prevent attentional capture by the irrelevant color singleton.

The present results are consistent with the findings of Theeuwes (Theeuwes et al. 2006; Theeuwes & van der Burg, 2007), who showed that verbal top-down knowledge cannot affect early perceptual processes. As outlined by Theeuwes (2010b), a verbal cue may not be able to affect these early perceptual selection processes because a verbal instruction (e.g., "search for red") cannot directly affect the "cycling" of neurons. However, Theeuwes (2010a, 2010b) argued that neurons in early visual areas coding for a particular feature (e.g., the color red) can change their cortical representation after these neurons have been exposed to the actual stimulus feature (e.g., neurons need to "cycle" in order to get set for a particular feature). In other words, after processing the color "red" on one trial, selection of red things on the next trial is improved. This effect, known as "priming of pop out," was first described by Maljkovic and Nakayama (1994). For example, in one of their studies, observers searched for a feature singleton that was defined in one of two different ways: a red target among green distractors, or a green target among red distractors. They showed that even when the repetition of the feature value was at chance level (i.e., repetition was not more likely than alternation), repeating a target (but not the response) improved performance. Feature priming refers to the benefits in processing that result from the repetition of features that characterize a target in visual search. Similarly, Theeuwes and colleagues (Theeuwes et al., 2006; Theeuwes & van der Burg, 2007) showed that

presenting the actual stimulus as a cue before the search display (e.g., a circle that was colored red) did improve selection, a result that was also explained as the result of intertrial priming (Pinto et al., 2005).

The aforementioned analysis suggests that intertrial priming might change the selection weights so that selection becomes perfect. According to this notion, observers should be able to select the appropriate target feature when this feature happened to be presented at the previous trial. In Experiment 2, we tested this hypothesis.

Experiment 2

Experiment 2 was basically the same as Experiment 1, except that we used only two target colors (red and green). Furthermore, on half of the trials, the color of the target was repeated on the next trial, and on the other half, it switched. As in Experiment 1, observers received a verbal cue prior to the search display, which informed them about the color of the upcoming target singleton.

Method

Observers Ten students (eight females; mean age = 21.2 years, ranging from 18 to 25 years) participated in Experiment 1 as paid volunteers. Each observer received €7 an hour. All observers were naive as to the purpose of the experiment.

Stimuli, Apparatus, and Design These were identical to those in Experiment 1, except that the color singletons were always green and red.

Results

The results of Experiment 2 are presented in Fig. 2. We discarded the data from both the practice blocks and the

Fig. 2 Experiment 2. Mean correct reaction time and error rate as a function of target color (repetition vs. switch), and target-distractor congruency. The error bars represent the .95 confidence intervals for within-subjects designs, following Loftus and Masson (1994). The confidence intervals are those for the Target Color × Congruency interaction
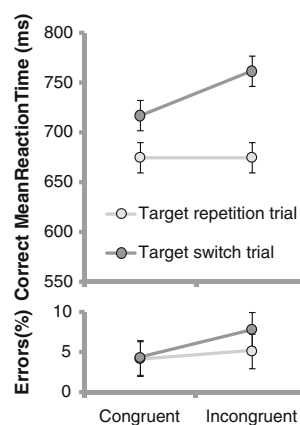
first three trials of each block. RT outliers were also discarded (0.3%). Finally, we discarded trials if responding on the preceding trial was incorrect. All data were subjected to a repeated-measures univariate ANOVA, with target color (repetition vs. switch) and congruency (congruent vs. incongruent) as within-subjects variables.

RT data Observers were significantly faster when the target color on the current trial was identical to the target color on the previous trial (675 ms) than when the target color on the current trial was different from the target color on the previous trial (739 ms), $F(1, 9) = 31.7$, $p < .001$. Furthermore, observers were faster when the orientation of the target line was identical to the orientation of the distractor line (696 ms) than when the orientations were different (718 ms), F(1, 9) = 10.7, p = .01. Importantly, this effect of target-distractor congruency was dependent on whether the target color was repeated or switched, as indicated by a reliable Target–Distractor Congruency × Target Color (switch/repeat) interaction, F(1, 9) = 11.1, p < .01. The effect of target-distractor congruency was further examined by separate two-tailed *t* tests. The *t* test revealed a reliable target–distractor congruency effect (congruent, 717 ms vs. incongruent, 761 ms) when the target color on the current trial was different from the target color on the previous trial (switch trial), $t(9) = 4.9$, $p = .001$. In contrast, no such congruency effect (congruent, 675 ms vs. incongruent, 675 ms) was observed when the target color on the previous trial was similar to the target color on the current trial (repetition trial), $t(9) = 0.008$, $p = 1$.

Error data Overall mean error rate was 5.3%. There was a trend toward a main effect of target color (switch/repeat) , F(1, 9) = 4.4, p = .06: Observers made fewer errors when the target color was repeated (4.7%) than when it switched (6.0%). There were no other effects.

Like in Experiment 1, a separate ANOVA with color (red and green) as a within-subjects variable was included to investigate the difference between the physical colors of the targets. This analysis yielded no reliable main effect of color.

Discussion

The present results show that when a color of the target is repeated on the next trial, there is no longer a congruency effect, which suggests that selection is perfect. In this condition, observers selected only the target and not the distractor item. Notably, when the target color switched, there was a robust congruency effect suggesting that observers selected the distractor before selecting the target item at least on a subset of trials. Even though in the switch

condition observers received a cue that told them with 100% certainty that the color of the target would switch on the upcoming trial, they could not make use of it and failed to show selective processing.

Note that in this experiment, in the condition in which the color of the target is repeated, there is no sign of a congruency effect, providing evidence for highly selective processing. This is important, because in a recent study, Folk, Remington, and Wu (2009) criticized our congruency manipulation that we have used in previous studies (see Schreij et al. 2008; Theeuwes, 1995; Theeuwes & Burger 1998). Folk et al. (2009) argued that one always finds a congruency effect because in our design, there is always parallel processing of the two popping-out singletons. Clearly, here is a condition in which there are two popping-out singletons, yet there is no sign of a congruency effect, which invalidates the idea that there is always parallel processing of popping-out elements (see also Schreij, Theeuwes, & Olivers, 2010, for other counter arguments).

In Experiment 1, we showed that the verbal cue did not result in selectively attending the target singleton. The results indicate that preparing in a top-down way for the upcoming color singleton was not effective in preventing attentional capture by the irrelevant color distractor. However, in one condition of Experiment 2, we found perfect selection: When the target color was repeated from one trial to the next, observers selectively attended the target singleton, as evidenced by the absence of a congruency effect. However, one may argue that Experiment 1 was not effective in creating perfect selection because the verbal cue could not be translated into an effective attentional set. To test this argument, we presented the actual color target singleton as a cue in the center of the display. Such a symbolic cue may help observers to prepare for the upcoming target. It should be noted, however, that by presenting the actual target singleton as a cue, we created a condition that also will result in priming (see Theeuwes et al., 2006; Theeuwes & van der Burg, 2007). Clearly, by looking at the cue—for example, the red circle—the color red is primed, which may give benefits in attentional selection.

Experiment 3

Experiment 3 was basically the same as Experiment 1, except that we used a symbol cue instead of a verbal cue. This symbolic cue was the actual target singleton participants had to search for in the upcoming search display. As in Experiment 1, there was never a repetition of the target or distractor color to avoid any influence of intertrial priming.

Method

Observers Twelve students (10 females; mean age = 21.0 years, ranging from 19 to 24 years) participated in Experiment 3 as paid volunteers. Each observer received €7 an hour. All observers were naive as to the purpose of the experiment.

Stimuli, Apparatus, and Design These were identical to those in Experiment 1, except that we replaced the word cue by a symbol cue, which was always identical (i.e., same size, color, and luminance) as the target color.

Results

We discarded the data from both the practice blocks and the first three trials of each block. RT outliers were also discarded (0.4%). We determined whether there were effects of congruency on RTs and errors.

RT data Even though numerically there appears to be a congruence effect (735 vs. 749 ms), the effect was not statistically reliable, $t(11) = 1.0$; $p = .3$. The absence of the congruency effect indicates that observers were able to selectively attend the target singleton indicated by the cue, without interference from the distractor singleton

Error data Overall mean error rate was 3.9%. In the congruent condition, the error rate was 3.4%; in the incongruent condition, it was 4.3%. This difference was not reliable.

Discussion

The present findings indicate that when there is no intertrial priming from one trial to the next, symbolic cuing is effective. Even though there was a small congruency effect, this turned out not to be reliable. On the basis of these data, one may argue that the verbal cue (Experiment 1) was not effective, whereas the symbolic cue (Experiment 3) was effective in guiding attention in a top-down way. However, the results may also be explained in terms of feature priming (Theeuwes, 2010a, 2010b). Previous studies have shown that presenting the object of search (in this case, a color singleton) as a cue in the center of the display primes the features of the target, resulting in a benefit in selecting the target singleton (e.g., Theeuwes & van der Burg, 2007; Theeuwes et al., 2006).

In the next experiment, we examined the effect of a centrally presented symbolic cue with conditions in which there was also intertrial (trial to trial) priming. We repeated

Experiment 2, but instead of a verbal cue, we used a symbolic cue that displayed the color target singleton that observers had to search for. Note that on half of the trials, the target color was repeated, and on the other half, the target color was switched from red to green, or vice versa. If processing the symbolic cue would result in efficient selection (as Experiment 3 suggests), then we should find no compatibility effect in either the condition in which there is intertrial priming (the target color is repeated from the next trial) or in the condition in which the color switches from one trial to the next.

Experiment 4

Experiment 4 was basically the same as Experiment 3, except we used two target colors (red and green). Furthermore, on half of the trials, the color of the target was repeated on the next trial, and on the other half, it switched. As in Experiment 3, observers received a symbol cue prior to the search display, which informed them about the color of the target singleton.

Method

Observers Nine students (six females; mean age = 21.2 years, ranging from 18 to 25 years) participated in Experiment 4 as paid volunteers. Each observer received €7 an hour. All observers were naive as to the purpose of the experiment. One particpant was excluded from further analyses because of an overall high error rate (~15%).

Stimuli, Apparatus, and Design These were identical to those in Experiment 3, except that the target color was always green or red.

Results

The results of Experiment 4 are presented in Fig. 3. We discarded the data from both the practice blocks and the first three trials of each block. There were no RT outliers. Finally, we discarded trials if responding on the preceding trial was incorrect. All data were subjected to a repeated-measures univariate ANOVA with target color (repetition vs. switch) and congruency (congruent vs. incongruent) as within-subjects variables.

RT data Observers were significantly faster when the target color on the current trial was identical to the target color on the previous trial (594 ms) than when the target color on the current trial was different from the target color on the previous trial (622 ms), $F(1, 8) = 12.3$, $p < .01$.

Fig. 3 Experiment 4. Mean correct reaction time and error rate as a function of target color (repetition vs. switch), and target–distractor congruency. The error bars represent the .95 confidence intervals for within-subjects designs, following Loftus and Masson (1994). The confidence intervals are those for the Target Color × Congruency interaction
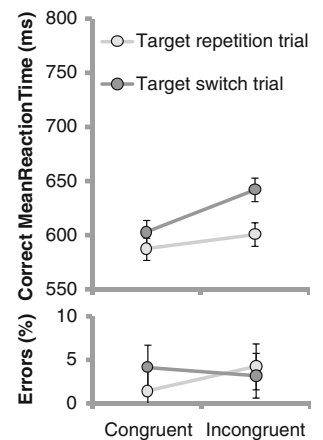
Furthermore, observers were faster when the orientation of the target line was identical to the orientation of the distractor line (595 ms) than when the orientations were different (621 ms), $F(1, 8) = 6.4$, $p < .05$. Importantly, this effect of target–distractor congruency was dependent on whether the target color was repeated or switched, as indicated by a reliable Target-Distractor Congruency × Target Color (switch/repeat) interaction, $F(1, 8) = 7.8$, $p < .05$. The effect of target–distractor congruency was further examined by separate two-tailed t tests. The t test revealed a reliable target–distractor congruency effect (congruent, 602 ms vs. incongruent, 642 ms) when the target color on the current trial was different from the target color on the previous trial (switch trial), $t(8) = 3.3$, $p = .01$. In contrast, no such congruency effect (congruent, 587 ms vs. incongruent, 601 ms) was observed when the target color on the previous trial was similar to the target color on the current trial (repetition trial), $t(8) = 1.2$, $p = .26$.

Error data Overall mean error rate was 3.2%. The ANOVA on errors yielded no reliable effects, all Fs < 2.7, ps > .14.

Discussion

The results of the present study show that intertrial priming (the color that was processed as a target on the previous trial) has a stronger impact than the color of the symbolic cue that was presented just before the search display. Unlike in Experiment 3, the symbolic cue did not result in efficient selection: When the cue indicated a color that was different than the color that observers processed on the previous trial, there was a congruency effect suggesting attentional capture by the irrelevant singleton. If we assume that the cuing effect, as found in Experiment 3, was due to cue–target priming, then the present results would indicate that intertrial priming is much stronger than cue–target priming. Note that because of the design of the experiments, which always uses a 100% valid cue (the cue is needed to indicate

which color singleton to respond to) we are not able to assess the independent contributions of cue-to-target priming and intertrial priming.

Experiment 5

Experiment 5 served as a control experiment to assess the nature of the congruency effect. One may claim that in all previous experiments, attention may have gone directly to the target singleton (suggesting perfect selection), and that after selecting the target color singleton, but before responding, observers attended to the distractor singleton, which then may also have resulted in a congruency effect. If this were the case, it would imply that our congruency manipulation would not be an adequate diagnostic for selection. To determine whether this is a viable possibility, we presented the displays very briefly, followed by a mask. If the congruency effect is indeed the result of selecting the distractor singleton *after* selecting the target singleton, one does not expect a congruency effect. Indeed, with very brief displays, one expects that a second shift of attention (i.e., to the distractor singleton) would not be possible and therefore that one would not expect to see a congruency effect. One the other hand, if attention is always captured immediately by the distractor singleton before attending the target, then one would expect this to happen also when the display is presented relatively briefly.

Experiment 5 was identical to Experiment 2, except that we varied the display duration. Furthermore, instead of measuring RT, we asked participants to make an unspeeded response toward the orientation of the line segment located within the color target singleton.

Method

Observers Six observers (three females; mean age = 21.3 years, ranging from 19 to 26 years) participated in Experiment 5 as paid volunteers. Each observer received €7 an hour. All observers were naive as to the purpose of the experiment.

Stimuli, Apparatus, and Design These were identical to Experiment 2, except the display duration was varied (8, 16, 24, 32, 48, 56, 64, and 300 ms). The search display was immediately followed by a mask until particpants made an unspeeded response toward the orientation of the line segment in the target. The mask consisted of 16 white line segments (length 1.2˚; 29.3 cd/m^2) of various orientations (including a horizontal and vertical line), which would appear over the line segments. Observers practiced one block of 128 trials. Then, they completed 12 experimental blocks of 128 trials each.

Results

The results of Experiment 5 are presented in Figs. 4 and 5. We discarded the data from both the practice blocks and the first trial of each block. Finally, we discarded trials when an error was committed on the previous trial. All data were subjected to a repeated-measures univariate ANOVA with target color (repetition vs. switch), congruency (congruent vs. incongruent), and display duration as within-ubjects variables.

Proportion correct Overall proportion correct was .839. The ANOVA yielded a reliable effect of display duration: Partcipants performed better when the display duration increased, $F(8, 40) = 28.4$, $p < .001$. Furthermore, observers were better when the target color on the current trial was identical to the target color on the previous trial (.855) than when the target color on the current trial was different from the target color on the previous trial (.820), $F(1, 5) = 6.4$, $p = .052$. Importantly, the ANOVA yielded a reliable Target–Distractor Congruency × Target color (switch/repeat) interaction, $F(1, 5) = 14.2$, $p = .013$. The effect of target–distractor congruency was further examined by separate two-tailed t tests. Figure 5 illustrates the overall mean proportion correct as a function of target color (repetition vs. switch) and congruency (collapsed over display duration).

As is clear from Fig. 5, the effect of a target–distractor congruency effect (congruent, .857 vs. incongruent, .856) was not reliable when the target color on the current trial was identical to the target color on the previous trial (repetition trial), $p = .99$. In contrast, the t test revealed a reliable target–distractor congruency effect (congruent, .842 vs. incongruent, .798) when the target color on the current trial was different from the target color on the previous trial
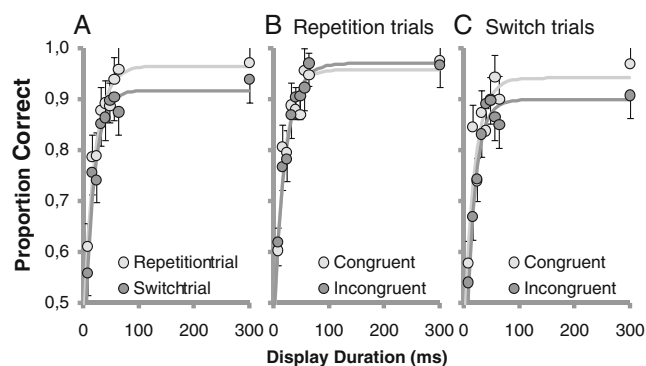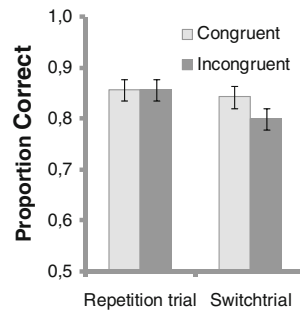

Fig. 4 Results from Experiment 5. **a** Proportion correct as a function of display duration and target color (repetition vs. switch). **b** Proportion correct as a function of display duration and congruency for target color repetition trials. **c** Proportion correct as a function of display duration and congruency for target color switch trials. The error bars represent the .95 confidence intervals for within-subjects designs, following Loftus and Masson (1994)

(switch trial), $t(5) = 2.7$, $p < .05$. Importantly, for the
purpose of the present experiment, the three-way interaction
among target color, target–distractor congruency, and
display duration was not relaible, $F < 1$, indicating that
the observed effects were independent on the presentation
time of the search display. None of the other effects were
reliable.

Discussion

The results of the present experiment show a congruency
effect on accuracy for trials in which the color to search for
switched, but not for trials in which the color stayed the
same. This effect basically replicates the results of
Experiment 2, with accuracy as a measure. More impor-
tantly, this effect was not modulated by display duration,
indicating that capture occurred regardless of the display
duration.[1] On the basis of these findings, one can argue that
the congruency effect in all previous experiments is not the
result of shifting attention to the distractor singleton *after*
selecting the target, but instead reflects the mandatory
processing of the distractor singleton *before* attention is
switched to the target singleton. Consistent with previous
studies, our identity intrusion manipulation represents a true
measure of attentional capture (Gibson, & Bryant, 2008;
Schreij et al., 2008; Schreij et al., 2010; Theeuwes, 1995;
Theeuwes & Burger 1998).

General Discussion

We created a very simple task in which observers had to
selectively attend to either one of two possible colored pop-
out singletons. Even though observers had to process the
cue in order to know what they had to respond to and had
enough time to selectively prepare for the color of the

upcoming target singleton (indicated by the cue), they were
not able to selectively search for the target singleton. The
congruency analysis revealed that attention was captured by
the irrelevant distractor singleton. This study demonstrates
the limits of selecting on the basis of a top-down set.
Because the two singletons present in the display were
about equally salient (there was no reliable difference in the
absolute RTs between the various colors in experiments),
the smallest increase in top-down weight on the target color
should have tipped the weight in favor of the target and
should have selectively biased attention to the target
singleton only. Our results show that this top-down bias
was not able to prevent the processing of the irrelevant
distractor singleton.

Crucially, Experiments 2, 4, and 5 present conditions in
which perfect selection was obtained: Only if the target
color was repeated from one trial to the next, did observers
selectively attend the target singleton only. The absence of a
congruency effect indicates that the irrelevant color single-
ton did not summon attention and was completely ignored.
Note that it is unlikely that this selective processing is the
result of top-down set because the effect occurred only in
conditions in which the color of the target was repeated on
the next trial. Consistent with earlier findings (Belopolsky,
Schreij, & Theeuwes, 2010; Pinto et al., 2005; Theeuwes et
al., 2006), we attribute this effect to automatic bottom-up
intertial priming. The processing of the color singleton on
a given trial drives the processing of that very same color
on the next trial. It is clear that top-down set as established
by the cue cannot result in selective processing; however,
bottom-up priming can.

Experiment 3, in which we prevented intertrial priming
by using a new color on every trial, showed that the
presentation of the actual target singleton as a cue also
resulted in selective processing. We attribute this effect to
cue to target priming. A previous study (Theeuwes & van
der Burg, 2007) that looked at cross-dimensional cuing
showed that this type of cuing had an effect on perceptual
sensitivity (measured in A-prime), regardless of whether the
cue was highly valid (83% validity) or whether the
cue predicted the target singleton at change level (50%).
Theeuwes et al. (2006) pushed this effect even further and
made the cue counter-predictive. For example, when a red
circle was shown as a cue, there was a high chance (83%)
that the target would be a green diamond. In other words, a
red circle as a cue indicated that observers should prepare
for a green diamond because in the majority of trials, a
green diamond was the target. The results indicated that
when a red circle served as cue even when it was counter-
predictive, it still had an effect on RT so that observers were
faster when the "unlikely" red (but primed) target singleton
was presented than when the "likely" green diamond was
presented as a target. The same was true for the reverse

---

[1] We also analyzed whether the congruency effect in Experiment 2
only occurred for slow responses. A median split analysis in which we
examined the congruency effect for slow and fast responses showed
that the critical interaction "speed of reponse" (fast vs. slow repsonses) ×
"congruency" was not reliable ($F(1, 9) = 0.2$, p = .65) indicating that the
observed congruency effect was present across the whole range of fast
and slow responses.

(diamond cue, red circle target). In fact, the RT benefits resulting from the cue were the same regardless of whether cue was counter-predictive or highly predictive. These findings strongly suggest that the effect of a symbolic cue showing the target to search for on the upcoming trial is not the result of top-down volitional control but is the result of bottom-up priming.

The results of Experiment 4 show that the cue–target priming effect cannot overcome the priming effect occurring between trials. The data suggest that after processing say, for example, a red target singleton, observers cannot prevent the selection of a red distractor on the next trial even when the symbolic cue shows a green circle. The intertrial priming is particularly strong (and stronger than cue–target priming) because the target singleton color on one trial becomes the distractor singleton color on the next. Previous research has shown that this type of priming (where the target on one trial becomes the distractor on the next) generates strong capture effects (Hickey, Chelazzi, & Theeuwes, 2010; Pinto et al. 2005; Theeuwes, 1991)

The present findings indicate that top-down guidance toward the target singleton is limited and does not always lead to selectively attending the target singleton only, a conclusion that is inconsistent with theories of visual search that assume a large role for top-down control. For example, the contingent capture hypothesis predicts that observers can set themselves on each trial for a particular color (Folk et al., 1992). Other theories argue that visual selection is biased toward an attentional template that ensures that objects that match the template are selected over those that do not match the template (Duncan & Humphreys, 1989). Guided search argues that a top-down set can increase the salience of the relevant feature dimension (in this example the feature "red" or "green") so that attention is (mostly) guided to relevant features only (Wolfe, 1994; Wolfe et al., 2003). Note, however, that most of these top-down attentional theories (including GS and DW) argue that top-down selection is never perfect (e.g., Müller et al., 2003; Wolfe, 1994) and that exceptions may occur.

However, the contingent capture account argues that selection is always fully contingent on the top-down control setting (Folk et al., 1992). More specifically, according to this view, selection always depends on the explicit or implicit goals held by the observer at any given time. Clearly, because the top-down set (what to look for) was made explicit by the verbal cue on each trial, according to contingent capture, top-down control should have been evident. For example, in a recent study, Lien et al. (2010) used a version of the spatial cuing paradigm of Folk et al. (1992) and also verbally cued the color of the upcoming target singleton (i.e., the letter "R" for red and "G" for green) at the center of the display. Their Experiment 3, in

which the color to search varied randomly from trial to trial, showed a cue validity effect even when the color to look for switched from one trial to the next. This led the authors to conclude that top-down set could be applied in a highly flexible way. It is clear that this conclusion is opposite of ours. However, it should be noted that there are many differences between the Lien et al. study and ours. For one, the Lien et al. study is a spatial cuing paradigm, and attentional selectivity is deduced on the basis of cue validity effects. It is feasible that the crucial validity effect is primarily driven by the invalid cue condition resulting from trouble in disengaging attention from the invalid location that has the same color as the target on the previous trial (see also Belopolsky et al., 2010; and Theeuwes, 2010b, for a similar argument).

Also, we used seven display elements, whereas Lien et al. (2010) used only four display elements (as is typical for contingent capture). A recent study by Yeh and Liao (2008) showed that contingent capture no longer holds when one runs the classic Folk et al. (1992) spatial cuing task with eight elements. Yeh and Hiao argued that with more nontarget elements, the target and distractor singleton become more salient (see also Theeuwes, 2004, for a similar argument). This implies that in the Lien et al. (2010) study with only four elements, of which two were considered "singletons" (e.g., a red and a green item among two gray items), the singletons may not have been salient enough for pop-out search, causing selection to be slow and serial. When there are no salient elements in the display, there is a lot of room for top-down serial guidance of attention, as, for example, was shown for conjunction search (Kaptein, Theeuwes, & van der Heijden, 1995). As such, the Lien et al. (2010) study may not be comparable to the present study in which two color singleton clearly popped out from the background.

Our findings are consistent with theories that assume a large role for bottom-up priming in visual search (e.g., Kristjánsson & Campana, 2010; Kristjánsson, Wang, & Nakayama, 2002; Maljkovic & Nakayama, 1994; Theeuwes et al., 2006; Theeuwes & van der Burg, 2008). Our results show that observers selectively searched for the color that was the target color on the previous trial, irrespective of their top-down set. Only this condition resulted in perfect selective search. We consider intertrial priming the result of passive automatic processing, which is impervious to prior knowledge and/or top-down processing (see Theeuwes, 2010a, 2010b for a detailed discussion of this account).

Consistent with this idea, Belopolsky et al. (2010) provided convincing evidence that the notion of contingent capture as advocated by Folk et al. (1992) may not represent an example of top-down control but simply may be the result of intertrial priming. Belopolsky et al. used

exactly the same spatial cuing paradigm as Folk et al. (1992), but instead of keeping the target fixed over a whole block of trials (as is done in the original contingent capture experiments), observers had to adopt a top-down set before the start of each trial. In other words, observers were cued at the beginning of each trial to look for either a unique color or the unique onset. Contrary to the typical Folk et al. contingent capture results, Belopolsky et al. showed that, irrespective to the cue, both relevant and irrelevant cue properties captured attention. Instead of top-down contingent capture, they showed that the cue of the premask display (irrelevant or irrelevant) that matched the target on the previous trial captured attention, a result that can easily be explained in terms of intertrial priming.

One final question is how interpriming changes the selection priority. As outlined in Theeuwes (2010a, 2010b), the idea is that priming may change the salience of the feature within the priority map. In this sense, salience is not necessarily a physical characteristic of a stimulus as it exists in the outside world. Instead, prior history—as that, for example, found in intertrial priming—may change the salience of a stimulus. For example, Desimone (1996) suggested that repeated processing of a stimulus produces a "sharpening" of its cortical representation, possibly making it more salient within its environment. Bichot and Schall (2002) showed that repeating a stimulus changes responses of neurons in the frontal eye field, a region that has been implicated to be the neural substrate of the salience map (Thompson & Bichot, 2005). Our notion is that salience is not solely defined by the physical appearance of a stimulus in the outside world, but depends on its representation in the priority map. In this view, the processing of a stimulus (as in intertrial priming) leads to a change in the representation of that stimulus in the priority map, and this change occurs independently of top-down intentions. This type of selection is considered to be automatic and passive. A recent finding of Olivers and Hickey (2010) provides converging evidence for this notion. They showed that intertrial priming results in latency shifts and amplitude differences in the P1 component of the ERP signal—a signal that is seen 80 to 130 ms following display onset. Since priming affects visual processing so early, it is unlikely to be the result of top-down processing. These findings also make it very improbable that priming is an example of implicit top-down guidance as, for example, was suggested by Wolfe et al. (2003). Because the effects occur so early in time and cannot be counteracted by strategy in a volitional way, it is more appropriate to label priming as automatic and bottom-up in nature (see Theeuwes, 2010a, 2010b, for a detailed discussion).

In summary, the results of the present study demonstrate that when two equally salient singletons are simultaneously present in the visual field, top-down attentional set cannot change the attentional weights in such a way that observers exclusively attend the target singleton, without interference from the distractor singleton. These attentional weights can be altered only by automatic intertrial priming, a process that is not under volitional top-down control.

References

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research, 49,* 1154–1165.

Belopolsky, A., Schreij, D., & Theeuwes, J. (2010). What is top-down about contingent capture? *Attention, Perception, & Psychophysics, 72,* 326–341.

Bichot, N. P., & Schall, J. D. (2002). Priming in macaque frontal cortex during popout visual search: Feature-based facilitation and location-based inhibition of return. *Journal of Neuroscience, 22,* 4675–4685.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences, 93,* 13494–13499.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

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

Folk, C. L., Remington, R. W., & Jonhnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044.

Folk, C. L., Remington, R. W., & Wu, S. C. (2009). Additivity of abrupt onset effects supports nonspatial distraction, not the capture of spatial attention. *Attention, Perception, & Psychophysics, 71,* 308–313.

Geyer, T., & Müller, H. J. (2009). Distinct, but top-down modulable color and positional priming mechanisms in visual pop-out search. *Psychological Research, 73,* 167–166.

Gibson, B. S., & Bryant, T. A. (2008). The identity intrusion effect: Attentional capture or perceptual load? *Visual Cognition, 16,* 182–199.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience, 30,* 11096–11103.

Kaptein, N. A., Theeuwes, J., & van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1053–1069.

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of priming in visual search. *Attention, Perception, & Psychophysics, 72,* 5–18.

Kristjánsson, Á., Wang, D., & Nakayama, K. (2002). The role of priming in conjunctive visual search. *Cognition, 85,* 37–52.

Lien, M.-C., Ruthruff, E., & Johnston, J. C. (2010). Attention capture with rapidly changing attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1–16.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology, 33,* 64–87.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. *Role of features. Memory & Cognition, 22,* 657–672.

Müller, H. J., & Krummenacher, J. (2006). Locus of dimension weighting: Pre-attentive or post-selective? *Visual Cognition, 14,* 490–513.

Müller, H. J., Reimann, B., & Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: Stimulus and expectancy-driven effects in dimensional weighing. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1021–1035.

Olivers, C. N. L., & Hickey, C. (2010). Priming resolves perceptual ambiguity in visual search: Evidence from behaviour and electrophysiology. *Vision Research, 50,* 1362–1371.

Pinto, Y., Olivers, C. N. L., & Theeuwes, J. (2005). Target uncertainty does not lead to more distraction by singletons: Intertrial priming does. *Perception & Psychophysics, 67,* 1354–1361.

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Perception & Psychophysics, 70,* 208–218.

Schreij, D., Theeuwes, J., & Olivers, C. N. L. (2010). Abrupt onsets capture attention independent of top-down control settings II: Additivity is no evidence for filtering. *Attention, Perception, & Psychophysics, 72,* 672–682.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics, 50,* 184–193.

Theeuwes, J. (1995). Perceptual selectivity for color and form: On the nature of the interference effect. In A. F. Kramer, M. G. H. Coles,

& G. D. Logan (Eds.), *Converging operations in the study of visual attention* (pp. 297–314). Washington, DC: American Psychological Association.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11,* 65–70.

Theeuwes, J. (2010a). Top-down and bottom-up control of visual selection. *Acta Psychologica, 123,* 77–99.

Theeuwes, J. (2010b). Top-down and bottom-up control of visual selection: Reply to commentaries. *Acta Psychologica, 123,* 133–139.

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology-Human Perception and Performance, 24*(5), 1342–1353.

Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition, 14,* 466–489.

Theeuwes, J., & van der Burg, E. (2007). The role of spatial and nonspatial information in visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1335–1351.

Theeuwes, J., & van der Burg, E. (2008). The role of cueing in attentional capture. *Visual Cognition, 16,* 232–247.

Theeuwes, J., Van Der Burg, E., & Belopolsky, A. (2008). Detecting the presence of a singleton involves focal attention. *Psychonomic Bulletin & Review, 15,* 555–560.

Thompson, K. G., & Bichot, N. P. (2005). A visual salience map in the primate frontal eye field. *Progress in Brain Research, 147,* 251–262.

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

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

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 singeltons. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 483–502.

Yeh, S. L., & Liao, H. I. (2008). On the generality of the contingent orienting hypothesis. *Acta Psychologica, 129,* 157–165.